Chapter

Multilingual Chatbots to Collect Patient-Reported Outcomes

Matej Rojc, Umut Ariöz, Valentino Šafran and Izidor Mlakar

Abstract

With spoken language interfaces, chatbots, and enablers, the conversational intelligence became an emerging field of research in man-machine interfaces in several target domains. In this paper, we introduce the multilingual conversational chatbot platform that integrates Open Health Connect platform and mHealth application together with multimodal services in order to deliver advanced 3D embodied conversational agents. The platform enables novel human-machine interaction with the cancer survivors in six different languages. The platform also integrates patients' reported information as patients gather health data into digital clinical records. Further, the conversational agents have the potential to play a significant role in healthcare, from assistants during clinical consultations, to supporting positive behavior changes, or as assistants in living environments helping with daily tasks and activities.

Keywords: embodied conversational agents, multimodal sensing, artificial intelligence, spoken language interfaces, cancer survivors

1. Introduction

An important type of patient-gathered health data (PGHD) represents so-called patient-reported outcomes (PROs). They are in general collected from patients in order to help address a health concern [1] and represent self-reports from everyday life. Therefore, in healthcare, they are also important data sources [2]. Further, PROs have become a complementary data source to telemonitoring [3], data mining, and imaging-based AI techniques [4–8]. Nowadays, the knowledge domains of clinical specialties are expanding rapidly, while due to the sheer volume and complexity of data, clinicians often fail to really exploit its potential [9]. Firstly, patient outcomes were collected mostly face to face, using paper-written forms [10–12]. Forms were added to paper-form health records (HRs), and only after the advances of information and communication technologies (ICT), the HRs are slowly being digitalized. Several studies already showed the efficiency of electronic questionnaire apps on, e.g., smartphones [13, 14]. Thus, electronic PROs, supported by artificial intelligence techniques, can further improve dropout and acceptance-rates. Further, they are also able to improve clinical and patient "satisfaction" [15–17]. A perfect example of how patient gathered health data (PGHD) and PROs are able to improve quality of life (QoL) is, e.g., ambient assisted living (AAL). Namely, AAL environments already exploit mobile devices, smart home products, software applications, and other wearable devices in the individual's everyday environment [17, 18].

Significant advances in speech and natural language processing (NLP) technologies already offer more personalized and human-like interaction, i.e., symmetric multimodality. Therefore, several spoken language interfaces, chatbots, and enablers, and the conversational intelligence became an emerging field of research in man-machine interfaces based on artificial intelligence techniques. Thus, embodied conversational agents (ECAs) can play an important role in healthcare, e.g., assistants in AAL environment in order to help with activities and daily tasks, or assistants during clinical consultations, in order to support positive behavior changes [19, 20]. These advanced interactive systems may certainly have a major impact on long-term sustainable quality of results and patient adherence over time.

The main challenges represent interoperability, integration of PGHD data, and lack of standardization [21, 22]. Namely, in healthcare, the integration of PGHD data in clinical decision-making still presents a big problem. Further, in the interoperability of electronic health records (EHRs), the unified representation of electronic health records (EHRs) still represents an issue. In order to get the highest contribution from PROs and PGHD, we considered the following: (i) "how to integrate data into clinical workflow?", (ii) "the cost and time for collecting PROs?", (iii) "how to efficiently collect data from patients?", and (iv) "how to enable proper interpretation by the clinicians?"

Within a Horizon 2020 project (PERSIST, https://projectpersist.com/, last accessed 19 June 2021), therefore, we propose a holistic system for collecting PROs remotely via both multilingual chatbots and ECAs. Further, the integration of PROs into the clinical workflow by using FHIR has been proposed. The FHIR server is located at the Open Health Connect (OHC) platform, and all traffic is orchestrated by a so-called multimodal sensing network (MSN) that runs several microservices, such as PLATTOS text-to-speech (TTS) system, ECA, RASA-based chatbot system, and SPREAD automatic speech recognition (ASR) system. In this way, we offer a fully symmetric model of interaction supporting speech, gesture, and facial expression on input and output. Further, the FHIR methodology is delivered as an enabler for efficient integration and a fully functional FHIR server [23].

The paper is structured as follows: in Section 2, related works and the ideas of our study will be presented. The PERSIST platform is described in Section 3, and fully symmetric ECA-based interaction model in Section 4. The results are presented in Section 5. In Section 6, the contributions of the PERSIST system are discussed, and the paper ends with the conclusions.

2. Related works

The paradigm of value-based healthcare represents a shift toward more efficient and more effective medical care. However, it requires additional sources of data to improve shared decision-making and enable more personalized decision-making. Therefore, conversational intelligence can significantly contribute to patient activation and engagement [24]. The technology is based on spoken language technologies (SLT), i.e., NLP, ASR, chatbot, and TTS, that enables machines to interact with humans in very natural way, using mobile or web platforms [25]. In healthcare, this started already in 1966 with ELIZA [26]. Nowadays, conversational agents have been used to solve much more complex tasks, such as booking tickets and acting as customer service agents [27]. In healthcare, conversational agents can provide patients with, e. g. personalized health and therapy information and relevant products and services. Additionally, they can connect them with healthcare providers, suggest diagnoses, and even recommended treatments based on patient symptoms and reports. Namely, multilingual communication, cost-effectiveness, and 24/7 availability make embodied conversational agents (ECAs) very useful for all those patients who have major medical concerns outside of doctor's operating hours. Several studies show that patients can perceive ECAs as interaction partners instead of human physicians and are able to trust them. Thus, they are willing to disclose medical information report more symptoms, etc. [28]. In oncology setting, CI (conversational intelligence) focuses mostly on (speech-enabled) chatbots [29]. They can contribute to lifestyle changes [30], to screening (i.e., iDecide [31]) and improving mental health state through managing psychological distress [32–34]. Therefore, chatbots are already well recognized as an enabler for adherence, active patient engagement, and satisfaction increase [35, 36]. However, the chatbots still tackle the long-term adherence with sustainable quality of the reported data [37]. In [36], they reported that active use of this technology drops already after 14 days. Namely, patients' understanding, their ability to remember the details, and perceived trustworthiness are the main factors of patient adherence [38]. Therefore, in the system of the PERSIST project, an ECA is additionally introduced. ECAs can undoubtedly increase this long-term adherence by engaging with users in interaction that is enriched by incorporating nonverbal communication [37]. Since ECA is autonomous and intelligent software entity with an embodiment used to communicate with the user [39], it can provide a system with symmetric multimodality based on speech, gesture, and facial expression. Embodiments can be designed as virtual human characters, animals, or robots [40–42]. Such fully symmetric interaction opens up the opportunity to introduce human-like qualities and significantly improves the believability of the human-machine interfaces [43]. ECAs in healthcare can be used for the treatment of mood disorders, anxiety, psychotic disorders, autism, substance use disorders, etc. [44]. In [17], ECAs already proved a promising tool for persuasive communication in healthcare. While in [42], technological and clinical possibilities of less complex ECAs were investigated, and ECAs are also shown to be a solution for routine applications in the means of rapid development, testing, and application. Stal in [45] also found out that the agents' textual output and/or speech as well as its gaze and facial expressions are the most important features. In general, for healthcare, ECA studies focused mainly on physical activity [46–48], stress [30], nutrition [49, 50], blood glucose monitoring [41], and sun protection [51]. However, there are several other studies that focus on speech, facial, and gaze expressions as the main design features [45]. ECAs in healthcare are mostly 2D-based, since gestures and appearance are not considered as main design features, and only a few studies addressed gestures.

