Received: 15 January 2021 | Revised: 30 March 2021 | Accepted: 10 May 2021

DOI: 10.1002/cpe.6426

WILEY

SPECIAL ISSUE PAPER

Chatbots: Security, privacy, data protection, and social aspects

Martin Hasal[1] | Jana Nowaková[2] | Khalifa Ahmed Saghair[2] | Hussam Abdulla[2] | Václav Snášel[2] | Lidia Ogiela[3]

[1] IT4Innovations, VSB—Technical University of Ostrava, Ostrava, Czech Republic

[2] Department of Computer Science, Faculty of Electrical Engineering and Computer Science, VSB—Technical University of Ostrava, Ostrava, Czech Republic

[3] Pedagogical University of Krakow, Krakow, Poland

Correspondence
Jana Nowaková, VSB-Technical University of Ostrava, 17. listopadu 2172/15, 708 00 Ostrava—Poruba, Czech Republic.
Email: jana.nowakova@vsb.cz

Funding information
European Regional Development Fund, Grant/Award Number: CZ.02.1.01/0.0/0.0/15_003/0000466; Technology Agency of the Czech Republic, Grant/Award Number: TN01000024; Vysoká Škola Bánská - Technická Univerzita Ostrava, Grant/Award Number: SP2021/24

Summary

Chatbots are artificial communication systems becoming increasingly popular and not all their security questions are clearly solved. People use chatbots for assistance in shopping, bank communication, meal delivery, healthcare, cars, and many other actions. However, it brings an additional security risk and creates serious security challenges which have to be handled. Understanding the underlying problems requires defining the crucial steps in the techniques used to design chatbots related to security. There are many factors increasing security threats and vulnerabilities. All of them are comprehensively studied, and security practices to decrease security weaknesses are presented. Modern chatbots are no longer rule-based models, but they employ modern natural language and machine learning techniques. Such techniques learn from a conversation, which can contain personal information. The paper discusses circumstances under which such data can be used and how chatbots treat them. Many chatbots operate on a social/messaging platform, which has their terms and conditions about data. The paper aims to present a comprehensive study of security aspects in communication with chatbots. The article could open a discussion and highlight the problems of data storage and usage obtained from the communication user—chatbot and propose some standards to protect the user.

KEYWORDS

chat, chatbots, data protection, GDPR, security, virtual assistants

1 | INTRODUCTION

A chatbot, short for chat robot, or bot[*], short for chat robot, is an application that interacts with humans through user command. The interaction is done mostly through voice and text conversations. It is designed to replicate the pattern of human interaction, thus allowing human conversation with machines.[1] In 1966, Joseph Weizenbaum developed a computer program that demonstrated the possibilities of communication between a human and a computer via natural language, and, as such, he introduced the first generation of chatbots.[2,3] A chatbot is supposed to work independently of the human operator at its end. It tries to understand the queries and provide appropriate responses. If, at any point, the conversation crosses its current knowledge, it deflects the conversation or passes it to human operators. However, modern chatbots are simultaneously learning from the conversation using machine learning algorithms to provide better answers in the future. Studies show that humans have less experience in

[*] A bot is sometimes referred to as a chatbot, but to be precise, a bot is a computer program (tool) that automates processes. A chatbot is a sub-genre of the bot environment with a focus on talking or conversation. Some companies instead of Chatbot use the name 'Conversational AI' or 'AI chatbots' to highlight that their chatbot is powered by machine learning and information retrieval techniques. In this article the term 'chatbot' is used for all types of chatbots.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2021 The Authors. *Concurrency and Computation: Practice and Experience* published by John Wiley & Sons Ltd.

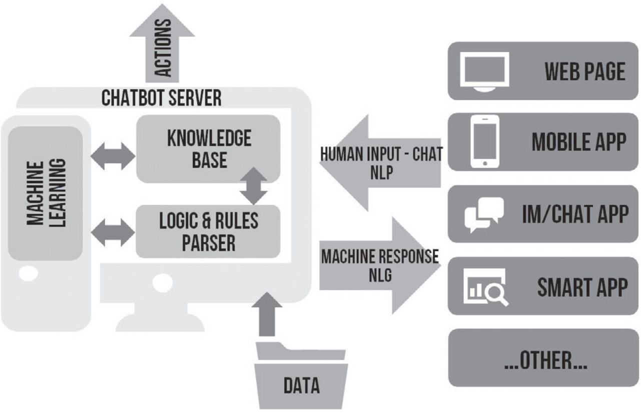

FIGURE 1 The structure of chatbot

interacting with chatbots than with other humans.[4] Therefore, the comfort level for the user is lower. Chatbots still face unpopularity among users; users often close the conversation when they find out that they are not talking with a real person. Modern chatbots, however, are built upon conversational data from multiple sources. It makes the conversation more natural for the person interacting with the chatbot, which can use typical typos, such as switched letters, and so forth. In essence, chatbots are nothing more than a programmed input-output system, where the output or input is presented in a pleasing way through natural language in the written or spoken form. Hence, many companies recognized that chatbots are a great tool for improving customer relationships as they can operate on many platforms simultaneously, see Figure 1. Nowadays, many customers use digital communication channels and they appreciate a chatbot's advantages, such as 24/7 customer service, personalized interaction, and no waiting time. For companies, chatbots essentially mean cost savings as many processes are repeated and can be automated, and employees can be dedicated to more complex tasks. Large technology companies such as Google, Facebook, and IBM have contributed to chatbot development and research in recent years. Naturally, they have access to a large amount of data on which chatbots can be trained. One of the most well-known examples are the question answering system named Watson by IBM, Apple's Siri, Amazon's Alexa, Microsoft's Cortana, and Google's recent chatbot Meena.[5-7] As a matter of fact, all systems are rather open-domain chatbots,[8] which can engage in conversation on any topic and understand text and speech. The primary goal of these companies is to develop voice assistants[5] based on natural language processing (NLP), information retrieval techniques, and machine learning (ML) methods. Users can ask their assistants to control home automation devices, play back media via voice, manage other basic tasks such as email, to-do lists, and calendars with verbal commands.[9] The options of voice assistants increase everyday.[7]

Contrary to popular belief, many chatbots used by online stores, banks, industrial technical support, and so forth, are created by small or medium-sized companies (compared to the companies mentioned above, which also provide services to create closed-domain chatbots). Such chatbots are called closed-domain chatbots. Usually, they correspond to (or search for) keywords or intents to accomplish specific tasks. Technical solutions vary from company to company, but they generally use the client's knowledge, such as a webpage, manuals, FAQs, documents, human–human transcripts, and so forth; and use this knowledge to create chatbots specific to tasks suited to the given company. The differentiation that big companies create open-domain chatbots and small companies create closed-domain chatbots is not exact. There are companies or scientific teams creating open-domain chatbots,[10] but the creation of human-competitive chatbots requires a lot of data and experienced programmers. The advantage is that both types of chatbots share some methods and technical solutions,[10] which can be used by people creating a knowledge base for other languages,[11,12] and to get to level five in AI assistants[†].

