Towards a Conversational-Based Agent for Health Services

Michael Soprano[1,*], Kevin Roitero[1], Vincenzo Della Mea[1] and Stefano Mizzaro[1]

[1]University of Udine, Udine, Italy

Abstract

Conversational agents provide new modalities to access and interact with services and applications. Recently, they saw a backfire in their popularity, due to the recent advancements in language models. Such agents have been adopted in various fields such as healthcare and education, yet they received little attention in public administration. We describe as a practical use case a service of the portal that provides citizens of the Italian region of Friuli-Venezia Giulia with services related to their own Electronic Health Records. The service considered allows them to search for the available doctors and pediatricians in the region's municipalities. We rely on the use case described to propose a model for a conversational agent-based access modality. The model proposed allows us to lay the foundation for more advanced chatbot-like implementations which will use also alternative input modalities, such as voice-based communication.

Keywords

Public Administration, Electronic Health Record, Conversational Agents

1. Introduction

Conversational agents are software-based systems designed to interact with humans using natural language. They are well-known for some time; indeed, their first mention can be dated back to 1966, with the ELIZA system [1]. Coming back to more recent times, Dale [2] described, already in 2016, the demographics of chatbots, voice assistants and other agents in terms of thousands of them, arguing about a backfire of their popularity. As of today, they have been studied and adopted within various fields, such as healthcare [3] and education [4]. The recent surge of advanced language models has brought further attention to the usage of such agents. Many of us have probably heard the buzzword "ChatGPT", during early 2023. Such a noun refers to a machine learning model released in November 2022 by OpenAI which can be described as a smaller and more focused version of GPT-3, one of the largest language models constructed by far. Even though ChatGPT is not free from criticism, in particular about the quality of the generated content [5], its increasing popularity can not be ignored. For instance, Microsoft decided to propose a ChatGPT-powered version of its own Bing search engine.[1] OpenAI published a technical report about GPT-4 [6] in March 2023. Such a model is, indeed, the latest iteration of the GPT family. They claim that GPT-4 outperforms ChatGPT on nine categories of internal adversarially-designed factual evaluations [6, Figure 6]. While at the time of writing is too early to draw particular remarks, the evolution of such models looks promising.

Turning back to conversational agents in general, they can have a great impact on society; they provide new modalities to interact with various services and may help the elderly, disabled, visually-impaired people, etc. Also, the unexpected popularity of chatbot-like conversational agents could be beneficial also to less explored fields, such as public administration. The available implementations are mostly designed to ease the communication between administrative bodies and people, by allowing the latter to obtain answers for their issues without the involvement of staff [7]. For instance, Anastasiou et al. [8] developed a multilingual chatbot service to provide expatriates who enter a new country some help in dealing with procedures such as application for residence.

Among all the services that a public administration can offer in a digital fashion, one of particular interest is providing citizens with access to their own Electronic Health Record (EHR). The EHR represents and collects demographic, administrative, and clinical patient-centred data [9]. Access by citizens to their own EHRs is becoming an integral part of healthcare systems worldwide. For instance, the Italian Government published, back in 2014, the guidelines for the presentation of regional projects plans for the creation of the EHR, known as "Fascicolo Sanitario Elettronico".[2] As of today, every Italian region provides some kind of portal to allow citizens to access their EHRs, and the government's agency for the digital innovation provided in February 2023 the technical

Ital-IA 2023: 3rd National Conference on Artificial Intelligence, organized by CINI, May 29–31, 2023, Pisa, Italy

*Corresponding author.
michael.soprano@uniud.it (M. Soprano); kevin.roitero@uniud.it (K. Roitero); vincenzo.dellamea@uniud.it (V. D. Mea); stefano.mizzaro@uniud.it (S. Mizzaro)
https://michaelsoprano.com/ (M. Soprano); https://kevinroitero.com/ (K. Roitero)
0000-0002-7337-7592 (M. Soprano); 0000-0002-9191-3280 (K. Roitero); 0000-0002-0144-3802 (V. D. Mea); 0000-0002-2852-168X (S. Mizzaro)

[1]https://news.microsoft.com/the-new-Bing/

[2]https://www.fascicolosanitario.gov.it/en

specifications for interoperability between the regional systems. We believe that relying on a conversational agent-based access modality for citizens to their own EHRs is an important matter to address.

