dHealth 2023
B. Pfeifer et al. (Eds.)

doi:10.3233/SHTI230012

Designing a Digital Medical Interview Assistant for Radiology

Kerstin DENECKE[a,1], Nikola CIHORIC[b,c] and Daniel REICHENPFADER[a]
[a] *Bern University of Applied Sciences, Biel/Bienne, Switzerland*
[b] *Department of Radiation Oncology, Bern University Hospital, Bern, Switzerland*
[c] *University of Bern, Switzerland*

Abstract. Radiologists rarely interact with the patients whose radiological images they are reviewing due to time and resource constraints. However, relevant information about the patient's medical history could improve reporting performance and quality. In this work, our objective was to collect requirements for a digital medical interview assistant (DMIA) that collects the medical history from patients by means of a conversational agent and structures as well as provides the collected data to radiologists. Requirements were gathered based on a narrative literature review, a patient questionnaire and input from a radiologist. Based on these results, a system architecture for the DMIA was developed. 37 functional and 17 non-functional requirements were identified. The resulting architecture comprises five components, namely Chatbot, Natural language processing (NLP), Administration, Content Definition and Workflow Engine. To be able to quickly adapt the chatbot content according to the information needs of a specific radiological examination, there is a need for developing a sustainable process for the content generation that considers standardized data modelling as well as rewording of clinical language into consumer health vocabulary understandable to a diverse patient user group.

Keywords. Medical History Taking, Natural Language Processing, Patients, Radiology

1. Introduction

Radiological examinations are one of the most common diagnostical tests performed worldwide. In 2020, over 1.6 million computer tomography (CT) exams and more than 1.3 million magnetic resonance imaging (MRI) exams were performed in Austria alone [1]. Based on the results of these examinations – the medical images – radiologists create diagnostic reports. However, due to time and resource constraints, radiologists usually do not interact with the patients in person [2]. In case more information on a patient`s medical history is needed, it must be manually searched for in the Hospital Information System (HIS) or requested from the referring physician.

Ideally, radiologists are presented with the most crucial patient information while generating a report, as this could enhance reporting efficiency and patient satisfaction [3]. One possibility of obtaining this information is enabling the patient to do a self-anamnesis using a medical conversational agent (CA) [4]. CAs, also known as chatbots,

[1] Corresponding Author: Kerstin Denecke, Bern University of Applied Sciences, Institute for Medical Informatics, Biel/Bienne, Switzerland, E-Mail: kerstin.denecke@bfh.ch.

are agents that interact with users via written or spoken natural language. Their application is prevalent in several areas of life, including e-commerce, education, personal assistance and healthcare [5]. Specifically in the healthcare domain, CAs have been applied to mental health interventions, appointment scheduling as well symptom checking [6]–[8]. The application of CAs to collect the medical history is still subject to research. However, due to technological advancements in the field of natural language processing (NLP), CA adaptation in general has seen a rise.

According to IBM, NLP refers to "the branch of computer science […] concerned with giving computers the ability to understand text and spoken words in much the same way human beings can" [9]. While rule-based CAs are limited with respect to their functionalities, NLP-enhanced CAs might provide a better, more secure, more effective and more usable user experience.

Without appropriate integration into the clinical workflow, in our case into the radiology workflow, collecting information using a CA from patients is redundant. Therefore, in this paper, we describe the requirement engineering process of a system to support the radiological reporting process by collecting information using a CA from a patient and the processing of the data for presentation to the radiologist during the reporting process. We refer to this system as a Digital Medical Interview Assistant (DMIA) - the chatbot is one component among others forming together the DMIA.

2. Methods

To ensure that the needs and perspectives of the target end-users (radiologists and patients) are considered in the DMIA, a "person-based approach" (PBA) was used as theoretical framework to guide the requirement engineering process. PBA's main objective is to incorporate iterative, in-depth qualitative research into every stage of development to guarantee that the intervention fits with the end-users' psycho-social environment [10]. Additionally, we considered the relevant recommendations of the conceptual framework DISCOVER. This framework provides a step-by-step guide for the design, development, evaluation and implementation of rule-based CAs delivered via smartphone [11]. Therefore, we defined core intervention objectives that serve as guiding principles during DMIA development. Based on these objectives, requirements were specified by means of a narrative literature review and a patient survey, augmented by requirements defined by a radiologist.