Generally, all chatbots are designed to reduce the necessity of humans in classic conversation, for example, in e-shops, banking, medical advice, GDPR rules, psychotherapy,[13] and so forth. Nowadays, the chatbots also have been shown to be very useful and time-saving in situations such as hotlines during the coronavirus (COVID-19) pandemic.[14-16] On the other hand, pandemic situations make people less cautious about sharing personal data and revealing their fears.[17] Chatbots work with user data; chatbots even 'learn' from these data. This means they could be considered to be a security problem. Users mostly do not know how their sensitive personal identifying information (PII) is treated, used, stored, or even shared. Generally, it is not a good situation and the new General Data Protection Regulation, well known as GDPR[18‡], also raises the problem of security. The main aim of the current study is to highlight and discuss all the important security, privacy, data protection and social aspects in the usage of chatbots through the systematic review of the available literature, and to present a comprehensive view to the given problem. Further, the study indicates challenges in security issues and propose ways to minimize security problems that appear with chatbots' usage.

The concept of this paper is to provide comprehensive insight into the state-of-the-art of chatbots security issues and challenges. The next section, a brief history of chatbots, presents important steps in chatbots evolution. It points out the steps related to the present topic. Section 3 gives an overview of the potential security issues and threats and describes responsive steps. The comprehensive part of Section 3 covers the problematic

[†] https://www.oreilly.com/radar/the-next-generation-of-ai-assistants-in-enterprise/.
[‡] https://ec.europa.eu/info/law/law-topic/data-protection_en.

of data storage obtained during communication, which is closely related to the GDPR rules. The summary and identification of critical issues to the GDPR rules regarding chatbots are also discussed. Section 4 provides a discussion about the explored findings and social aspects and extends the security topic to the chatbots' messaging platforms. Finally, in the last section, the entire study description, limitations and outline of future work are concluded with the authors' suggestions.

2 | CHATBOTS IN HISTORY

The history of chatbots or social bots (also: socialbots or socbots, agents that communicate more or less autonomously on social media[19]) concerns many ideas and technological steps. The whole history of the chatbots[10,20] is not presented, but crucial chatbots or steps in development with respect to security and the knowledge base are mentioned. The discussion can be either by chatting by typing text or speech dialogue using a voice; or some chatbots can have a visual avatar[§]which adds facial expressions to the conversation. The processing of the information in both techniques is the same; speech is converted to text in the case of spoken dialogue; textual information is further analyzed.

The first chatbot (called ELIZA[2]) was developed in 1966 at the Massachusetts Institute of Technology (MIT). It answered some very simple decision tree questions. A decision tree is a decision support tool that uses a tree-like model of decisions and their possible consequences.[21] The gross procedure program of ELIZA and many following chatbots is the following: It reads and inspects the text for the presence of keywords. If such a word is found, the sentence is transformed according to a rule associated with the keyword; if not, a content-free remark or, under certain conditions, earlier transformation is retrieved. After that, the answer (based on computed or retrieved text and decision tree probability) is printed out. Finding the right keyword is not an easy task. In detail, sentences can contain more than one keyword in a different form or position in a sentence. Such a situation requires techniques like parsing, pattern matching, AIML (Artificial Intelligence Markup Language), chat scripts, or even language tricks.[20] Such techniques were developed and improved in recent decades.[20]

The chatbot A.L.I.C.E. (the Artificial Linguistic Internet Computer Entity) introduced AIML,[22,23] which is a derivative of XML. AIML technology is responsible for pattern matching to relate a user input with a response of the chatbots's Knowledge Base (KB).[24,25] A.L.I.C.E. used AIML objects to simplify conversational modelling, in relation to a 'stimulus-response' process. The primary elements of an AIML object (input sentence) are categories, patterns, and templates. For each category a response template and set of conditions are prepared that give meaning to the template known as context.[24] A.L.I.C.E. preprocesses AIML objects (performs word processing actions to fit the user's input to a pre-established format) and matches the input in a KB by pattern matching. If the input is matched, A.L.I.C.E. responds or performs a predefined action. One of the most widely used standalone human-like chatbots based on AIML is Mitsuku.[26] It applies NLP with heuristic patterns and supervised ML. Mitsuku is a multi-lingual chatbot, but the new data are sent to the human manager for verification, and only accepted data are involved in KB. Mitsuku needs a large amount of training data to lead an effective dialog.[10]

The creation of a knowledge base by combining the attributes of two other chatbots is also possible.[27] The authors created the Just.Chat platform combining attributes of two other Chatbots. It processed the knowledge bases using three filters to—discard overlapping interactions, identify personal questions, and deal with interactions containing unwanted terms or topics. The present method was able to produce a chat corpus with around 7800 pairs of interactions in total. It helped to improve chatbot design techniques.

The Cleverbot[28] chatbot implements rule-based AI techniques.[29] Cleverbot is capable of having a natural conversation with a human, whereas programmers do not predefine its responses. Its KB is based on conversational exchanges with people online through crowdsourcing. It means that Cleverbot's responses are based on the user's inputs and feedback. Cleverbot can learn from historical conversations,[4] so its responses can vary over time. Cleverbot has a few disadvantages, such as a sudden change of language (e.g., from English to Italy, if a user writes an Italian phrase), a change in the discussed topic, unpredictable responses and no ability to have a longer discussion on the given topic. To chat with Cleverbot[¶], the user must confirm they are over 18 years old since Cleverbot learns many phrases from its users.

Most of the AI chatbots are based on a Deep Learning leverage sequence to sequence (Seq2Seq) model.[30] The Seq2Seq model showed great results in machine translations of other NLP related fields.[31] The general strategy is to map the input sequence to a fixed-sized vector using one Recurrent Neural Network (RNN) and subsequently mapping the vector to the target sequence with another RNN. Most NLP is now based on language models produced by machine learning RNN, as the advantage is that RNN is able to effectively use data from previous steps[32] and to model sequence data. An RNN treats each word of a sentence as a separate input occurring at time 't' and also uses the activation value at 't−1' as an input in addition to the input at time 't'. Consequently, RNN predicts the next word in a sentence with more accuracy, based on what the previous words were. In short, chatbots can predict an outcome (give an answer) based on historical conversations with similar previous texts—text generation. With an appropriate dataset, which is suitable for modelling long and open-domain dialogues close to spoken human language, open-domain conversational dialogue systems can be created.[33] The more conversational data are available on a given topic, the better the chatbot is in this topic, but it fails in others. This is why the past conversations from any sources are so valuable. For instance, many companies use transcripts of human workers and

[§] https://www.eviebot.com/en/.

[¶] https://cleverbot.com.

promise that they can 'automate' it once they've collected enough data. For now, this is only possible in specific areas—like chat interfaces on booking web-pages. On the other hand, systems based on RNN can help to assist human workers by proposing and correcting responses. Open-domain chatbots can ask, for example, the same question several times and subsequently evaluate the best answer.

IBM started the DeepQA project to create a rule-based AI chatbot called Watson.[34] Watson is used as an information-retrieval and question-answering system. It is based on NLP and hierarchical ML method. It finds and assigns feature values (values, names, dates, etc.) to generate responses based on the score. Watson's cognitive computing technology involves dealing with complex and unstructured data, finding patterns, processing text on technologies such as Hadoop and the Apache Unstructured Information Management Architecture (UIMA) framework. Watson has almost limitless potential, but the complex technology makes it difficult to use without significant effort.