In the PERSIST system, therefore, we use two 3D embodied conversational agents, female or male that can interact with patients in the following six languages: Slovenian, English, Spanish, French, Russian, and Latvian. ECAs are able to represent facial expressions and exploit gestures in order to enhance user experience. Namely, in this way, it is possible to better support verbal counterparts, regulate communicative relationships, and maintain clarity in the discourse.

In [52], the conversational agents are designed as a prototype, while the contribution to health-related outcomes is evaluated without relevant statistical significance. Further, Sayeed et al. in [53] describe an approach to create a patient-centered health system that is based on the FHIR standard and applications that can make requests and reports of HL7 FHIR resources.

3. The multilingual ECA-based PERSIST platform

3.1 The multilingual sensing network (MSN)

In **Figure 1**, we present the building blocks of the MSN network. The MSN consists of Apache Camel module. This module implements ActiveMQ Artemis, REST API, and Apache Kafka. In this way, we implemented specific machine-to-machine (M2M) communication between several services. The ActiveMQ Artemis module is then used for the MQTT broker. And Apache Kafka module is used for microservice architecture. The Apache Camel module is like a router in the system, since it has the ability to convert asynchronous to synchronous messages, or vice versa. We can run Apache Camel module also as a Spring Boot application in order to provide REST API end points for all HTTP requests. The MQTT broker in the system represents a link between mHealth app and OHC. Namely, the mHealth app is MQTT client that is just subscribed to ActiveMQ Artemis module. Further, microservices are using HTTP APIs and Kafka topics. For microservices, asynchronous communication is used. All predefined topics for dedicated language are supported. The synchronous communication is then used for RASA chatbot. In this case, HTTP REST requests are used and performed via Camel REST end points API.

In **Figure 2**, we can recognize two types of connections. The first one represents the synchronous connection used for communication over the secured application protocol HTTPS REST. It is needed for questionnaires, responses, and requests. The second one is then asynchronous connection. It is needed for the MQTT protocol, where we use MQTT topics. Established connections with the OHC platform can use

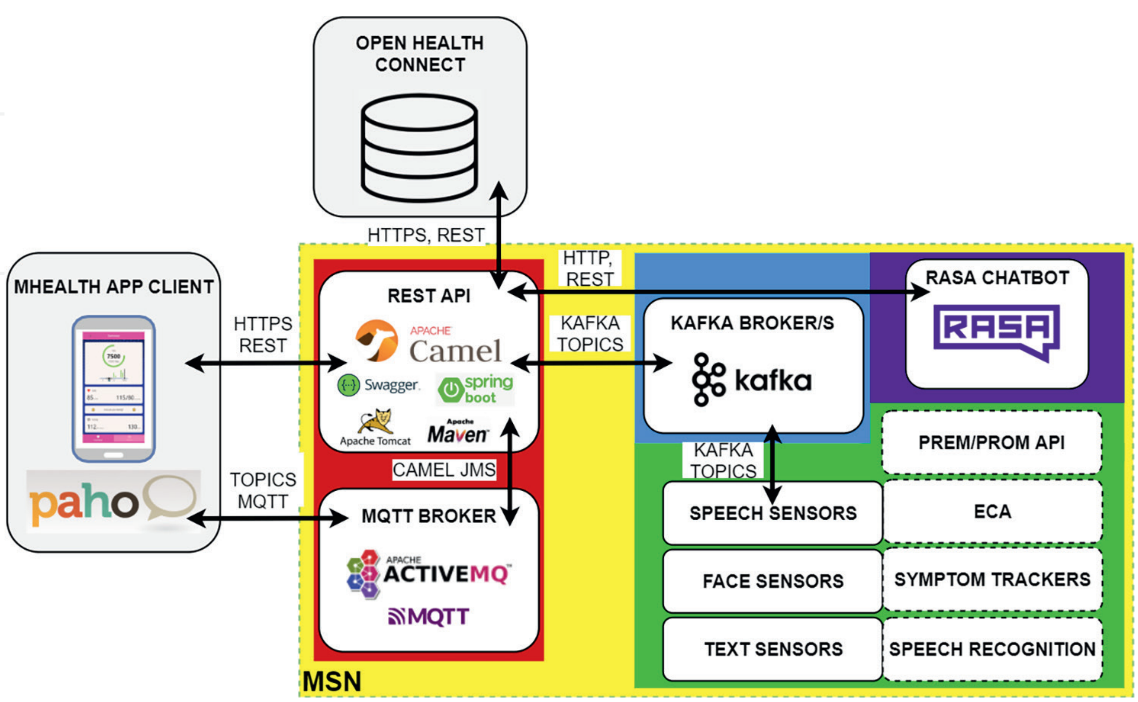

Figure 1.
The architecture of the PERSIST system.

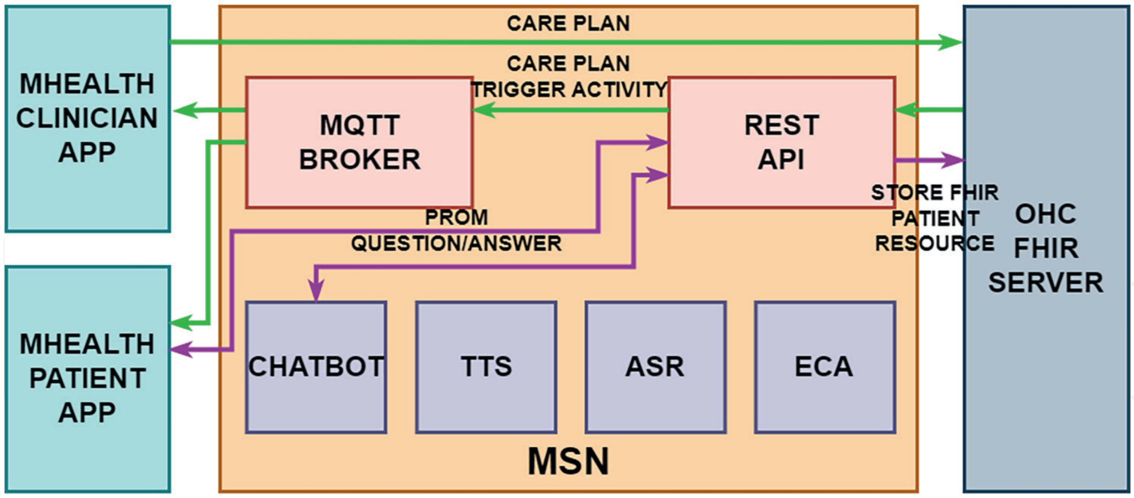

Figure 2.
Machine-to-machine communication (M2M) platform for the PERSIST system.

