Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 219 (2023) 1289–1296

Procedia Computer Science

www.elsevier.com/locate/procedia

CENTERIS – International Conference on ENTERprise Information Systems / ProjMAN – International Conference on Project MANagement / HCist – International Conference on Health and Social Care Information Systems and Technologies 2022

Investigating conversational agents in healthcare: Application of a technical-oriented taxonomy

Kerstin Denecke[1*], Richard May[2]

[1]Bern University of Applied Sciences, Quellgasse 21, 2502 Biel, Swotzerland
[2]Harz University of Applied Sciences, Friedrichstrasse 54, 38855 Wernigerode, Germany

Abstract

Conversational agents (CA) are increasingly applied to realize health applications that collect patient data, provide information or even deliver health interventions. We developed a taxonomy focusing on technical characteristics of health CA with the purpose of creating a reporting guideline towards health CA and of building technical-oriented archetypes. The taxonomy comprises 18 dimensions which can be grouped into four perspectives. In this work, we wanted to find out whether the taxonomy is complete and can be applied appropriately by researcher to describe the technical characteristics of their health CA. Through a literature review, we identified 103 unique health CA for which publications have been published in 2021 and 2022. We contacted the corresponding or first authors of those papers asking for providing the information along our taxonomy for the CA described in their paper. For this purpose, our taxonomy was transformed into a questionnaire. To study applicability and understandability of the taxonomy, we also extracted the requested information from the papers using the taxonomy and compared the results to those of the participants. 95 E-Mails could be delivered. 26 persons out of 95 replied to our request resulting in a return rate of 27.3%. Results show that the majority of CA is simple in terms of CA personality; visualized as avatar or without embodiment. Systems are mainly rule-based, domain-specific and support one language. We recognized several differences between replies given by the participants and what has been extracted from the publications on the CA by us. We conclude that in order to apply the taxonomy as reporting guideline clear definitions must be given for the single characteristics. Some additional characteristics have to be added.

* Corresponding author. Tel.: +41 32 321 67 94. *E-mail address:* kerstin.denecke@bfh.ch

10.1016/j.procs.2023.01.413

Keywords: conversational agents; health chatbots; taxonomy; technology

1. Introduction

Conversational agents (CA) have existed for decades (1), however they became an increasing research topic in the last years: Interacting with a system in a human-like manner gained momentum. CA are dialog systems often integrating techniques from computer linguistics and allow interacting with users by means of natural language, speech or other modalities (2). Focusing on simulating the communication with a human, CA have many potential application areas. In healthcare, they have been used among other things for patient education (3), supporting behavior change (4) or to deliver cognitive behavior therapy (5). CA are considered to be useful for implementing (mobile) health applications because of their intuitive natural language interaction and their ability to deliver personalized and situation-dependent information and services (6). With respect to their technical implementation, we can roughly distinguish embodied CA from not-embodied CA and those with constrained user input from those with unconstrained user input. However, there are many other technical aspects leading to a highly diverse CA landscape. In this work, we focus on CA with applications in healthcare. We develop a technical-oriented taxonomy for health CA. The purpose of this work is to study its understandability. More specifically, we want to find out whether researcher can use the taxonomy to provide details on their health CA and whether it covers the relevant technical characteristics of current health CA.

In related work, some researchers already developed CA taxonomies. We found taxonomies dealing with a specific area within the health CA research field. For example, Welivita et al. (7) introduced a taxonomy of empathic chatbot responses. There are also reviews of chatbots and descriptions of specific characteristics without any systematic consolidation of the findings in terms of a taxonomy or archetypes. For example, Vaidyam et al. (8) reviewed chatbots and their particularities in mental health. Car et al. (9) conducted a scoping review on the characteristics of general health CA. Each of these research works characterize CA along some dimensions and characteristics. Overall, characterizing the landscape of healthcare CA in terms of technical aspects is still missing. From technical perspective other issues are important. For example, when a CA includes emotion recognition or sentiment analysis to generate empathy this would require additional technical evaluations. For these reasons, we developed a technical-oriented taxonomy, aiming at defining a reporting guideline, and forming technical archetypes of healthcare CA which in turn can be used to specify set of evaluation metrics per technical archetype.