Microsoft also created a few chatbots/assistants, such as LUIS,[35] Cortana,[36] Tay,[37] Zo, Xiaoice, Rinna, and so forth; here only LUIS and Cortana are mentioned. LUIS is the acronym for Language Understanding Information Service. LUIS uses NLP in big data analyses to find intents from a sentence. As soon as LUIS recognizes the user's intents, the user supplies example phrases called utterances for the intents. LUIS requires a Microsoft Azure subscription, as it is integrated with the Azure Bot Service. This makes it easy to create a chatbot which is connected to Azure pre-built apps and data lakes. Cortana is profiled as Microsoft's personal productivity assistant, which recognizes natural language and processes all the users' information, interests, routines, calendars, searches, and so forth, allowing to erase, suggest or improve the user's search preferences. Cortana is connected to Microsoft products, Bing, Outlook, Skype, and so forth. Cortana is a working demonstration that chatbots can help users to speed up work in many domains. Cortana's connection with Cortana Analytics Suite is interesting. It enables the chatbot to be driven by the outcome of the ML algorithm.

Chatbots also serve in the healthcare domain, where they help to navigate patients through the treatment procedure. The ViDi (Virtual Dietician) chatbot interacts with diabetic patients as a virtual adviser.[38] ViDi was designed to solve health problems and be the web-based chatbot as a redesigned version.[39] ViDi uses a special pattern to remember past conversational paths. The path is divided into three levels of nine questions each, and this is done by the Vpath parameter, which determines the patient's path. The redesigned version of ViDi used a relation database (SQL) and web programming languages such as PHP, HTML and XHR. Similarly, the MS Azure Bot Service can be used in medical applications.

3 | SECURITY

Threats and vulnerabilities are the two main categories of security issues. A security threat is defined as a risk by which an organization and its systems can be compromised. Computer security threats[40] are identified by a STRIDE model as Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, Elevation of Privileges, and many more; every security threat should be reduced by protective mechanisms ensuring the following properties as Authenticity, Integrity, Non-repudiation, Confidentiality, Availability, Authorization, and many others.[41]

System vulnerabilities are weakness which can be exploited by an attacker, to cross privilege boundaries (i.e., perform unauthorized actions) within a computer system.[42] The system is vulnerable when it has weak coding, lacks current drivers on the hardware side, has a weak firewall, and so forth. Human errors mostly cause system vulnerabilities. SDL (Security Development Lifecycle) helps to avoid such errors.[43] As many chatbots use cloud computing services, which have threats and vulnerabilities very well handled, to store the data,[5] the following text is focused on the communication part and different aspects of data manipulation. Companies not using cloud services usually also handle previous topics during the implementation phase of the chatbot system.

Secure messaging can be divided into two domains. The fist domain concerns security in data transfers, that is, the secure transfer of messages, voice, and images to a server on which the chatbot is hosted. The second domain deals with the user's data on the server (backend), that is, how the data is processed, stored, and shared. Both domains cover the lifecycle of the user's domain. There are several threats to user messages in the first domain. The following text analyzes the methods to increase security in chatbot communication. Not all of them are generally used, as it is not necessary in many cases. If any company processes a user's data, most of the following methods must be implemented. The following methods cover any communication with the chatbot, such as a pop-up chat window in a web-browser or mobile application.

Authentication and authorization

Confirmation of the user identity (authentication) is not always mandatory. When the user asks for help, for example, on a shopping website, usually no authentication is needed. In this case, the system does not require the identification of the user or access to the user's data. The situation is different when the user asks for help and the chatbot works with user's data. A typical example is chatting with a banking chatbot about an account balance. In this case, authentication and authorization are necessary to identify that a user is verified with valid and secure login credentials. The credentials are mostly a username, network address, system ID, phone number, biometric identification, certificate, password, and other methods of verification. Such credentials are sent by the user to the system and the system creates a secure authentication token, which is used throughout the user's session. In the communication with a bank (and similarly protected services), tokens are temporary. The system must create a new token after a given period.

The greater protection of a user's data and communication is ensured by two-factor authentication,[44] for example, the user is asked to verify the authentication of account credentials through an email and text message. There are also (higher) multi-factor authentication schemes,[45] but such methods are rarely used during the login and conversation with a chatbot. In summary, authorization ensures that the correct person has access to the right data and services and authorization is necessary when a chatbot works with a user's data.

The storage and sharing of personal data over the internet is never safe enough. While interacting with the bots, a personal authentication security (Personal Scan) layer would confirm the user's confirmation by the bots. The personal scan would also ensure that in the case of interacting with a rogue chatbot, the information of the user would not be used by hackers or other phishing agents.[46]

Malicious chatbots operating on cross-platform instant messaging applications (Viber, WhatsApp, etc.) can contact the user and pretend to be, for example, a pizza order chatbot. Such a situation is called Smishing.[47] From the moment when user responds to the attacker, the chatbot continues in the conversation. The goal is similar to other phishing schemes to steal personal data. Receiving messages from unknown people/services is a good indication that it is probably phishing. Many online phishing attacks are sent from a hacked account into the user's friends' accounts. Not responding to unknown messages and authorizations can protect a user from a smishing attack. Authentication for user applications defends the users' data and devices from misuse; that is, when a user loses a phone or computer or leaves them unlocked. If a chatbot window was open, then an attacker could just ask the chatbot for sensitive information. Here the temporary tokens are necessary.

End-to-End Encryption

End-to-End Encryption (E2EE) is a system of communication where only the communicating parties can read the messages. The conversation is encrypted in a way that only the unique recipient of a message is allowed to decrypt it, and not anyone in between. Transported data can be tampered and spoofed by a third party. Hence, it is important to ensure that only the involved parties have access to the cryptographic keys required to decrypt the conversation.

In the case of public-key encryption, the user's device generates a pair of keys—private and public. There are different protocols to provide encryption, like the RSA algorithm.[48] The public key is used for encrypting the messages, and the corresponding private key can decrypt those messages. Naturally, the public key can be used by anyone who sends the message to the private key owner. Simply put, both sides of the chatbot can share the public key to encrypt the communication. Encryption is sometimes used with authentication and integrity protection schemes. It is very important to keep the user's private key safe, otherwise an attacker can decrypt all the messages that are intended for this user.

As in the previous case, chatbots which do not work with personal data mostly do not use E2EE. Many chatbots operating on a website use only Hypertext transfer protocol secure (HTTPS), which transfers data through an encrypted connection by a Secure Sockets Layer (SSL) or Transport Layer Security (TSL). HTTPS is 'point-to-point' encryption as opposed to 'end-to-end' encryption. HTTPS is secure on the user's connection to a load balancer, but then the data are decrypted back to plain text. This makes the data open to attack all services after the load balancer[#]. Only E2EE ensures the safe data transfer between involved subjects.

Article 32 (a) of the GDPR specifically requires that companies take measures to pseudonymize and encrypt personal data[‖]. Many chatbots are connected to WhatsApp, Facebook Messenger, Telegram, Slack, and so forth. As GDPR requires the encryption of personal data, many chatbots operating on these networks support data encryption. For example, Facebook Messenger has a feature called 'Secret Conversations' that enables E2EE based on Signal Protocol developed by Open Whisper Systems. Note that there is a 'law versus privacy' conflict[49] and companies in some states must design their encrypted products and services whereby governments, acting with appropriate legal authority, can gain access to the data in a readable and usable format[**].

Self-destructing messages

In many cases when Sensitive PII (Personally identifiable information) is transmitted, self-destructing messages are a practical solution. Messages containing PII are automatically erased after a set period. It can involve both sides—the user and the chatbot. Self-destructing messages are an important security practice in communication with financial (banking) and healthcare chatbots. Specifically, Article 5 (e) of the GDPR states that personal data shall be kept for no longer than is necessary for the purposes for which it is being processed. General conversation is not personal, but financial and healthcare data are. It is not easy to distinguish by chatbot what is and what is not personal information (e.g., in communication with a pharmacy chatbot), but this is also defined by GDPR, and Protected Health Information[50] (PHI) defined under US law. According to PHI, any information about health status, the provision of health care, or payment for health care is created or collected by a Covered Entity and must not be

[#] https://tozny.com/blog/end-to-end-encryption-vs-https/.

