CHATBOT ACCEPTANCE IN HEALTHCARE: EXPLAINING USER ADOPTION OF CONVERSATIONAL AGENTS FOR DISEASE DIAGNOSIS

Sven Laumer  
*Friedrich-Alexander University Erlangen-Nuernberg, sven.laumer@fau.de*

Christian Maier  
*University of Bamberg, christian.maier@uni-bamberg.de*

Fabian Tobias Gubler  
*University of Bamberg, mail@fabiangubler.de*

Follow this and additional works at: [https://aisel.aisnet.org/ecis2019_rp](https://aisel.aisnet.org/ecis2019_rp)

**Recommended Citation**  
[https://aisel.aisnet.org/ecis2019_rp/88](https://aisel.aisnet.org/ecis2019_rp/88)

This material is brought to you by the ECIS 2019 Proceedings at AIS Electronic Library (AISeL). It has been accepted for inclusion in Research Papers by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
CHATBOT ACCEPTANCE IN HEALTHCARE: EXPLAINING USER ADOPTION OF CONVERSATIONAL AGENTS FOR DISEASE DIAGNOSIS

Research paper

Laumer, Sven, Friedrich-Alexander Universität Erlangen-Nürnberg, Germany, sven.laumer@fau.de

Maier, Christian, University of Bamberg, Germany, christian.maier@uni-bamberg.de

Gubler Fabian, University of Bamberg, Germany, mail@fabiangubler.de

Abstract

In this research, we develop a research model explaining the adoption of conversational agents for disease diagnosis. Healthcare is challenged by a parallel increasing demand for healthcare services and a decreasing supply of healthcare professionals. Mobile Health is proposed to overcome geographical, temporal, and organizational barriers of healthcare services. Conversational agents (CA), i.e. software programs that interact with users through natural language, are developed that are even able to diagnose a disease based on an individual’s input using a chat interface. However, these systems face an adoption challenge. To understand that, we use UTAUT2 as theoretical lens and 35 semi-structured interviews with potential users of a CA for disease diagnosis. Based on that we propose a research model that contains (1) well-known UTAUT2 factors (performance and effort expectancy, facilitating conditions), (2) re-defined other factors to better fit the context (social influence, price value, habit), and (3) newly identified ones (privacy risk expectancy, trust in provider and system, compatibility, experience in e-diagnosis, access to health system). We also reveal that hedonic motivation is not relevant for CA adoption. The newly proposed model addresses research gaps in CA research in general, but also in mHealth and especially the use of CA in healthcare research in particular. We also discuss rather general implications for technology acceptance research and provide some suggestions for providers of CA in healthcare to increase the diffusion rates of these systems.

Keywords: Conversational agents, disease diagnosis, adoption, chatbot, mHealth, UTAUT

1 Introduction

The increasing demand for healthcare service challenges healthcare providers as particularly the shortage of healthcare professional cause that they have to overcome geographical, temporal, and organizational barriers to provide a good healthcare service to patients (Deng et al. 2014; Samhan and Joshi 2015). So-called mobile health services (mHealth) are discussed to be useful to overcome these barriers and provide patients better access to health services at low cost, regardless of place and time (Samhan 2017; Silva et al. 2015). While the first mHealth applications provide only simple features, such as an app for counting calories, apps are now more complex and among others provide diagnostic suggestions to patients (Miner et al. 2017; Silva et al. 2015).

Conversational agents (CA), i.e. software programs that interact with users through natural language using either a text- or voice-based interface, are a new type of mHealth service (Carroll et al. 2017; Guo et al. 2013). One example is Ada (Ada Health GmbH 2018) that enables its users to respond to several
questions such that Ada provides a diagnostic suggestion based on the information provided using a natural language interface. Ada can identify over 1,500 clinical pictures and 200 rare diseases (Flick 2018). Since legal requirements have so far forbidden definitive remote diagnoses by apps, Ada currently offers a suggestion, which (specialist) physician should be consulted. Although, Ada is rated 4.7 out of 5 stars in the Google Play Store, the following comments from the Google Play Store illustrates that there is an adoption challenge of CA for private healthcare, meaning that low number of patients start using Ada, although it is shown that it provides value to its users:

“I thought it would be total nonsense.” (female, 34),

“I was really sceptical and unsure whether to use this app.” (male, 46),

“My first thoughts were: This will never work and replace my doctor.” (female, 54)

This is in line with general conclusions, that mHealth service have found little support among physicians and patients (Fitzpatrick et al. 2017; Guo et al. 2013). Especially persons who would benefited most from mHealth adopt them only inadequately (Deng et al. 2014; Guo et al. 2013; Samhan 2017). Hence, to overcome barriers and to provide patients access to health services (Samhan and Joshi 2015) it is important to explain how and why patients would use a CA for healthcare related tasks and especially diagnostic suggestion. So, we will discuss the following research question:

**What are the factors that influence whether individuals adopt conversational agents for disease diagnosis?**

From a theoretical perspective, we focus on technology acceptance research and especially the Unified Theory of Acceptance and Use of Technology (UTAUT2) for a consumer context (Venkatesh et al. 2012, see section 2). This theoretical lens is used as it is originally developed for consumer contexts, such as mHealth or the specific consumer app Ada, which is focused in that research. Based on UTAUT2 we develop a model to explain adoption of CA in a healthcare context. From a methodology perspective (see section 3), we do so by conducting 35 interview. Based on these interviews we propose a research model explaining the adoption of CA by patients for disease diagnosis (see section 4). The model provides then implications for both theory and practice (see section 5).