Health-related CA are increasingly made accessible in app stores or studied in clinical trials. However, the aspects that are reported in literature on healthcare CA differs significantly. In particular concerning technical aspects, the descriptions lack relevant information. A recent paper studied how security and privacy is considered in healthcare CA (10). The authors found out that very limited information is presented on these aspects. In contrast to general domain CA, e.g., service hotlines, health CA have to address specific and often more complex requirements. Trust and reliability are essential since negative experiences with a system could negatively affect a patient's future help-seeking behavior (11). Health CA often store and process health data of individuals; data, which in the standard healthcare setting is considered highly sensitive, requiring high standards of security and privacy. Furthermore, health agents might require long-term engagement and pursue a user-specified goal (e.g., behavior change) which requires building a bond of trust between agent and user. In addition, the agent must be accurate and comprehensible to avoid harming the user through inappropriate, inaccurate or poorly communicated information. The objective of this work is to assess whether the technical-oriented taxonomy of healthcare CA developed by us, is complete and whether it can be applied by developers to describe their CA from a technical perspective.

2. Technical-oriented taxonomy of healthcare CA

Inspired by existing work on health CA (e.g. (3, 10, 12–16)) and discussions among the authors, we came up with a technical-oriented taxonomy, comprising 18 unique technical dimensions (see Table 1). The taxonomy comprises four perspectives: agent appearance, setting, interaction and data processing. From a technical perspective these are the four main facets of a healthcare CA. Each perspective integrates several dimensions which in turn have specific characteristics: Agent appearance (personality of CA, embodiment, application technology, intelligence framework, sentiment detection), setting (context, service duration, human involvement), interaction (input, output mode, service channel, device, language, integration mode) and data processing (Internet access, hosting, data exchange with 3rd party devices and services, data privacy). One possible application of that taxonomy is to apply it as reporting guideline for healthcare CA. This way it can be ensured that all technical information relevant for system quality and that might potentially impact on patient safety are reported. For this purpose, we want to assess whether our technical-oriented taxonomy is complete and is self-explanatory to be applied by others.

Table 1. Technical-oriented taxonomy of health CA

Agent appearance	Setting	Interaction	Data processing
Personality of CA 1) simple, 2) complex	**Context** 1) general purpose, 2) domain specific	**Input mode / output mode** 1) written, 2) spoken, 3) visual, 4) hybrid, 4) haptic	**Internet access** 1) online, 2) offline
Embodiment 1) no, 2) avatar, 3) physical	**Service duration** 1) ad-hoc supporters, 2) persistant companions, 3) temporary advisors	**Service channel** 1) smartphone embedded software, 2) social media, 3) website (web-based), 5) smart speaker	**Hosting** 1) local, 2) outsourced, 3) both
Application technology 1) virtual reality, 2) augmented reality, 3) vocal, 4) normal	**Human involvement** 1) diad, 2) triad, 3) quadriad	**Device** 1) PC, 2) mobile device, 3) both, 4) other	**Data exchange with 3rd party device or service** 1) access, 2) storing, 3) both, 4) none
Intelligence framework 1) rule-based, 2) self-learning		**Language** 1) single language, 2) multi language	**Data privacy** 1) privacy policy, 2) data encryption, 3) both, 4) nothing
Sentiment/emotion detection 1) yes, 2) no		**Integration mode** 1) stand-alone, 2) part of a system	

3. Material and methods

We first conducted a literature search to identify recent papers published in 2021 and 2022 in which implemented health CA are presented. We asked the first or corresponding authors of the resulting papers to apply our taxonomy. Second, based on our taxonomy, we extracted information from the papers found in the literature search from which we obtained responses. The results were analyzed in the context of a systematic literature review (SLR) to study: 1) the technical landscape of current health CA, and 2) study the applicability and completeness of our taxonomy.