First, we carried out a narrative, non-exhaustive literature review with the goal to collect general requirements for the CA, based on existing implementations of CAs and studies thereof. Although a fully systematic approach was not applied due to cost-benefit-considerations, we still applied a structured search approach as follows: For the review, the four search strings "Anamnesis chatbot", "Medical interview chatbot", "Anamnesis conversational agent" and "Medical interview conversational agent" were entered into the literature database Google Scholar. For each query, we considered the first thirty search results and decided based on title and abstract if a search result is deemed of relevance. We then scanned the remaining search results for new insights in CA development and for possible requirements to be added to this specification.

Second, we developed a questionnaire comprising twelve questions to assess requirements of the patients. The first question described the nature of a DMIA, showed an example interaction, and asked the participant whether he or she ever used such system before. We assessed whether the participant could imagine answering questions on the personal medical history when being asked by such DMIA and how many questions would be desirable. On a 5-point Likert scale we asked to judge the importance of certain characteristics of the DMIA (e.g., reacting as a human, explains what happens to the data etc.). We also wanted to know when the participant would consider interacting with the DMIA and which experiences they have made with medical interviews in radiology. The questionnaire was shared by e-mail with the members of the patient lobby group (Patientenrat) from a collaborating hospital on June 13th, 2022. The lobby group comprises 25 patients and relatives of patients, having personal experience with ambulatory or stationary healthcare providers. The questionnaire is available at OSF (DOI: 10.17605/ OSF.IO/49NQV).

Furthermore, we interviewed a radiologist to obtain information on which questions radiologists ask a patient to collect the medical history, which answers they expect and additional scenarios in which the DMIA could be applied. We addressed this for mammography as an example. Last, based on the beforementioned research activities, functional as well as non-functional system requirements for the DMIA were specified. These requirements in turn enabled us to draft a system architecture.

3. Results

Four core intervention objectives according to the PBA approach were defined: A DMIA should collect relevant aspects of the medical history, foster adherence and engagement for patients, present the obtained information in a useful way to radiologists and address peculiarities of a diverse patient population. Based on these principles, we continued the requirement engineering process according to the methods described in the previous section.

3.1. Literature review

From the literature review, we gathered information on requirements towards the DMIA and interaction with the user as well as on modules relevant to design a DMIA. Wang et al. assessed feasibility of a CA for COVID-19 screening before radiology assessment [12]. Particularly, they studied the readability of the CA messages with the Flesch Reading Ease Score and concluded that patient-centred language, reading level and conversation length should be considered for developing an inclusive and accessible radiology communication tool. Ni et al. developed a proof-of-concept CA system intended to be used in primary care [13]. The system consists of a CA that carries out the medical interview as well as a NLP module that extracts the patient's symptoms, reported in lay language. The authors proposed the generation of a thesaurus for synonyms of symptoms to improve performance. Seitz et al. investigated trust-building factors for interacting with CAs, including user-, environment- and software-related factors [14]. Only the latter can be influenced and should be therefore considered by software designers. Rapp et al. carried out a systematic review of research on text-based CAs [15]. Their findings show that in the healthcare domain, empathy might be a relevant factor to determine acceptability. When the CA fails to interpret a user's utterance, the misunderstanding should be acknowledged by the CA itself instead of providing a potentially wrong answer. We conclude that for an effective DMIA, considering the language of the user in terms of readability and understandability is essential. Moreover, empathy matters.

Safi et al. performed a scoping review on technical aspects of developing CAs for medical applications [16]. The authors identified four main components of a CA system, namely text understanding module, a dialogue manager, database layer and text generation module. Denecke et al. developed an evaluation framework for conversational agents in healthcare, comprising concrete metrics, heuristics and checklists [17]. The authors clarify that certain aspects must be defined already in early phases of intervention design and development such as data privacy, accessibility or security. Regarding the DMIA architecture, we derived rather general requirements including aspects related to security and data privacy.