## 2 Related Research

### 2.1 Conversational Agents

CA refer to both text-based and voice-based automated dialog systems that can interact with a human user via natural language and answer questions on specific topics. The architecture of a CA has three levels. First, the ‘interface’ as the first processing level, which captures the user’s language either as text or spoken language (Sarikaya 2017; Yan et al. 2016), for example, when the user provides a description of symptoms or pictures of some body reactions using a chat interface. Second, within the further processing levels (‘algorithm’), the user input is translated into a set of actions for processing the action requested by the user. Third, this is done using the data required to process the request (‘data’), for example, when an algorithm uses clinical pictures and compare them with the input to identify a possible disease. The processing level produce the output, which is converted into either text- or voice-based text output where possible, for example, when a suggestion of a diagnosis and which (specialist) physician one should consult, is provided to the user via the chat interface. While CA used to be simple input-output patterns, many CA can be classified as artificial intelligence (AI) (Gnewuch et al. 2017; Gnewuch et al. 2018; Sarikaya 2017).

There are few research studies on individual acceptance and use of CA. Among others, these address different aspects including design principles (Gnewuch et al. 2017; Nunamaker et al. 2011; Seymour et al. 2017), users’ information disclosure (Schroeder and Schroeder 2018), the impact of the CA on user’s cognitive abilities or performance (Siddike et al. 2018), the relationship between humans and CA (Elson et al. 2018; Schuetzler et al. 2018; Wuenderlich and Paluch 2017), privacy concerns (Saffarizadeh et al. 2018).
2017), and trust (Elson et al. 2018; Saffarizadeh et al. 2017). However, we observe a lack of user acceptance studies discussing individual adoption and use of CA in a healthcare context, as most CA studies in general are limited to systems that serve entertainment purposes and neglect systems relevant for utilitarian purposes.

2.2 mHealth and Conversational Agents

Research on mHealth focuses on mobile applications for prevention, discovery of health problems, diagnosis of diseases, treatment of diseases, ensuring a good end of life, and healthcare administration (Bergman et al. 2011). Regarding the use of CA, some research has been conducted to identify the requirements for CA (Coyne et al. 2017; Miner et al. 2016), to discuss potential and challenges of CA in healthcare (Miner et al. 2017; Riccardi 2014) and to investigate the use of cross-domain CA (e.g. Siri) (Rammohan et al. 2017; Wilson et al. 2017). The majority of these prior works focus on design characteristics and ethical challenges of using CA in healthcare, whereas no study investigated the patients’ adoption of CA for private healthcare purposes.

The focus of this paper is on the diagnosis of diseases, which is the process to identify the causes of health problems and to initiate the subsequent treatment. From the individual’s point of view, this also includes self-diagnosis on the Internet (Bergman et al. 2011). It is different to the discovery of health issues, which is about the proactive monitoring of the individual's own state of health (Bergman et al. 2011). Prior research that focuses explicitly on CA for the diagnosis of diseases report design characteristics or proof-of-concepts demonstrating that it is possible to diagnose a disease using a CA (Denecke et al. 2018; Divya et al. 2018; Minutoloa et al. 2017; Mishra et al. 2018; Ni et al. 2017). However, these studies do also not investigate the described adoption challenge of CA, that low number of individuals adopt a CA for disease diagnosis, although it is shown that these type of systems provide value to the users and are able to overcome barriers in the delivery of healthcare services. Hence, no study has published the type of study that is presented in this paper.

2.3 Technology Acceptance Research

To address the described adoption challenge of CA and the identified gaps in the literature, we build on technology acceptance research. Several models explaining individual technology adoption have been suggested, generalized, contextualised and revised (Venkatesh et al. 2003). One refinement is the Unified Theory of Acceptance and Use of Technology for the consumer context (UTAUT2) (Venkatesh et al. 2012), which has been discussed as the most recent advancement for explaining consumer adoption and use in a private context (Gao et al. 2015; Wong et al. 2014). As the focus of this paper is on a private (healthcare) context as well and we are intended to research consumers (patients) adoption of CA for disease diagnosis we select UTAUT2 as the most appropriate model to develop a model explaining the adoption of CA in a private healthcare context by patients.

Another reason to focus on UTAUT2 as the underlying model for our research is that it has also been applied several times in the mHealth context (Gao et al. 2015; Yuan et al. 2015). However, it has also been highlighted that a total of 94 different predicators for acceptance of mHealth services, 62 of which were significant in at least one study, has been discussed, whereas the majority of these approaches did not use a theoretical base for investigating mHealth acceptance (Or and Karsh 2009). By reviewing prior mHealth acceptance studies, we have to conclude that these studies focus on different applications such that mHealth research does not provide a model for CA adoption by patients for disease diagnosis. Hence, we will introduce UTAUT2 in the following to use it as our theoretical base to develop a model explaining the adoption of CA for disease diagnosis by patients.