Let us now focus again on conversational agents. Given that a public administration portal is used by people of varying ages, digital skills, etc., building an effective conversation model is not an easy task. A way to cope with such difficulty is relying on all the social cues typical of human conversation such as small talk, gender, age, gestures, facial expressions to which humans react, and many others. The agent should thus display and employ such cues whenever possible. Feine et al. [10] provide a taxonomy for such social cues in conversational agents, finding 48 of them, and Amatulli et al. [11] show that they have an impact on the tendency of older consumers' choice of contemporary over traditional products. We also hypothesize that the usage of an advanced ChatGPT-like language model to improve part of the interaction (if not the whole) of the citizen with the EHR portal could be interesting to implement and study.

2. Aims

In this paper, we propose a conversational agent-based access modality for citizens to their own EHR in the form of a conversation model for the implementation of a first working prototype. Such a model allows interaction with the relevant databases and APIs of the public administration services in a seamless and efficient manner, providing users with a more user-friendly and streamlined experience by allowing them to input text in a natural and conversational manner, rather than requiring them to follow a rigid set of rules.

The model we propose is focused on implementing a specific and simple use case of the portal implemented by the Italian region Friuli-Venezia Giulia to provide EHR access to their citizens, known as SeSaMo,[3] whose name is derived from the Italian translation of "mobile health services".

3. Related Work

ChatGPT is a sibling model of InstructGPT [12], which is trained to follow instructions in a prompt and provide a detailed response. They are derived from GPT-3 and then optimized using human feedback in the form of mid-training reinforcement learning [13]. When building a general language model for various types of tasks its size seems to matter [14], as happened for GPT-3 over its predecessor GPT-2 and other models [15]. However, when restricting the focus to a narrower scope, like chatbots in the case of ChatGPT, other aspects such as the data focus or the training procedure are more important [16].

EHRs systems can assist providers in delivering high quality care to patients. Indeed, Kruse et al. [17] assessed in 2018 the validity of EHRs to improve the quality and efficiency of healthcare. Furthermore, Menachemi and H. Collum [18] describe in detail in their recent work why society needs such systems along with their advantages and clinical outcomes. Tapuria et al. [19] reviewed 74 papers about providing patient access to their own electronic health records. They found out that the majority of papers (54 out of 70) showed positive outcome or benefits by accessing to their EHRs via patient portals, and de Mello et al. [20] propose a taxonomy for semantic interoperability in EHRs.

Turning to conversational agents, their usage can rely on natural language vocal cues. The Conversational Information Seeking [21] (CIS) research area involves interaction sequences between one or more users and an information systems where the possible interactions are primarily based on natural language dialogue, while other types of interaction can still be included. Understanding the characteristics of people that use conversational agents can be useful to improve an existing implementation. Gkinko and Elbanna [22] propose a taxonomy to describe chatbot users according to four different types: early quitters, pragmatics, progressives and persistents. Furthermore, Parmar et al. [23] show that several health-focuses apps that use chatbots exist and they are often used to address gaps in healthcare quality.