3.1. Identifying healthcare CA: Conducting the literature search and review

To identify CA in healthcare, we used the SLR guidelines of Kitchenham et al. (17). We searched for relevant scientific papers on PubMed, ACM Digital Library, and IEEE Xplore. In contrast to searching for CA in app stores, a certain quality of the presented CA can be expected since these literature databases only list peer-reviewed publications. To identify appropriate literature, we defined the following search string: *(application OR app OR approach OR implementation) AND (chatbot OR bot OR conversation OR conversational user interface) AND (health OR healthcare)*. For PubMed we did not use the keywords *health* and *healthcare* since PubMed only lists publications from this domain. We included publications that were a peer-reviewed conference paper or journal article of original work dealing with a concrete CA applied in healthcare. Being aware that researchers quickly move to another research topic once papers are published, we only invited the authors of papers published in 2021 and 2022. We excluded papers not dealing with a concrete healthcare-related CA or only describing the design process, reviews or meta-analyses. The search was conducted on April 4 and 5, 2022, resulting in 80 papers from IEEE Xplore, 150 papers from ACM Digital Library and 215 papers from PubMed. After a first exclusion process, 138 retrieved items were assessed for eligibility by examining title and abstract where each reviewer looked at half of the papers. In a second round, the full texts were carefully considered by the reviewers to confirm eligibility. 103 papers on unique CA were considered in our taxonomy application process.

3.2. Questionnaire for applying the taxonomy

We contacted the corresponding or first authors of those 103 papers, asking a questionnaire that reflects our taxonomy. Thus, they characterized their CA based on our taxonomy. We transformed our taxonomy into a questionnaire, i.e. the technical dimensions were formulated as questions (e.g. "How would you describe the personality of your CA?") and answer options were the possible characteristics of this dimension. Correspondingly, the questionnaire comprised 18 questions, for each technical dimension of our taxonomy one question. The characteristics defined in the taxonomy were the only possible answers to be chosen from. At the beginning of the questionnaire, participants were asked to enter a code that enables us to link the reply to the corresponding paper from our review. Answers to the questionnaire were compared to CA classifications made by the authors of this paper based on the corresponding scientific publication. A last question asked for the technical evaluations that were already realized with the CA. E-Mails were sent out between June 17 and 21, 2022.

4. Results

4.1. Sample based on questionnaire

The questionnaire was open to fill from June 17-30, 2022. 103 persons were invited, but 8 mails could not be delivered. 26 persons out of 95 replied to our request resulting in a return rate of 27.3%. 4 replies refer to papers from 2022; 22 from 2021. Figure 1 shows the results.
Agent appearance: 92.3% of the systems included in this validation are claimed to have a simple personality. 53.8% have no embodiment, 38.5% have an avatar and 7.7% a physical embodiment. 76.9% are text-based; 1 system applies augmented reality and 19.2% uses vocal application technologies. 73% are rule-based; 27% are based on machine-learning. Interestingly, 42.3% integrate sentiment analysis (see Figure 1).
Setting: The context of the CA are 1) domain-specific (80.8%), (18–38), or 2) general purpose (19.2%), (39–43). Regarding service duration, medium- and long-term interaction seem to be prevalent: 46.2% of the systems are claimed as temporary advisor, 38.5% as persistent companion and only 15.4% are considered ad-hoc supporter. The CA interacts mainly with one user (80.8%). However, some systems involve additional persons (23.1%) or allow persons to access the conversation protocol.
Interaction: 69.2% of the CA, i.e. the majority, allow for written input and output. 19.2% claimed to have a hybrid input more; 7.7% allow for verbal input and 3.8% (1 CA) enables visual input. Regarding output modes, things differ slightly: 26.9% have a hybrid output mode and 3.8.% allow for spoken output. Half of the systems are provided as smart-phone embedded software; 23.1% through social media and 26.9% were web-based. 50% of the CA support PC and mobile device; 46.2.% run only on a mobile device; 1 system only on a PC. 73.1% support a single language, while 26.9% support multiple languages. 57.7% are stand-alone systems and 42.3% of the CA are part of a system.
Data processing: Privacy was not yet considered in 38.5% of the CA, 34.6% claim of having a privacy policy. Privacy and data encryption is included in 23.1% of the 26 CA.1 system states data encryption. 46.2% of the agents are hosted locally and the same percentage of the agents are outsourced. 88.5% require the internet. 23.1% of the 26 CA access 3rd party services or devices and another 19.2% access and store data with 3rd party services and devices. 57.7% do not interact with such services.
Aspects of evaluation: In addition, we asked the participants about which aspects have been evaluated on their CA. Most stated usability, ease of use, relevance of the output. Some evaluated healthcare outcome, acceptability, helpfulness, engagement. One group assessed aesthetics, enjoyment, privacy concerns, control, perceived usefulness. Only one participant answered (44), that their CA's Information Security Management System (ISMS) and Privacy Information Management System (PIMS) has been audited and certified by BSI for ISO 27001 and ISO 27701. Further, the developers regularly evaluate and update the AI models of this CA to ensure their clinical safety and performance.