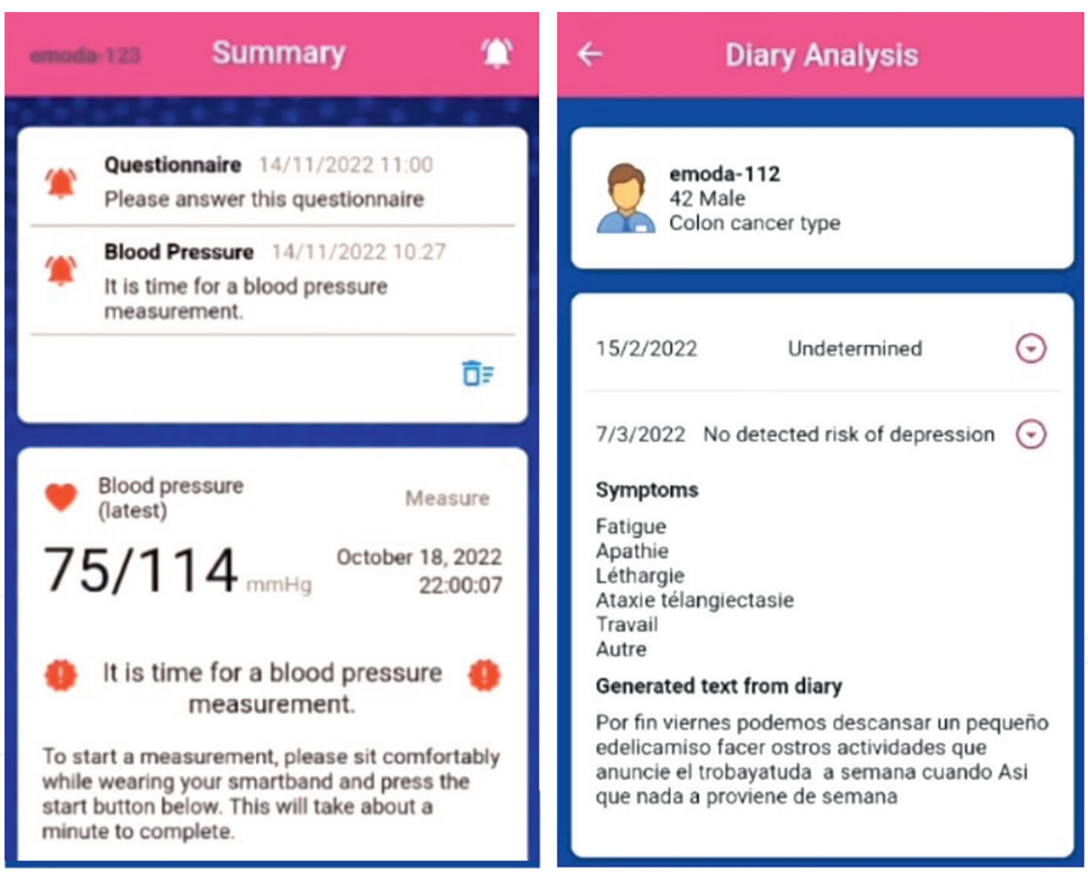

Figure 3.
Patient (left) and clinician (right) mHealth app interface.

in this way synchronous HTTPS REST protocol. Further, MSN internal connections uses MQTT, Camel Java Messaging Service (or JMS), Kafka topics, and REST.

3.2 mHealth application

As can be seen in **Figure 3**, patients and clinicians have separate mHealth applications. One is patient mHealth application that is used for data gathering and trends monitoring, while the clinician mHealth application is used for patient monitoring and specifying the patient's care plans (developed by company Emoda). The first one enables mood selection, diary recordings, reading of specific articles advised by clinicians, etc. And the second one has options to see the patients' lists and their clinical details. It is also possible to delete or edit existing patient records, or create a new one. Further, new appointments can be created by clinicians, receive notifications from patients, see the calendar, or just send/receive messages from patients. Thus, this application uses both asynchronous and synchronous protocols. We use the REST protocol for communication with the MSN REST OpenAPI (Swagger) and OHC end points, and for receiving notifications the MQTT protocol is used.

3.3 OHC FHIR server

The OHC platform has been provided by Dedalus. Basically, this is a streaming and integration platform that can be used for large--scale distributed environments. This digital health platform can also unlock isolated data. Further, OHC enables all the interfaces to be connected to and make decisions across disparate data sources in real time. It comprises a set of components, as depicted in the conceptual/logical architectures, is flexible, and can be deployed on private data center, or via cloud in environments like Azure or AWS. It provides the latest version of HAPI FHIR R4 [54].

4. The fully symmetric ECA-based interaction model

4.1 End-to-end multilingual text-to-speech synthesis system PLATTOS

Text-to-speech (TTS) PLATTOS in **Figure 4** is the first microservice in the PERSIST system. It is used for generating speech from text for the ECA agents that communicate with the patients. The PLATTOS system follows ideas presented in [55, 56] and enables real-time generation of speech in several languages, with practically human-like quality. It is basically the combination of two complex network models: a feature prediction NN model and a flow-based neural-network-vocoder WaveGlow.

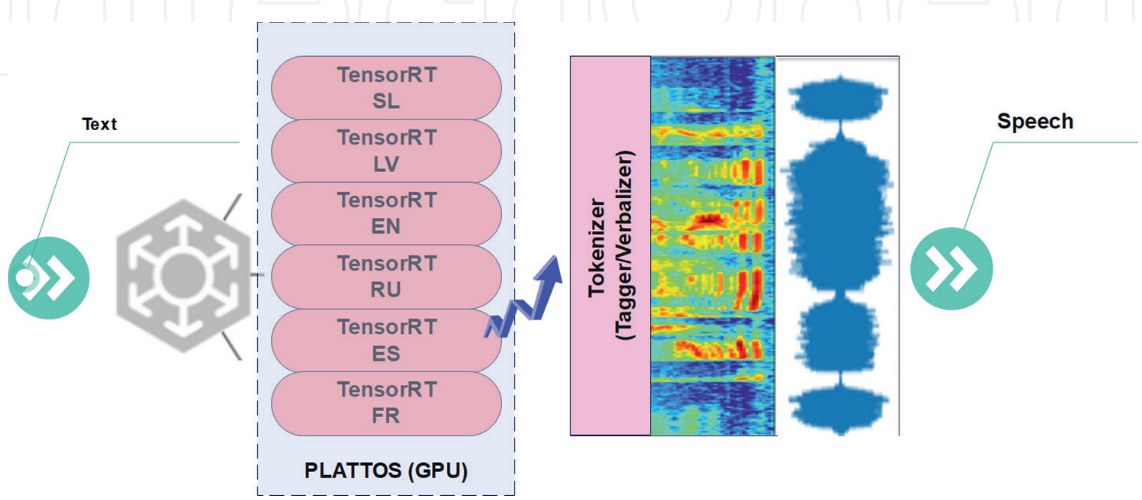

Figure 4.
TTS system PLATTOS.

4.2 End-to-end multilingual speech recognition system SPREAD

This microservice is developed to support the spoken language-based interface in the Health app and to feed the survivor's answers to the dialog management component (i.e., RASA chatbot) for several languages. E2E ASR system SPREAD in **Figure 5** follows some ideas from Jasper model [57–59], where the training has been improved by NovoGrad optimizer.

4.3 Embodied conversational system and embodied conversational agent

A RASA NLU [60] and ECA framework [61] are a core framework for an Embodied Conversational System (ECA). In this way, multilingual ECAs are capable of creating responses in natural language. All responses can also be visualized. Namely, multilingual chatbots are used to manage the more natural discourse between the system and patient. They are implemented as an API. Here, the NLU is the main engine of the chatbots and is programmed in Python and YAML language. Chatbots are all running on a Linux server. It implements standardized patient-reported outcomes (PROs) as storylines in six languages used in the PERSIST Clinical Study [62]. For storing the data, SQLite database within RASA is used, while POST and GET requests are used to store information, such as patients' answers, questionnaires, and other events that are triggered in a specific conversation.

