2020 IEEE International Conference on E-health Networking, Application & Services (HEALTHCOM) | 978-1-7281-6267-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/HEALTHCOM49281.2021.9399013

DocPal: A Voice-based EHR Assistant for Health Practitioners

Vidisha Bhatt
Computer Science Department
North Dakota State University
Fargo, USA
vidisha.n.bhatt@ndsu.edu

Juan Li
Computer Science Department
North Dakota State University
Fargo, USA
j.li@ndsu.edu

Bikesh Maharjan
Computer Science Department
North Dakota State University
Fargo, USA
bikesh.maharjan@ndsu.edu

***Abstract*— Electronic health record (EHR) systems have been widely adopted across healthcare organizations. While there are many benefits of using EHR such as improved accessibility and secure sharing of patient data, a shortcoming is that its manual data input is time-consuming and error prone. Physicians spend as much as 49.2% of their office time on EHR. In this paper, we present the design, development, and evaluation of a voice-based assistant, DocPal, to assist healthcare practitioners to access and update EHR through their voice. User survey and experimental evaluation illustrate that DocPal has good usability, time efficiency, and accuracy. When applied in the healthcare industry, we expect it to reduce data entry time and provide better patient care.**

Keywords-voice assistant; electronic health record; Alexa; workflow

I. Introduction

The importance of having quality patient-provider time has long been recognized by the healthcare industry [1]. The evolution of electronic health record (EHR) systems is an example project of realizing this goal. EHR systems enable healthcare organizations to store and retrieve electronic records of health-related information of patients [2]. They have been used to provide substantial benefits to physicians, clinic practices, and health care organizations. The benefits include but not limited to: providing accurate, up-to-date, and complete information about patients at the point of care; enabling quick access to patient records for more coordinated, efficient care; securely sharing electronic information with patients and other clinicians; and helping providers more effectively diagnose patients, reduce medical errors, and provide safer care [3]. Despite the many befits, one weakness of EHR is that its manual data input is time consuming [4]. Based on a recent study, during the office day, physicians spend as much as 49.2% of their time on EHRs [5]. When a patient goes to her clinical appointment, she would not like the experience of spending long time waiting for her healthcare practitioners entering all the information. For the health practitioners, instead of spending more time in understanding the patient needs and providing patients with the care plan, they are frustrated by taking lots of time entering data in EHRs.

In order to improve patients' experience during a clinic visit, increase patient-provider interaction time, and save doctor from tedious manual EHR input, our research focuses on an alternative EHR management solution by using voice assistant technologies. A voice-based virtual assistant is a software agent that can perform tasks or services for an individual based on voice commands or questions. One implementation of a voice assistant is through smart speaker technology. According to the voicebot.ai survey of 2019, starting from January 2018 to the end of the year of 2018 the population accessing the smart speakers went up from 47.3 million to 66.4 million in the United States [6]. Out of that, Amazon Alexa had a market share of about 61.1% by brand as compared to other smart speakers in the market [6]. The top use cases in the smart speaker world are related to personal music, news, and how-to instructions, followed by history, products, restaurants, and shops.

The smart speaker adoption in healthcare industry is still falling behind other industries. One factor that cause this lag is the concern towards Health Insurance Portability and Accountability Act (HIPAA) compliance [7] given that all the use cases would be consuming health records in some way. Another concern is the user's trust towards communicating to a virtual agent. In 2019, Amazon's Alexa became the first voice assistant that allows their skills to be HIPAA compliant. This created an open opportunity for creating Alexa skills for healthcare. Hospitals such as Mayo Clinic, Boston's Children Hospital have already partnered with several companies to be early adopters of voice assistant [8].

Because of Alexa's popularity and its HIPAA compliance, we implement our voice assistant based on Alexa. Our Alexa assistant aims to help health practitioners spend less time in data entry and more time in patient care. We implement the voice assistant as an Alexa skill that interacts with health practitioners, taking their commands and then acting on them. Advantages of the voice assistant include efficient data entry, increased patient-provider time, smaller learning curve for EHR system, and potential reduction of contamination risk.

This research has three main contributions:

1. Collected and analyzed the key requirements for having a voice-based assistant for EHR management in a clinic setting.

2. Designed and developed an Amazon Alexa skill, DocPal, for health practitioners for efficient EHR management.

3. Evaluated the usability and performance of the voice assistant with extensive experiments.

The rest of the paper is organized as follows. Section II surveys related work on voice assistant and smart speaker technology in healthcare domain. Section III describes our proposed methodology in detail. Section IV presents our evaluation through use case studies, user experiences studies, and qualitative evaluations. Finally, in Section V, we provide conclusions and future work directions.

II. Related Work

The acceptability of voice-enabled devices in medical field is dramatically increasing given the technological advances and its feasibility.

Many voice-based applications and tools have been used for wellbeing and healthcare of users (e.g. [9], [10], [11]). In their research, Lordon et al. proposed a smartphone voice assistant to help surgery patients [12]. The investigators created an "Action" using Google Assistant for patients who had hernia operation and were seeking medical help after being discharged from the hospital. They evaluated the "Hernia Coach" based on personas and scenarios created by the researchers. The voice assistant can help users with low healthcare literacy levels to comprehend medical instructions. As another example, Greene et al. proposed fall detection system that uses voice to activate devices to monitor and generate help calls and texts to caregivers of the elderly [13].

Motalebi and Abdullah proposed a conversational chatbot for couples suffering from Post-traumatic stress disorder (PTSD) [14]. In a home setting for a couple, the chatbot gives users a private and intimate environment to practice and focus on their therapy. Duke Health Urgent Care [15] is a skill available on Amazon Skill market. It is a voice-enabled skill which can help patients find urgent care near their location and book an appointment (given that the urgent care provider is associated with this application's backend). Along with that, patients can retrieve wait times, urgent care hours and phone numbers as required. According to a survey [16], 69% of patients are comfortable using voice assistants to search and locate urgent care facilities nearby. By having this skill enabled on a patient's Alexa device, this saves the time and effort to open and search for urgent care hours nearby through laptop, computer, or mobile phone.

kBot was designed as an android voice application on the smartphone [17]. This chatbot was developed for patients, specifically children, suffering from Asthma. KBot allows users to ask questions about the weather that affects Asthma. It also provides interface for users to report their conditions. Based on a knowledgebase, KBot can give users suggestions. According to the researchers, their holistic approach to chronic, multi-factorial asthma seemed to address the self-management issue in the patients. The app was designed for children to self-manage asthma symptoms on their day-to-day basis in association with Dayton Children's Hospital.

Compared with the many voice assistants for patients [18] [19][20] [21], there are relatively fewer applications for medical providers. But increasingly, more doctors start to turn to smart speakers during medical procedures, so that they can use voice commands where they cannot use a traditional computer for help. For example, when a doctor is in the middle of a procedure, she needs to remain sterile, so she cannot use a computer during the procedure. Seals et al. [22] created a virtual radiology assistant that offers clinicians many non-interpretive radiology skills. The proposed skills can automatically communicate with referring clinicians and quickly provide evidence-based answers to frequently asked questions. This allows the referring physician to provide real-time information to the patient about the next phase of treatment, or basic information about an interventional radiology treatment.

Another Alexa skill was developed for Emergency medical technicians (EMTs) to get medical information assistance. During their emergency medical response, EMTs may need information assistance, but it is impractical to go through medical procedure and reference manuals in such situation. The Alexa skill developed by Brewster Ambulance Service [23] can assist EMTs on this. Through this skill, EMTs can ask Alexa for reference info or to recite the procedure steps to them while they focus on the patient.

To conclude the discussions above, voice assistants are soon to become ubiquitous due to its vast integration with smartphones, smart speakers, smart cars, and other specialized devices developed to utilize its capabilities. Health in general is a very important aspect of human life and thus needs more effective technological advancements to provide better care [20]. Therefore, having a voice-enabled assistant may benefit for the public including both doctors as well as patients.

III. System Design

Recognizing the complexity of healthcare data and its privacy rules and regulations, our system needs to choose a platform which can provide the safety and security to handle EHR data. We chose Amazon Alexa as our virtual assistant's device, as Alexa is HIPAA complaint. We follow the Software Development Life Cycle (SDLC) [24] to design and develop the voice assistant, DocPal. In particular, we used the waterfall model [25], a sequential task based SDLC, to provide a rigid schedule with deliverables at the end of each phase. It also helps in documenting the processes as the project progresses.

A. Requirements Analysis

Requirements analysis is a fundamental part of our system development. It collects and determines the needs or conditions of the final product. As part of the requirements analysis, we created a set of questions as surveys sent to health practitioners. The goal of the survey is to confirm that there is a need from medical practitioners to use voice to improve EHR management. The survey consisted of 10 questions

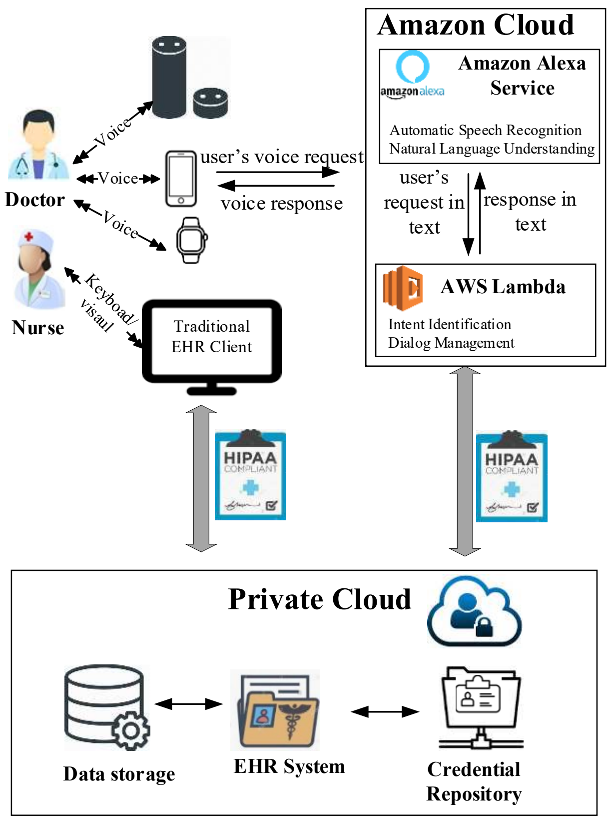

Figure 1. System Architecture

where the first 9 questions were multiple choice with the last one being an open-ended suggestion request. We leveraged SurveyMonkey to create and send out the survey to the potential survey takers. The survey takers included current staff from Family HealthCare, Essentia Health, and Sanford in Fargo, North Dakota. Our survey was a simple questionnaire designed to gauge the interest of potential users into this system as well as understand their idea of voice-enabled devices in healthcare settings. Below list a few example questions in the survey:

- Do you use any smart speakers at home? Yes, all the time; Yes, sometimes; No.
- *How would you feel about using a smart speaker at work aka Healthcare Environment? Yes, I would like to use its capabilities, if done correctly; No, I am not interested; Maybe. I would like to explore the idea.*
- *Roughly, on a regular day, what percentage of your time goes typing on a keyboard? A scale from 1-100%.*
- *Do you agree that manual input is one of the pain points of using EHR data? Yes; No; Don't know.*
- *Do you think using voice would save your time on typing? Yes; No; Maybe.*
- *Considering a smart speaker (e.g. Alexa) which is HIPAA complaint and can retrieve information from EHR, and allows users to add/revise EHR information through voice, would you consider having it in the patient room to eliminate typing patient information? Yes; No; Maybe.*

Based on the responses we gathered through the survey, there are two main discoveries: First, more than 80% respondents agreed that manual input is one of the pain points in using EHR data. Second, more than 80% respondents believed that using voice would help them save time.

B. Design and Development

The goal of the system is to design an Amazon Skill to complement the manual operation of EHR. The system is developed using Amazon Alexa Software Development Kit (SDK), which provides Automatic Speech Recognition (ASR) and Natural Language Understanding (NLU) functionality.

Figure 1. depicts the system's architecture and information flow. Users of the system are healthcare practitioners such as doctors, nurse, technicians. They use the system's EHR functions through voice commands and conversations with the system. For example, they can schedule a patient's appointment or add a patient's clinical document or note to the EHR. To use the system's EHR function, a health practitioner will speak a voice command to an Alexa device such as Echo Dot, smart watch, or a smartphone. Alexa then sends the sounds of that voice command to the remote Amazon voice process cloud (Amazon Alexa Service) via its connected Wi-Fi network. The Alexa voice service converts the speech input into text using the Automatic Speech Recognition and Natural Language Understanding. Based on the predefined schema, Alexa can match with the intent defined in the AWS lambda. Then, dialog management in AWS lambda examines input, executes the dialogue policy and up-dates the dialogue state. Based on the user's intent, AWS Lambda service will call the backend EHR services located at the private cloud using HIPAA compliant protocol. In the following, we present the details of three major components of the system.

1) Intent Identification. In our system, an intent represents an EHR operation that fulfills a user's spoken request. Intents can optionally have arguments called slots. When we create our customized intent, we provide a list of utterances that users would say to invoke this intent. To improve the usability and flexibility of the skill, we provide a wide range of sample utterances and custom slot values to represent the real-world language use. Providing these different phrases as sample utterances will help improve voice recognition. To obtain a wide range of relevant sample utterances, we refered to synonames defined in WordNet [26], concepts in Wikipedia. Morever, you plan to observe the users during our real world testing, so that we can note the phrases that healthcare practitioners speak to invoke each intent. Then, we can continually update the custom values and sample utterances file to ensure that it includes instances of the users' most common phrasings.

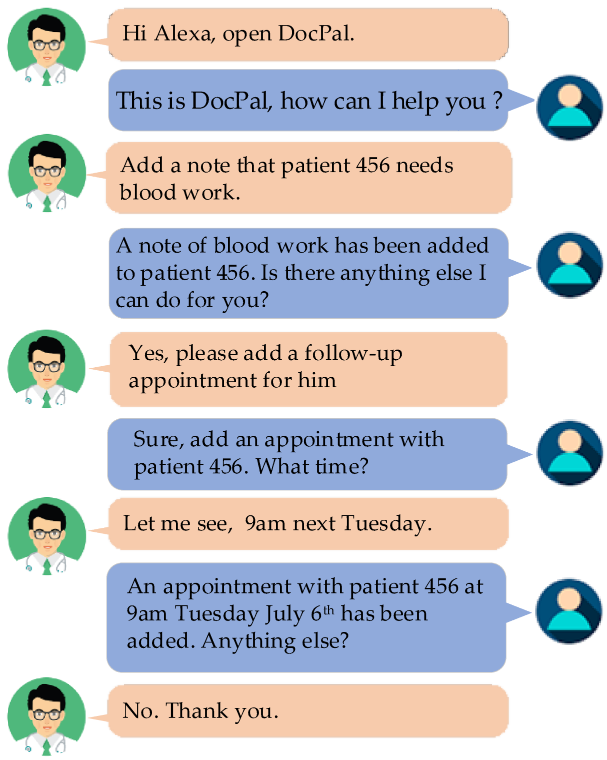

Figure 2. An example conversation between a doctor and the DocPlan Skill

2) Dialog Management: The main role of Dialog Management is to decide what action the voice assistant should take at each point of the dialog. Our system's function is straightforward: user says a request, providing relevant information, and the system calls correnspoding services provided by the EHR system. Most of the intents are command-based and do not involve complex conversaton flow. This componet mainly deals with the following function:

- Acknowledgement. This is to confirm the user that the system understood what he or she just said. It also helps the user to check if the voice assistant has any mistake or misunderstanding.
- Slot filling. Sometimes when a user says a request, the system may not have enough information to execute the request. For example, when a doctor asks the system to arrange an appointment for a patient, the system will ask for the date and time for the appointment.
- Context management. When chatting with a user, the system needs to collect and manage all kinds of conversation context, which is built while chatting with the user and contains all the short- and long-term information the system needs to handle a conversation.

3) EHR Acceess: After the system has figured out the appropriate action to take, it calls EHR web service API provided by the cloud based EHR systems, which store and maintains the patient's medical data such as prescriptions, testing data, pathology data, and nursing charts. REST (REpresentational State Transfer) is used to access the service via HTTP protocol. With REST, we can access and execute CRUD (Create, Read, Update and Delete) operations through APIs implemented on the web. JSON (JavaScript Object Notation), a simple and light-weight data representation format, is used to for data exchange.

IV. EVALUATION

We have implemented a prototype system as an Amazon skill, DocPal, that fulfills the goals mentioned above while maintaining a certain quality and efficiency. We have performed a set of preliminary evaluations on the DocPal.

A. Use Case Evaluation

Based on the backend services, our skill can support a variety of EHR functions such as scheduling appointment, order lab test, providing patient's past vital information, medication list and past note history, etc. In particular, we have tested on the following cases: (1) scheduling appointment, (2) adding notes, (3) getting nearest appointment, (4) getting notes, and (5) getting vitals. These cases are not independent, based on the conversation flow and user's intents, multiple cases can be integrated and used in one conversation.

Figure. 2 demonstrates one example conversation. The user is a doctor and he uses the voice assistant in his office. In this conversation, the user asks the virtual assistant to add a blood work note to a patient's record and add a follow-up appointment. Utterance (1) shows how users invokes the skill. Once the DocPal skill is activated, it greets the user in utterance (2). In utterance (3) the user requests the virtual assistant to add a note of lab work. This will trigger the backend service to call the EHR cloud API to add a note in the database. In utterance (4) the user requests to add an appointment for the patient. The virtual assistant first acknowledges this request, then it asks for missing information, in this case time information, from the user. After the user provides the time information at utterance (5), in utterance (6) the virtual assistant acknowledges the appointment with the patient's ID, the time and date. In this turn, the virtual assistant uses the conversation context. Although the user does not specify the patient's ID, the virtual assistant figure out that from previous conversation. In addition, it includes other context, for example, the current date to figure out the appointment date that is next Tuesday.

This use case showed the flow of the conversation. It also demonstrated how the assistant makes acknowledgement, fills missed information, uses context information to make smooth and efficient conversation with a user.

B. User Expeirence Evaluation

As we did not have the opportunity to observe users' interaction with DocPal, we use surveys to gauge user experience. As surveys are not generally considered as usability test, we call this test as user experience test. We surveyed on users' satisfaction on three different types of

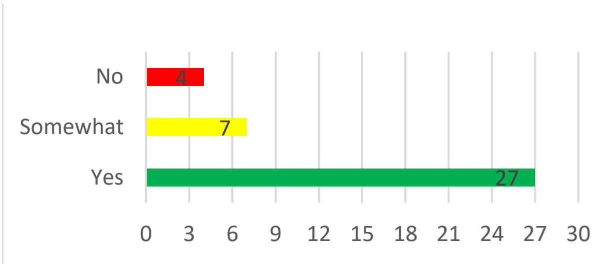

Fig. 3. Survey result on the satisfaction of usability of adding notes case

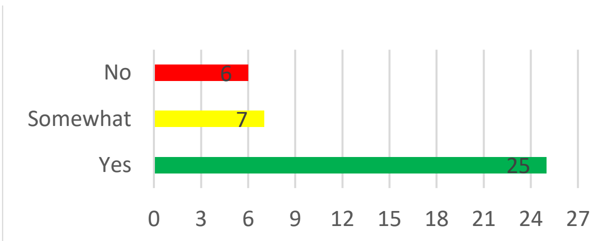

Fig. 4. Survey result on the satisfaction of usability of scheduling appointment

cases, namely adding notes, scheduling appointment, and reading vitas in operation rooms. We got feedback from 38 health practitioners ranging from doctors, nurses, and physicians, in which 30 were collected at January 2020, and extra 8 more were collected at April 2020. The survey takers included current staff from Family HealthCare, Essentia Health, and Sanford in Fargo, North Dakota. Figure 3 shows the result on the satisfaction of usability of adding notes. We can see that among the 38 survey takers, 27 believed that this function was very useful, 7 believed that this function was somewhat useful, and 4 did not think they were useful or they would like to use this in their office. Similarly, Figure 4 illustrates the result on the satisfaction of usability of scheduling appointment. We can see that most testers are satisfied with this function. This is an ongoing project; we are working on improving the system to improve user's experience.

C. *Qualitative Evaluation*

We have performed a set of qualitative evaluations in our research laboratory environment by our research team. In the experiment session, a total of 200 conversations were attempted. A conversation starts when the tester invokes the skill and ends when the user's request has been addressed. For simplicity, all the conversations are single intent conversation. For each conversation, one intent was randomly chosen from fiver categories, namely scheduling appointment, adding notes, getting nearest appointment, getting notes, reading vitals. The tester has the maximum flexibility to talk to the virtual assistant using their own ways.

TABLE I
DIALOG PERFORMANCE TEST

Metrics\Purpose	cat 1	cat 2	cat 3	cat 4	cat 5
Avg Dialog time (s)	36.65	24.89	15.4	18.44	22.5
Time Per Turn (s)	5.02	5.61	7.7	9.22	10
Number of Turns per Dialog	7.3	4.44	2	2	2.25
Successful Rate	100%	100%	100%	100%	100%
Accuracy	100%	100%	100%	100%	100%

Table I shows the experimental results. In these experiments, we set an upper bound of the number of dialog turns. If a dialog lasts longer than the upper bound and the user has not got a satisfied response, we treat the conversation as a failed one. In our system, most of the intents are command-based and the conversations are straightforward and short. we set the threshold of maximal dialog turns as 12. Based on this upper bound, a conversation success rate is defined in Equation (1).

$$\text{Success Rate} = \frac{\#of\ successful\ conversations}{\#\ of\ total\ conversations} \quad (1)$$

We also check the accuracy (or correctness) of the system's operation to make sure that the system does truly understand user's request and there is no misunderstanding happened.

$$\text{Accuracy} = \frac{\#\ of\ correct\ operations}{\#\ of\ total\ operations} \quad (2)$$

We can see from Table I that the conversations are generally short, as the intent is command-based and straightforward. The conversations are always successful in our experiments. And the accuracy rate is 100%. This means that our system can accurately understand user's command and execute the right operation at the backend. Sometimes the system misread user's voice. Fortunately, acknowledgement from the voice assistant helps users to find the mistake and then correct it in the following conversation.

V. CONCLUSIONS

Increasingly, voice assistant is being used in healthcare settings to empower patients and assist doctors. In this paper, we present the design, development, and evaluation of a voice assistant, DocPal, to help healthcare practitioners to easily access and manage EHR records and benefits their daily routine task. Voice is a natural and convenient way of communication for humans. The voice-based assistant provides a natural interactive interface with users and it does not require any technical background.

The voice assistant has been evaluated by different types of use cases, user experience surveys, and qualitative evaluations. The experimental results demonstrate a satisfying user experience, and good dialog success rate and EHR operation accuracy.

This project is ongoing. In the future, we plan to deploy the voice assistant in a real setting in clinicals or hospitals. More comprehensive user studies will be performed to evaluate the voice assistant's usability, the satisfaction rate of the users, and the workflow improvement outcomes. Moreover, we are working on improving the voice assistant's conversation flexibility and intelligence using machine learning-based techniques.

ACKNOWLEDGMENT

This work was partially supported by the National Science Foundation (NSF) with award number: 1722913 and 1921576.

REFERENCES

[1] E. Salt, G. D. Rowles, and D. B. Reed, "Patient's perception of quality patient-provider communication," *Orthop. Nurs.*, 2012.

[2] S. Bowman, "Impact of electronic health record systems on information integrity: quality and safety implications.," *Perspectives in health information management / AHIMA, American Health Information Management Association*. 2013.

[3] "Official Website of The Office of the National Coordinator for Health Information Technology (ONC)." [Online]. Available: https://www.healthit.gov/faq/what-are-advantages-electronic-health-records.

[4] N. Menachemi and T. H. Collum, "Benefits and drawbacks of electronic health record systems," *Risk Manag. Healthc. Policy*, 2011.

[5] S. Hingle, "Electronic health records: An unfulfilled promise and a call to action," *Annals of Internal Medicine*. 2016.

[6] Vivint, "Smart Home Use Cases Are Key to Driving Mass Market Adoption," *Business Wire (English)*. 9AD.

[7] B. B. Frey, "Health Insurance Portability and Accountability Act," in *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*, 2018.

[8] L. M. Orlov, "Voice, Health and Wellbeing 2020."

[9] N. A. Ikhu-Omoregbe and A. A. Azeta, "A Voice-based Mobile Prescription Application for Healthcare Services (VBMOPA)," *Int. J. Electr. Comput. Sci. IJECS-IJENS*, 2011.

[10] S. J. Shiny Prakash and K. Sekar, "An intelligent home centric healthcare system based on the internet-of-things," *Int. J. Appl. Eng. Res.*, 2015.

[11] H. M. Abu-Dalbouh, "A questionnaire approach based on the technology acceptance model for mobile tracking on patient progress applications," *J. Comput. Sci.*, 2013.

[12] R. J. Lordon, "Design, Development, and Evaluation of a Patient-Centered Health Dialog System to Support Inguinal Hernia Surgery Patient Information-Seeking," 2019.

[13] S. Greene, H. Thapliyal, and D. Carpenter, "IoT-Based fall detection for smart home environments," in *Proceedings - 2016 IEEE International Symposium on Nanoelectronic and Information Systems, iNIS 2016*, 2017.

[14] N. Motalebi and S. Abdullah, "Conversational agents to provide couple therapy for patients with PTSD," in *ACM International Conference Proceeding Series*, 2018.

[15] "Duke Health Urgent Care." [Online]. Available: https://www.amazon.com/Paragon-Consulting-Inc-Health-Urgent/dp/B07XRHCXQ2/ref=sr_1_1?keywords=duke+health&qid=1568746777&s=digital-skills&sr=1-1.

[16] "Are Patients Ready for Amazon Alexa, MD?" [Online]. Available: https://www.softwareadvice.com/resources/alexa-healthcare-skills/.

[17] D. Kadariya, R. Venkataramanan, H. Y. Yip, M. Kalra, K. Thirunarayanan, and A. Sheth, "KBot: Knowledge-enabled personalized chatbot for asthma self-management," in *Proceedings - 2019 IEEE International Conference on Smart Computing, SMARTCOMP 2019*, 2019.

[18] A. E. Chung, A. C. Griffin, D. Selezneva, and D. Gotz, "Health and fitness apps for hands-free voice-activated assistants: Content analysis," *JMIR mHealth uHealth*, 2018.

[19] M. Hadian, T. Altuwaiyan, X. Liang, and W. Li, "Efficient and Privacy-Preserving Voice-Based Search over mHealth Data," in *Proceedings - 2017 IEEE 2nd International Conference on Connected Health: Applications, Systems and Engineering Technologies, CHASE 2017*, 2017.

[20] A. Ilievski, D. Dojchinovski, and M. Gusev, "Interactive voice assisted home healthcare systems," in *ACM International Conference Proceeding Series*, 2019.

[21] E. De Joode, C. Van Heugten, F. Verhey, and M. Van Boxtel, "Efficacy and usability of assistive technology for patients with cognitive deficits: A systematic review," *Clin. Rehabil.*, 2010.

[22] K. Seals *et al.*, "Utilization of deep learning techniques to assist clinicians in diagnostic and interventional radiology: Development of a virtual radiology assistant," *J. Vasc. Interv. Radiol.*, 2017.

[23] "Brewster Ambulance Service." [Online]. Available: https://www.brewsterambulance.com/.

[24] J. Broad, "System Development Life Cycle (SDLC)," in *Risk Management Framework*, 2013.

[25] S. L. Models, "Object-Oriented and Classical Software Engineering LIFE-CYCLE," *Development*. 2010.

[26] G. A. Miller, "WordNet: A Lexical Database for English," *Commun. ACM*, 1995.