UTAUT2 explains consumers’ voluntary adoption and use of a technology by a consumer’s intention to adopt or use this technology. As factors influencing the intention, UTAUT2 theorizes the following factors:
- **performance expectations**, which is defined as “the degree to which using a technology will provide benefits to consumers in performing certain activities” (p.159),
- **effort expectations**, which is defined as “the degree of ease associated with consumers’ use of technology” (p.159),
- **social influences**, which is defined as “the extent to which consumers perceive that important others believe they should use a particular technology” (p.159),
- **facilitating conditions**, which is defined as “consumers’ perceptions of the resources and support available to perform a behaviour” (p. 159),
- **hedonic motivations**, which is defined as “the fun or pleasure derived from using a technology” (p. 161),
- **price values**, which is defined as “consumers’ cognitive trade-off between the perceived benefits of the applications and the monetary cost for using them” (p. 159), and
- **habits**, which is defined as the “extent to which people tend to perform behaviours automatically because of learning” (p.159) (Venkatesh et al. 2012).

3  **Research Design**

The overall approach of this paper is to develop a research model based on UTAUT2 for explaining the acceptance of CA for disease diagnosis. We choose a three-stage qualitative approach (Elliott and Timulak 2005). Initially, relevant concepts were identified by examining the state of research in the context of technology acceptance and the use of CA in healthcare (see section 2). Based on that an interview guideline was developed. Subsequently, the recruitment of the interview participants and the execution of semi-structured individual interviews took place, before the research was completed by coding and analysing these interviews to identify factors explaining the acceptance of CA in a private healthcare context.

3.1  **Interview-guideline**

The interview guideline (see Appendix, Table 1) ensured that respondents were given sufficient freedom to describe their experiences with and opinions regarding CA for disease diagnosis. This made it possible to analyse UTAUT2 factors, but also to identify additional acceptance factors, which had not been taken into account in UTAUT2. The interview guideline was designed as a semi-structured one as this format enables consistency across the different interviews and allows the interviewer to respond to what is new in the dialogue and to clarify unclear or ambiguous answers by asking questions (Bryman 2016; Myers 2010). We refined the guidelines on the basis of two pilot interviews.

3.2  **Interviews**

Overall, 35 (18 female, 17 male, age mean=24.8 years, age min=19, age max=40, 34 students, 1 professional, 11 computer science or information system students, 10 business students, 13 social science students; 31 having experiences in using CA, 4 with no experience) semi-structured interviews were conducted. We chose students as participants for our study because they belong to a younger age group and are more likely to belong to the group of early adopters. Early adopters are individuals who have a higher level of education and higher socioeconomic status (Rogers 1983). We ended our data collection when we saw that only redundant aspects were identified in subsequent interviews and that additional interviews would not provide new insights (Lapointe and Rivard 2005). Our sample is above the recommended number of twelve interviews for a homogeneity group of interviewees (Guest et al. 2006).

The interviews were conducted exclusively in person and took place in an office of our university to guarantee freedom from disturbance. The interview participants were recruited through an open call via student groups in Facebook as well as through personal contact. As an incentive to participate, three Amazon vouchers worth 20 euros each were raffled among the participants. A reference was made to the audio recording and the introduction to the interview. Here, it was emphasized to reduce a possible
response bias as it was mentioned all answers and opinions are treated anonymously and strictly confidentially, that expressions in both positive and negative directions are possible, and that the recording of the interview is used only for study purposes.

3.3 Interview analysis

The transcription in preparation for the data analysis was carried out after each interview to ensure that thematic aspects were not lost, for example through incomprehensible passages in the interview recordings. To systematically analyse and categorize the information from the interviews, qualitative content analysis (Mayring 2014; Schreier 2013) was chosen as this enables to build on existing acceptance research and to generate new propositions from the interviews. The interviews were coded in a parallel deductive and inductive approach, in which the code categories were continuously adapted on the basis of the interview material (see Gruzd et al. 2012 for a similar approach) using MAXQDA.

The deductive coding focused on the theoretical foundation of our research such thematic codes based on UTAUT2 has been identified (Mayring 2014). For example, statements like “I don't see any gain yet with Ada” were coded as performance expectancy, as the statement reflects the definition of the factor performance expectancy proposed by UTAUT2. In a similar deductive approach, we searched for statements reflecting the factors proposed by UTAUT2.

The parallel inductive coding focused on factors that explain the adoption of CA for disease diagnosis, but which were not focused by UTAUT2 so far. For example, when statements like "And then my personal data. I think the subject of data protection is somehow too delicate for me.” occurred they were labelled as new factor (here: privacy). Based on the identified codes reflecting new factors they were grouped afterwards and a definition was proposed to reflect the overall theme of the codes identified (Boyatzis 1998; Schreier 2013).