4. Searching For A Doctor

The SeSaMo portal provides a set of so-called "quick services" (English translation of "servizi fast"), which complement the EHR access and browsing features and do not require explicit user authentication. Such services allow citizens, for example, to search for doctors and pediatricians, surgeries and pharmacies. Let us consider the first case. A citizen of the Italian Friuli-Venezia Giulia region may search for a doctor or a pediatrician by using the corresponding quick service.[4] Such a service is of critical importance for Italian citizens because, for instance, each of them must choose the general practitioner among the doctors available in the municipality of residence and they is the one a citizen should contact first to address any health-related problem. The service, thus, allow citizens to find every detail needed about their general practitioner of choice, including the working hours and alternative professionals to contact if they is not available. Figure 1 summarizes the service's usage by providing a high-level use case diagram. We use it in

[3]https://sesamo.sanita.fvg.it/

[4]https://sesamo.sanita.fvg.it/sesamo/#/il_mio_medico

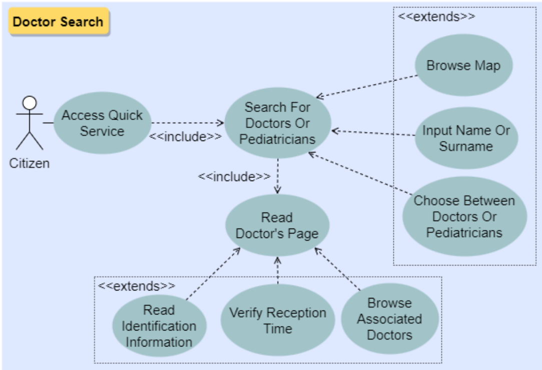

Figure 1: Use case diagram of the quick service.

the following as a practical use case to illustrate a first example of a conversation model for our agent.

Upon accessing the service's web page, the user is shown a map of the Friuli-Venezia Giulia Italian region and a search box. In the bottom part of the service's page, the full list of more than 800 doctors and pediatricians of the region is provided. The user can zoom into the map or manually skim the list to search for a given doctor or refine the list by writing the name or surname of the doctor they is looking for. Eventually, the user can also input the name of a municipality to filter the list and browse only its doctors. Also, a checkbox can be used to view doctors or pediatricians only. Figure 2 shows a glimpse of the service's user interface, where the user is searching explicitly for doctors only in the municipality of Udine. A total of 62 doctors and pediatricians are found and pin-pointed on the map, zoomed on the municipality's territory. Upon finding a doctor or pediatrician of choice the user can click its record and be shown the doctor's page on the SeSaMo portal. The page shows three main information sections. The top one the full name along with the identifier and the regional healthcare service unit of reference. Indeed, the doctors of the Friuli-Venezia Giulia Italian region are grouped into three main regional healthcare service units, whether they are

Figure 2: Interface of the quick service.

doctors, pediatricians or else. Moving on, the middle section describes the doctors's surgeries. Indeed, a doctor can examine the patients in more than one surgery; this may happen when, for instance, the capital of the municipality is large enough, or when the municipality itself is composed of multiple villages. The portal reports information for each surgery in a tabular fashion, where each row describes a given doctor's surgery. The left part of each row shows its address, phone number and the assigned regional healthcare service unit, together with the sanitary district of reference, which groups a subset of municipalities that are part of a given regional healthcare service unit. The right part of each row, on the other hand, shows a small table that indicates the surgery's working hours. It has two columns (morning and afternoon) and five rows, one for each working day from Monday to Friday. If the doctor examines citizens within the current time of the day, the corresponding table cell is marked in bold. Lastly, the bottom information section shows a list of doctors "associated" with the current one. A doctor is associated with another one if they can examine a citizen when the general practitioner assigned is not available. The list may not be shown in the portal's interface if a given doctor is not associated with other doctors.

5. The Conversation Model

Figure 3 proposes a graph, as a conversation model between a citizen who is using the quick service of the SeSaMo portal described in Section 4 and the implementation of a chatbot-like conversational agent. Each color-coded block of the graph corresponds to a message prompted to the user by the conversational agent. Each color has a different interpretation. A white block implies that a message is simply shown to the user, while a yellow one means that they is shown a button-based choice to be performed to proceed. A green block means that the user is required to provide a free textual answer. A red block in the graph implies that the chatbot performs internal business logic, whose outputs are used in the following messages. On the other hand, a violet block indicates a conditional logic step that can branch the conversation like when a yellow block is shown, but without prompting any message to the user. Lastly, a blue block means that the user is prompted with multiple looped messages. The messages which are shown within each block aim to use the textual social cues described by Feine et al. [10, Table A1].