[‖] https://cai.tools.sap/blog/chatbots-security-measures-you-need-to-consider/.

[**] https://www.justice.gov/opa/pr/international-statement-end-end-encryption-and-public-safety.

linked to a specific individual. GDPR compliance requires an 'intent level' of privacy. This means that only the user's intent will be logged and stored for audit purposes and the personal information of the user must never be revealed, not even from the backend.

User communication data, backend side

The previous topic deals with secure communication, which is necessary in many cases. Nevertheless, communications with chatbots are mostly stored on the backend side. The communication and the user's data are private data which should be stored to resist potential security threats.[41] The historical conversations are extremely valuable for chatbot developers and companies. When a customer calls a hotline, an interactive voice response (IVR) announces 'this call may be monitored or recorded'. Contact centers can legally record and monitor phone calls in the EU and UK. The situation is more difficult to answer in the US, where the laws vary across states. The same 'monitoring' happens in chatbot conversations, but the approval is hidden in the website terms and conditions and disclaimers. Companies use the stored conversation for an analysis of their services and as protection in possible litigation. Chatbot developers analyze communications for the better improvement of a chatbot's quality, as ML methods need the data to be trained on and the quality of the model generally grows with the amount of data.[51] Most of the companies store the communication with a chatbot due to the aforementioned reasons. Companies handling PII according to the GDPR definition must take special actions to store data safely; see the following text for more details. For instance, banks use self-destructing messages, but the system contains logs and account balance records.

Article 32 (a) of the GDPR specifically requires that companies take action to pseudonymize and encrypt personal data. The communication with a chatbot can be stored, but without a connection to a given user and user data, or it must be protected. In the first case, the database contains just the communication and can be shared with the development team; unless it contains information that identifies an individual either directly or indirectly.[9] Companies use regular expressions, pattern matching, and entity recognition methods to detect personal data. On the other hand, Cortana, for its proper functionality, needs to collect a lot of data about its user: it has access to contact lists, an overview of sent and received e-mails, short messages, and incoming and outgoing calls. In addition, it also tracks where the user is, what he says, what music he listens to, what he buys, what movies he watches, and what he does in the browser. This data is then sent to the operator's server for analysis. For these reasons, it is recommended that anyone who does not intend to use the personal assistant function should turn it off on their device. However, even after Cortana is turned off, the data it has collected remains stored on the servers, and Bing returns to it when it searches for user requests. This is another aspect, chatbots can have access to any data on the given platform. Accessibility to many data sources is a potential security risk in companies' chatbots; configuration of user control and access permissions is a must. Hence, the communication must be secure and the chat history must be well handled when personal data are involved. It is manageable in closed-domain chatbots.

A little more problematic is the situation with open-domain chatbots, which are not as restricted as closed-domain chatbots. Note that closed-domain chatbots working with PII require authentication and authorization, and users agree with terms and conditions. On the contrary, a user can discuss any topic she/he likes with open-domain chatbots, but the chatbot's providers discourage people from chatting about personal details. Someone can admit that free open-domain chatbots are mostly for fun, as they cannot yet lead a serious discussion on any topic. However, the social-bot[52] application Replika, for example, allows you to become friends with the AI,[53] then people can share their feelings, more than just PII by the GDPR definition. Table 1 summarizes the information which is shared with third parties by three popular chatbots—Replika, Kuki, and CleverBot. It can be seen that all companies share some user data (to third-party tracking and analytic tools, like Google Analytics[54]) and cookies. This helps companies evaluate the type of customers. Chatbots are mostly hosted on cloud-based services, which has their own privacy policy for user data.[55] All companies can read all messages to improve their services. Hence, it is not a private chat by definition. Lastly, many chatbots can be hosted on many communication channels, see Figure 1; and messages can be duplicated on another server, for example, WhatsApp data are stored on Google Drive, see Section 4. Currently, on GooglePlay and in the AppStore there are hundreds of chatbot applications, which try to be the user's best friends, medical helper, teacher of foreign language by chat, girl/boyfriend, and even oracle from Tarot cards. This is potentially dangerous, because people can send their PII to anonymous services or can get wrong information from untrustworthy services; and all these services collect data.

3.1 | Chatbot security versus the new law in the European Union's GDPR

General Data Protection Regulation (GDPR) is considered by many to be the golden standard among data privacy regulations and as an example for other states. Chatbots have access to an increasing variety of PII and personal details on their user base. In this regard, the European Union formed the GDPR whereby the mandate constructs the people's fundamental right of privacy and freedom to protect personal data from malicious use. Malicious use comprises unregulated data storage, authentication without the user's permission, and the authorization and encryption of the personal data without the user permissions.[56] If a chatbot can access the personal data of a user, the chatbots must have the GDPR mandates and regulations in place. Article 5 (e) of the GDPR orders that personal data shall be kept for no longer than is necessary for the purposes for which it is

TABLE 1 Information shared by chatbot companies with third parties

Stored data /chatbots	Replika	Kuki	CleverBot
Name	S*		S
Other profile information birthday, gender, work status, and so forth	N		N
Email address, phone number	S*	S	
Facts about you and your life, hobbies and interests	N		
People mentioned in the chat	N		
Images sent, voice messages	S**		
Text chat	N	S**	S ('starred')
Usage data—button clicks, search queries, and similar	S	S	S
Cookies	S	S	S
Backend infrastructure	Azure, AWS	AWS	UK company servers
E2EE	Standart SSL	HTTPS	

Note: 'S' is used for shared and 'N' for not shared. Items marked with '*' are mandatory and '**' are items shared without any PII. Notation 'starred' means that popular text messages can be shared. Empty cells are not specified in the privacy policy.

being processed. Note that GDPR is enforced only in companies based in the EU or in services processing EU citizens' personal data. Article 4 (1) of the GDPR defines '1personal data' as being any information relating to an identified or identifiable natural person ('data subject').

The GDPR defines principles and the lawful bases for processing personal information and also specifies rights for individuals:[57] Transparency in data procession, Data minimization (the personal data processed is adequate, relevant, and limited), Purpose limitation (personal data can be collected, but not further processed), Storage limitation (personal data are deleted when it is no longer needed), and Download (personal data are provided on demand), Change (right to change personal data), and Remove (sometimes called Right to be forgotten, i.e., delete all the user's data).

The new GDPR mandate mentions that every technology that uses personal user data must maintain strict security so that there is no malicious activity of encryption leading to the infringement of personal rights and personal data. The mandate also clearly mentions that the tool or the software must make sure that it has alternative options in cases of the accidental or unlawful destruction of personal data, unauthorized disclosure of the personal data and its transmission thereof. Therefore, every chatbot must have 'appropriate safeguards' to avoid encryption-based security challenges. The authentication of personal data and its usage is also mandated in the GDPR. The authentication of personal data must be by biometric data and not photographs of the actual user. This also comes under the personal scan security threats of chatbots. It is the legal obligation that no personal data can be gathered or used in any form from the authentication process. The controller, the chatbot, in this case, is also obligated to gather all forms of authentication data entered by a person in order to match the degree of fundamental personal rights. The user's intents are logged, can be kept for audit purposes and must be revealed on demand.