The ECA framework is then used to transform plain text generated by the chatbot into ECA's multimodal responses incorporating gestures. The proprietary algorithm proposed in [61] has been used (**Figure 6**). It uses proprietary EVA-Script notations.

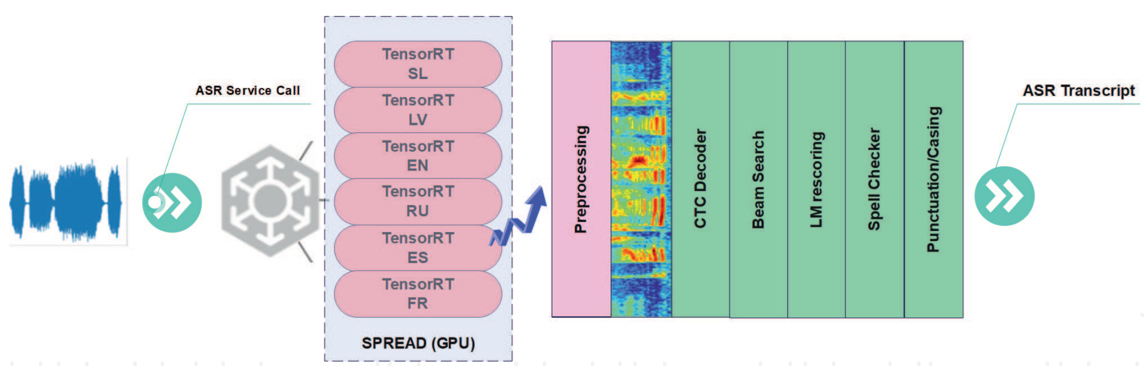

Figure 5.
ASR system SPREAD.

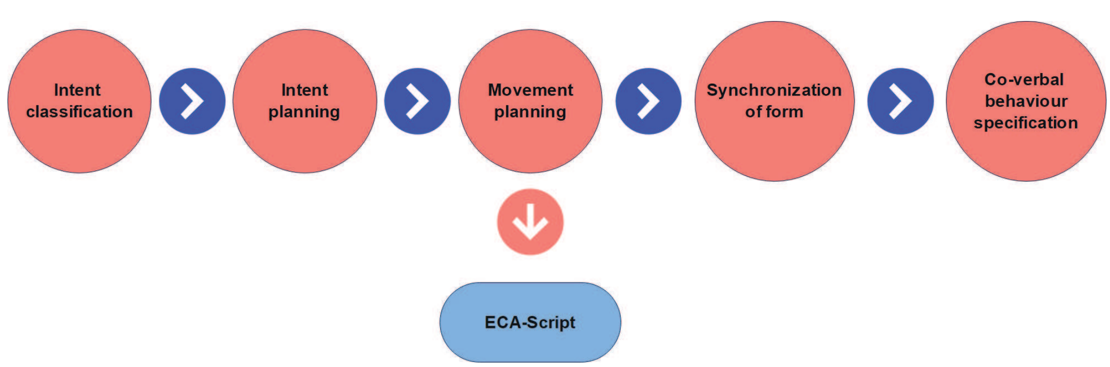

Figure 6.
Generation of expressive co-verbal behavior.

Each movement is formalized as a simultaneous execution within the block <bgesture>. The poses are described then within stroke phases, where the preparation phases are defined by <unit> blocks. Each <unit> also contains the complete configuration of individual movement controllers that are used in the representation of the specific pose. The retraction and hold phases then represent the shape being withheld or just retracted into some neutral state. They are both added within the <unit> by using attributes DurationHold and DurationRetraction.

5. Results

The PERSIST platform was deployed on two physical servers at the University of Maribor, FERI. The functional scheme of the system is highlighted in **Figure 7**. The PERSIST system is used mainly by the clinician. Namely, they have to define and schedule activities as part of patient's care workflow (phase 1). On the other hand, the patients execute activities (phase 3). MSN and OHC are the main services within the system. The MSN service is used to implement activities and make their execution more natural by delivering the symmetric model of interaction, and the OHC service is used to store data and automate the execution of the clinical workflow.

Questionnaires are available in six different languages: Slovenian, English, Russian, Latvian, French, and Spanish. On the output side, the system represents the information generated by chatbot as female ECA Eva and the male ECA Adam (**Figure 8**). In this way, in the output also non-verbal elements are associated with synthesized speech. In this way, raw texts are presented to the user as a multimodal output, which combines a spoken communication channel and synchronized visual communication

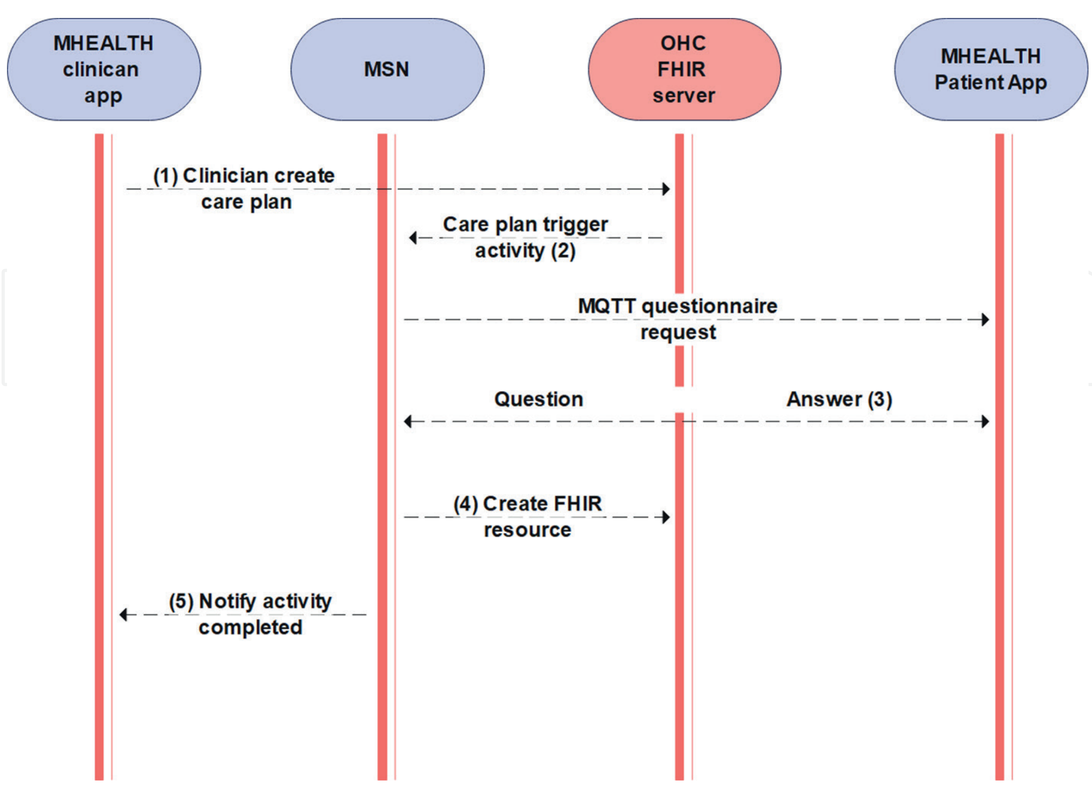

Figure 7.
Functional flow: Integration phases—Allocation of an activity (1), request for execution of the activity (2), implementation of the activity (3), creation of resource (4), and completion of the activity (5).