3.2. Patient survey

8 out of 25 requested patients answered the questionnaire resulting in a return rate of 32%. The results show that at least some users already interacted with a CA before. Since CAs become available for customer support on many websites, this trend will continue, and more people are expected to become familiar with these systems. People are willing to answer even more than ten questions. Interaction should be domain-specific (no small talk) and in a human-like manner. Users would like to have the chance to ask questions or make comments. The wording should be easy understandable. Information about the purpose of the DMIA and the use of the data is essential. It should be possible to interact with the DMIA at home as well as in the waiting room. Only a limited amount of free text answers should be asked; predefined answer options are preferred. A voice user interface seems to be unnecessary according to the survey results but might be relevant in terms of accessibility. Predefined answers and interaction with images (e.g. to point to the location of pain on an avatar) are highly desired. Font size must be adaptable or at least not too small. Furthermore, the survey showed that people might reject interacting with a DMIA since they prefer personal contact with a health professional.

3.3. Requirements of radiologists

According to the radiologist, specialized in the domain of mammography, three use-cases can be distinguished for the application of a DMIA: Screening programs, investigation preparation and breast symptom check. The first two use-cases focus on obtaining important information from the patient and providing a structured report of the conversation to the care providers. The third use-case, breast symptom check, is intended to determine how urgent a patient should attend a care provider. It is apparent that for each use-case, the conversation content differs, although many questions are used across multiple domains and use-cases. Therefore, we conclude that a DMIA should enable physicians to easily create new conversation flows and to reuse existing elements. For the second use-case, investigation preparation, such an example conversation flow was defined and consists of 72 questions, comprising general information (regarding e.g. allergies, weight, size) and domain-specific information (regarding e.g. breast implants, previous mammography, breast-related diagnoses). The complete questionnaire is available upon request.

Regarding the preferred structure of the provided information, a Common Data Element (CDE)-based approach was suggested by the radiologist. A CDE defines “the attributes and allowable values of a unit of information” and facilitates the exchange of structured information [18]. In the context of mammography, a CDE could represent e.g., the shape, margin or density of identified masses [19]. Therefore, each question within a conversation flow and its answer corresponds to one structured CDE. This implies that CDEs must be defined together with the conversation flow, enabling reusability across multiple domains and use-cases. For this study, custom CDEs were defined.

3.4. Aggregation of requirements: DMIA system architecture

The above-mentioned requirements are aggregated into 37 functional and 18 non-functional requirements. Based on these requirements, the system architecture was drafted, see Figure 1. Our DMIA comprises five system components: Patients are onboarded to the DMIA via the administration module, which might be integrated into an existing Radiology Information System (RIS). Patients either need to verify their identity via an SMS-confirmation if they interact with the DMIA at home, or they receive a tablet which is preconfigured by administrative staff or, when interacting with a robot-based interface, identify themselves with a barcode and/or date of birth.

The chatbot module handles the complete interaction with the patient, who accesses the DMIA via smartphone, tablet, PC or service robot. Upon completion of the conversation, structured information is exported to external systems as a FHIR QuestionnaireResponse resource, including the original conversation protocol. Also, the content of each conversation flow is defined in a standardized way as a FHIR Questionnaire resource, realized in the content definition module. This module provides a graphical user interface and drag-and-drop functionalities to facilitate questionnaire design. Each item of this questionnaire corresponds to a question to be asked to the patient and at the same time to a CDE which defines allowed answer values. Both resource types serve as an exchange format between the components as well as between the DMIA and external systems.

To extract structured, clinical information from the patient's free text responses, the chatbot module forwards the user's input, together with the CDE definition including possible answer values to the NLP module to be analyzed by an information extraction model and extracted according to the FHIR SDC Implementation Guide [20]. A workflow engine initializes DMIA instances and keeps track of conversation statuses and patient references.