There are frequent warnings against relying biased on existing theory when conducting qualitative research, as this would prevent new findings and theoretical breakthroughs (Andersen and Kragh 2010; Glaser and Strauss 2009; Maxwell 2013). However, the aim of this work is not to develop a completely new theory, but rather to adapt previous technology acceptance research to the new application context of CA (Hong et al. 2014, Elliott und Timulak 2005; Andersen and Kragh 2010). For this reason, an approach is chosen here in which theory formation takes place as interpolation and existing theoretical considerations are related to the empirical data (Andersen and Kragh 2010). Hence, we coded and analysed the data first. When the inductive approach revealed a new factor, prior research focusing on this factor in a different context was considered as well to derive the propositions explaining patients’ adoption of CA for disease diagnosis. We present the results of this approach in the following.

4 Research Results

In our analysis we identified factors explaining adoption of CA in healthcare, which correspond with the factors suggest by UTAUT2 (see section 4.1, grey shaded in Figure 1) or were slightly redefined to fit the context of CA adoption for disease diagnosis (see section 4.2., dark grey shaded in Figure 1). Moreover, we identified new factors, which are specifically relevant for the context of CA in healthcare (see section 4.3, black shaded in Figure 1). We also did not find support for factors that were suggested by UTAUT2 (see section 4.4, white shaded in Figure 1). Hence, a synthesis of the acceptance factors of a CA for disease diagnosis and their impact relationships described in more detail in the following subchapters results in the research model shown in Figure 1. We will discuss the resulting implications after the results presentation.
Figure 1. Acceptance of Conversational Agents for Disease Diagnosis

4.1 UTAUT2 Factors

4.1.1 Performance Expectancy

The interviews show that the intention to adopt a CA for disease diagnosis depends to a large extent on performance expectancy, which is patients’ expectations that using a CA for disease diagnosis will result in a faster and/or better diagnosis of one’s symptoms. The following statements exemplary for similar statements was found in the interviews:

"I could imagine using it because [...] an initial consultation with the doctor is then omitted and, in the best case, the doctor then gets the information directly." (#32)

The interviews confirmed the results of earlier studies using UTAUT2 for explaining acceptance of mHealth (Dwivedi et al. 2016; Hoque et al. 2017; Sun et al. 2013). However, the interviews indicate that the majority of the interviewees do not want to use a system like Ada, because they believe that the diagnostic capability is not yet sufficient or even leads to false diagnoses. In the sense of these remarks, the first proposition based on UTAUT2 and our analysis is:

**P1**: The better (worse) the performance expectancy the higher (lower) the intention to adopt conversational agents for disease diagnosis.

4.1.2 Effort Expectancy

The interviews also show that effort expectancy is an important factor explaining the intention to adopt a CA for disease diagnosis. It reflects the extent to which the potential user thinks that much or little effort is necessary to use an appropriate CA and to learn how to deal with it. In our interviews we found:

"It would be too complicated for me, and it would be too time-consuming for me to type it in or look it up." (#23)

This observation is supported by related research (Bickmore et al. 2009; Coskun-Setirek and Mardikyan 2017) supposing that effort expectancy is important for whether individuals adopt CA, such that we propose:

**P2**: The better (worse) the effort expectancy the higher (lower) the intention to adopt conversational agents for disease diagnosis.
4.1.3 Facilitating Conditions

We also found support in our interviews for the UTAUT2 factor facilitating conditions, which reflects patients’ perceptions that they are equipped with the resources and the support necessary to use a CA for disease diagnosis:

"You first need such a device. So if I had such a device and then maybe I would ask them something about a disease, too." (#24)

Since CA are used to diagnose disease in private contexts, the interviewees refer to supporting conditions as having necessary resources or that he can get help from others if necessary. Based on these remarks the following hypotheses are formulated, taking into account that research has shown both the direct and indirect influence of facilitating conditions on intention and use behaviour (Venkatesh et al. 2012):

P3a: The better (worse) the facilitating conditions, the higher (lower) the intention to adopt conversational agents for disease diagnosis.

P3b: The better (worse) the facilitating conditions, the higher (lower) the actual use of conversational agents for disease diagnosis.

4.2 Redefined UTAUT2 Factors

4.2.1 Social Influence

A factor regularly mentioned in the interviews is the influence of friends, acquaintances or work colleagues. In previous acceptance research social influence has been understood as the perception of a concrete call to action by the social environment (Venkatesh et al. 2012) or the expectation that the social environment demands a specific behaviour (Eckhardt et al. 2009). The statements identified in our research show that not only the call to or the expectation for a specific behaviour is important, but also the recommendations and experience reported by persons whom the individual trusts:

"If, for example, someone of my friends, acquaintances and families had very positive experiences and said: ‘Yes, come on, you have to try’ then I would definitely think about using it." (#26)

Especially in the health care system, the interview statements suggest the social influence is important for the final adoption decision. Some interviewees can imagine adopting a CA simply based on experience reports or recommendations from their social environment. Hence, social influence is an important factor in the acceptance of CA in healthcare, however, it definition needs to be refined to better fit the context of CA adoption for disease diagnosis. Based on our interviews we define social influence as the extent to which consumers perceive that important others believe they should use a particular technology, that important others say they should use a particular technology, that important others recommend to use a technology or that individuals observes that others are using a specific technology. Based on these remarks we formulate the following proposition:

P3: The higher (lower) the social influence, the higher (lower) the intention to adopt conversational agents for disease diagnosis.