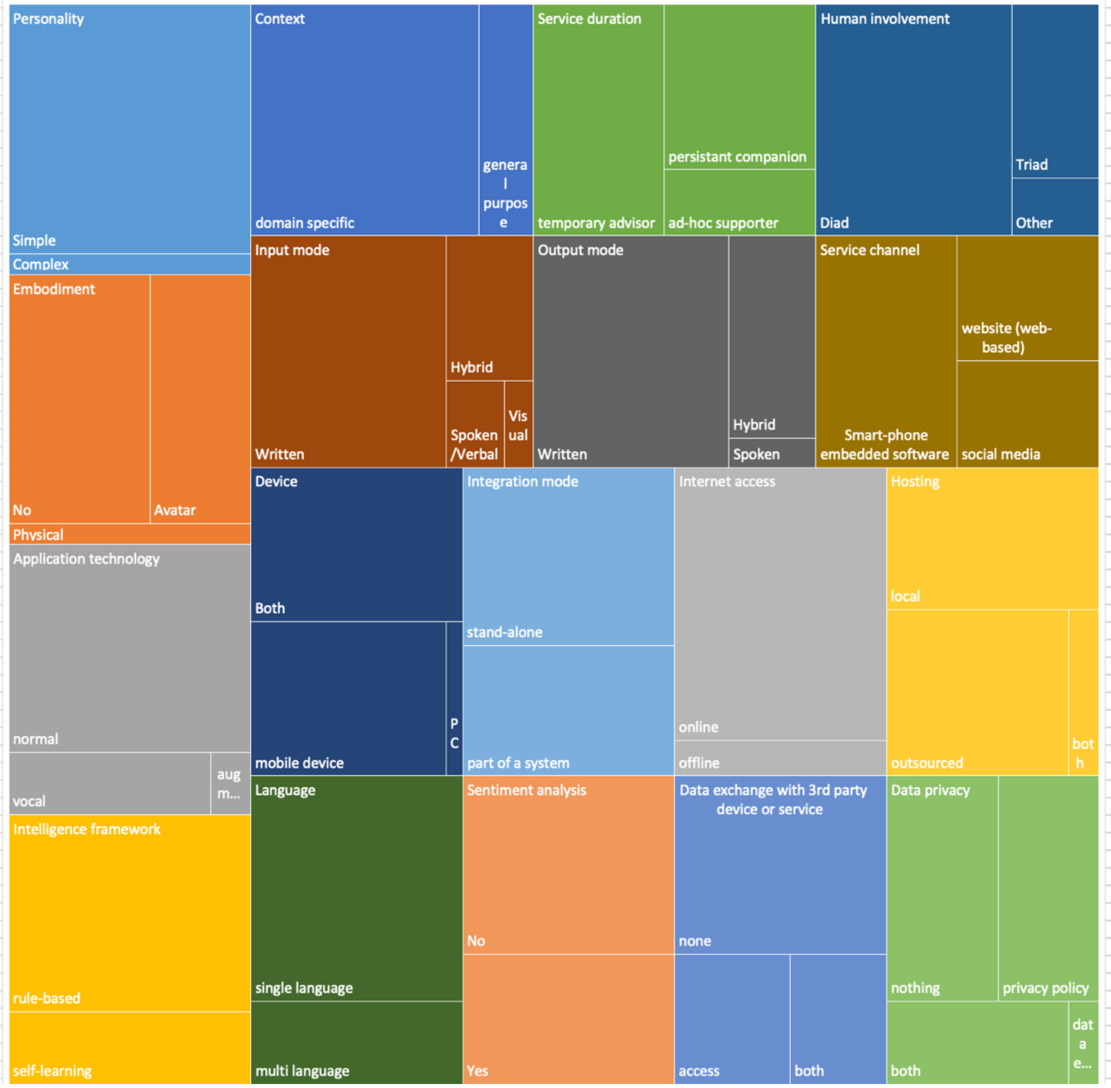

Fig. 1. Results from the questionnaire. Each color represents a technical dimension. Sizes of the rectangles correspond to the amount of CA with this particular characteristic for a dimension, n=26