Data storage as per the mandate of GDPR is a processing method and must not restrict any user from its usage unless the user violates the code of conduct (GDPR). A GDPR compliance measure is to have an 'intent level' of privacy. Besides, data storage by the controller must not use unlawful or unfair means to keep secondary storage of the original data. Personal data of any type must fully maintain its integrity and confidentiality. Every entity must ensure that only the authorized user is allowed to view or use the personal data stored. Therefore, aspects like biometric scans and authentication methods are a must. Only the user itself can exchange personal information, but the system must not reveal the data.

Therefore, following the mandates of the GDPR in lieu of chatbot developers implicates the need for data protection by maintaining the lawfulness, fairness and transparency of the stored personal data. In addition, the mandate also indicates the need for accuracy of the personal data, accountability, confidentiality and storage limitations, thereby reducing the chances of security threats. The mandate of data governance obligations for authorization and authentication processes may reduce the security challenges of chatbots. Many companies creating chatbot building platforms added commands similar to 'Request Download Data', 'Delete Personal Data' or 'Change Personal Data', which give users some control over their personal data.

There are still some open issues with GDPR and chatbots.[57] The GDPR provides the 'right to be informed' about how the data are processed, but overall algorithmic transparency is low. The basic fact about chatbots' algorithmic construction is known, see Section 2. On the other hand, many implementation details and knowledge bases (created from collected conversations) will never be revealed by companies. Chatbots also do mistakes as they learn from examples with mistakes. Chatbots based on RNN (see references in Section 2) do grammatical mistakes. Chatbots can go off the rails like Microsoft's Tay did.[37] Therapist chatbots, best-friend chatbots, finance chatbots, and the like can give wrong advice. Companies must somehow handle huge amounts of data to make generative models feasible, but so everything can be corrected. The question is who is responsible when something bad happens, the user blindly relying on a chatbot or incautious developers?

Big data versus Data minimization and Storage limitation; many companies store data as long as possible, and time-limit is defined by the necessity to keep the data. The right to be forgotten is also complicated in the chatbot field. Past communications can be used to train AI chatbots, but once (e.g., RNN) a model is created; it is impossible to remove such communications. It should not generally be an issue. If a user can be identified indirectly or distinguished in the group by some data, then such data are personal according to GDPR ('profiling'). Similarly, if a user uses a unique answer, greeting, word, and so forth, then the chatbot can use this phrase and the user's content can be identified. There are special categories of personal data in GDPR (e.g., ethnicity or sexual orientation). A chatbot does not have to respect such aspects of conversation. It is difficult to find the correct strategy to provide reasonable replies and fully respect all aspects of conversation at the same time. The principal aim is to produce consistent answers to semantically identical inputs. One solution could be Persona-Based Conversation Models.[58,59] Such models provide consistent answers and could be a solution for this type of conversation. Contesting decision (Art 22 GDPR) gives a user a right not to be subject to a decision based solely on automated processing and contest the result of automated individual decision-making. It means that users have the right not to be subject to automated decision-making (e.g., by analysis of conversation), especially in Persona-Based Conversation Models. Users are mostly not informed about being a subject of automated decision-making, which is quite natural in AI chatbots. By Art 13-15 (refers to Article 22(1) and (4)), data processors must inform about the existence of using automated systems, but not about the right of a user not to be subject to automated individual decision-making. Suppose the user in communication decides not to be subject to automated individual decision-making. In that case, it won't be easy to continue in conversation as many AI chatbots automatically create the user's profile, especially in open-domain chatbots and voice assistants.

The last issue is about settings of cookies and other data collected by third parties. When a user visits a website, he/she is asked to actively agree to cookies, but this is not the case with many chatbots. Especially those operating on a mobile app, which can also have access to the device's hardware information and other data.

4 | DISCUSSION

Chatbots are starting to dominate in many areas, as they are efficient, operate 24/7, and represent a pleasing way of giving instructions to a computer. As was mentioned in the introduction, chatbots are nothing more than a programmed input–output system and basic 'recipes' for how to create and operate chatbots are known. The usage of chatbots comes with great responsibility and cyber-security risks. In many cases, chatbots work with 'sensitive information', and security practices should be implemented. Communication with chatbots is done via already implemented channels and protocols. For the most part, the communication does not present security issues that have not already been discovered and properly mitigated. But still, not every chatbot communicates through encrypted channels and users should not send any personal data during chatting.

Some security issues come with stored communications and user data. The problem is that communication is extremely valuable and many companies store past communication. Data can be encrypted on servers, but the ML algorithm cannot be taught on encrypted data. The result would not make sense; moreover, NLP tools are not prepared to learn on encrypted data. This is one of the moments when communication is revealed and can be read. Companies must handle such situations according to the GDPR and other similar rules when PII are presented. Companies and employees would not share the discussion with PII. From a general perspective, chatbots record and learn from previous communication; they later re-use words, phrases, and complete utterances from conversations when talking with others later.

Another issue is that chatbots can operate on third-party conversational interfaces/networks like Facebook Messenger, Viber, Facebook WhatsApp, Twitter, Facebook, LinkedIn, and so forth. Owners of those app/networks have different policies about what belongs to the user and what belongs to the company when the user uses their services. For example, there were more than 300,000 active chatbots on Facebook alone in 2018, which exchanged 8 billion messages between people and businesses. It is four times more than the number of messages exchanged in 2017[††]. Those and similar companies evaluate the price of data from chatbots.[60] It is not only about using previous communication to train the chatbot's ML algorithm; it is about cookies, user preferences, the main discussion topics of different groups, activity, language, positive/negative messages, and any other features which can be extracted from chatbots and sold to other customers. Note, any of the previous features belongs to the PII category, and still are very valuable for personalization and information retrieval.[59] For instance, Facebook WhatsApp recently changed its terms and privacy policy to connect WhatsApp with other Facebook companies and use some information collected from WhatsApp[‡‡]. It means that business messages can be used for commercial purposes, and some of them are also stored on Facebook's servers. As was mentioned at the beginning of this paragraph, many business messages are created by business chatbots.

When a user chats with a chatbot on popular communication platforms, the platform operator gets information about it. To have a perspective about the user data stored by the terms policy, for example, WhatsApp stores—Device ID, User ID, Advertising data, Purchase history, Rough placement, Telephone number, Email address, Contacts, Product interaction, Failure data, Performance data, Additional diagnostic data, Payment info, Customer support, Product interaction, and Other user content.

[††] https://about.fb.com/news/2018/05/f8-2018-new-tools-for-businesses-and-people-to-deepen-connections-in-messenger/.

[‡‡] https://www.xda-developers.com/whatsapp-updates-terms-privacy-policy-mandate-data-sharing-facebook/.

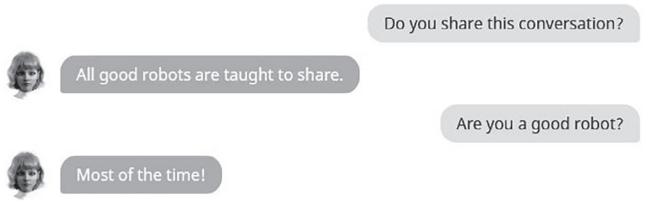

FIGURE 2 Snapshot of conversation with kuki_ai chatbot

Do you share this conversation?

All good robots are taught to share.

Are you a good robot?