The conversation works as follows. Initially, the user is asked whether they knows the name of a given doctor. If that is the case, they is required to provide it; otherwise, the agent proposes to filter the available doctors by municipality. This can be useful if, indeed, the user

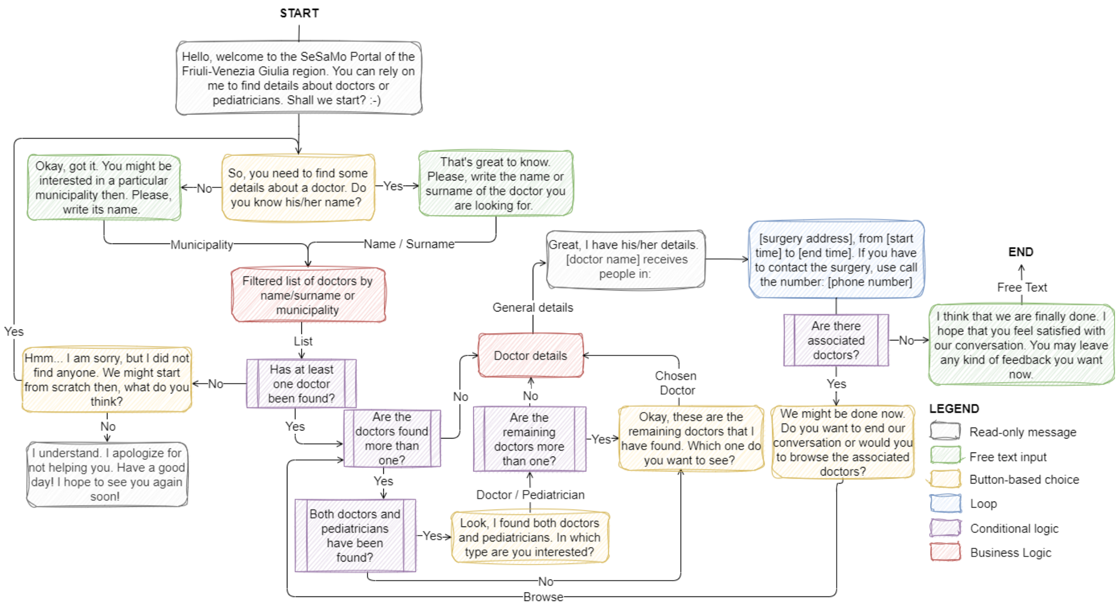

Figure 3: Conversation graph between a user of the quick service and a chatbot-like conversational agent.

is not aware of the doctors available in the municipality of residence. Then, the agent fetches and filters the list of available doctors. If none of them is found, it acknowledges the user by asking whether they wants to start searching from scratch. On the other hand, if the agent finds at least one doctor the conversation continues. If a single doctor is found, the details are shown to the user; otherwise, the agent checks if there are both doctors and pediatricians in the resulting set. In such a case, the user is asked whether they is interested in the former category or the latter. Then, the agent filters the result set and checks once again the number of remaining doctors. If there is a single one of them only, the conversation goes on and the details are shown; otherwise, the agent prompts an additional message and asks the user to explicitly select the desired doctor.

To summarize, the conversation can reach the point where the details of a single doctor are shown by following the three possible branches described. Then, showing the doctor's details involves reporting the information about each surgery of the one chosen, including its address, working hours and phone number, as happens in the portal's interface shown in Figure 2. As a last step, the agent checks if a list of doctors associated with the selected one exists. The agent thus asks explicitly the user whether they is interested in browsing them and if they agrees the conversation loops, as can be seen in the bottom right part of Figure 3. If the user refuses, the conversation can end and they is allowed to write some kind of textual feedback or comment.