4.2. Comparison to assignments based on the SLR

When comparing our annotations made based on the information available in the papers to the annotations from the paper authors, we can recognize differences. For dimensions within the perspective *data processing* corresponding information was very often missing in the papers. Therefore, only the questionnaire revealed the details how data processing is realized. We could not compare our results to the results received through the questionnaire since too many values were missing. When considering the remaining 13 dimensions; 102 of the 338 items differed, resulting in an agreement of 70% of the values. Furthermore, not all information needed to answer the questions were included in the papers. For example, for 2 CA we were unable to decide whether the CA applies machine learning or is rule-based.

We labelled 4 systems as based on machine-learning even though their developers replied they are rule-based. The *personality* was 5 times differently annotated by us; for dimension *embodiment* our assignments differed even 9 times. From the paper description it was obviously not clear that the CA use an avatar. Regarding *application technology*, we labeled 4 systems as “normal” while the participants selected “vocal”. We labelled two systems as based on virtual reality, which are not. Also integration of *sentiment analysis* in the CA was not made clearly in the papers: 5 systems more than we identified claimed to include sentiment analysis. 4 systems we considered domain-specific were annotated as general purpose.

Judging the *service duration* turned out to be challenging as well. For this dimension, our annotation differed for 15 out of 26 CA (57.7%). Distinguishing between ad-hoc supporter, temporary advisor and persistant companion is quite vague which obviously led to these differences. Our classifications also differed a lot from the external ones for the dimension *service channel* (50% of the CA were differently classified by us). The corresponding information was often not explicitly described which resulted in assumptions based on the given information. The same holds true for the dimension *device* where it was unclear to us that the CA is running on a mobile device and a PC.

4. Discussion

4.1 Principal results and practical implications

The main result of this study is that relevant technical details on data processing are missing in current publications on healthcare CA. Privacy and data encryption remains often unconsidered by researchers, not only ignored in the publications. The CA for which we received answers are mainly simple in terms of CA personality; visualized as avatar or without embodiment. Systems are mainly rule-based, domain-specific and support one language. We recognized several differences between replies given by the participants and what has been extracted from the publications on the CA. We conclude that in order to apply the taxonomy as reporting guideline clear definitions must be given for the single characteristics. One additional characteristic has to be added which is “stand-alone software” for dimension *service channel*. The results also show a clear need of harmonized reporting of technical details on healthcare CA. Without such information, patient safety and possible risks through the use of technology cannot be judged adequately. For some technical dimensions, we can recognize a clear trend on the most frequently used technology; for other dimensions it seems to be well distributed (see Figure 1).

4.2 Practical implications

Reasons for the differences in classifications by us and the developers of the CA are manifold. One reason might be that the information in the paper was insufficient. For example, one participant claimed, their CA is rule-based, but the paper clearly presents a machine learning model. Another reason might be that we did not provide any explanations on how to decide for a characteristic in the single dimensions. This gives room for interpretation which in turn can result in inconsistencies. For example, the different interpretations between us and the participants whether the CA has a simple or complex personality. However, without any definition, the taxonomy would be more easy to apply. A third reason might be that characteristics are missing (as one participant mentioned in the comment field) and since we forced participants in making a decision, we have to assume that a random value or a characteristic most close to reality was chosen. We conclude that for using the taxonomy as reporting guideline, concrete definitions must be provided and options have to be given for missing values. The study demonstrated once more that information on data processing of CA is often not reported in scientific literature. May et al. (10) found out already in their SLR that privacy and security in health CA is still not well considered by researchers. As stated by many participants, privacy and encryption is not even considered. A reason might be that these systems are developed for clinical trials and focus is on providing the functionalities. However, to ensure security and not to harm the trust of individuals in health IT in general and conversational agents in particular, we recommend to consider these aspects right from the beginning and to communicate them clearly to the users of such systems.