Figure 8.
Multimodal conversational response with ECAs.

channel. At the input, the system accepts speech or text. Additionally, a word-to-concept mapping is delivered as part of spoken language understanding. This is needed in order to properly map user responses into answers expected by PROs.

We deployed the system on a server hosting five virtual machines over the Proxmox VE 6.3–2. Further, the server is running the Xubuntu 20.04 LTS operating system. On the other platform, named PERSIST_INFERENCE, there are the Ubuntu Server 20.04 LTS OS, and microservices for ASR, TTS, and ECA. Microservices are integrated using predefined topics, and Kafka producers and consumers. To evaluate the hardware performance of the system, we simulated the load on the system by measuring CPU usage, memory usage, and average response time for both Camel and RASA chatbot. The results are outlined in **Figures 9-11**.

As seen in **Figure 9**, with the duplication of active users in tests the CPU usage is rising linearly from 11.65% with 25 active users to 56.04% with 1000 active users in the case of Camel, and mostly linear from 5.86% with 25 active users to 30.44% with 1000 active users for Rasa chatbot. The volatile memory was stagnating on both

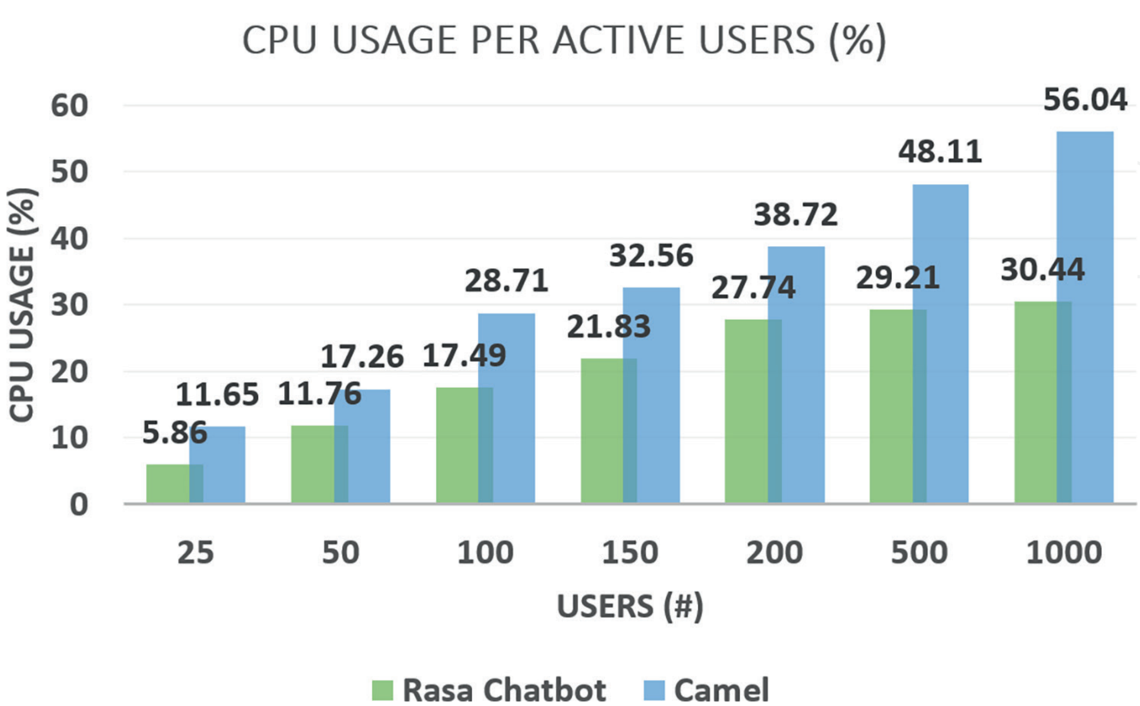

Figure 9.
CPU use (%) per active users.

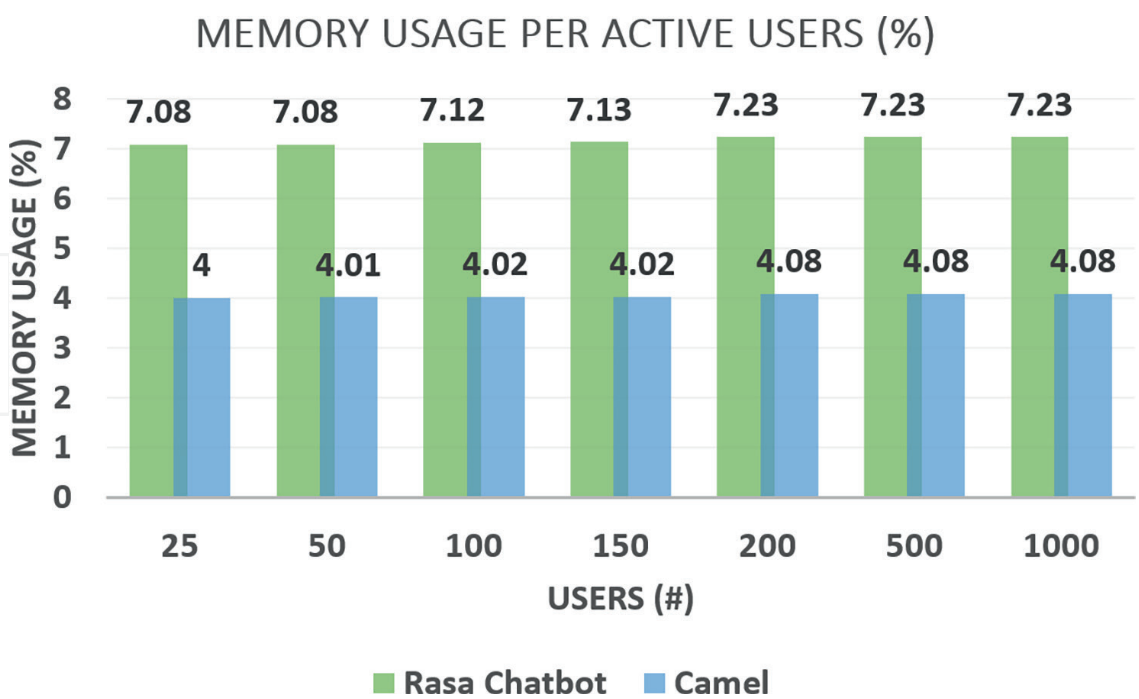

Figure 10.
Memory consumption (GB) per active users.