The model shown in Figure 3 does not consider more advanced interfaces or input methodologies, such as voice-based communication with the agent. This is due to the fact that the initial prototype of the chatbot-like conversational agent is going to be implemented to support the traditional text-guided conversation loop described by the graph. The contribution that a ChatGPT-like language model could provide to improve the conversation will be considered in future prototypes and for, most likely, more advanced and complex use cases related to EHR access, browsing and usage. However, we can already argue that ChatGPT-like models will likely allow us, for instance, to implement the green and yellow blocks of the graph by letting the users write messages using natural language only.

Another aspect that must be considered to address the evolution of the conversation in the future is the usage of alternative input modalities, such as voice-based communication provided by devices such as Amazon Alexa [24]. Indeed, voice-based devices are ubiquitously available and there are studies concerning the design of effective interfaces [25]. Even though there are security concerns that need to be mitigated [26], such voice-based devices have demonstrated their usefulness, for instance, by helping prevent and manage chronic and mental health conditions [27]. We thus believe that using voice-based interfaces can fruitfully improve the quality and effectiveness of our conversational agent proposal for EHR access and browsing in public administration.

6. Implementation And Evaluation Principles

Technologies that allow implementing a chatbot-like conversational agent already exist. Several social media platforms such as Telegram, Slack and others offer native APIs and interfaces to build such agents, given that they will be available only for the users of such a platform. To address such a limitation, integrated development environments for building conversational agents exist, such as the Microsoft Azure Bot Services.[5] Indeed, one of the most interesting features is that it allows deploying easily the developed agents on multiple platforms such as Telegram, Slack, and more. Another aspect to consider is how to evaluate the performances of the conversational agent. Kohavi and Longbotham [28] describe several approaches to evaluate prototypes using online controlled experiments. The most simple experimental setup is the A/B testing procedure; the default version of a system is evaluated against its updated or new version. We thus argue that the initial prototype that implements the model shown in Figure 3 could be evaluated in such a fashion against a subsequent iteration of the conversational agents that uses to some extent a ChatGPT-like model.

Evaluating the performances of our conversational agents will require effectiveness metric which should consider multiple factors such as engagement, conversation length, and others. Irvine et al. [29] define, in their recent work, four intuitive evaluation metrics as proxies to measure the level of engagement of conversational agents. The mean conversation length measures the average number of user queries over multiple conversation sessions, defined as ordered sets of user and agent response pairs. This metric could be particularly useful if we decide, for instance, to track the previous conversations of each user. A common functionality of conversational agents is allowing the user to regenerate another response to their message. The graph proposed in Figure 3 is rather simple, yet the loop shown in the left part of the graph, triggered when the agent does not retrieve any doctor, could be seen in such a way. When the conversation with the user ends with the green block shown in the right part of the graph, the agent could ask the user to provide feedback by rating their answers using a given rating scale and thus measure user satisfaction. Lastly, the user retention of the conversational agent could be computed to evaluate their engagement after the first conversation; however, we believe that such a metric is less important for evaluating our agent, given that we aim to target a public administration service and not a commercial platform and we are not interested in its monetization or similar aspects.

[5]https://azure.microsoft.com/en-us/products/bot-services/

7. Conclusions

In this paper, we propose a first conversation model for prototyping a chatbot-like conversational agent to allow citizens of the Friuli-Venezia Giulia Italian region to access the quick service of the SeSaMo portal and search for doctors. The considered use case allows us to lay the foundations for future versions which will provide more advanced EHR-related features. The recent and popular language models could be used to enhance future prototypes, for instance by asking the users to state their goal, into a set of actions that the conversational agent can understand using natural language processing techniques. Such models can also capture the nuances of natural language, including paraphrasing and ambiguity; this will allow us to handle a wider range of user inputs and provide more accurate responses.

Lastly, it will be essential in the future to ensure that the conversational agent adheres to ethical principles, such as transparency, fairness, and privacy. This can be achieved through various techniques, such as providing clear explanations of how the agent works, ensuring that it is not biased against certain groups, and protecting user data from unauthorized access or misuse.