This is the first study where developers of healthcare CA are asked to provide technical details on their systems. Several reviews on healthcare CA exist where the authors extract features on the CA (1,8,9). All of these reviews specify their extraction schemes which include technical aspects to a certain extent. Laranjo et al. for example identified input and output modality, dialogue initiative, dialogue management (1). Our study concentrated on the technical aspects of healthcare CA and involved the actual developers in information gathering. To the best of our knowledge, this had not yet been done before in the healthcare domain.

4.3 Limitations

The questionnaire provides only a snapshot of the current technical landscape. We cannot say for sure that this reflects the complete landscape, thus it is only a trend. Correctness of the answers given in the questionnaire could only be approved in conversations with the participants asking for reasons why they stated something different that is written in the paper. Regarding methodology, we missed to ask the participants to give feedback on the clarity of the characteristics. Thus, we cannot conclude properly whether definitions are really required.

5. Conclusions

In this paper, we asked to apply a technical-oriented taxonomy to provide details on healthcare CA. Our taxonomy was almost complete, but clarity has to be improved by providing definitions for the single characteristics. Using it as reporting guideline or at least checklist when publishing on healthcare CA could help in getting correct details on the technical implementations. From a research perspective this would allow for easier comparison of systems. From a user perspective it would enable to better judge the technology a user is supposed to interact with. However, it would require technical knowledge to understand the characteristics. From a developer perspective, the application of the taxonomy could help in identifying design archetypes. It would also help to possibly associate technical implementations to outcomes and compare outcomes of different CA relating to the technical criteria. This would help in answering questions such as: for which use cases machine learning integrated in a health CA is useful.

Acknowledgements

This research was partially funded by Swiss National Science Foundation. We thank all participants who answered the questionnaire.

References

[1] Laranjo, L., 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. (2018) "Conversational agents in healthcare: a systematic review." *Journal of the American Medical Informatics Assocociation* **25 (9)**: 1248–1258.

[2] McTear, M. F., Z. Callejas, D. Griol. (2016) "The conversational interface." Springer.

[3] May, R., and K. Denecke. (2020) "Extending Patient Education with CLAIRE: An Interactive Virtual Reality and Voice User Interface Application". *EC-TEL*: 482–486.

[4] Larbi, D., E. Gabarron, K. Denecke. (2021) "Social media chatbot for increasing physical activity: usability study." *Studies in Health Technology and Informatics* **285**: 227–232.

[5] Prochaska, J. J., E. A. Vogel, A. Chieng, M. Kendra, M. Baiocchi, S. Pajarito, A. Robinson. (2021) "A therapeutic relational agent for reducing problematic substance use (Woebot): development and usability study." *Journal of Medical Internet Research* **23 (3)**: e24850.

[6] Knote, R., A. Janson, M. Söllner, J. Leimeister. (2019) "Classifying smart personal assistants: an empirical cluster analysis." *HICSS*: 1–10.

[7] Welivita, A. and P. A. Pu. (2020) "Taxonomy of Empathetic Response Intents in Human Social Conversations." *COLING*: 4886–4899.

[8] Vaidyam, A. N., H. Wisniewski, J. D. Halamka, M. S. Kashavan, J. B. Torous. (2019) "Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape." *The Canadian Journal of Psychiatry* **64 (7)**: 456–464.

[9] Tudor C. L., D. A. Dhinagaran, B. M. Kyaw, T. Kowatsch, S. Joty, Y. L. Theng. (2020) "Conversational Agents in Health Care: Scoping Review and Conceptual Analysis." *Journal of Medical Internet Research* **22 (8)**: e17158.

[10] May, R. and K. Denecke. (2021) "Security, privacy, and healthcare-related conversational agents: a scoping review." *Informatics for Health and Social Care*: 1–17.

[11] Miner, A. S., A. Milstein, J. T. Hancock. (2017) "Talking to machines about personal mental health problems." *JAMA* 318 (13): 1217–1218.