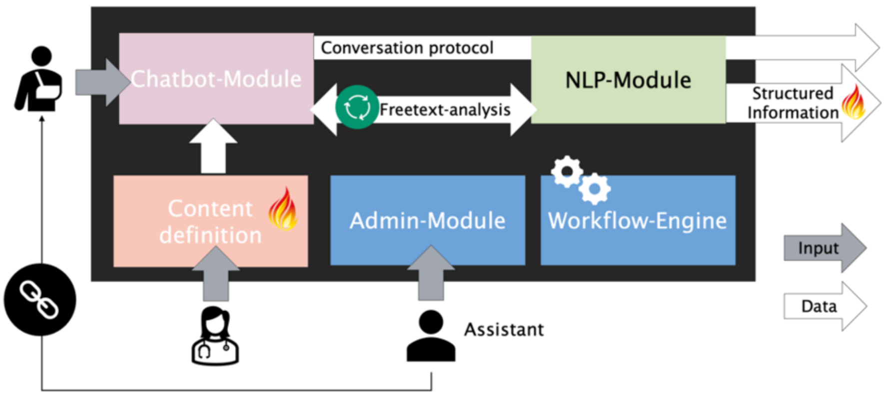

Figure 1. Reference architecture of a Digital Medical Interview Assistant

4. Discussion

In this paper, we described the development process of the conceptual architecture of a DMIA designed based on a three-dimensional requirement engineering process, comprising a literature review as well as retrieving input from patients and radiologists. The resulting 55 requirements provide a detailed guideline during implementation and can be used for a summative evaluation, as defined by [17]. Input and output data formats of the DMIA are defined based on FHIR resources, facilitating development and enhancing interoperability.

Our research process is based on two conceptual frameworks, which we adapted: The PBA approach has been used for other digital applications such as health promotion and disease self-management but has not yet been applied to the development of a DMIA. Therefore, this approach was adapted based on an extended definition of intervention, referred to as an activity undertaken to determine, prevent, improve, or stabilize a medical condition. Hence, we also consider a DMIA as an intervention although in its current design, it only collects information. The DISCOVER framework was originally intended to support the development of smartphone-delivered and rule-based CAs. In contrast, our DMIA might also be accessed via tablet, PC or service robot and provides NLP functionalities which requires considering additional requirements as suggested in DISCOVER. Further, our DMIA comprises additional components than just a CA. DISCOVER addresses in turn the phases of evaluation and implementation which was not the focus of our work.

It is important to acknowledge the limitations of this study, as they may impact the generalizability and applicability of the findings: First, while the DMIA was designed for outpatients as the main target group, the system could also be used by inpatients. However, this requires handling of possible duplicate information already available in the HIS, e.g., by prepopulating answers and asking the patient for any changes. Next, the sample size of the qualitative methods is small. To reduce any bias, more patients and radiologists must be included in the ongoing development process. This also applies to ensuring generalizability of the designed system to other domains beyond mammography-related use cases. Last, we only used custom CDEs only for the first questionnaire. As there are several public repositories of CDEs available, a method needs to be developed to assess whether an already existing CDE is suitable for usage within a DMIA questionnaire.

Next steps include the technical implementation, the definition of functionalities and the generation of a concept for the NLP module. Moreover, work has already started to specify a sustainable process for defining the conversation content: For different radiological examinations, different questions should be asked by the DMIA. Thus, the content must be adapted or extended easily. However, physicians formulate their information needs (or anamnesis questions) not in a way that is understandable to every patient. We are envisioning a process, where CDEs are defined by radiologists and the actual chatbot questions are generated in a semi-automatic process to ensure that the questions are comprehensible for diverse patient groups. After completion of implementation, the application of DMIA should be validated regarding its efficacy to improve the radiological process in clinical context. We assume that in future, a DMIA is a vital part of the treatment process, improving clinical outcome, patient satisfaction and resource allocation.

Acknowledgements

This research was conducted as part of the research project Smaragd, funded by Innosuisse.