Acknowledgments

This work has been partially supported by the “Department Strategic Plan of the University of Udine - Interdepartmental Project on Digital Governance and Public Administration” (Jan. 2021 – Dec. 2025). We thank Paolo Coppola, who provided helpful comments and advice.

References

[1] J. Weizenbaum, ELIZA—a Computer Program For The Study Of Natural Language Communication Between Man and Machine, Comm. of the ACM 9 (1966). doi:10.1145/365153.365168.

[2] R. Dale, The Return Of The Chatbots, Natural Language Engineering 22 (2016). doi:10.1017/S1351324916000243.

[3] L. Laranjo, A. G. Dunn, H. L. Tong, A. B. Kocaballi, J. Chen, R. Bashir, D. Surian, B. Gallego, F. Magrabi, A. Y. S. Lau, E. Coiera, Conversational Agents In Healthcare: A Systematic Review, Journal of the AMIA 25 (2018). doi:10.1093/jamia/ocy072.

[4] C. Rodrigues, A. Reis, R. Pereira, P. Martins, J. Sousa, T. Pinto, A Review of Conversational Agents in Education, in: Technology and Innovation in Learning, Teaching and Education, Springer, 2022. doi:10.1007/978-3-031-22918-3_37.

[5] H. H. Thorp, ChatGPT Is Fun, But Not An Author, Science 379 (2023). doi:10.1126/science.adg7879.