4.2.2 Price Value

In our interviews, we found also support for the relevance of the price value for explaining the adoption of CA for disease diagnosis. Price value was discussed in the literature as a comparison of the costs for using an app such as Ada to the benefits obtained from using the app (Venkatesh et al. 2012). In our interviews, we also found comparison of the cost of using an app such as Ada to the costs using other healthcare services such as visiting a doctor:

"I don't know if it would be worth the money to me, because these things always cost something after all. And in my opinion, the benefit is not enough for me to use something like this." (#17)

"When people would have to pay for each visit at the doctor, it might be costless alternative to ask the app instead of actually going to a doctor." (#2)
Especially, the second aspect is an extension of the definition of price value as provided by Venkatesh et al. (2012), which is especially relevant in the healthcare section according to our interviewees. While a patient with statutory health insurance in a country such as Germany has to pay little for a visit to the doctor in any case, in other countries such as the USA, depending on the insurance status, considerable costs are incurred for each visit to the doctor. Presumably, this possible cost saving will be a decisive factor in countries with poor insurance coverage or with high costs of using other healthcare service. Therefore, it is not only important to compare the direct costs of using an CA for disease diagnosis to the benefits obtained, but also the costs of using a other healthcare service to the cost of using a CA for disease diagnosis by obtaining similar benefits. Therefore, we re-define price value as a consumers’ trade-off between the perceived benefits of the applications and the monetary cost for using them and the trade-off between the costs for using the applications and the costs of using other services. The following proposition is formulated:

**P5:** The better (worse) the price value, the higher (lower) the intention to adopt conversational agents for disease diagnosis.

### 4.2.3 Habit

We also found support for a habit that influences the adoption of CA for disease diagnosis in our interviews. However, it is not related to the use of CA for disease diagnosis, but to the use of CA in general. A habit is the “extent to which people tend to perform behaviours automatically because of learning” (Venkatesh et al. 2012). The interviews did not mention a habit to use a CA for disease diagnosis, as it would correspond with the definition provided, but the effect that when they have a habit to use a CA for any question, they will also use a CA for disease diagnosis. Hence, the habit of using CA in general can also trigger the use of a CA for a specific purpose:

> "And maybe habit, so when people automatically talk to a chat bot they will also start asking questions about their health." (#24)

We therefore extend the definition of habit such that it also covers the extent to which people tend to perform automatically because of been used to perform a specific action that is very close to the behaviour of interest. So in the sense of the arguments of Limayem et al. (2007) it is assumed that the habit of using a CA in general directly influences the actual use and also indirectly influences it via intention:

**P6a:** The higher (lower) the habit of using conversational agents in general, the higher (lower) the intention to adopt conversational agents for disease diagnosis.

**P6b:** The higher (lower) the habit of using conversational agents, the higher (lower) the actual use of conversational agents for disease diagnosis.

### 4.3 Additional Factors

Besides the well-established UTAUT2 factors, we also identified factors that has not been discussed so far, but revealed to be important in our interviews. These are presented next.

#### 4.3.1 (Data) Privacy Risk Expectancy

One factor that was mentioned in numerous interviews is the data protection risk perceived by users and the risk of impairment of personal privacy when using CA for a healthcare purpose. A large number of respondents stated that they do not want to use CA to inadequate data protection. The following quote shows exemplary for a large number of similar statements made by the interview participants that their privacy risk expectancy influences their intentions to use CA for disease diagnosis:

> "That's just my data protection concerns. In the health sector, of course, it is particularly sensitive data. I think the subject of data protection is somehow too delicate for me such that I would not provide too much information about myself. But the app would require this information to be useful, thus, there is no sense for me to use such an app." (#15)
Some interviewees said that data protection is even more important for using CA in the healthcare sector, as it involves much more sensitive data than, for example, the everyday use of systems such as Amazon Alexa or Apple Siri. In the health context, the danger of unauthorized use of this data by third parties is particularly high (Wirth et al. 2019) such that the interviewees mentioned both reduced performance expectancy perceptions and low intentions to adopt.

Thus, the proposed model for explaining acceptance of CA for disease diagnosis also includes a "privacy risk expectancy" (Pavlou 2003), which is not covered by UTAUT2 so far. It is defined as the risk perceived by the user of possible impairment of one’s privacy due to inadequate data protection when using CA for disease diagnosis. In the sense of the statements from the interviews that their privacy risk expectancy influences both their performance expectancy and intention to adopt, the argumentation of Cenfetelli (2004) on the effect of an inhibitor (e.g. perceived privacy risk) on enablers (e.g. performance expectancy), the following propositions are formulated:

P7a: The higher (lower) the privacy risk expectancy, the lower (higher) the intention to adopt conversational agents for disease diagnosis.