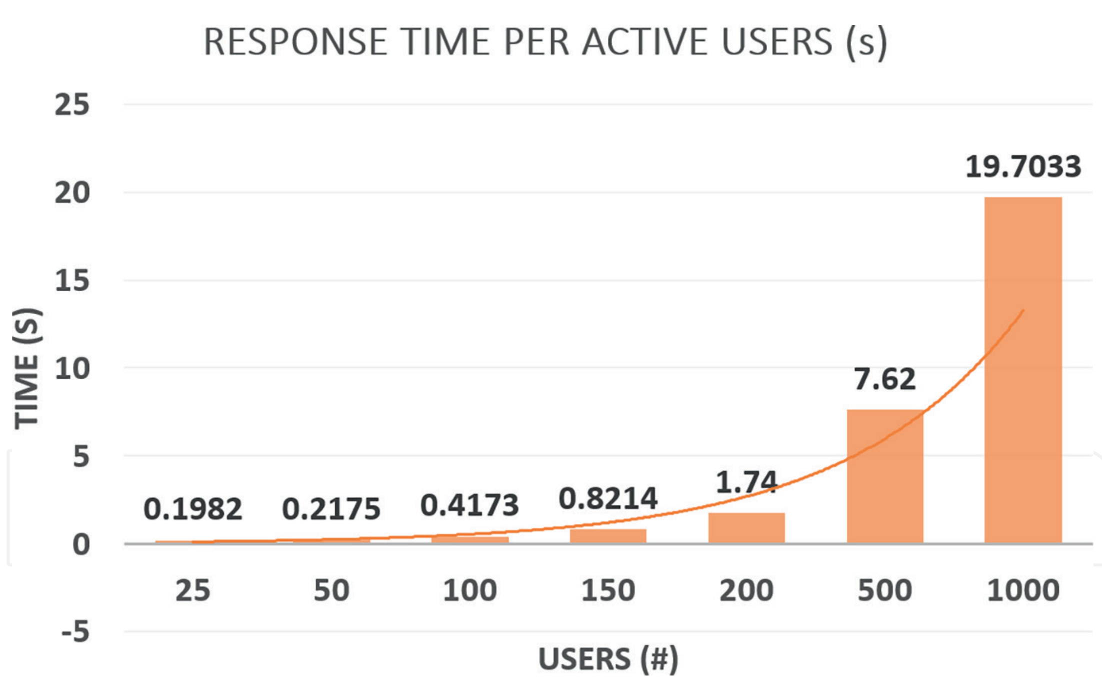

Figure 11.
Graphical results of average response time per active user.

the Camel and the Rasa chatbot and proved independent of the increase of users (**Figure 10**). In the case of the Camel, the memory usage was near 50%, while on the Rasa chatbot near 25%. Further, **Figure 11** presents the MSN's internal average response time on requests between 25 and 1000 active users. The response time in this case is 0.1982 s with 25 active users and is increasing linearly as the number of users is increasing. We have 1.74 s response time with 200 active users. Then it starts rising more exponentially to 197,033 s delay, with 1000 active users.

The models for the end-to-end ASR system SPREAD for six languages were trained on DGX-1, 8 × V100, 8 × 32 g GPU MEM, while the inference engine had 2 RTX8000, with 2 × 48 g GPUMEM. The audio datasets size used was minimal 1700 h of speech. The best model reached 2.6% WER, and all other models reached below 9% WER. The quality of the end-to-end TTS system PLATTOSand MUSHRA listening tests [63] were performed by PERSIST consortium partners. In this way, 21 consortium members participated, all in general with background knowledge in this field. Different TTS architectures were evaluated, while the architecture based on Tacotron and Waveglow was best rated. PLATTOS for all six languages was evaluated with score around 82 on 100 level scale. The results show that speech generated is highly intelligible and understandable. Further, the evaluation of the multimodal conversational response was reported in [61], where 30 individuals assigned an average score of 3.45 on the five-level Likert scale. The results show that the system produces a very viable and believable natural user interface.

6. Discussion

The main challenges for wide adaptation of PGHD in clinical practice include usability and sustainable quality of results (i.e., patient motivation and adherence) [21, 37]. The presented system includes patient/clinician mobile applications, OHC FHIR server, and the MSN server. OHC FHIR server provides interoperability between all components. The framework provides several tools that can be used for ingestion, indexing, storage, integration, and surfacing of patient information. In this way, the PERSIST system represents an open digital integration hub that can deliver scale, speed, and flexibility to securely gain value through the integration of health systems. Further, the OHC enables innovation through near-real-time access to longitudinal patient records, where the APIs provide opportunities to flexibly design services that can seamlessly ingest discrete data from the source into a third-party application. The FHIR has also been recognized as an approach suitable for citizen developers, since it also supports "low-code/no-code" solutions [21]. Our future efforts will be directed toward transformation and ingestion of EHRs from existing IT platforms into FHIR ready server. Based on the studies, the main activities will involve the definition of an ontology that will correlate existing fields with specific FHIR resources. The information in existing EHRs is mostly stored as partially structured or unstructured text; therefore, a specific focus will be directed toward extracting information by using modern NLP techniques and data to concept mapping.

The other challenge relates to the patient's perspective and long-term sustainability and quality of collected information [36, 37]. Perceived complexity and trustworthiness represent also the main drivers of patient adherence [38]. Therefore, MSN delivers the necessary microservice infrastructure, where the services are distributed among the servers and can be replicated if needed. A fully articulated ECA was deployed for all six languages in order to implement more natural human-machine interaction, where the EVA realization framework transforms the co-verbal descriptions contained in EVA events into articulated movement generated by the expressive virtual entity. The EVA-Script language is actually applied onto the articulated 3D model EVA in the form of animated movement [43]. Trustworthiness is a clinical value, which has a significant impact on adherence mitigating pervasive threats to health [64]. The symmetric multimodal model for dialog systems enables the ECAs to deliver and to understand input/output modes, including speech, gestures, and facial expressions. This makes the interfaces more familiar and trustworthy [38], where trustworthiness is one of the building blocks of patient compliance and responsiveness [65].

The RASA chatbot API is using PREMs and PROMs to see the patients' health status and the patients' perceptions of their experience while receiving treatment. In this case, we created several stories that contained probable conversations with patients. These are basically the intents that have to be executed, based on patient's responses [66]. Inclusion of multilingual ECAs have positive effect on patient adherence, as also several experiments imply. Further, ECAs contribute to long-term sustainability and familiarity [29] and decrease the complexity of user interfaces. Namely, having a virtual body that shows the nonverbal cues can provide easier understanding of the context, coherence for information exchange, and an increase for believability and trustworthiness to the virtual entity.

However, the phenomenon of "uncanny valley" may have significant negative impact on the overall user experience with articulated entities compared to "disembodied" agents as suggested in [67]. Thus, in the future, we will focus specifically on the synchronization issues of nonverbal behavior with speech.

7. Conclusions

In this paper, a multilingual holistic approach toward sustainable collection of PGHD and PROs and their efficient integration into clinical workflow has been presented. Namely, the PGHD may contribute to personalized care and early identification related to psychological and physiological symptoms and negative health outcomes. The PERSIST system represents an opportunity to integrate the benefits and deliver them to the patients. The system consists of patient/clinician mobile applications, an OHC FHIR server, and a MSN server. The research and this study address several technologies from the prototype (proof-of-concept) perspective. The used technology was evaluated on modular basis, statistically, and on a short-term-use basis.

Acknowledgements

This study is part of the project "PERSIST: Patient-centered survivorship care plan after cancer treatment" that has received funding from the European Union's Horizon 2020 research and innovation program (GA No. 875406). This work is partially financed by the HosmartAI project, funded by European Union's Horizon 2020 research and innovation programme under grant agreement No 101016834.