[6] OpenAI, GPT-4 Technical Report (2023). URL: https://cdn.openai.com/papers/gpt-4.pdf.
[7] K. K. Nirala, N. K. Singh, V. S. Purani, A Survey On Providing Customer And Public Administration Based Services Using AI: Chatbot, Multimedia Tools and Applications 81 (2022). doi:10.1007/s11042-021-11458-y.
[8] D. Anastasiou, A. Ruge, R. Ion, S. Segărceanu, G. Suciu, O. Pedretti, P. Gratz, H. Afkari, A Machine Translation-Powered Chatbot for Public Administration, in: Proc. of the 23rd EAMT Conf, 2022. URL: https://aclanthology.org/2022.eamt-1.54.
[9] M. R. Cowie, J. I. Blomster, L. H. Curtis, S. Duclaux, I. Ford, F. Fritz, S. Goldman, S. Janmohamed, J. Kreuzer, M. Leenay, A. Michel, S. Ong, J. P. Pell, M. R. Southworth, W. G. Stough, M. Thoenes, F. Zannad, A. Zalewski, Electronic Health Records To Facilitate Clinical Research, Clinical Research in Cardiology 106 (2017). doi:10.1007/s00392-016-1025-6.
[10] J. Feine, U. Gnewuch, S. Morana, A. Maedche, A Taxonomy of Social Cues for Conversational Agents, Int. Journal of Human-Computer Studies 132 (2019). doi:10.1016/j.ijhcs.2019.07.009.
[11] C. Amatulli, A. M. Peluso, G. Guido, C. Yoon, When Feeling Younger Depends On Others: The Effects of Social Cues On Older Consumers, Journal of Consumer Research 45 (2018). doi:10.1093/jcr/ucy034.
[12] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell, P. Welinder, P. Christiano, J. Leike, R. Lowe, Training Language Models To Follow Instructions With Human Feedback (2022). doi:10.48550/ARXIV.2203.02155.
[13] P. Christiano, J. Leike, T. B. Brown, M. Martic, S. Legg, D. Amodei, Deep Reinforcement Learning From Human Preferences (2017). doi:10.48550/ARXIV.1706.03741.
[14] J. Kaplan, S. McCandlish, T. Henighan, T. B. Brown, B. Chess, R. Child, S. Gray, A. Radford, J. Wu, D. Amodei, Scaling Laws For Neural Language Models, arXiv (2020). doi:10.48550/arXiv.2001.08361.
[15] R. Dale, GPT-3: What's It Good For?, Natural Language Engineering 27 (2021). doi:10.1017/S1351324920000601.
[16] A. Farseev, Is Bigger Better? Why The ChatGPT Vs. GPT-3 Vs. GPT-4 'Battle' Is Just A Family Chat, Forbes Technology Council (2023). URL: https://www.forbes.com/sites/forbestechcouncil/2023/02/17/is-bigger-better-why-the-chatgpt-vs-gpt-3-vs-gpt-4-battle-is-just-a-family-chat/.
[17] C. S. Kruse, A. Stein, H. Thomas, H. Kaur, The Use of Electronic Health Records to Support Population Health: A Systematic Review of the Literature, Journal of Medical Systems 42 (2018). doi:10.1007/s10916-018-1075-6.
[18] N. Menachemi, T. H. Collum, Benefits And Drawbacks Of Electronic Health Record Systems, Risk Management and Healthcare Policy 4 (2011). doi:10.2147/RMHP.S12985.
[19] A. Tapuria, T. Porat, D. Kalra, G. Dsouza, S. Xiaohui, V. Curcin, Impact Of Patient Access To Their Electronic Health Record: Systematic Review, Inform. Health Soc. Care 46 (2021).
[20] B. H. de Mello, S. J. Rigo, C. A. da Costa, R. da Rosa Righi, B. Donida, M. R. Bez, L. C. Schunke, Semantic Interoperability In Health Records Standards: A Systematic Literature Review, Health and Tech. 12 (2022). doi:10.1007/s12553-022-00639-w.
[21] J. Dalton, S. Fischer, P. Owoicho, F. Radlinski, F. Rossetto, J. R. Trippas, H. Zamani, Conversational Information Seeking: Theory and Application, in: Proc. of the 45th Int. ACM SIGIR Conf., ACM, 2022. doi:10.1145/3477495.3532678.
[22] L. Gkinko, A. Elbanna, The Appropriation Of Conversational AI In The Workplace: A Taxonomy Of AI Chatbot Users, IJIM 69 (2023). doi:10.1016/j.ijinfomgt.2022.102568.
[23] P. Parmar, J. Ryu, S. Pandya, J. Sedoc, S. Agarwal, Health-focused Conversational Agents In Person-centered Care: A Review Of Apps, NPJ Dig. Med. 5 (2022). doi:10.1038/s41746-022-00560-6.
[24] M. B. Hoy, Alexa, Siri, Cortana, And More: An Introduction To Voice Assistants, Med. Ref. Services Quarterly 37 (2018). doi:10.1080/02763869.2018.1404391.
[25] S. Reeves, M. Porcheron, J. E. Fischer, H. Candello, D. McMillan, M. McGregor, R. J. Moore, R. Sikveland, A. S. Taylor, J. Velkovska, M. Zouinar, Voice-Based Conversational UX Studies and Design, in: Ext. Abstracts of the 2018 CHI Conf., ACM, 2018. doi:10.1145/3170427.3170619.
[26] Y. Lit, S. Kim, E. Sy, A Survey On Amazon Alexa Attack Surfaces, in: IEEE 18th Annual CCNC Conf., 2021. doi:10.1109/CCNC49032.2021.9369553.
[27] C. Bérubé, T. Schachner, R. Keller, E. Fleisch, F. Wangenheim, F. Barata, T. Kowatsch, Voice-Based Conversational Agents For The Prevention And Management Of Chronic And Mental Health Conditions: Systematic Literature Review, J. Med. Internet Res. 23 (2021). doi:10.2196/25933.
[28] R. Kohavi, R. Longbotham, Online Controlled Experiments And A/B Testing, Springer, 2017. doi:10.1007/978-1-4899-7687-1_891.
[29] R. Irvine, D. Boubert, V. Raina, A. Liusie, V. Mudupalli, A. Korshuk, Z. Liu, F. Cremer, V. Assassi, C.-C. Beauchamp, X. Lu, T. Rialan, W. Beauchamp, Rewarding Chatbots For Real-World Engagement With Millions Of Users (2023). doi:10.48550/ARXIV.2303.06135.