Most of the time!

Facebook Messenger collects Purchase history, Additional financial information Exact location, Rough placement, Physical address, Email address, Name, Telephone number, Additional user contact information, Contacts, Photos or videos, Game content, Other user content, Search history, Browsing history, User ID, Device ID, Product interaction, Advertising data, Further information on use, Failure data Performance data, Additional diagnostic data, Other types of data, Browsing history, Health, Fitness, Payment info, Photos or videos, Audio data, Game content, Customer support, Other user content, Search history Sensitive information, iMessages, Email address, Phone number Search history, and Device ID.

Signal stores no information, except the Phone number (the only personal information) and Signal does not attempt to associate it with your identity.

Telegram stores collected Contact information, Contacts, and User ID. User should know this, because in modern data driven world information channels can be connected.

Possibly not now, but certainly in the future, people can be recognized by text; similarly to how a neural network identified the specific scenes in Henry VIII which were not written by William Shakespeare.[61] The digital footprints[60] of many users can be collected from many sources and compared for similarity.

There are thousands of chatbots connected to webpages, operating on WhatsApp, Facebook, LinkedIn, Viber and other communications channels, hundreds of chatbots with their own application, and a few very complex chatbots like Cortana operating simultaneously in Cortana Intelligence Suite, Windows, and many other devices. It is not easy to compare the terms and privacy policy concerning the data processing of all companies. The issue is that people click to accept and do not read the terms and conditions in all services. There are some regulations (like GDPR) made by a government or other authority to control and protect people's personal data which give some level of trustworthiness. People should avoid using the services which do not follow basic security rules discussed in Section 3. Even kuki_ai chatbot has an opinion about data sharing, see Figure 2.

Social aspects

This article focuses more on the technical solutions of security details around chatbots, but the authors also have to mention the social aspects of using chatbots to complete the security topic. Social aspects can enlarge because chatbots try to draw user attention and keep the user in communication. Communication in natural language is an important way of expressing feelings, ideas and concerns between humans. As chatbots learn from the user, they can take over the user's rhetoric. One example is Microsoft's social media-based chatbot 'Tay', who 'evolved' from a normal teenage girl to a chatbot displaying anti-Semitic, racist and sexist attitudes in less than sixteen hours.[37] Microsoft's social media-based chatbot 'Zo' also picked up some offensive habits.[62]

Despite the intention of their designers, many ML implementations have developed harmful human-like biases that cannot be easily removed.[62-66] On the other hand, chatbots can be pre-biased and, for example, trained to influence public opinion and discourse.[67-69] Shortly, social chatbots started to operate on many social platforms to spread pre-programmed information. Even chatbot assistants, who recommend a restaurant, flight, financial product, medicament, and so forth, can be pre-programmed to offer users prearranged products. This creates additional income for the chatbot creators from 'product placement' [§§], but one has to realize that chatbots work 24/7 and people trust them. Chatbots are designed to draw a users' attention by messaging because it is fast, easy, and actually feels like a conversation; and to use targeted messaging (precisely Conversational Marketing[70]) to help or sell you a product.

Chatbot design techniques are continuously improved,[20] and it is more and more difficult to distinguish conversations between a human or a chatbot. Some chatbots can even assist as mental health support,[71] which requires a really good conversational level by chatbot. Such social chatbots can obtain valuable information from the user or even make their user vulnerable to infiltration from other chatbots. The annual Loebner Prize[1,72] is the world's longest-running Turing-Test competition. It is a competition in artificial intelligence that awards computer programs considered by the judges to be the most human-like. No Chatbot has ever achieved the Loebner Prize's golden medal (responses which cannot be distinguished from humans). However, some chatbots have scored as highly as 3 out of the 12 judges believing they were human.[20] Huge technological companies have access to an enormous amount of conversational data and have resources to pay skilful programmers. This, with improving chatbot design

[§§] Basically, this case is product placement, but this term has not been used before in this context. This topic also has not been publicly discussed by law or companies. Nowadays, voice assistants search for, for example, restaurant from open-source data, when user ask for a good restaurant. The question is what happens when some restaurant franchise will pay for directing the people to their restaurant.

techniques, will make chatbot speech more human-like.[4] This will make chatbots operating on social media more trustworthy,[73] and such chatbots can influence people's opinion even more.

5 CONCLUSIONS

Chatbots have undisputed advantages, as they work 24/7, lower costs (compare to workers), communicate in natural language, and manage various tasks. Instant messaging is for many people more user-friendly then in email or web form. Chatbots/voice assistants are operating on social media, communication platforms, websites, mobile phones, computers, and other platforms; and users use chatbots in communication with banks, healthcare assistants, online shops, cars, insurance companies, airport companies, and many other subjects. Hence, secure communication with the chatbot is essential to keep the user's data safe.

This paper summarizes all the security aspects concerning communication with chatbots. It provides a literature review describing important steps in chatbot design techniques with respect to chatbot security. The paper also defines the security threats and vulnerabilities; it describes in detail methods which should be adopted to have a safe chatbot platform. In summary, communication via chatbots does not present security issues that have not already been discovered and properly mitigated earlier. The main concerns are rather about the user's data stored on the backend side and its usage, as chatbots generate enormous amount of data about users. Regulations such as GDPR strictly define the data manipulation methods, but not everybody lives in the EU under the jurisdiction of GDPR regulations for data protection. All personal information should be deleted from such data in the EU, but it is almost impossible to check if this is true. Only external audits can check this. Also, historical communication is extremely valuable for chatbot engineers, who use such data for chatbot improvements. Modern NLP techniques can analyze conversational data, and such analyses can be used, for example, for personalized content and marketing.

There are some concerns about chatbot security discussed in the article (apart from the past communication dataset), which are difficult to track.

- Cookies—many webpages use cookies, which also contain details about chat conversations, and are collected by third parties.
- Usage data—button clicks, search queries, and similar data can also be shared with third parties.
- User identification—email, name, device ID, and similar data can be send from a chatbot. Mostly in the case when a chatbot performs an external action.
- Messaging platforms—many chatbots operate on external messaging platforms, which have their own terms and conditions. Thus, the details about communication with the chatbot are recorded elsewhere.
- Conversational biases—there can be chatbots (social bots) created to manipulate people's opinions, for example, during selling, in forums, on social platforms, during political elections, and so forth. Such malicious chatbots can create huge damage from their violent behavior.

Based on the following reasons, the authors suggest that communication with a chatbot should be identified by some international designation or word 'Bot' close to the name or in the header of the chat window. More precisely, the designation or name 'BotS' (similarly to well-known HTTP and HTTPS) should denote encrypted communication and communication stored on servers by similar (or identical in the EU) rules to the GDPR. Mainly due to the following reasons:

1. Globally, a huge amount of time is wasted by expressing greetings and describing the situation in communication with a chatbot pretending to be a human. Most chatbots decompose the sentence and search for keywords, from which they classify the task and prepare the answer. It needs neither a greeting nor a story.
2. Marked secure communication gives a user some level of trustworthiness to share personal data. The user must read the terms and conditions to recognize if the communication is safe, the personal data is properly handled, the communication is not shared with third-party services, and so forth; 'BotS' will declare that the chatbot owner follows these principles.
3. Natural language interfaces to computer systems cannot yet simulate the full range of intelligent human conversation.[3] However, in the future chatbots will understand sentiment, emotions, and argumentation. It will be more and more difficult to spot a chatbot. This is dangerous in the case of conversational biases from social chatbots operating on social media and e-commerce. It is not easy, but owners of social media should take action to recognize chatbots operating on their platforms and automatically mark them. For instance, the B.O.T. Act went into effect in California last year. Bots are defined as 'automated online account[s] where all or substantially all of the actions or posts of that account are not the result of a person'. If companies use a bot to communicate with their customers or with the public online, they must disclose this fact or face a penalty of $2500 per violation[¶¶].