[12] Nißen, M., D. Selimi, A. Janssen, D. R. Cardona, M. H. Breitner, T. Kowatsch, F. von Wangenheim. (2022) "See you soon again, chatbot? A design taxonomy to characterize user-chatbot relationships with different time horizons." *Computers in Human Behavior* **127 (1)**: 107043.

[13] Diederich, S., A. Brendel, L. Kolbe. (2019) "Towards a taxonomy of platforms for conversational agent design." AIS eLibrary: 1100–1119.

[14] Gnewuch, U., S. Morana, A. Maedche. (2017) "Towards Designing Cooperative and Social Conversational Agents for Customer Service." *ICIS*: 1–13.

[15] Zierau, N., T. Wambsganss, A. Janson, S. Schöbel, J. M. Leimeister. (2020) "The anatomy of user experience with conversational agents: a taxonomy and propositions of service clues." *ICIS*: 1–17.

[16] Denecke K., S. Hochreutener, A. Pöpel, R. May. (2018) "Self-Anamnesis with a Conversational User Interface: Concept and Usability Study." *Methods of Information in Medicine* **57 (5/6)**: 243–52.

[17] Kitchenham, B. A., D. Budgen, O. P. Brereton. (2015) "Evidence-Based Software Engineering and Systematic Reviews." CRC Press.

[18] Badlani S., T. Aditya, M. Dave, S. Chaudhari. (2021) "Multilingual Healthcare Chatbot Using Machine Learning." INCET: 1–6.

[19] Chan W. W., E. E. Fitzsimmons-Craft, A. C. Smith, M. L. Firebaugh, L. A. Fowler, B. DePietro, et al. (2022) "The Challenges in Designing a Prevention Chatbot for Eating Disorders: Observational Study." *JMIR Formative Research* **6 (1)**: e28003.

[20] Gabrielli S., S. Rizzi, G. Bassi, S. Carbone, et al. (2021) "Engagement and Effectiveness of a Healthy-Coping Intervention via Chatbot for University Students During the COVID-19 Pandemic: Mixed Methods Proof-of-Concept Study." *JMIR mHealth & uHealth* **9 (5)**:e27965.

[21] García-Queiruga M., C. Fernández-Oliveira, M. J. Mauríz-Montero, Á. Porta-Sánchez, L. Margusino-Framiñán, I. Martín-Herranz. (2021) "Development of the @Antidotos_bot chatbot tool for poisoning management." *Farmacia Hospitalaria* **45 (4)**:180–183.

[22] Dhinagaran D. A., T. Sathish, A. Soong A, et al. (2021) "Conversational Agent for Healthy Lifestyle Behavior Change: Web-Based Feasibility Study." *JMIR Formative Research* **5 (12)**: e27956.

[23] Tian X., Z. Risha, I. Ahmed, A. B. Lekshmi Narayanan, J. Biehl. (2021) "Let's Talk It Out: A Chatbot for Effective Study Habit Behavioral Change." CSCW1: 1–32.

[24] Bézie A., V. Morisseau, R. Rolland, A. Guillemassé, B. Brouard, B. Chaix. (2022) "Using a Chatbot to Study Medication Overuse Among Patients Suffering From Headaches." *Frontiers in Digital Health* **4** :801782.

[25] Valtolina S., L. Hu. (2021) "Charlie: A chatbot to improve the elderly quality of life and to make them more active to fight their sense of loneliness." *CHItaly*: 1–5.

[26] Weber P., K. Krings, J. Nießner, S. Brodesser, T. Ludwig. (2021) "FoodChattAR: Exploring the Design Space of Edible Virtual Agents for Human-Food Interaction." *DIS*: 638–650.

[27] Kramer L. L., B. C. Mulder, L. van Velsen, E. de Vet. (2021) "Use and Effect of Web-Based Embodied Conversational Agents for Improving Eating Behavior and Decreasing Loneliness Among Community-Dwelling Older Adults: Protocol for a Randomized Controlled Trial." *JMIR Research Protocols* **10 (1)**: e22186.