P7b: The higher (lower) the privacy risk expectancy, the lower (higher) the performance expectancy.

4.3.2 Trust

In the interviews, trust was another important factor for the adoption decision of CA. In this respect, the participants frequently referred to existing relationships of trust with physicians and compared it to their potential relationship to a CA for disease diagnosis. They also mentioned their trust in the provider of a CA for health diagnosis to be a relevant factor to use such as system.

"Well, I guess I just wouldn't really trust it. I think the diagnosis cannot be reliable. Well, I don't think I would trust that and in case of doubt I would always rather contact a doctor." (#15)

"It depends on the provider. This app is only provided by this company. So, this seems to be ok. But if Amazon itself would provide this app, I would not use it because I do not trust Amazon at all." (#11)

Hence, in line with prior research the interviews show that trust is not a monolithic phenomenon and it should be distinguished into "trust in the provider" and "trust in the technology" (McKnight et al. 2002; Söllner et al. 2016).

**Trust in the provider** plays a decisive role as to whether the interviewee would ultimately use the dialogue system or not. In our interviews, we identified that trust was mentioned in relation to other acceptance factors such as privacy risk ("I don't know what security measures this app is based on. Instead of me going to a doctor, and I know it's definitely in safe hands."), performance ("One somehow overdiaognose oneself and then immediately assume the worst, so I have more confidence in the actual doctor than this app") or effort expectations ("If I know that I can trust them than there would be not much efforts to use it").

In the sense of our interview results that trust in the provider is related to privacy expectations, the arguments of Gefen (2002) – proposing that trust is an important factor for the reduction of uncertainties or perceived risks in social or economic interactions – and of Krasnova et al. (2010), who already proved this relationship for the context of social networks, we formulate the following propositions:

P8a: The higher (lower) the trust in the providers of a conversational agent for disease diagnosis, the lower (higher) the privacy risk expectations.

Our interview results highlighting the relation between trust and performance and effort expectancy reference to the work of (Söllner et al. 2016), who has theorized about the relationships between trust in the provider and performance as well as effort expectancy in a different context. Hence, we assume:

P8b: The higher (lower) the trust in the providers of a conversational agent for disease diagnosis, the higher (lower) the performance expectancy.
P8c: The higher (lower) the trust in the providers of a conversational agent for disease diagnosis, the higher (lower) the effort expectancy.

Trust in the technology. The user of the CA has to trust the underlying ability for correct diagnosis. Analogous to how patients usually trust a physician with regard to their diagnosis, similar trust of the user must be established for dialogue systems. In the sense of the work of McKnight et al. (2011), trust in the diagnostic system is understood to mean the belief that it has the capacity to diagnose diseases. Our interviewees also mention this aspect in relation to the performance expectancy ("If I can trust the results, than this app is really useful") and the effort expectancy. However, regarding the latter as already indicated by prior research (Benbasat and Wang 2005; Söllner et al. 2016) as a reverse effect ("If I understand how to use the app, I would rather trust it easier") such that we assume:

P8d: The higher (lower) the trust in the conversational agent for disease diagnosis itself, the higher (lower) the performance expectancy.

P8e: The higher (lower) the effort expectancy, the higher (lower) the trust in the conversational agent for disease diagnosis itself.

With regard to the relationship between the two aspects of trust, it was found that higher trust in the provider also strengthens initial trust in the CA. A person who trusts Amazon, for example, is also more willing to trust a system developed by Amazon (Pavlou and Dimoka 2006). Thus follows:

P8f: The higher the trust in the providers of conversational agent for disease diagnosis, the higher the trust in the conversational agent for disease diagnosis itself.

4.3.3 Compatibility

The respondents also stated that the compatibility of the dialogue system with their environment is important. In the health context, compatibility of the CA with the existing health environment can increase the perception of the usefulness of the CA for disease diagnosis:

"So, if the connection to my doctors were really there then it would be useful. So if the compatibility is really given, this would make sense." (#33)

Thus, the proposed model for explaining acceptance of CA for disease diagnosis also includes a "compatibility" factor, which is not covered by UTAUT2 so far. It is defined as the perception that the CA for disease diagnosis is well integrated in the healthcare environment the patient is based in. In the sense of the statements from the interviews that this perception influences the performance expectancy of a CA for disease diagnosis it follows:

P9. The higher (lower) the compatibility of CA with the user's health environment, the higher (lower) the performance expectancy.