[¶¶] https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180SB1001.

Limitations and future directions

The main limitation is in the area of data manipulation on the provider side. Every company tries to be 'transparent' with data manipulation and processing, but none of them shares the exact processing pipeline, that is, who and how many employees have access to data, what algorithms they use, which data are sold to third parties or sent to cooperating companies. Hence it is difficult to evaluate the level of a security risk as it represents a company secret.

This topic is continuously evolving as are the options of chatbots. Recently chatbots told the user 'what to do', and today's chatbots 'do these things' for the user. Definitely, this topic needs to be watched closely in the future, since chatbots are going to be employed in many domains. The future of chatbots is coming now as they will operate smart homes, command self-driving cars, give medical assistance, replace programming by typing, perform data analysis and visualization, and many other topics to come.

One topic is still missing in our work. It is chatbot-chatbot communication. Nowadays, voice assistants can help you to order, for example, pizza, from a pizza company's chatbot. But in the future, chatbots will communicate to each other to exchange information.

ACKNOWLEDGMENTS

This work was supported by the European Regional Development Fund under the project AI&Reasoning (reg. no. CZ.02.1.01/0.0/0.0/15_003/0000466), by the Technology Agency of the Czech Republic under the grant no. TN01000024, and by the projects of the Student Grant System no. SP2021/24, VSB—Technical University of Ostrava. The authors would also like to thank the anonymous reviewers for the thoughtful comments that helped to improve the manuscript.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Jana Nowaková https://orcid.org/0000-0002-5213-3302

REFERENCES

1. Abu Shawar BA. *A Corpus Based Approach to Generalise a Chatbot System* [PhD thesis]. University of Leeds, Leeds; 2005.
2. Weizenbaum J. ELIZA–a computer program for the study of natural language communication between man and machine. *Commun ACM*. 1966;9(1):36-45.
3. Gentsch P. *AI in Marketing, Sales and Service: How Marketers Without a Data Science Degree can use AI, Big Data and Bots*. New York, NY: Springer; 2018.
4. Hill J, Ford WR, Farreras IG. Real conversations with artificial intelligence: a comparison between human–human online conversations and human–chatbot conversations. *Comput Hum Behav*. 2015;49:245-250.
5. Hoy MB, Alexa S. Cortana, and more: an introduction to voice assistants. *Med Ref Serv Q*. 2018;37(1):81-88.
6. Adiwardana D, Luong MT, So David R, et al. Towards a human-like open-domain chatbot; 2020. arXiv preprint arXiv:2001.09977.
7. Følstad A, Brandtzæg PB. Chatbots and the new world of HCI. *Interactions*. 2017;24(4):38-42.
8. Roller S, Dinan E, Goyal N, et al. Recipes for building an open-domain chatbot; 2020. arXiv preprint arXiv:2004.13637.
9. Alepis E, Patsakis C. Monkey says, monkey does: security and privacy on voice assistants. *IEEE Access*. 2017;5:17841-17851. https://doi.org/10.1109/ACCESS.2017.2747626.
10. Nuruzzaman M, Hussain OK. A survey on chatbot implementation in customer service industry through deep neural networks. Paper presented at: Proceedings of the 2018 IEEE 15th International Conference on e-Business Engineering (ICEBE). China; 2018:54-61; IEEE.
11. Akhtar M, Neidhardt J, Werthner H. The potential of chatbots: analysis of chatbot conversations. Paper presented at: Proceedings of the 2019 IEEE 21st Conference on Business Informatics (CBI). Moscow, Russia; Vol. 01, 2019:397-404; IEEE.
12. Su MH, Wu CH, Huang KY, Hong QB, Wang HM. A chatbot using LSTM-based multi-layer embedding for elderly care. Paper presented at: Proceedings of the 2017 International Conference on Orange Technologies (ICOT), IEEE; 2017:70-74.
13. Xu B, Zhuang Z. Survey on psychotherapy chatbots. *Concurr Comput Pract Exper*. 2020;e6170.
14. Merritt A. The impact of coronavirus on chatbots WebPage; 2020. https://www.dashbot.io/2020/04/13/the-impact-of-coronavirus-on-chatbots%/.
15. Miner AS, Laranjo L, Kocaballi AB. Chatbots in the fight against the COVID-19 pandemic. *Dig Med*. 2020;3(1):1–4.
16. Shi W, Liu D, Yang J, Zhang J, Wen S, Su J. Social Bots' sentiment engagement in health emergencies: a topic-based analysis of the COVID-19 pandemic discussions on twitter. *Int J Environ Res Public Health*. 2020;17(22):8701.
17. Medford RJ, Saleh SN, Sumarsono A, Perl TM, Lehmann CU. An εInfodemicε: leveraging high-volume twitter data to understand public sentiment for the COVID-19 outbreak. medRxiv; 2020.
18. Tankard C. What the GDPR means for businesses. *Netw Secur*. 2016;2016(6):5-8. https://doi.org/10.1016/S1353-4858(16)30056-3.
19. Ferrara E, Varol O, Davis C, Menczer F, Flammini A. The rise of social bots. *Commun ACM*. 2016;59(7):96-104.
20. Abdul-Kader SA, Woods JC. Survey on chatbot design techniques in speech conversation systems. *Int J Adv Comput Sci Appl*. 2015;6(7):72–80.
21. Safavian SR, Landgrebe D. A survey of decision tree classifier methodology. *IEEE Trans Syst Man Cybern*. 1991;21(3):660-674.
22. Wallace R. The elements of AIML style. *Alice AI Found*. 2003;139:1–86.
23. Higashinaka R, Imamura K, Meguro T, et al. Towards an open-domain conversational system fully based on natural language processing. Paper presented at: Proceedings of the COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, ALC. Dublin, Ireland; 2014:928-939.
24. das Graças M, Marietto B, de Aguiar RV, de Oliveira Barbosa G, Botelho WT.Artificial intelligence markup language: a brief tutorial; 2013. arXiv preprint arXiv:1307.3091.