References

[1] OECD, 'OECD.Stat', 2023. https://stats.oecd.org/ (accessed Jan. 25, 2023).
[2] A. R. Margulis and H. D. Sostman, 'Radiologist-patient contact during the performance of cross-sectional examinations', *J. Am. Coll. Radiol.*, vol. 1, no. 3, pp. 162–163, Mar. 2004, doi: 10.1016/j.jacr.2003.12.011.
[3] A. G. Rockall, C. Justich, T. Helbich, and V. Vilgrain, 'Patient communication in radiology: Moving up the agenda', *Eur. J. Radiol.*, vol. 155, p. 110464, Oct. 2022, doi: 10.1016/j.ejrad.2022.110464.
[4] K. Denecke, P. Lombardi, and K. Nairz, 'Digital Medical Interview Assistant for Radiology: Opportunities and Challenges', in *Studies in Health Technology and Informatics*, G. Schreier, B. Pfeifer, M. Baumgartner, and D. Hayn, Eds. IOS Press, 2022. doi: 10.3233/SHTI220345.
[5] M. Allouch, A. Azaria, and R. Azoulay, 'Conversational Agents: Goals, Technologies, Vision and Challenges', *Sensors*, vol. 21, no. 24, p. 8448, Dec. 2021, doi: 10.3390/s21248448.
[6] N. Ben-Shabat *et al.*, 'Assessing data gathering of chatbot based symptom checkers - a clinical vignettes study', *Int. J. Med. Inf.*, vol. 168, p. 104897, Dec. 2022, doi: 10.1016/j.ijmedinf.2022.104897.
[7] A. A. Abd-alrazaq, M. Alajlani, A. A. Alalwan, B. M. Bewick, P. Gardner, and M. Househ, 'An overview of the features of chatbots in mental health: A scoping review', *Int. J. Med. Inf.*, vol. 132, p. 103978, Dec. 2019, doi: 10.1016/j.ijmedinf.2019.103978.
[8] P. Amiri and E. Karahanna, 'Chatbot use cases in the Covid-19 public health response', *J. Am. Med. Inform. Assoc.*, vol. 29, no. 5, pp. 1000–1010, May 2022, doi: 10.1093/jamia/ocac014.
[9] IBM, 'What is Natural Language Processing?' https://www.ibm.com/topics/natural-language-processing (accessed Jan. 25, 2023).
[10] L. Yardley, L. Morrison, K. Bradbury, and I. Muller, 'The Person-Based Approach to Intervention Development: Application to Digital Health-Related Behavior Change Interventions', *J. Med. Internet Res.*, vol. 17, no. 1, p. e4055, Jan. 2015, doi: 10.2196/jmir.4055.
[11] D. A. Dhinagaran *et al.*, 'Designing, Developing, Evaluating, and Implementing a Smartphone-Delivered, Rule-Based Conversational Agent (DISCOVER): Development of a Conceptual Framework', *JMIR MHealth UHealth*, vol. 10, no. 10, p. e38740, Oct. 2022, doi: 10.2196/38740.
[12] W. T. Wang, N. Tan, J. A. Hanson, C. A. Crubaugh, and A. K. Hara, 'Initial Experience with a COVID-19 Screening Chatbot Before Radiology Appointments', *J. Digit. Imaging*, May 2022, doi: 10.1007/s10278-022-00650-7.
[13] L. Ni, C. Lu, N. Liu, and J. Liu, 'MANDY: Towards a Smart Primary Care Chatbot Application', in *Knowledge and Systems Sciences*, Singapore, 2017, pp. 38–52. doi: 10.1007/978-981-10-6989-5_4.
[14] L. Seitz, S. Bekmeier-Feuerhahn, and K. Gohil, 'Can we trust a chatbot like a physician? A qualitative study on understanding the emergence of trust toward diagnostic chatbots', *Int. J. Hum.-Comput. Stud.*, vol. 165, p. 102848, Sep. 2022, doi: 10.1016/j.ijhcs.2022.102848.
[15] A. Rapp, L. Curti, and A. Boldi, 'The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots', *Int. J. Hum.-Comput. Stud.*, vol. 151, p. 102630, Jul. 2021, doi: 10.1016/j.ijhcs.2021.102630.
[16] Z. Safi, A. Abd-Alrazaq, M. Khalifa, and M. Househ, 'Technical Aspects of Developing Chatbots for Medical Applications: Scoping Review', *J. Med. Internet Res.*, vol. 22, no. 12, p. e19127, Dec. 2020, doi: 10.2196/19127.
[17] K. Denecke, A. Abd-Alrazaq, M. Househ, and J. Warren, 'Evaluation Metrics for Health Chatbots: A Delphi Study', *Methods Inf. Med.*, vol. 60, no. 05/06, pp. 171–179, Dec. 2021, doi: 10.1055/s-0041-1736664.
[18] D. L. Rubin and C. E. Kahn, 'Common Data Elements in Radiology', *Radiology*, vol. 283, no. 3, pp. 837–844, Jun. 2017, doi: 10.1148/radiol.2016161553.
[19] C. D'Orsi, 'ACR BI-RADS atlas: breast imaging reporting and data system; mammography, ultrasound, magnetic resonance imaging, follow-up and outcome monitoring, data dictionary', *ACR Am. Coll. Radiol.*, 2013.
[20] HL7, 'FHIR Structured Data Capture Implementation guide: Form Data Extraction. 3.0.0 - STU 3', Oct. 01, 2022. https://build.fhir.org/ig/HL7/sdc/extraction.html (accessed Mar. 14, 2023).