Conflict of interest

The authors declare no conflict of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

Author details

Matej Rojc*, Umut Ariöz, Valentino Šafran and Izidor Mlakar
Faculty of Electrical Engineering and Computer Science, University of Maribor, Slovenia

*Address all correspondence to: matej.rojc@um.si

References

[1] National Health Council. What are Clinician-Reported Outcomes (ClinROs)? National Health Council. 2019. Available online: https://nationalhealthcouncil.org/coa-series-what-are-clinician-reported-outcomes-clinros/ [Accessed: June 19, 2021]

[2] Health IT, Office of the National Coordinator for Health Information Technology (ONC), US Department of Health Human Services. What are patient-generated health data?. Available online: http://healthit.gov/topic/otherhot-topics/what-are-patientgenerated-health-data [Accessed: October 15, 2019]

[3] Fauzana N, Gulcharan BI, Azhar MA, Daud H, Mohd NN, Taib I. Integrating emerging network technologies to heart rate monitoring system to investigate transmission stability and accuracy: Preliminary results. International Journal of Electrical Engineering Computer Science (EEACS). 2021;**3**:21-22

[4] Sawssen B, Okba T, Noureeddine L. A mammographic images classification technique via the Gaussian radial basis kernel ELM and KPCA. International Journal of Applied Mathematical Computer Science and System Engineering. 2020;**2**:92-98

[5] Zheng Q, Yang L, Zeng B, Li J, Guo K, Liang Y, et al. Artificial intelligence performance in detecting tumor metastasis from medical radiology imaging: A systematic review and meta-analysis. EClinical Medicine. 2021;**31**:100669

[6] Inès A, Zgaya H, Slim H. Workflow tool to model and simulate patients paths in Pediatric Emergency Department. International Journal of Electrical Engineering Computer Science. 2020;**2**:73-78

[7] Abdelnabi MLR, Jasim MW, El-Bakry HM, Taha MHN, Khalifa NEM, Loey M. Breast and colon cancer classification from gene expression profiles using data mining techniques. Symmetry. 2020;**12**:408

[8] Austin E, LeRouge C, Hartzler AL, Segal C, Lavallee DC. Capturing the patient voice: Implementing patient-reported outcomes across the health system. Quality of Life Research. 2020;**29**:347-355

[9] Groccia MC, Guido R, Conforti D. Multi-classifier approaches for supporting clinical decision making. Symmetry. 2020;**12**:699

[10] Ellwood PM. Outcomes Management. The New England Journal of Medicine. 1988;**318**:1549-1556

[11] Tarlov AR, Ware JE, Greenfield S, Nelson EC, Perrin E, Zubkoff M. The medical outcomes study: An application of methods for monitoring the results of medical care. JAMA. 1989;**262**:925-930

[12] Bielli E, Carminati F, La Capra S, Lina M, Brunelli C, Tamburini M. A wireless Health outcomes monitoring system (WHOMS): Development and field testing with cancer patients using mobile phones. BMC Medical Informatics and Decision Making. 2004;**4**:1-13

[13] Tran C, Dicker A, Leiby B, Gressen E, Williams N, Jim H. Utilizing digital health to collect electronic pa-tient-reported outcomes in prostate cancer: Single-arm pilot trial. Journal of Medical Internet Research. 2020;**22**:e12689

[14] Wright AA, Raman N, Staples P, Schonholz S, Cronin A, Carlson K, et al. The HOPE pilot study: Harnessing patient-reported outcomes and biometric data to enhance cancer care. JCO Clinical Cancer Information. 2018;**2**:1-12

[15] Rajguru P, Ryan S, McLaurin E, Wirta D, Grieco J. A novel method for collecting patient reported outcomes (PROs): Developing and validating electronic PROs on a mobile smartphone platform. Investigative Ophthalmology & Visual Science. 2020;**7**:110

[16] Van Egdom LSE, Pusic A, Verhoef C, Hazelzet JA, Koppert LB. Machine learning with PROs in breast cancer surgery; caution: Collecting PROs at baseline is crucial. The Breast Journal. 2020;**26**:1213-1215

[17] Kramer LL, Ter Stal S, Mulder B, De Vet E, Van Velsen L. Developing embodied conversational agents for coaching people in a healthy lifestyle: Scoping review. Journal of Medical Internet Research. 2020;**22**:e14058

[18] Queirós A, Dias A, Silva AG, Rocha NP. Ambient assisted living and health-related out-comes-a systematic literature review. Inform. 2014;**4**:19

[19] Alosaimi W, Ansari TJ, Alharbi A, Alyami H, Seh A, Pandey A, et al. Evaluating the impact of different symmetrical models of ambient assisted living systems. Symmetry. 2021;**13**:450

[20] Laranjo L, Dunn A, Tong HL, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: A systematic review. Journal of the American Medical Informatics Association. 2018;**25**:1248-1258

[21] Jim HSL, Hoogland A, Brownstein NC, Barata A, Dicker AP, Knoop H, et al. Innovations in research and clinical care using patient-generated health data. CA: A Cancer Journal for Clinicians. 2020;**70**:182-199

[22] Rehman A, Naz S, Razzak I. Leveraging big data analytics in healthcare enhancement: Trends, challenges and opportunities. Multimedia Systems. 2021;**2021**:1-33

[23] Resourcelist—FHIR v4.0.1. Available online: http://hl7.org/fhir/resourcelist.html [Accessed: April 1, 2021]

[24] Wald JS, Sands DZ. Transforming Health care delivery through consumer engagement, Health data transparency, and patient-generated Health information. Yearbook of Medical Informatics. 2014;**23**:170-176

[25] Shawar B, Atwell E. Chatbots: Are they really useful? LDV Forum. 2007;**22**:29-49

[26] Weizenbaum J. ELIZA—A computer program for the study of natural language communication between man and machine. Communications of the ACM. 1966;**9**:36-45

[27] Sharma RK, Center NI. An analytical study and review of open source Chatbot framework, Rasa. International Journal of Engineering Research. 2020;**9**:060723

[28] Palanica A, Flaschner P, Thommandram A, Li M, Fossat Y. Physicians' perceptions of Chatbots in Health care: Cross-sectional web-based survey. Journal of Medical Internet Research. 2019;**21**:e12887

[29] Bibault J-E, Chaix B, Nectoux P, Pienkowski A, Guillemasé A, Brouard B. Healthcare ex Machina: Are conversational agents ready for prime time in oncology? Clinical Translational and Radiational Oncology. 2019;**16**:55-59

[30] Gardiner PM, McCue KD, Negash LM, Cheng T, White LF, Yinusa-Nyahkoon L, et al. Engaging women with an embodied conversational agent to deliver mindfulness and lifestyle recommendations: A feasibility randomized control trial. Patient Education and Counseling. 2017;**100**:1720-1729

[31] Owens OL, Felder T, Tavakoli AS, Revels AA, Friedman DB, Hughes-Halbert C, et al. Evaluation of a computer-based decision aid for promoting informed prostate Cancer screening decisions among African American men: iDecide. American Journal of Health Promotion. 2019;**33**:267-278

[32] Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. JMIR Mental Health. 2017;**4**:e19

[33] Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: Real-world data evaluation mixed-methods study. JMIR mHealth and uHealth. 2018;**6**:e12106

[34] Ly KH, Ly A-M, Andersson G. A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. Internet Interventions. 2017;**10**:39-46

[35] Girgi A, Durcinoska I, Levesque JV, Gerges M, Sandell T, Arnold A, et al. The PROMPT-care program group eHealth system for collecting and utilizing patient reported outcome measures for personalized treatment and care (PROMPTCare) among cancer patients: Mixed methods approach to evaluate feasibility and acceptability. Journal of Medical Internet Research. 2017;**19**:e330

[36] Kneuertz PJ, Jagadesh N, Perkins A, Fitzgerald M, Moffatt-Bruce SD, Merritt RE, et al. Improving patient engagement, adherence, and satisfaction in lung cancer surgery with implementation of a mobile device platform for patient reported outcomes. Journal of Thoracic Disease. 2020;**12**:6883-6891