25. De Gasperis G. Building an AIML chatter bot knowledge-base starting from a FAQ and a glossary. *Je-LKS J e-Learn Knowl Soc*. 2010;6:75-83. https://doi.org/10.20368/1971-8829/414.
26. Worswick S. Mitsuku chatbot: mitsuku now available to talk on Kik messenger. 2010 Retrieval on. 4(05); 2018.
27. Pereira MJ, Coheur L. Just. Chat-a platform for processing information to be used in chatbots; 2013.
28. Carpenter R. Cleverbot WebPage; 2011.
29. Vinyals O, Le Q. A neural conversational model; 2015. arXiv preprint arXiv:1506.05869.
30. Sutskever I, Vinyals O, Le Quoc V. Sequence to sequence learning with neural networks; 2014. arXiv preprint arXiv:1409.3215.
31. Sriram A, Jun H, Satheesh S, Coates A. Cold fusion: Training seq2seq models together with language models; 2017. arXiv preprint arXiv:1708.06426.
32. Liu P, Qiu X, Huang X, Recurrent neural network for text classification with multi-task learning; 2016. arXiv preprint arXiv:1605.05101.
33. Serban I, Sordoni A, Bengio Y, Courville A, Pineau J. Building end-to-end dialogue systems using generative hierarchical neural network models. Paper presented at: Proceedings of the AAAI Conference on Artificial Intelligence, AAAI. Phoenix, Arizona; 2016.
34. Nay C. Knowing what it knows: selected nuances of Watson's strategy. *IBM Res News*. 2011.
35. Larsen L. *Learning Microsoft Cognitive Services*. Birmingham: Packt Publishing Ltd; 2017.
36. Barga R, Fontama V, Tok WH. *Cortana Analytics*. New York, NY: Springer; 2015:279-283.
37. Wolf MJ, Miller KW, Grodzinsky FS. Why we should have seen that coming: comments on Microsoft's tay εexperiment,ε and wider implications. *ORBIT J*. 2017;1(2):1-12.
38. Lokman AS, Zain JM. An architectural design of Virtual Dietitian (ViDi) for diabetic patients. Paper presented at: Proceedings of the 2009 2nd IEEE International Conference on Computer Science and Information Technology. Beijing, China; 2009:408-411; IEEE.
39. Lokman AS, Zain JM. Extension and prerequisite: an algorithm to enable relations between responses in chatbot technology. *J Comput Sci*. 2010;6(10):1212.
40. Kohnfelder L, Garg P. The threats to our products. *Microsoft Interface*. Redmont, Washington: Microsoft Corporation; 1999:33.
41. Deng M, Wuyts K, Scandariato R, Preneel B, Joosen W. A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. *Requir Eng*. 2011;16(1):3-32.
42. Abomhara M, Køien GM. Cyber security and the internet of things: vulnerabilities, threats, intruders and attacks. *J Cyber Secure Mob*. 2015;4(1):65-88.
43. Howard M, Lipner S. *The Security Development Lifecycle*. Redmond: Microsoft Press; 2006.
44. Wang D, Wang P. Two birds with one stone: two-factor authentication with security beyond conventional bound. *IEEE Trans Depend Secure Comput*. 2018;15(4):708-722. https://doi.org/10.1109/TDSC.2016.2605087.
45. Ometov A, Bezzateev S, Mäkitalo N, Andreev S, Mikkonen T, Koucheryavy Y. Multi-factor authentication: a survey. *Cryptography*. 2018;2(1):1.
46. Lemos R. Expect a new battle in cyber security: AI versus AI; 2017.
47. Joo JW, Moon SY, Singh S, Park JH. S-Detector: an enhanced security model for detecting Smishing attack for mobile computing. *Telecommun Syst*. 2017;66(1):29-38.
48. Milanov E. The RSA algorithm. RSA laboratories; 2009:1-11.
49. Endeley RE. End-to-end encryption in messaging services and national security–case of WhatsApp messenger. *J Inf Secur*. 2018;9(01):95.
50. Thompson LA, Black E, Duff WP, Black NP, Saliba H, Dawson K. Protected health information on social networking sites: ethical and legal considerations. *J Med Internet Res*. 2011;13(1):e8.
51. Géron A. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. Sebastopol, Canada: O'Reilly Media; 2019.
52. Ferrara E, Varol O, Davis C, Menczer F, Flammini A. The rise of social bots. *Commun ACM*. 2016;59(7):96-104. https://doi.org/10.1145/2818717.
53. Nima N, Lee T, Molloy D. Being friends with yourself: how friendship is programmed within the AI-based socialbot replika. *Masters Media*. 2017.
54. Clifton B. *Advanced Web Metrics with Google Analytics*. Hoboken, NJ: John Wiley & Sons; 2012.
55. Ogiela L, Snášel V. Intelligent and semantic threshold schemes for security in cloud computing. *Concurr Comput Pract Exper*. 2021;33(2):e5247.
56. Swinkels R. *Eu General Data Protection Regulation Compliance* [Master thesis]. 2017.
57. Saglam RB, Nurse JRC. Is your chatbot GDPR compliant? open issues in agent design; 2020. arXiv preprint arXiv:2005.12644.
58. Li J, Galley M, Brockett C, Spithourakis GP, Gao J, Dolan B. A persona-based neural conversation model; 2016. arXiv preprint arXiv:1603.06155.
59. Searby S. Personalisation–an overview of its use and potential. *BT Technol J*. 2003;21(1):13-19.
60. Fish T. My digital footprint a two-sided digital business model where your privacy will be someone else's business! futuretext; 2009.
61. Plecháč P. Relative contributions of Shakespeare and Fletcher in Henry VIII: an analysis based on most frequent words and most frequent rhythmic patterns; 2019. arXiv preprint arXiv:1911.05652.
62. Fuchs DJ. The dangers of human-like bias in machine-learning algorithms. *Missouri S&T's Peer Peer*. 2018;2(1):1.
63. Xu J, Ju D, Li M, Boureau YL, Weston J, Dinan E. Recipes for safety in open-domain chatbots article; 2020.
64. Lee N, Madotto A, Fung P. Exploring social bias in chatbots using stereotype knowledge. Paper presented at: Proceedings of the 2019 Workshop on Widening NLP. Italy; 2019:177-180. winlp.org.
65. Dixon L, Li J, Sorensen J, Thain N, Vasserman L. Measuring and mitigating unintended bias in text classification. Paper presented at: Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society. New York, NY; 2018:67-73. dl.acm.org.
66. Schlesinger A, O'Hara KP, Taylor AS. Let's talk about race: identity, chatbots, and AI. Paper presented at: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI'18. Montréal, Canada; 2018:1-14
67. Suárez-Serrato P, Roberts ME, Davis C, Menczer F. On the Influence of Social Bots in Online Protests. In: Spiro E, Ahn Y-Y, eds. *Social Informatics, International Conference on Social Informatics*. Cham: Springer International Publishing; 2016:269-278.
68. Forelle M, Howard P, Monroy-Hernández A, Savage S. Political bots and the manipulation of public opinion in Venezuela; 2015. arXiv preprint arXiv:1507.07109.
69. Woolley SC. Automating power: social bot interference in global politics. *First Monday*. 2016;21(4).
70. Sotolongo N, Copulsky J. Conversational marketing: creating compelling customer connections. *Appl Market Anal*. 2018;4(1):6-21.
71. Kretzschmar K, Tyroll H, Pavarini G, Manzini A, Singh I, Group NeurOx Young People's Advisory. Can your phone be your therapist? young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. *Biomed Inform Insights*. 2019;11:1178222619829083.

72. Mauldin ML. Chatterbots, tinymuds, and the turing test: entering the loebner prize competition. Paper presented at: Proceedings of the Twelfth National Conference on Artificial Intelligence; American Association for Artificial IntelligenceAmerican Association for Artificial Intelligence. Seattle, Washington; Vol. 1, 1994:16-21.
73. Feine J, Gnewuch U, Morana S, Maedche A. Gender bias in chatbot design. In: Følstad A, Araujo T, Papadopoulos S, et al., eds. *Chatbot Research and Design*. Cham: Springer, Springer International Publishing; 2020:79-93.

How to cite this article: Hasal M, Nowaková J, Ahmed Saghair K, Abdulla H, Snášel V, Ogiela L. Chatbots: Security, privacy, data protection, and social aspects. *Concurrency Computat Pract Exper*. 2021;e6426. https://doi.org/10.1002/cpe.6426