[28] Hunt M., S. Miguez, B. Dukas, O. Onwude, S. White. (2021) "Efficacy of Zemedy, a Mobile Digital Therapeutic for the Self-management of Irritable Bowel Syndrome: Crossover Randomized Controlled Trial." *JMIR mHealth and uHealth* **9 (5)**: e26152.

[29] Dosovitsky G., E. Kim, E. L. Bunge. (2021) "Psychometric Properties of a Chatbot Version of the PHQ-9 With Adults and Older Adults." *Frontiers of Digital Health* **3**: 645805.

[30] Scholten M. R., S. M. Kelders, J. E. W. C. Van Gemert-Pijnen. (2021) "Can a monologue-style ECA more effectively motivate eHealth users in initial distress than textual guidance?" *Heliyon* **7 (3)**: e06509.

[31] Chatzimina M., H. Papadaki, C. Pontikoglou, L. Koumakis, K. Marias, M. Tsiknakis. (2021) "Designing a conversational agent for patients with hematologic malignancies: Usability and Usefulness Study." *BHI*: 1–4.

[32] Gupta M., T. Malik, C. Sinha. "Delivery of a Mental Health Intervention for Chronic Pain Through an Artificial Intelligence-Enabled App (Wysa): Protocol for a Prospective Pilot Study." *JMIR Research Protocols* **11 (3)**: e36910.

[33] Kowatsch T., T. Schachner, S. Harperink, F. Barata, U. Dittler, G. Xiao, et al. (2021) "Conversational Agents as Mediating Social Actors in Chronic Disease Management Involving Health Care Professionals, Patients, and Family Members: Multisite Single-Arm Feasibility Study." *Journal of Medical Internet Research* **23 (2)**: e25060.

[34] Barreto I. C. de H. C., N. B. S. Barros, R. L. Theophilo, V. F. Viana, F. R. de V. Silveira, O. de Souza, et al. (2021) "Development and evaluation of the GISSA Mother-Baby ChatBot application in promoting child health." *Cien Saude Colet* **26 (5)**: 1679–90.

[35] Maenhout L., C. Peuters, G. Cardon, S. Compernolle, G. Crombez, A. DeSmet. (2021) "Participatory Development and Pilot Testing of an Adolescent Health Promotion Chatbot." *Frontiers of Public Health* **9**: 724779.

[36] To Q. G., C. Green, C. Vandelanotte. (2021) "Feasibility, Usability, and Effectiveness of a Machine Learning-Based Physical Activity Chatbot: Quasi-Experimental Study." *JMIR mHealth and uHealth* **9 (11)**: e28577.

[37] van Cuylenburg H. C., T. N. D. S. Ginige. (2021) "Emotion Guru: A Smart Emotion Tracking Application with AI Conversational Agent for Exploring and Preventing Depression." *UCET*: 1–6.

[38] Roca S., S. Rosset, J. García, Á. Alesanco. (2022) "A Study on the Impacts of Slot Types and Training Data on Joint Natural Language Understanding in a Spanish Medication Management Assistant Scenario." *Sensors* **22 (6)**: 2364.

[39] Bhangdia Y., R. Bhansali, N. Chaudhari, D. Chandnani, M. L. Dhore. (2021) "Speech Emotion Recognition and Sentiment Analysis based Therapist Bot." *ICIRCA*: 96–101.

[40] Mokmin N. A. M., N. A. Ibrahim. (2021) "The evaluation of chatbot as a tool for health literacy education among undergraduate students." *Education and Information Technologies* **26 (5)**: 6033–6049.

[41] Pavic K., A. Oker, M. Chetouani, L. Chaby. (2021) "Age-related changes in gaze behaviour during social interaction: An eye-tracking study with an embodied conversational agent." *Quarterly Journal of Experimental Psychology* **74 (6)**: 1128–39.

[42] Ollier J., S. Neff, C. Dworschak, A. Sejdiji, P. Santhanam, R. Keller, et al. (2021) "Elena+ Care for COVID-19, a Pandemic Lifestyle Care Intervention: Intervention Design and Study Protocol." *Frontiers of Public Health* **9**: 625640.

[43] Chkroun M., A. Azaria. (2021) "A Safe Collaborative Chatbot for Smart Home Assistants." *Sensors* **21 (19)**: 6641.