[37] Tellols D, Lopez-Sanchez M, Rodríguez I, Almajano P, Puig A. Enhancing sentient embodied conversational agents with machine learning. Pattern Recognition Letters. 2020;**129**:317-323

[38] Martin LR, Williams SL, Haskard KB, DiMatteo MR. The challenge of patient adherence. Therapeutics and Clinical Risk Management. 2005;**1**:189-199

[39] Isbister K, Doyle P. The blind men and the elephant revisited evaluating interdisciplinary ECA research. In: Ruttkay Z, Pelachaud C, editors. From Brows to Trust Evaluating Embodied Conversational Agents. Dordrecht, The Netherlands: Springer; 2004. pp. 3-26

[40] Bickmore T, Gruber A, Picard R. Establishing the computer–patient working alliance in automated health behavior change interventions. Patient Education and Counseling. 2005;**59**:21-30

[41] Klaassen R, Bul KCM, Akker ROD, Van Der Burg GJ, Kato PM, Di Bitonto P. Design and evaluation of a pervasive coaching and gamification platform for young diabetes patients. Sensors. 2018;**18**:402

[42] Provoost S, Lau HM, Ruwaard J, Riper H. Embodied conversational agents in clinical psychology: A scoping review.

Journal of Medical Internet Research. 2017;**19**:e151

[43] Rojc M, Kăčič Z, Mlakar I. Advanced content and Interface personalization through conversational behavior and affective embodied conversational agents. In: Fernandez MAA, editor. Artificial Intelligence Emerging Trends and Applications. London, UK: IntechOpen; 2018

[44] Brinkman WP. Virtual health agents for behavior change: Research perspectives and directions. In: IVA2016 Workshop: Graphical and Robotic Embodied Agents for Therapeutic Systems - GREATS16. Los Angeles, CA, USA: Institute for Creative Technologies, USC; 2016

[45] Stal S, Kramer LL, Tabak M, Akker HOD, Hermens H. Design features of embodied conversational agents in eHealth: A literature review. International Journal of Human Computer Studies. 2020;**138**:102409

[46] Friederichs S, Bolman C, Oenema A, Guyaux J, Lechner L. Motivational interviewing in a web-based physical activity intervention with an Avatar: Randomized controlled trial. Journal of Medical Internet Research. 2014;**16**:e48

[47] Bickmore TW, Caruso L, Clough-Gorr K, Heeren T. 'It's just like you talk to a friend' relational agents for older adults. Interacting with Computers. 2005;**17**:711-735

[48] Ellis T, Latham NK, DeAngelis TR, Thomas CA, Saint-Hilaire M, Bickmore TW. Feasibility of a virtual exercise coach to promote walking in community-dwelling persons with Parkinson disease. American Journal of Physical Medicine & Rehabilitation. 2013;**92**:472-485

[49] Henkemans BOA, van der Boog PJ, Lindenberg J, van der Mast CA, Neerincx MA, Zwetsloot-Schonk BJ. An online lifestyle diary with a persuasive computer assistant providing feedback on self-management. Technology and Health Care. 2009;**17**:253-267

[50] Bickmore TW, Schulman D, Sidner C. Automated interventions for multiple health behaviors using conversational agents. Patient Education and Counseling. 2013;**92**:142-148

[51] Sillice AM, Morokoff PJ, Ferszt G, Bickmore T, Bock BC, Lantini R, et al. Using relational agents to promote exercise and sun protection: Assessment of participants' experiences with two interventions. Journal of Medical Internet Research. 2018;**20**:e48

[52] Benze G, Nauck F, Alt-Epping B, Gianni G, Bauknecht T, Ettl J, et al. PROutine: A feasibility study assessing surveillance of electronic patient reported outcomes and adherence via smartphone app in advanced cancer. Annals of Palliative Medicine. 2019;**8**:104-111

[53] Sayeed R, Gottlieb D, Mandl KD. SMART markers: Collecting patient-generated health data as a standardized property of health information technology. NPJ Digital Medicine. 2020;**3**:1-8

[54] Versions—FHIR v4.0.1. Available online: https://www.hl7.org/fhir/versions.html [Accessed: April 1, 2021]

[55] Shen J, Pang R, Weiss RJ, Schuster M, Jaitly N, Yang Z, et al. Natural tts synthesis by condi-tioning wavenet on mel spectrogram predictions. In: Proceedings of the 2018 IEEE

International Conference on Acoustics, Speech and Signal Processing. (ICASSP), 15-20 April 2018. Calgary, AB, Canada: IEEE Press; 2018. pp. 4779-4783

[56] Oord AVD, Dieleman S, Zen H, Simonyan K, Vinyals O, Graves A, et al. Wavenet: A generative model for raw audio. arXiv 2016, arXiv:1609.03499.57

[57] Li J, Lavrukhin V, Ginsburg B, Leary R, Kuchaiev O, Cohen JM, et al. Jasper: An End-to-End Convolutional Neural Acoustic Model. 2019:71-75. DOI: 10.21437/Interspeech.2019-1819

[58] Graves A, Jaitly N. Towards end-to-end speech recognition with recurrent neural networks. In: Proceedings of the International Conference on Machine Learning (ICML 2014), 21-26 June 2014. Vol. 32. Beijing, China : JMLP, W&CP; 2014. pp. 1764-1772

[59] Chorowski J, Bahdanau D, Cho K, Bengio Y. End-to-end continuous speech recognition using attention-based recurrent nn: First results. In NIPS 2014 Workshop on Deep Learning. Dec 2014

[60] Bocklisch T, Faulkner J, Pawlowski N, Nichol A. Rasa: Open source language understanding and dialogue management. arXiv 2017. preprint arXiv:1712.05181

[61] Rojc M, Mlakar I, Kacic Z. The TTS-driven affective embodied conversational agent EVA, based on a novel conversational behavior generation algorithm. Engineering Applications of Artificial Intelligence. 2017;**57**:80-104

[62] Mlakar I, Smrke U. Clinical Study to Assess the Outcomes of a Patient-Centred Survivorship Care Plan Enhanced with Big Data and Artificial Intelligence Technologies. 2021. Available online: https://www.isrctn.com/ISRCTN97617326 [Accessed: June 19, 2021]

[63] Schoeffler M, Bartoschek S, Stöter FR, Roess M, Westphal S, Edler B, et al. Web MUSHRA—A comprehensive framework for web-based listening tests. Journal of Open Research Software. 2016:6. DOI: 10.5334/jors.187

[64] H2020 Project PERSIST. Available online: https://projectpersist.com/ [Accessed: May 31, 2020]

[65] Sofer C, Dotsch R, Wigboldus DH, Todorov A. What is typical is good: The influence of face typicality on perceived trustworthiness. Psychological Science. 2015;**26**(1):39-47

[66] Singh A, Ramasubramanian K, Shivam S. Introduction to Microsoft Bot, RASA, and Google Dialogflow. In: Building an Enterprise Chatbot: Work with Protected Enterprise Data Using Open Source Frameworks. Apress. Berkeley, CA, USA: Springer; 2019. pp. 281-302

[67] Ciechanowski L, Przegalinska A, Magnuski M, Gloor P. In the shades of the uncanny valley: An experimental study of human–chatbot interaction. Future Generation Computer Systems. 2018;**92**:539-548