4.3.4 Previous Experience

The user’s previous experience with online diagnoses in general is another important factor mentioned in our interviews. When the interviewees were asked about previous experiences in the field of diagnosing diseases through IT support, all persons mentioned that they had already searched the internet for symptoms and diseases. However, most of them described negative experiences:

"You google something, and you look at ten pages, and on nine of them you are terminally ill and have something terrible. If the CA provides the same, it is useless." (#22)

The interviews revealed a clear connection between previous experiences with online diagnosis and the performance expectations of CA for disease diagnosis. Thus, the proposed model for explaining acceptance of CA for disease diagnosis also includes a ‘previous experience’ factor, which is not covered by UTAUT2 so far. It is defined as the experiences made with disease diagnosis on the Internet in general. In the sense of the statements from the interviews that this perception influences the performance expectancy of a CA for disease diagnosis we assume:

P10: The better (worse) the previous experience with IT systems or online platforms for disease diagnosis the higher (lower) the performance expectancy.
4.3.5 Perceived Access to Health System

Individuals with at very limited access to public health services will assess the value and usefulness of such a technology differently (Steele et al. 2009). This was also shown in the interviews. Some participants said that rapid access to doctors has so far prevented them from using systems such as Ada. But distance from doctors or their non-availability could lead them to use a CA for diagnosis:

"We have a relatively large number of doctors relatively close to our community. However, if one lives somewhere in Wallachia, it might actually be more useful than driving a car for two hours to see a doctor." (#35)

Thus, the user's perceived access to the health care system, especially to diagnostic services significantly determines ones impression of the usefulness of a CA for diagnosis. This access can be limited by local, financial or institutional factors (e.g. long waiting times). The more satisfied the user is with ones access to diagnostic services from physicians, the lower the benefits of the CA for diagnosis. Hence, we propose

P11: The better (worse) the perceived access to health systems the lower (higher) the performance expectancy.

4.4 UTAUT2 factors not discovered in the interviews

In addition to the factors described above, which were revealed in the interviews as possible factors influencing an individual's decision to adopt an CA for the diagnosis of diseases, technology acceptance research has analysed many other factors (Venkatesh et al. 2016). We are not able to find support for these factors to be included in our proposed research model based on our interviews. One of the factors is hedonic motivation, which is proposed by the UTAUT2 (Venkatesh et al. 2012), but could not be revealed as important to explain the acceptance of CA in the healthcare sector. Nonetheless, this is in line with other studies in the healthcare context which have also not identified hedonic motivation to be an important factor (e.g. Tavares and Oliveira 2016; Vassli and Farshchian 2018). Consequently, according to the interviews conducted, the factor hedonic motivation is excluded.

5 Discussion

The overall contribution of this work is to propose a research model that theorizes acceptance factors relevant for explaining the acceptance of CA for disease diagnosis. By doing so, several UTAUT2 factors are identified as being relevant in this context or redefined to better fit the context of CA for disease diagnosis, whereas one factor (hedonic motivation) could not be revealed as important. Furthermore, evidence was provided for factors not considered by UTAUT2 so far, but which are relevant according to the interviews conducted. It is a step towards addressing the identified adoption challenge of a specific mHealth service that enable to overcome barriers of healthcare services (Deng et al. 2014) and to provide patients access to health services (Samhan and Joshi 2015). Hence, the proposed research model explains the adoption of CA for disease diagnosis, which focuses on the research gaps discussed in section 2, and also proposes three rather general implications as follows.

5.1 Implications for Research

Research on CA in general (see section 2.1) has identified fun as an important predictor of using CA for an entertainment propose (Gnewuch et al. 2017; Nunamaker et al. 2011; Seymour et al. 2017). Our study provides a first general CA adoption study in a more serious context identifying factors that are specifically relevant in this context (e.g. privacy risk expectancy) and by illustrating that hedonic motivation is not important in a more serious context. This indicates that some factors identified in one context, might not be relevant in another one, whereas the seriousness of the context is a trigger for the importance of hedonic motivation in explaining the acceptance of CA. Based on our results in comparison to the results document in the literature, we propose that if the CA serves an entertainment purpose hedonic motivation is important, whereas when it serves a more serious purpose hedonic motivation is not important. Regarding research on personal healthcare services and CA in this context
(see section 2.2), we provide a first study revealing factors that are important to explain the intention to adopt CA for a healthcare purpose in general and disease diagnosis in particular. The majority of these prior works (e.g., Denecke et al. 2018; Divya et al. 2018; Minutoloa et al. 2017; Mishra et al. 2018) focus on design characteristics, and ethical challenges of using CA in healthcare, whereas no study investigated the patients acceptance of CA for private healthcare purposes, even though a large number of individuals do not adopt them. Regarding research focusing on the acceptance of personal healthcare services (see section 2.3), we are able to extend research on mHealth services (Or and Karsh 2009). We contribute to this stream of research by focusing on a different application that has not focused so far and we identify the factors relevant for explaining the adoption of a specific healthcare service.

Besides addressing these specific gaps in the literature, proposing a research model for the adoption of CA for disease diagnoses reveals also some more general implications. First, we show that certain beliefs require extensions or adaptations to the research context and we provide three revised construct definitions that might also be relevant for other studies applying UTAUT2 (Venkatesh et al. 2012). In our study we redefined the constructs price value, social influence and habit to be more appropriate to explain the adoption of CA for disease diagnosis. First, regarding price value we conclude that the costs of a technology need to be compared to costs of other services and not only to the benefits obtained from the technology itself. Second, social influence can be provided by different referent groups using different forms of influences. Third, habituational influences of using a specific class of technology (here: CA in general) might influence the use of a specific instance of this class of technology (here: CA for disease diagnosis). Future research can focus on developing revised measurement items for the redefined constructs derived from our research approach to empirically test the effects of these constructs on adoption behaviour in this, but also in other contexts (e.g., Laumer et al. 2013; Laumer et al. 2015).

