(Why) Do We Trust AI?: A Case of AI-based Health Chatbots

Ashish Viswanath Prakash
Indian Institute of Management Tiruchirappalli, Tamil Nadu, India
ashish@iimtrichy.ac.in

Saini Das
Indian Institute of Technology Kharagpur, India

Abstract

Automated chatbots powered by artificial intelligence (AI) can act as a ubiquitous point of contact, improving access to healthcare and empowering users to make effective decisions. However, despite the potential benefits, emerging literature suggests that apprehensions linked to the distinctive features of AI technology and the specific context of use (healthcare) could undermine consumer trust and hinder widespread adoption. Although the role of trust is considered pivotal to the acceptance of healthcare technologies, a dearth of research exists that focuses on the contextual factors that drive trust in such AI-based Chatbots for Self-Diagnosis (AICSD). Accordingly, a contextual model based on the trust-in-technology framework was developed to understand the determinants of consumers' trust in AICSD and its behavioral consequences. It was validated using a free simulation experiment study in India (N = 202). Perceived anthropomorphism, perceived information quality, perceived explainability, disposition to trust technology, and perceived service quality influence consumers' trust in AICSD. In turn, trust, privacy risk, health risk, and gender determine the intention to use. The research contributes by developing and validating a context-specific model for explaining trust in AICSD that could aid developers and marketers in enhancing consumers' trust in and adoption of AICSD.

Keywords: Artificial Intelligence, Health Chatbot, Trust in Technology, Explainability, Contextualization, Free Simulation Experiment.

1 Introduction

The popularity of autonomous chatbots/conversational agents (CAs) has grown substantially in recent years, owing to recent advancements in artificial intelligence (AI) technology and the growing digitalization of industries (Araujo, 2018; Chi et al., 2021). An AI-based health chatbot is an autonomous system that can converse intelligently about healthcare-related issues in either audio or textual format (Laumer et al., 2019; Prakash & Das, 2020). These chatbots are posited to address the most pressing problems of healthcare services, such as misinformation, information overload, and timely access to quality healthcare (Denecke et al., 2019; Wang & Siau, 2018). An exciting innovation in this area is the AI-based chatbots for self-diagnosis (AICSD) (Laumer et al., 2019). These user-facing intelligent applications engage with individuals in real-time, collect information about the condition from the user, make personalized diagnostic inferences from the data, and offer recommendations. Ada Health, Microsoft Healthbot, NHS chatbot, Babylon Health, etc., are popular AICSDs offering direct-to-consumer diagnosis/triage services (Siwicki, 2018). The global healthcare chatbot market size is growing at a CAGR of 21.56% and is expected to be US$ 967.7 million by 2027 (PR Newswire, 2021).

AICSDs are seen as credible alternatives to traditional search engines and can potentially improve access and alleviate the overburdened healthcare resources (Wang & Siau, 2018). However, the benefits promised come with significant risks to the safety and life of patients as there are doubts regarding the efficacy and accuracy of such chatbots (Bickmore et al., 2018; Laranjo et al., 2018). Further, handling personal health information raises a plethora of issues relating to data privacy, security, consent, and ownership (Powell, 2019). Moreover, regardless of how effective and reliable they are, most of these bots, by virtue of their inscrutable algorithms, provide little hints as to how they arrive at their decisions, raising concerns about transparency, accountability, and responsibility – which are fundamental to building trust in a system (Gille et al., 2020; Powell, 2019; Rai, 2019). Further, it is observed that the perceived humanness of these intelligent agents elicits paradoxical responses from the users (Ho & MacDorman, 2017). This may further exacerbate the consumers' trust in these applications. Thus, the adoption of this promising healthcare technology faces acute challenges owing to its unique characteristics and the sensitivity of the context (O'Connor et al. 2021).

Prior research has argued that trust in specific technology is central to understanding individual technology use behavior (Mcknight et al., 2011). Trust in AICSD is the belief that it has the essential qualities to provide competent, reliable, and truthful symptom-checking and triage services. Further, studies on AI-enabled systems in general (Bach et al., 2022; Hengstler et al., 2016) and those specific to AI-based health chatbots (Laranjo et al., 2018; Nundy et al., 2019) indicate that lack of trust in the system could be one of the major impediments to the adoption. Hence, it is valuable for practitioners to know the comprehensive set of interconnected factors that predict the trust in AICSD to debottleneck its adoption.

However, academic research into this specific area of patient/consumer trust in AI applications in healthcare (including chatbots) has been sparse (Asan & Choudhury, 2021; Laumer et al., 2019; O'Connor et al., 2021; Prakash & Das, 2020). Research on the specific issue of what constitutes consumers' trust in AICSD and the conditions for achieving it is yet to be explored (O'Connor et al., 2021; Prakash & Das, 2020). Moreover, AICSD, due to their unique characteristics, raises an array of new concerns, such as patient safety, lack of transparency, and trust manipulation (through deceptive use of anthropomorphic features) (Bach et al., 2022; Hengstler et al., 2016; Miner et al., 2019; Prakash & Das, 2020; Rai, 2019; Scorici et al., 2022). Consequently, existing trust models that work in the contexts of non-AI systems/generic information technologies (IT) may not be relevant to AI-based systems, given these unique characteristics (Lockey et al., 2021). Additionally, the context adds another layer of complexity as the AI chatbot usage in a non-health setting (e.g., retail, banking, etc.) cannot be compared to the unique healthcare context in which AICSD is used (O'Connor et al., 2021). Therefore, we must rethink and appropriately contextualize the existing trust models, which were developed for generic technologies and non-health settings, before using them to understand the complex nature of trust in AI-based technology in healthcare services (chatbot for self-diagnosis). Such contextualization may generate rich theoretical and practical insights (Johns, 2006; Hong et al., 2014).

We, therefore, aim to develop a contextualized research model to understand the determinants of consumers' trust in AICSD and its behavioral consequences. In line with this purpose, we propose the following research questions: (1) What factors determine consumers' trust in AICSD? (2) How does consumers' trust in AICSD impact the behavioral intentions to use it? To this end, a contextualized model was formulated by integrating relevant contextual factors into the trust in technology framework (Mcknight et al., 2011). The model was tested empirically using a free simulation experiment (Fromkin & Streufert, 1976) in India.

The findings of this study indicate that perceived anthropomorphism, perceived information quality, perceived explainability, disposition to trust technology, and perceived service quality predict trusting beliefs in AICSD and trusting beliefs, risk beliefs (related to privacy and health), in turn, determine intention to use. Results also reveal the interplay between trust and risk perceptions wherein trust in AICSD contributes to assuaging the risk perceptions, thus indirectly helping to improve the intention to use. The study contributes to the theory and practice by developing and validating a contextual research model that captures the complexity of the novel technology and the setting in which it is deployed (healthcare). The insights from the study will aid designers and developers in creating trustworthy and purpose-fit AICSD, which could lead to a greater return on investment for organizations.

2 Literature review

2.1 Prior research on AI-based health chatbots

Chatbots are software programs that can converse with humans via text or voice via interactive interfaces (McTear et al., 2016). Technology has improved considerably in the more than six decades since the invention of the first chatbot, ELIZA, from being able to recognize a few keywords to having real-time voice conversations with people. Much of this progress can be attributed to the development of advanced Machine Learning (ML) algorithms and computational linguistics (Shah et al., 2016). New-generation chatbots are being widely deployed across various industries to automate interaction with customers (Følstad & Brandtzæg, 2017).

In the domain of healthcare services, AI-based chatbots are poised to play a significant role in supporting/complementing human service providers. There is an increasing body of research evidence favoring the effectiveness of the use of chatbots for supporting behavioral changes, mental health therapy, support for the elderly, etc. (Laranjo et al., 2018). AI-powered chatbots that can check symptoms and triage patients (Wang & Siau, 2018) with their dynamic learning capability and Natural Language Processing (NLP), represent a huge leap from the earlier generation of symptom checkers, which simply used programmed decision trees to match symptoms with diseases (Razzaki et al., 2018; Wang & Siau, 2018). Despite the immense value, the literature on AICSD is scant (Laumer et al., 2019; O'Connor et al., 2021; Prakash & Das, 2020). Previous research on AICSD has focused on issues such as design and development (Minutolo et al., 2018), estimation of efficacy (Razzaki et al., 2018), and patient risk (Bickmore et al., 2018); research on consumer/patient perspectives on the use is scant. The table below (Table 1) summarizes the available literature on the use of AICSD.

Study	Objective	Theory	Methodology	Important findings
Razzaki et al. (2018)	To validate the accuracy and safety of AICSD.	NIL	Experimental study	AICSD was able to detect the medical condition presented by a clinical scenario with accuracy similar to human doctors.
Bickmore et al. (2018)	To ascertain the extent and nature of the harm that could be caused by using CAs for medical information.	NIL	Scenario-based experiment	Identified failure modes of CAs in the scenarios tested. Found that subjects were led to take actions that could have caused harm (12.4%) or death (6.9%).

Wang & Siau (2018)	To formulate a theory on trust in health chatbots.	Keeney's Value-focused thinking	Conceptual	Reviewed factors affecting trust-building in health chatbots. Proposed a study to formulate a theory on trust in health chatbots.
Laumer et al. (2019)	To develop a conceptual model that explains consumer adoption of AICSD.	UTAUT2	Qualitative – Semi-structured interviews (n=35)	Developed a research model that explains the adoption of CAs for disease diagnosis by extending the UTAUT2 model with factors such as privacy risk, trust in provider and system, compatibility, experience in e-diagnosis, and access to the health systems.
Nadar-zynski et al. (2019)	To explore the participant's willingness to engage with an AI-based health chatbot.	TFA	Mixed-Method: interviews (n=29), Survey (n=216), Binary regression.	Analysis of the interviews revealed user concerns about cyber-security, the accuracy of the bot, and the inability of AI-based chatbots to empathize. Perceived IT skills, dislike for talking to computers, perceived utility, attitude, and perceived trustworthiness were correlated with acceptability.
Prakash & Das (2020)	To explain consumer's trust in AICSD.	TTM, Risk Theory, SRT	Quantitative – Survey (n= 107), SEM	Social presence, usefulness, safety risk, and propensity to trust predicted trust in AICSD. The effect of ease of use, privacy risk, and third-party endorsement was insignificant. Trusting beliefs determined trusting intention.
Mesbah & Pumplun (2020)	To investigate the factors that influence seniors' adoption of health chatbots.	UTAUT2	Qualitative – Semi-structured interviews (n=21)	Proposed an extended UTAUT2 model with additional factors such as technology anxiety, privacy risk, trust, resistance to change, etc., to explain the acceptance of health chatbots by seniors.
Fan et al. (2021)	To understand the usage of health chatbots in the actual world, as well as the difficulties and limitations of their usage.	NIL	Case study – Data-driven approach in analyzing the system log 47,684 consultation sessions over six months	Users consulted on a wide variety of issues, especially those that entail privacy and social stigma. A significant percentage of users left their consultation in the middle. Identified user concerns such as insufficient actionable information and inaccurate diagnostic suggestions.
Seitz et al. (2021)	To understand the difference between trust and distrust toward diagnosis, chatbots	TTM	Qualitative – Interviews, brainstorm sessions, think-aloud, and participatory observation (n=8)	Findings reveal that distrust in chatbots arises affectively, whereas trust arises cognitively.

O'Connor et al. (2021)	To develop a theoretical framework that explains patients' trusting intentions towards robots in healthcare	TTM	Conceptual	Proposes a conceptual framework based on trust in technology model for investigating trusting intentions towards intelligent agents in healthcare.

Table 1. Prior research on the use of AICSD
Note. UTAUT2, Unified theory of acceptance and use of technology; TTM, Trust in technology model; SRT, Social response theory; TFA, Theoretical framework of acceptability of healthcare interventions; SEM, Structural equations modeling.

The current literature on consumer/patient perspectives majorly focuses on acceptance/ adoption (Laumer et al., 2019; Mesbah & Pumplun, 2020; Nadarzynski et al., 2019) and the actual use of AICSD (Fan et al., 2021). It is observed that the issue of trust in AICSD (O'Connor et al., 2021; Prakash & Das, 2020; Seitz et al., 2021; Wang & Siau, 2018) is an emerging area of interest. In terms of the methodologies used, most of these studies are conceptual or qualitative in nature (Laumer et al., 2019; Mesbah & Pumplun, 2020; O'Connor et al., 2021; Seitz et al., 2021; Wang & Siau, 2018) with the exceptions of Prakash and Das (2020) (quantitative), Nadarzynski et al. (2019) (mixed-methods) and Fan et al. (2021) (data-driven approach). With respect to the theoretical foundations of the studies, technology acceptance theories (UTAUT2 and TFA) and trust in technology model (TTM) are most frequently used.

We make a few important observations from our analysis of the available literature. First, despite the agreement that customer trust is critical for the successful adoption of AICSD (Laumer et al., 2019; Mesbah & Pumplun, 2020; Nadarzynski et al., 2019), research on this topic is not adequately represented in the literature. The first attempt in this direction was by Wang and Siau (2018), which proposed developing a trust framework but provided no empirical findings. The only empirical study that attempted to study the determinants of trust in AICSD was by Prakash and Das (2020). Although this study provides some preliminary insights into the consumer's trust in AICSD, it just re-applies well-known theories in a new context but fails to consider the unique context-specific AI-specific factors that might influence trust. Another recent study (Seitz et al., 2021), which explored the difference between trust and distrust towards diagnosis chatbots using a qualitative approach, found that distrust has affective roots, whereas trust arises cognitively. A more recent study on trust in healthcare robots (O'Connor et al. 2021) points out that healthcare possesses a unique context compared with its counterparts (e.g., manufacturing, retail) and argues for the need to explore contextual factors influencing patients' trusting intentions towards robots in healthcare.

Second, we argue that prior research on similar technologies like healthcare IT (Song & Zahedi, 2007; Xie et al., 2020) may not adequately explain the trust in and use of the novel AI-based chatbots for self-diagnosis. This is primarily due to its unique characteristics - autonomous nature, anthropomorphic features, self-learning ability, superhuman accuracy levels, inherent bias, and the non-transparent nature of the algorithms (Hengstler et al., 2016; O'Connor et al., 2021; Rai, 2019). Further, AI-based chatbots are an exception among software tools due to their position as highly responsive interaction partners. They exhibit specific social traits as compared to previous generations of information systems (IS) since they simulate human intelligence and are capable of independent decision-making (Glikson & Woolley, 2020). They are also anthropomorphized deliberately to make the interaction more natural (Feine et al., 2019).

Third, these unique attributes of AI-based systems raise a plethora of new concerns, such as patient safety (Prakash & Das, 2020; Miner et al., 2019), lack of transparency (Rai, 2019), and concerns of trust manipulation through deceptive use of anthropomorphic features (Bach et al., 2022; Scorici et al., 2022). The existing models of trust in technology (McKnight et al., 2011) or healthcare IT (Song & Zahedi, 2007; Xie et al., 2020) or recent ones on AI-based chatbots/personal assistant (Zierau et al., 2020) does not adequately capture the intricacies in the context of AI-based chatbot applied in a self-diagnosis context (O'Connor et al. 2021). Hence, we strongly believe that the existing models on trust are ineffective in capturing the phenomenon of trust in AICSD in its full essence. Therefore, we must rethink and contextualize existing theoretical models on trust in technology, which were initially developed for non-AI-based technologies, before we can use them to understand the trust formation in AICSD and its behavioral consequences. Such contextualization can lead to rich theoretical and practical insights (Johns, 2006; Hong et al., 2014).

2.2 Theoretical background

From a thorough review of the literature on trust, we believe the framework of trust in technology (TTM) (Mcknight et al. (2011) and Lankton et al. (2015) could be a useful model in explaining the complex phenomena of trust in AICSD. Trust generally represents "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other party will perform a particular action important to the trustor, irrespective of the ability to monitor or control the party" (Mayer et al., 1995, p.712). Prior research has established that trust is central to understanding human behavior in diverse domains ranging from workgroup dynamics to commercial relations (McKnight et al., 2011). Trust was studied originally in the context of interpersonal relationships (Mayer et al., 1995); later, it was extended to explain the human relationship with nonhuman entities (McKnight et al., 2011). Accordingly, trust (or, more accurately, trusting beliefs) in technology is defined as "the beliefs that a specific technology has the attributes necessary to perform as expected in a given situation in which negative consequences are possible" (McKnight et al., 2011, p.7).

The original model operationalizes the trust in technology constructs consisting of three sets of concepts: (a) disposition to trust general technology, (b) institution-based trust in technology, and (c) trust in a specific technology. Disposition to trust captures person-specific characteristics that can influence trust formation in a specific technology (Mcknight et al., 2011). Institution-based trust focuses on the belief that success is likely due to the presence of supportive situations and structures tied to a specific context or a class of trustees (e.g., structures and guarantees, such as third-party certifications/seals and privacy assurances). Unlike disposition to trust, institution-based trust beliefs are highly situation-specific. Furthermore, trust in technology shapes both users' attitudes and behavior toward the target technology (Lankton et al., 2015). It is an important psychological step that enables the users to rule out the probability of unintended consequences of using the technology and increase the intention towards using it (Gefen et al., 2003a).

Additionally, to adequately capture the specificity of the context, we did a thorough review of the literature on trust in AI-based technologies (Bach et al., 2022; Lockey, 2021) and closely related applications such as social robots (Naneva et al., 2020), intelligent conversational agents (Zierau et al., 2020) to understand the relevant factors that could impact the user's trust.

As per the social response theory (SRT), humans are predisposed to see computers as "social actors" unconsciously, even when they are aware that they lack feelings, motives, or intentions (Nass & Moon, 2000). This is particularly true when technology exhibits anthropomorphic (human-like) characteristics such as verbal skills or physical appearances (Troshani et al., 2020). It is widely agreed that anthropomorphism is one of the fundamental characteristics that separate AI systems from non-AI systems (Troshani et al., 2020). Anthropomorphism is defined as "the attribution of human-like characteristics, behaviors, and emotions to nonhuman agents" (Adam et al., 2020). Anthropomorphic design characteristics (social/emotional cues) of an agent trigger a sense of social presence ("awareness of the other person in the interaction" (Short et al., 1976, p. 65)) leading to a sense of emotional closeness and/or social connectedness with the agent, this ultimately increases trust (Adam et al., 2020; Qiu & Benbasat, 2009). Consequently, humans tend to apply social rules to anthropomorphically designed systems during an interaction (Nass & Moon, 2000). However, it is also observed that even though human likeness improves user affinity when it increases beyond a certain point, it can trigger a sense of eeriness (Ho & MacDorman, 2017). Moreover, anthropomorphic design cues can also create undue trust in the agents and, hence, could be used to manipulate the users into trusting untrustworthy systems (Culley & Madhavan, 2013). Hence, consumers' perception of anthropomorphism in the AICSD could play a substantial role in the formation of trust.

Similarly, another prominent concern related to the use of AI-based technologies in aiding decision-making is the non-transparent black-box nature of the algorithms (Rai, 2020). According to the literature, the "black box" problem is one of the key barriers to promoting users' trust and acceptance of AI systems. Decision-makers have difficulties comprehending how AI systems generate specific results (Adadi & Berrada, 2018), especially in healthcare settings (London, 2019). Building trust is hence considered necessary to deal with complexity and ambiguity since people cannot fully comprehend the inner workings of non-transparent AI systems (Rai, 2019). Transparency into the inner processes of AI systems and the explainability of why a certain outcome was generated is essential to building trust (Rai, 2019). Hence, the perceived explainability of the system is a critical factor to be considered while modeling trust in AICSD (Rai, 2019).

Further, as a recommender of health information, the perceived quality of both information, the system, and the overall service provided could influence trust formation. The IS success model by DeLone and McLean (2003) linked these quality factors to system use, user satisfaction, and, ultimately, to IS success. Empirical studies that followed revealed that these quality factors were the fundamental drivers of system use, user satisfaction, and trust formation process in various systems, specifically the e-health systems. E.g. Song and Zahedi (2007) argued that for health infomediaries to become customers' favorite information source, it is critical to foster trust in the quality of information they provide. The infomediary's interactivity and ease of use (key dimensions of system quality) were crucial to fostering loyalty and trust (Song & Zahedi, 2007). Moreover, a recent qualitative study investigating the requirements for trust for AI systems identified system quality and service quality as important prerequisites for building trust (Bedué & Fritzsche, 2022). Accordingly, we believe the perceived IS quality (DeLone & McLean, 2003) could drive trusting beliefs in the AICSD.

Further, trust implies the presence of uncertainties/risk factors in the usage context (McKnight et al., 2011). The users typically make a leap of faith in spite of the presence of these uncertainties due to their confidence in the attributes of the specific technology (Gefen et al., 2003b). From the literature, perceived risk generally denotes the probability of a loss as well as the perception or belief of unfavorable outcomes (Cunningham, 1967). It is usually measured as a consolidation of risk beliefs (e.g., financial, performance, physical, psychological, social, etc.) (Pavlou, 2003). Perceived risk is critical in establishing interpersonal, social, and economic relationships (Chiles & McMackin, 1996; Song & Zahedi, 2007). According to existing research, a negative relationship often exists between perceived risk and "intention to transact" (Pelaez et al., 2019). In the context of AICSD, two specific risk beliefs are pertinent: perceived privacy risk and health risk. In this case, the privacy risk perception stems from the possibility of losing privacy while using the bot for diagnosis, which typically requires disclosing private personal information (Laumer et al., 2019). Similarly, health risk refers to the user's perceptions of negative health consequences if the information provided by the chatbot is erroneous. According to the risk calculus perspective, users engage in a mental calculus of risk against benefits while attempting to depend on (trust) or use new technology.

From the literature, it becomes clear that in order to explain the intricacies of the phenomenon under consideration, i.e., trust in AICSD, the contextual factors that are linked to the phenomenon have to be weaved into the fabric of the generic theory of trust in technology which is ineffective in its own to capture the phenomenon in its entirety (Prakash & Das, 2020). Given that studying the phenomenon of trust in AICSD and its consequences continues to be underexplored in the literature (Prakash & Das, 2020; O'Connor et al., 2021), and without identifying the comprehensive set of interconnected factors that drive trust in such technologies, the adoption by users is likely to lag behind the rapid pace of technological advancement. Hence, we took a contextualized approach to modeling trust in AICSD.

2.3 Contextualization for trust in AICSD

Context is defined as "situational opportunities and constraints that affect the occurrence and meaning of organizational behavior as well as functional relationship between the variables" (Johns, 2006, p. 386). According to Whetton (2009), context is a set of factors around a phenomenon that exerts a direct or indirect impact. In information systems, the research context refers to the characteristics and usage context of the technology artifact (Orlikowski & Iacono, 2001). Prior research has observed that although embracing contextualization may necessitate researchers to sacrifice parsimony and generalizability (Hong et al., 2014), its impact on research outcomes can be powerful (Johns, 2006). It is observed that contextualization helps develop a comprehensive understanding of the interaction between technology, people, and context, without which the findings of the study might be incomplete (Johns, 2006). Furthermore, contextualization can make the research appealing to practitioners as the results are more relevant in solving practical problems (Breward et al., 2017; Johns, 2006).

Hong et al. (2014) provide some useful guidelines for context-specific theorizing in IS research. The paper outlines two major approaches for incorporating context into theory development. The first approach (single-context theory contextualization) typically starts with identifying some well-established general theories that are relevant to the domain of interest. Then, it incorporates contextual factors as antecedents of the core constructs, moderators of the relationship, or decomposing the core constructs into contextual factors (Hong et al., 2014). The identification of the context-specific factors can be based on past research on relevant technologies and/or an in-depth analysis of the technology under investigation using qualitative methods (Hong et al., 2014). The second approach (cross-context theory replication) aims to replicate theoretical models in different contexts and consolidate findings into a context-contingent theory using meta-analysis. In this study, we use the first approach (Hong et al., 2014), where the well-established trust in technology theory is taken as the general framework and refined by incorporating relevant contextual factors as antecedents to the core theory constructs.

Despite the growing interest in the use of AI-based tools for self-diagnosis, our literature review reveals that empirical studies concerning consumer trust in AICSD are limited both in terms of number and its examination of the antecedents that may influence the trust in AICSD (O'Connor et al., 2021). As discussed in section 2.1, a study by Prakash and Das (2020) is one of the first attempts to explain the drivers of trust in AICSD. The study followed Seitz et al. (2021), who explored the difference between trust and distrust towards AICSD without giving a comprehensive insight into the determinants. A more recent conceptual study (O'Connor et al. 2021) argued for a more contextualized model that incorporates contextual factors (such as individual characteristics/personality traits, explanation competency, health-related beliefs, etc.) to explain the complex phenomenon of trust in robots in healthcare. While these articles offer some insights into the intricate understanding of consumer trust in AICSD, they do not explicitly integrate the relevant contextual factors that could impact trust in AICSD.

To address this gap, we followed guidelines proposed by Hong et al. (2014) and chose the TTM (McKnight et al., 2011) as the initial generic theoretical framework for explaining the trust in AICSD and its behavioral consequences. Subsequently, we combed through the relevant studies in the field to identify the relevant contextual factors as the antecedents of the contextualized core constructs. The approach of identifying contextual factors by examining extant literature is supported by Hong et al. (2014). We identified four sets of factors, namely perceived IS quality, perceived anthropomorphism, risk beliefs, and individual characteristics, as the key contextual factors that could influence the core constructs in the general model through a thorough literature review on AICSD and related technologies. We followed the guidelines by Hong et al. (2014) to decompose the higher-level risk beliefs into two specific risk perceptions, namely, health risk and privacy risk, which are relevant to AICSD usage. Further, following the guidelines of Hong et al. (2014), we also examine the interdependence between the salient characteristics of technology (e.g., trust signs), characteristics of users (age, gender, prior experience), users' trusting beliefs, risk perceptions stemming from the sensitivity of usage context and behavioral outcomes (intention to use). The proposed theoretical model is explained in the following section.

3 Hypotheses development

The proposed theoretical framework (Figure 1) illustrates the hypothesized relationships between the constructs. As indicated earlier, trusting beliefs in a specific technology (AICSD) is the main dependent variable in the model, which refers to the user's perception that the specific technology (AICSD) possesses the necessary attributes to function as planned in a setting where adverse consequences are likely (McKnight et al., 2011). Intention to use AICSD was included as a behavioral outcome of trusting beliefs formed after the initial use. Intention to use also captures the essence of the concept of trusting intentions, defined as the intention to engage in trust-related behaviors (McKnight et al., 2002; Li et al., 2006, 2008). In the subsequent sections, the proposed hypotheses are rationalized with appropriate theory and supporting evidence from the related literature.

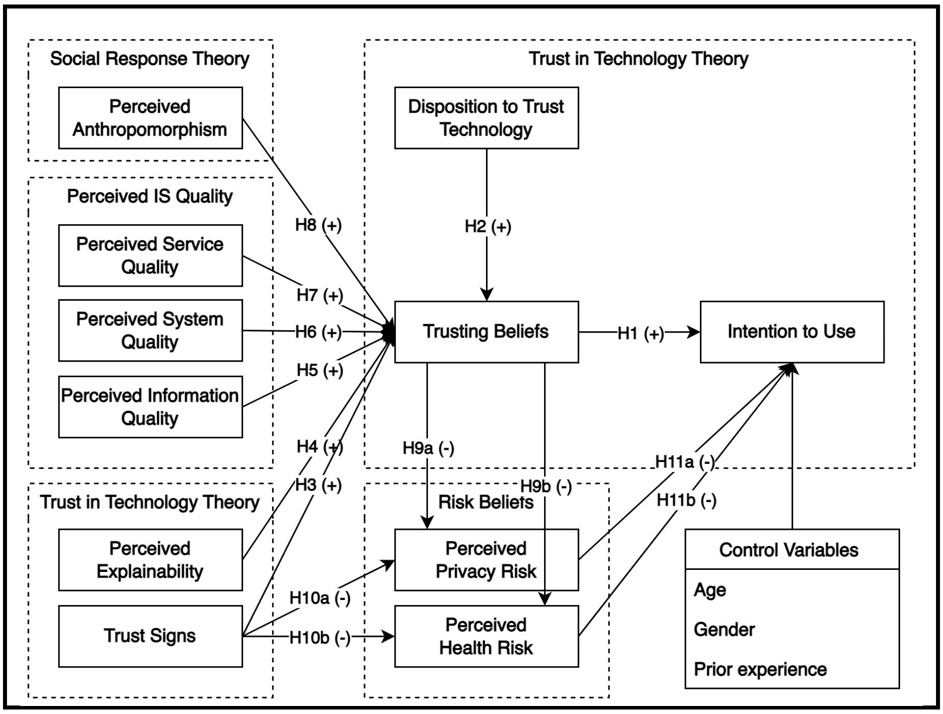

Figure 1. Theoretical framework

3.1 Trusting beliefs and behavioral intentions to use

Trusting beliefs (or trust) in technology refers to the user's perception that the specific technology possesses the necessary attributes to function as planned in a setting where adverse consequences are likely (McKnight et al., 2011). Regardless of whether an objective technology characteristic exists, users' views about performance could vary based on their experience or the context in which it is used (McKnight et al., 2011). Trusting beliefs are typically assessed using human-like trust beliefs or systems like trust beliefs depending on the degree of technology humanness (Lankton et al., 2015). An AI-based chatbot qualifies as a human-like technology due to its anthropomorphic features. Hence, human-like trusting beliefs could be used to measure trusting beliefs (Lankton et al., 2015). Competence is the notion that a person/entity possesses the abilities, competencies, and traits necessary to have an influence in the domain of interest. Benevolence is the presumption that a person/entity will want to do good for the trustor for reasons other than self-centered benefits. Finally, the belief that a person follows an appropriate set of standards is referred to as integrity (Lankton et al., 2015; McKnight et al., 2002). The higher the perceptions along the competence dimension about the technological artifact, the higher will be the perception that it can perform as intended. Similar higher benevolence will reduce the likelihood of opportunistic behavior from the part of the technology artifact (McKnight et al., 2002). Finally, higher integrity implies that the AICSD sticks to the ethical guidelines. Together, these would create confidence in using the system and hence promote usage intentions. We also know from the literature that trusting beliefs in technology drive behavioral intentions (Lankton et al., 2015; Li et al., 2008; McKnight et al., 2002, 2011). People who believe that the technology has the characteristics that make it trustworthy are more inclined to trust it and show intentions to rely on or use it (McKnight et al., 2011; Lankton et al., 2015). Hence, we suggest the following hypothesis,

H1: *Trust (trusting beliefs) will positively influence the intention to use AICSD.*

3.2 Disposition to trust technology and trust

Disposition to trust refers to a tendency to trust other people (Meyer et al., 1995; Rotter, 1971). When applied to the technology context, disposition to trust is the extent to which the person displays a consistent tendency to be willing to depend on technology in general across a broad spectrum of situations and technologies (McKnight et al., 2011). According to the literature on trust, it is neither technology-specific nor situation-specific, unlike trusting beliefs and intentions, which are object-specific but cross-situational (McKnight & Chervany, 2001; McKnight et al., 2011). This construct primarily derives from trait psychology, and it states that actions/behaviors are molded by specific childhood-derived characteristics that become more or less stable through time. (McKnight & Chervany 2001). People may grow up with a disposition or develop because of their life experiences (McKnight & Chervany 2001). Dispositional trust will affect trusting beliefs (in specific technology) when the situation and technology are unfamiliar. Thus, if the individual has a higher disposition to trust technology (DT) in general, in the case of novel technology in an unfamiliar situation, he or she is likely to rely on this dispositional propensity. As a result, he/she is likely to form favorable opinions and beliefs about a novel technology like AICSD. Moreover, DT has been identified as a positive driver of trust (trusting beliefs) in various technology use settings (Chi et al., 2021; McKnight et al., 2011; Lankton et al., 2015). Accordingly, we propose,

H2: *User's disposition to trust technology is positively related to their trust (trusting beliefs) in AICSD.*

3.3 Trust signs and trust

Trust signs/cues are generally referred to as characteristics of an online service that users evaluate when determining the trustworthiness of the online service (Hoffman et al., 2014; Wang & Benbasat, 2008). In the case of health infomediaries, these informational cues include self-regulating policies such as privacy and security declarations, as well as third-party seals that ensure the provider adheres to fundamental ethical standards while providing information on its website (Song & Zahedi, 2007). It could also include users' feedback/reviews about the services offered. Hence, we define trust signs in this context as "users' perceptions about privacy assurance, third-party seals/certificates, and user reviews/ratings on the AICSD." The effectiveness of trust cues/signs is based on "signaling theory" (Bacharach & Gambetta, 2001). The trust cue/signs activate previously engraved cognitive and affective associations in the mind, allowing for rapid information processing and decision-making (Hoffmann et al., 2014; Wells et al., 2011). There are several reasons why these trust signs could bolster trust. For example, endorsements from previous users showcase the previous use of the service and signal a social norm regarding the general acceptability to the users (Pennington et al., 2003; Rice, 2012). Similarly, third-party seals/certificates could signal adherence to the ethical standards and best practices, which the user interprets as a decreased likelihood of opportunistic behavior on the part of the provider (Hoffmann et al., 2014). Thus, trust signs are elements designed to convince consumers that there is no risk in using the system and reaffirm their perceptions about the integrity of the system (Pennington et al., 2003; Song & Zahedi, 2007). These cues could strengthen consumers' trusting beliefs in the bot's integrity, benevolence, and competence (Hoffman et al., 2014; Song & Zahedi, 2007). It has been demonstrated earlier that trust signs/cues directly impact trusting beliefs in the case of the website used for online services (Hoffmann et al., 2014; McKnight & Chervany, 2001). Hence, based on these arguments, we propose,

H3: *Trust signs will positively influence trust (trusting beliefs) in AICSD.*

3.4 Perceived explainability and trust

The deep learning approaches used by AI systems are typically inscrutable (Rai, 2020). Thus, the users of AI systems typically have little understanding of why a particular decision is being recommended by the system (Zierau et al., 2020). This inscrutability may exacerbate the issue of trust, particularly in settings where the implications are significant (Rai, 2020). Machines are beneficial to the extent their actions can be expected to achieve the user's objectives (Russell, 2019). Suppose no information is provided on why a particular recommendation is being suggested; then the user at that point has no logical reason to believe that the course of action suggested by the chatbot could be relevant and beneficial for him/her (the consequences may not be immediately apparent). The lack of visibility and uncertainty thereof is likely to undermine the user's beliefs about the competence of the AICSD. The inscrutability and the resultant ambiguity about the functionality may also lead to the system being rejected by the users (Rai, 2020; Zierau et al., 2020).

Literature on trust in AI systems argues that over and above the information about the accuracy and performance of the AI system, providing an explanation about the system's behavior is likely to make the recommendation more valuable/actionable and can thus boost the user's trust in the AI system (Rai, 2020). However, although users will be interested in knowing "why" a particular diagnosis is recommended, they are not likely to be interested in knowing "how" the complex AI algorithms arrived at the recommendation (Rai, 2020). High degrees of transparency about how AI models work could be counterproductive as it can entail high attention costs, create information overload, and annoy end-users (Rai, 2019). Thus, a simple "why" explanation instead of a complex "how" is likely to suffice (Rai, 2020). Moreover, the explanation-driven trust stream of research in chatbots argues for the significance of explanation in the formation of trust in AI-powered chatbots (Zierau et al., 2020). Additionally, recent research on personalized AI recommendation systems has identified and validated the role of explainability in determining users' trust (trusting beliefs) in the system (Shin, 2020). Hence, we propose,

H4: *Perceived explainability will positively influence the users' trust (trusting beliefs) in AICSD.*

3.5 Perceived information quality and trust

The most significant aspect for users of health information is the quality of information (Song & Zahedi, 2007). Perceived information quality in this context can be defined as the user's cognitive beliefs about the usefulness, adequacy/completeness, accuracy, relevance, currency, and understandability of the information provided by the AICSD. It has already been established as a crucial factor in the trust-building process in online interactions (Nicolaou & McKnight, 2006; Song & Zahedi, 2007). Since perceived information quality involves the user's perception of positive information attributes such as accuracy and completeness, it can impact the trustee's trusting belief–integrity (Nicolaou & McKnight, 2006). The users may interpret the accuracy and completeness of information as a sign that the health chatbot is committed to providing comprehensive and accurate information which enhances the user's beliefs about the integrity of the chatbot in question (Nicolaou & McKnight, 2006; Song & Zahedi, 2007).

Similarly, accurate, understandable, and reliable information could imply that the source of information is competent (Nicolaou & McKnight, 2006; Song & Zahedi, 2007). As a result, perceived information quality would have a positive relationship with the trusting belief–

competence. Finally, higher perceived information quality also implies that the information is relevant, useful, and reliable to the end-user; it suggests that the chatbot is concerned enough to offer the user helpful information (Song & Zahedi, 2007). Therefore, perceived information quality should also influence the benevolence dimension of trusting beliefs. Thus, all three dimensions of trust may be positively influenced by the perceived information quality of the AICSD. Based on these arguments, we suggest,

H5: *Perceived information quality is positively related to their trust (trusting beliefs) in AICSD.*

3.6 Perceived system quality and trust

Perceived system quality represents the features of an online system that is preferred in terms of usability, availability, responsiveness, and interactivity (DeLone & McLean, 2003). It reflects the user's perceptions of the system's technical ability to provide easy and quick access to information (DeLone & McLean, 2003). Usability or ease of use is a key construct that impacts consumer trust in the context of e-commerce (Gefen et al., 2003a). Suppose the users perceive difficulty in accessing the needed information. They are likely to conclude that the health chatbot and its designers lack the competence to create a user-friendly design (Song & Zahedi, 2007). Similarly, a lack of responsiveness and interactivity could cause the users to conclude that the health chatbot lacks the benevolence to address their queries (Song & Zahedi, 2007). Similarly, the feature that a health chatbot is always up and available signals the integrity and reliability of the system. In other words, the component dimensions of perceived system quality directly and positively influence the dimensions of trusting beliefs (i.e., benevolence, integrity/reliability, and competence). Hence, the following hypothesis is suggested,

H6: *Perceived system quality will positively influence their trust (trusting beliefs) in AICSD.*

3.7 Perceived service quality and trust

Perceived service quality is defined as the consumer's perceptions of "an entity's overall excellence or superiority" (Dagger et al., 2007, p. 124). In the current study context, we define perceived service quality as the user's perception of the excellence of services provided by the AICSD. Customers/users typically use five aspects of service, "tangibles," "reliability," "responsiveness," "assurance," and "empathy," while evaluating the service quality (Jiang et al., 2002; Parasuraman et al., 1988). Regarding its outcomes, service quality perceptions influence attitude and behavioral intention (Cronin et al., 2000). In this case, if the users perceive the chatbot to be highly reliable, they are likely to interpret that the chatbot is high on the integrity aspect, i.e., they consistently deliver what was promised. Similarly, assurance (the feeling of safety in transactions) adds to reliability, reinforcing the integrity dimension of trusting beliefs. Empathy and care shown by the chatbot, along with the promptness of the response provided (responsiveness), could signal that the chatbot is kind, thoughtful, and considerate of the requests made by the user. The agent's courteous, caring, and responsive behavior will inspire confidence in customers, especially in a high-involvement professional service context (Westinger, 1998). Finally, having a visually attractive user interface could signal significant investments on the provider's part, which could positively impact the user's perception of the entity's functionality (Eisingerich & Bell, 2008). Thus, perceived service quality dimensions could positively influence the trusting beliefs in the AICSD.

Additionally, prior research has discovered that in the context of professional, high-credibility services (e.g., e-healthcare), consumers' uncertainty may be reduced by providing consistent quality services, which will eventually lead to trustworthy relationships between consumers and providers (Akter et al., 2013; Eisingerich & Bell, 2008). Overall, customers will be more trusting of an entity that routinely meets or exceeds their technical or core performance expectations (Eisingerich & Bell, 2008). Furthermore, the link between perceived service quality and consumer trust has been confirmed by subsequent research (Akter et al., 2013; Chiou & Droge, 2006; