Developing Intelligent Interviewers to Collect the Medical History: Lessons Learned and Guidelines

Floriana GASHI[a,1], Selina F. REGLI[a,1] , Richard MAY[b] ,
Philipp TSCHOPP[c] and Kerstin DENECKE[a,2]
ª Bern University of Applied Sciences, Bern, Switzerland
ᵇ Hochschule Harz, Wernigerode, Germany
ᶜ Soccin22 - Zentrum für Hausarztmedizin, Basel, Switzerland

Abstract. Background: Physicians spend a lot of time in routine tasks, i.e. repetitive and time consuming tasks that are essential for the diagnostic and treatment process. One of these tasks is to collect information on the patient's medical history. Objectives: We aim at developing a prototype for an intelligent interviewer that collects the medical history of a patient before the patient-doctor encounter. From this and our previous experiences in developing similar systems, we derive recommendations for developing intelligent interviewers for concrete medical domains and tasks. Methods: The intelligent interviewer was implemented as chatbot using IBM Watson assistant in close cooperation with a family doctor. Results: AnCha is a rule-based chatbot realized as decision tree with 75 nodes. It asks a maximum of 44 questions on the medical history, current complaints and collects additional information on the patient, social details, and prevention. Conclusion: When developing an intelligent digital interviewer it is essential to define its concrete purpose, specify information to be collected, design the user interface, consider data security and conduct a practice-oriented evaluation.

Keywords. Medical history, dialog-based system, chatbot, conversational user interface, intelligent system

1. Introduction

Collecting information from a patient is an essential part in the diagnostic process. It normally comprises the present condition, personal history, family history, and a social anamnesis. The accuracy and completeness of this information significantly affects the quality of the diagnosis [1]. Normally, all relevant information is collected during the patient-doctor encounter. However, patients not always reveal all concerns and problems in these encounters due to language barriers [2], fear of stigmatism, nervousness, or other reasons [3,4]. Beyond, time in the encounter is limited. In order to improve the situation of incomplete and error-prone information on the medical history paper-based or electronic questionnaires were used [4] enabling "offline" reporting of complaints and medical history. Furaijat et al. implemented a digital communication assistance tool that

---

[1] Contributed equally

[2] Corresponding author: Kerstin Denecke, Bern University of Applied Sciences, Biel / Bienne, Switzerland; E-mail: kerstin.denecke@bfh.ch

lists complaints users can select and were questions with predefined answers are provided [2]. Another approach are chatbots, in this context often used in so-called conversational user interfaces that simulate a dialog and collect information from the user in this dialog. The first chatbot was developed in the 1960s. The computer program, named ELIZA, was designed to mimic a conversation with a psychotherapist [5]. Chatbots have also proven to be successful in today's COVID-19 pandemic to inform about symptoms, to encourage behavior change, or as monitoring tools [6].

There are chatbot systems available that make a presumptive diagnosis based on the information entered by the user. Ada (https://ada.com/) and the Healee (https://www.healee.com) are apps that collect symptoms from patients, derive a set of possible diseases from them, and show them to the user. A physician can be contacted if needed. However, we are pursuing another approach: We aim at supporting physician and patient in the diagnostic process by means of an intelligent interviewer. In previous work, we have already developed a concept for self-anamnesis and implemented a prototype for collecting the music biography of a person [7-9]. With the experiences gained from this work, we developed an intelligent interviewer for another medical specialty, which is family medicine. In this paper, we aim to contribute the following: 1) We present the developed prototype called AnCha, including its concept, technical specialties, as well as benefits and limitations. 2) Additionally, we aggregate our experiences from the development of AnCha and our previous developments in order to extract recommendations for designing and developing intelligent interviewer for the medical domain. Thus, we aim to facilitate the development of similar systems in the future.

2. Methods

We developed our concept and prototype in three steps: requirement analysis, concept development with mockup generation, and implementation of the prototype. We identified requirements regarding the chatbot AnCha based on two interviews with a general practitioner (GP). In this way, we determined required functionalities and we specified the dialog flow, i.e. which questions the chatbot has to ask a patient to collect information on the medical history and complaints. In addition, we assessed relevant literature to collect requirements and review functionalities of existing apps. Based on the requirements, a mockup was created using Axure RP 9. By means of the mockup and in further meetings with the GP as well as his feedback, the conversation flow of the chatbot was improved. The goal was to find the most effective way to gather all the important information of a patient from the GP'S perspective. The chatbot was implemented using the IBM Watson Assistant. Among other things, IBM Watson Assistant enables creating, training and implementing chatbots. Using integrated machine learning models, user questions posted in a chat can be understood and answered accurately.

In order to come up with recommendations for developing intelligent interviewers for the medical domain, we additionally considered our previous work on the chatbot Ana, who collects the music biography of a person [7-9]. Together with experiences gained in this work, we derive our recommendations for future developments of intelligent interviewers for the medical domain.

3. Results

3.1. Requirements and use case

The main objective of AnCha is to ask a patient questions on his or her medical history and to provide a physician access to the responses in a digital manner as preparation for the patient-doctor encounter. The chatbot is supposed to be rule-based, i.e. a set of specified questions has to be asked in a given order. This specification has the advantage that the patient is guided through the conversation without digressing and the doctor receives the information which is important to him. Several response types are required: predefined answers for selecting one or more options and free text. To avoid misunderstandings, the majority of questions will not require a free text answer. The collected information is not supposed to be a complete medical history. The aim instead is to have an initial set of information in order not to overlook relevant aspects and to focus the communication on the encounter. The information collected by the chatbot will help the physician to create a mind map of the patient's current problems. The patient should be given the opportunity to express all current worries and complaints. In addition, the chatbot should be accessible by patients through the website of the doctor's practice and data security has to be ensured.

The chatbot is supposed to be integrated into the workflow as follows: When a patient comes to a doctor's office with a complaint, he usually spends time in the waiting room before the actual medical appointment. During this waiting time, the patient accesses AnCha. In communication with the chatbot, the patient provides information on current complaints and if needed the medical history using a tablet provided or the patient's own smartphone. As soon as the patient has transmitted the data, it is available to the physician. Shortly before the actual appointment with the patient, the physician accesses the received information. This allows him to prepare for the consultation and to ask specific questions to the patient in the encounter. Another option is to ask the patient for interacting with the chatbot when she makes the appointment.

3.2. Conversation with AnCha

AnCha is running on a website, i.e. it is accessible by patients without installation. The main functionality of AnCha is the collection of current symptoms and complaints as well as aspects of the medical history. The current prototype of AnCha contains a total of 44 questions covering seven topics (see Table 1). At the beginning of the chatbot conversation, the patient is welcomed and receives initial instructions about the chatbot. Formalities for the interaction with the chatbot are clarified, i.e. the patient is asked for her name, date of birth and whether this is the first visit to this practice. In case it is the first visit, additional personal and social information is requested, general questions on medication, alcohol consumption etc. are asked as well as questions on the current status of prevention (topics 2 to 5 in Table 1). The questions about prevention are adjusted according to the entered gender and age. For female patients, questions about gynaecological check-ups are asked, which are omitted for male patients. Patients over 50 years of age are asked if they already possess a health care directive or if they regularly attend a colonoscopy. This means, the number of questions asked to a concrete patient depends on the basic demographic data. Minimum and maximum number of queries per topic are shown in Table 1.

In order to establish an emotional bond of trust to the patient, the chatbot has a name (AnCha, acronym of Anamnesis Chatbot) and is visualized by an avatar. Questions are formulated in a way that they could originate rather from a friend than from a doctor. Specific medical terminology that the patient might not understand is avoided. AnCha tries to motivate the patient to complete the entire dialog. For example, when the user claims, she is not yet willing to continue, AnCha asks a question out of topic or fun facts are posted as distraction. The chatbot not only asks specific questions, but encourages to provide more details. The aim of this is, that the patient feels well accepted and perceived with her concerns and complaints. Once all questions have been asked, the collected information is aggregated, grouped along topics (questions and corresponding answers of the patient) and transformed into a more readable format for the physician. The aggregated data is sent via secure E-Mail to the physician. The conversation with AnCha takes a maximum of ten minutes.

Table 1. Subjects covered by the chatbot as well as minimum and maximum number of questions

	Subject	Questions	Minimal number of questions	Maximum number of questions
(1)	Welcome	- Name - Date of birth - First visit	3	3
(2)	Personal information	- Gender - Age - Size and weight	0	4
(3)	Social information	- Occupation - Marital status and children - Leisure activities	0	4
(4)	General questions	- Alcohol and tobacco consumption - Recent hospitalization - Medication - Pre-existing conditions	0	11
(5)	Prevention	- Vaccinations - Medical check-up - Health care directive - Other doctors	0	10
(6)	Specific questions about the medical consultation	- Reason for visiting the doctor - Questions about the symptoms	7	10
(7)	Other concerns		0	2
	Total		**10**	**44**

3.3. Prototype implementation

Figure 1 shows the system architecture of AnCha. The chatbot is implemented with IBM Watson Assistant, a service of IBM Cloud. The content is arranged in a decision tree which consists of 75 nodes. Currently, the patient accesses AnCha via a separate website and not through the GP practice's website. In detail, the chatbot is integrated into a HTML file using a script. At the end of the questionnaire, the data is transmitted using Node.js. A new action is created in IBM Cloud which uses the module Nodemailer and allows to send e-mails via Node.js. To be able to send e-mails, the mail server, the sender e-mail address, and password have to be stored.
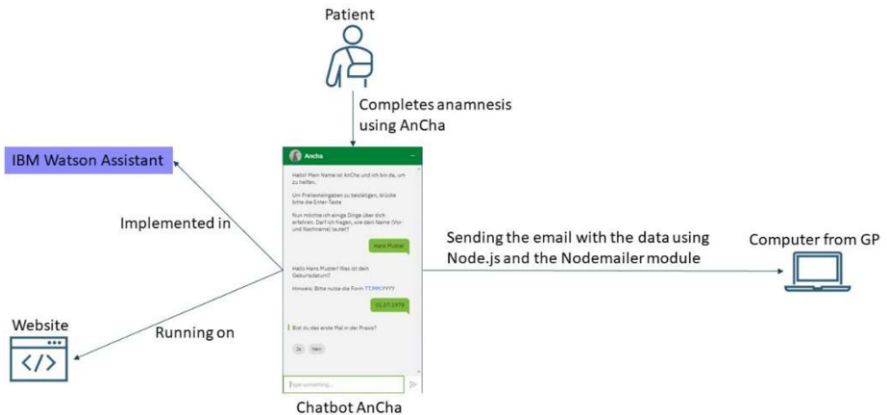
Figure 1. System architecture of AnCha

4. Limitations and benefits of AnCha

In this paper, we introduced the intelligent interviewer AnCha that collects information on the medical history from a patient. In contrast to apps like Ada or Heala, our application AnCha is not supposed to autonomously make any decisions in terms of deriving diagnoses or making a triage. Its aim is to support collection of data on the medical history of a patient and current complaints. The application offers several benefits: The patient can take time to think about the questions and to answer while staying in the waiting room. This approach has the potential to reduce feelings of nervousness which otherwise might lead to wrong or missing information. Further, the patient gets the opportunity to explain everything of relevance to him related to the current complaint. In the patient-doctor encounter, the physician normally has limited time and not everything can be discussed. The information collected by AnCha helps to clarify the patient's complaints in advance, so that the doctor already has an initial overview. This way, he can address these concerns. With the help of AnCha, the patient no longer forgets important information and the doctor is prepared in advance. AnCha thus fulfills the goal of helping the doctor, not replacing him. AnCha is supposed to be an offer to patients. It still has to be assessed, 1) how many patients possess the necessary technical competency for interacting with AnCha, 2) how many patients are willing to use the chatbot in advance, 3) whether 10 minutes of interaction are sufficient to generate a bond of trust to an intelligent assistant and to encourage a patient delivering the requested information. These research questions will be addressed in a future study.

The chatbot could also help in dealing with language barriers: The questions can be provided in multiple languages with appropriate translation. In addition, it is possible to provide explanations or to include images to facilitate communication with non-native speaking patients.

Since AnCha is a rule-based chatbot, the flow of questions is clearly defined. This ensures that all relevant data is collected. However, the communication flow is fixed and might be perceived as unusual by patients. Such perception can be reinforced by the fact that the actual content of the user input is neither analyzed nor interpreted by AnCha. This issue has the following reasons: First, analyzing free text using artificial intelligence

such as IBM Watson Assistant can be error-prone. Second, for our specific use case it is not relevant that the chatbot interprets the information. The primary purpose is to collect information that is then interpreted by the physician.

Since this work was realized during the COVID-19 pandemic, we could not test the application within a family medicine practice. We tested AnCha only with 4 people in the circle of acquaintances of the authors (age: 15, 19, 57 und 60). They had to go through the interview process with AnCha. Their feedback was positive; only the process of sending the e-mail was too complicated since the chatbot asked several times whether the mail should be sent. We are currently planning a study in daily routine with real patients of a medical practice to study patient and physician satisfaction with the adapted process. This evaluation will help to find out whether AnCha would be used by patients and whether it really supports the physician. Based on the results, we will improve our prototype. Depending on the user feedback, we will also extend the motivation capacities of AnCha.

In its current implementation, the data is not encrypted which has to be realized before the system is used in daily routine. At least secure e-mail such as HIN-Mail should be used (https://www.hin.ch/services/hin-mail/). We still have to assess how to securely transfer the chatbot protocol to the physician and whether an integration with the practice information system (or electronic medical records (EMR)) is helpful. Still an open issue is the data transfer and the connection with such system. The chatbot protocol could be stored as note in the EMR which would require a corresponding interface or the sent e-mail might be read by the physician in his mail browser and then stored in the document section of the patient's EMR. AnCha can also be applied in other areas of medicine. Since it is a rule-based chatbot based on a decision tree that does not have many branches, it can be easily extended. The existing questions can be modified, adapted for other medical specialties or additional questions could be added.

5. Recommendations for developing intelligent interviewers

Based on our experiences in designing and building intelligent interviewers, e.g. AnCha as described in this paper or Ana [7-9], we are able to derive several lessons learned. In order to successfully develop intelligent interviewers, in particular based on chatbot technology, it is essential to collect and document all requirements in terms of a comprehensive requirements engineering process [10]. The whole requirements engineering process should at least include the following steps that can be seen as essential guidelines for a successful development process:

First, the concrete purpose of the intelligent interviewer has to be defined. The precise definition depends on the specific use case which includes potential users, e.g. physicians or patients, the environment, e.g. a waiting room, and the medical specialty, e.g. music therapy in the context of mental health [7-9]. If the planned system is only expected to collect information, e.g. personal information, it is sufficient to build a rule-based system. The main advantages of these systems are that their conversations are reliable, easily reproducible and comparable. There are several frameworks or scripting languages available that are appropriate to build such systems, e.g. AIML, SIML, or RiveScript [11]. In our previous work, we developed chatbots based on these technologies and gained several in-depth insights into the technical advantages and disadvantages of these technologies [7-9,11,12,13].

Second, the information to be collected by the intelligent interviewer has to be precisely specified according to the use case. We recommend to use questionnaires or to conduct personal interviews, e.g. with experts. Based on the collected data, a questionnaire can be generated and transferred into one of the already mentioned rule-based chatbot languages. We also recommend to consider the motivation of the later user. In this context, the wording of the intelligent agent's responses should vary in order to create an immersion effect, e.g. in terms of gamification [14]. Moreover, we suggest to use short questions and as few questions as possible. Adding motivating words, funny comments and enabling a reasonable conversation flow, e.g. by creating a logic order without irrelevant questions, also motivates the user to answer all questions satisfactorily. It also makes the intelligent interviewer appear more natural and human.

Third, the user interface has to be designed according to the needs of the future users [15]. This design process includes the general interface design of the whole application and in particular the selection of the input type. The latter can be based on free text responses or pre-defined answers for selection, e.g. based on buttons or sliders. We recommend to provide at least a few free text responses to give the user the opportunity to deliver specific details in his or her own words. However, too many free text responses can be demotivating and confusing for the user. Regarding the app design, we recommend to use round shapes with smooth animations, e.g. for the speech bubbles, and comforting colors, e.g. green [16]. The general appearance of the application should also be based on already known messengers, e.g. WhatsApp or Telegram, in order to enable easy handling and to base upon experiences of users. By adding additional input or output technologies, e.g. speech recognition, the agent's immersion effect can also be increased [17]. However, the accuracy of such services should be considered before their implementation. Combining an intelligent interviewer with immersive technologies, e.g. virtual reality and voice user interfaces, can also be useful in some use cases, e.g. for patient education [13].

Fourth, data security has to be considered. Although often ignored in prototypes, data security, especially for personal patient data, has to be ensured. Therefore, encrypting the collected data, e.g. by using algorithms such as AES-256, can provide an appropriate security standard. In addition, implementing secure mail services, e.g. by using certificates or public and private keys, are essential to realize a secure information exchange of the data requested from the intelligent interviewer [18]. We do not recommend using third party services that can access (personal) data collected by the interviewer (e.g. Facebook messenger) or realize logins using Google or Facebook accounts. For instance, this includes the choice of chatbot technology, pre-built user interfaces, as well as input and output services. Fifth, by conducting a practice-oriented evaluation, e.g. based on a quantitative or qualitative questionnaire and participants according to the target group, potential issues or space of improvements can be identified. In this context, the evaluation should focus on both content-related tasks as well as general usability [8]. Accordingly, further requirements that were previously hidden can be identified and addressed. Based on the evaluation results, the gap between the theoretical concept of the intelligent interviewer in a specific use case and the actual practice can be minimized or even closed.

6. Conclusions

Providing intelligent digital interviewer in healthcare gains in interest. We suggest to consider 5 essential aspects that will help in faster development of reliable and secure intelligent interviewer for the medical domain: define concrete purpose, specify information to be collected, design user interface and address future user's needs, consider data security and conduct a practice-oriented evaluation. Regarding our chatbot AnCha, we will assess in future whether use of the artificial intelligence capabilities of IBM Watson can be used. This requires additional analysis of possible user questions and decisions on how the chatbot should react to them. AnCha is not a competitor to the patient-doctor interaction and thus the patient-doctor relationship should not be negatively impacted.

References

[1] P.J. Moonen, L. Mercelina, W. Boer, T. Fret, *Diagnostic error in the emergency department: follow up of patients with minor trauma in the outpatient clinic*. Scand J Trauma Resusc Emerg Med 2017; 25(01):13

[2] G. Furaijat, E. Kleinert, A. Simmenroth. et al. *Implementing a digital communication assistance tool to collect the medical history of refugee patients: DICTUM Friedland - an action-oriented mixed methods study protocol*. BMC Health Serv Res 2019(19):103

[3] T.M. Palermo, D. Valenzuela, P.P. Stork, *A randomized trial of electronic versus paper pain diaries in children: impact on compliance, accuracy, and acceptability*. Pain 2004;107(03):213–219

[4] D. Zakim, *Development and significance of automated history taking software for clinical medicine, clinical research and basic medical science*. J Intern Med 2016;280(03):287–299

[5] J. Weizenbaum, *Eliza - a computer program for the study of natural language communication between man and machine.* Commun ACM 1966;9(01):36–45

[6] A.S. Miner, L. Laranjo, A.B. Kocaballi, *Chatbots in the fight against the COVID-19 pandemic*. npj Digit. Med. 3, 65 (2020). https://doi.org/10.1038/s41746-020-0280-0

[7] K. Denecke, *A Mobile System for Music Anamnesis and Receptive Music Therapy in the Personal Home*, Stud Health Technol Inform. 2017;245:54-58

[8] K. Denecke, A. Pöpel, S. Lutz Hochreutener, R. May, *Self anamnesis with a Conversational User Interface: Concept and Usability Study*. Methods Inf Med. 2018 Nov;57(05-06):243-252

[9] K. Denecke, R. May, A. Pöpel, S. Lutz Hochreutener, *Can a chatbot increase the motivation to provide personal health information?* pHealth 2020, Studies in Health Technology and Informatics, Volume 273: pHealth 2020, 10.3233/SHTI200619, 85-90

[10] A. Van Lamsweerde, *Goal-oriented requirements engineering: A guided tour*. Proceedings of the IEEE International Conference on Requirements Engineering (RE), 249–261 (2001)

[11] K. Denecke, M. Tschanz, T. Dorner, R. May, *Intelligent Conversational Agents in Healthcare: Hype or Hope?* Studies in health technology and informatics. 2019;259:77-84

[12] K. Denecke, R. May, Y. Deng, *Towards Emotion-Sensitive Conversational User Interfaces in Healthcare Applications.* Studies in health technology and informatics. 2019;264:1164-1168

[13] R. May, K. Denecke, *Extending Patient Education with CLAIRE: An Interactive Virtual Reality and Voice User Interface Application*. Lecture Notes in Computer Science, Volume 12315: Springer, Cham 2020. 10.1007/978-3-030-57717-9_49.

[14] S. Deterding, D. Dixon, R. Khaled, L. Nacke, *From game design elements to gamefulness: defining "gamification"*. Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, 9–15 (2011)

[15] J. Nielsen, *10 usability heuristics for user interface design.* Nielsen Norman Group 1, 1 (1995)

[16] A. Akers, J. Barton, R. Cossey, P. Gainsford, M. Griffin, D. Micklewright, *Visual Color Perception in Green Exercise: Positive Effects on Mood and Perceived Exertion*. Environment Science Technology, 2012;46(1):8661–8666

[17] R. Nakatsu, N. Tosa, *Active immersion: the goal of communications with interactive agents.* IEEE International Conference on Knowledge-Based Intelligent Engineering Systems and Allied Technologies 2000;1(1):85–89

[18] A. Dan, S. Gupta, S. Rakshit, S. Banerjee, Toward an AI Chatbot-Driven Advanced Digital Locker. *Advances in Intelligent Systems and Computing*, Springer Singapore 2019: 811:37–46