Second, beside the general identification of new factors being important in the CA in healthcare context, one specific theme emerged by grouping the newly identified factors. Our study highlights the importance of environmental factors (compatibility, experiences in e-diagnosis, access to health system) for explaining the performance expectancy of CA in healthcare. Hence, we identified a specific set of factors that can also be further considered in other context to developed a general revised UTAUT2 model. This is in line with a call by Venkatesh et al. 2011, who called for research investigating environment-specific factors. Future research can develop measurement instruments for the identified environmental factors and test the relative importance to other factors using empirical approaches.

Overall, we observe that individuals are willing to adopt CA for disease diagnosis, but they often do not see them as a substitute for the existing health care system. Future research should not be based on the assumption of earlier acceptance research that new technologies will completely replace old practices (Moore and Benbasat 1991). Instead, it should be investigated under which conditions (e.g. severity of symptoms, lack of time, etc.) potential users prefer CA for diagnosis to the discussion with the physician and when not and how a combination of both can be designed to be accepted as an effective and efficient health service by patients. Our proposed research model is the base for these future research approaches.

5.2 Implications for Practice

There are also implications for CA providers in order to increase the adoption rates of CA for disease diagnosis. Performance expectations is one of the most important factor in our interviews. CA providers should therefore clearly explain how such a CA improve an existing health care situation. For example, all interviewees used Google at least once to find out about their own symptoms and even to diagnose them. However, the experience concerning this Internet search was predominantly described as negative. Here, healthcare CA providers can show the usefulness of their CA compared to Internet searches. For example, comparative studies in cooperation with doctors or health insurance companies can prove that a CA such as Ada can achieve significantly better results than a Google search. The interviews also showed that privacy risk expectancy influence the adoption decision. CA providers can create more transparency about the type and use of stored data in order to reduce the subjectively experienced data protection risk for potential users. Data protection concerns can be reduced by concise and well presented data protection declarations as well as by laymen's understandable descriptions of the
measures taken to protect the collected data from access by third parties (Featherman and Pavlou 2003; Yang et al. 2015). In addition, trust-building measures are also appropriate, since higher trust in the providers of a CA and their technology can reduce the subjectively perceived risk to the user's privacy. Providers can demonstrate how the system is externally protected. One possible measure is to show that the CA meets all official standards on health and data protection, for example that all user data is stored in encrypted form in Europe or that the function of a CA for diagnosing diseases are certified by insurance companies or other authorities ("kitemarking", Delamothe 2000). One can also convince important others such as doctors such that there are positive social influences.

5.3 Limitations
This work has some limitations that have to be considered. One of them is the limited generalizability of the findings. For example, only students from a German university are interviewed. Older persons as well as persons with a lower level of education are not taken into account. However, with an average age of 24.8 years, the interviewees represent an age group that is significantly promoting the adoption of CA (PwC 2018). Moreover, as a technology that is continuously developed and expanded by additional features, the knowledge gained in February 2018 cannot be projected into the future, since it can be assumed that the perception of use will continue to develop as the systems continue to develop. Furthermore, it should be noted that there is a possible distortion in the selection of participants. Since only a few interview participants were directly addressed and motivated to participate, most interviewees volunteered to participate via open calls in social networks such that persons having a certain interest in CA were presumably overrepresented in the study.

6 Appendix: Interview guideline

<table>
<thead>
<tr>
<th>Part A: General use and experience with CA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Which CA did you used the most so far? For what purposes? What are the most common functions you have used so far? When and why did you start using CA? What are experiences in the general use of CA? Are there any things you remembered positively or negatively?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part B: Use and experiences with CA in healthcare</th>
</tr>
</thead>
<tbody>
<tr>
<td>For what purposes CA have been used or can be used in the healthcare sector? Describe scenarios for what CA can be used in the health sector? What do you think are the benefits or could be the benefits of using CA in healthcare? What do you think are the disadvantages or could be the disadvantages of using CA in healthcare?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part C: Use of Ada for disease diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Considering “Ada” would you use such a CA? If so, why? If no, why not? For this purpose, Ada was shown on a mobile devise and the Google Play Store description of “Ada” and its features were shown to study participants. Would you use such a dialogue system? Why or why not? What advantages are coming to your mind considering the use of Ada? Do you see any disadvantages that such a system has? If so, explain them. What could convince you to use a CA like Ada?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part D: Individual characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questions about age, education, etc.</td>
</tr>
</tbody>
</table>

Table 1. Interview guideline

7 References


Laumer et al. / Conversational Agents for Disease Diagnosis


Rammohan, R., Dhanabalsamy, N., Dimov, V., and Eidelman, F. J. 2017. “Smartphone Conversational Agents (Apple Siri, Google, Windows Cortana) and Questions about Allergy and Asthma Emergencies,” Journal of Allergy and Clinical Immunology (139:2), AB250.


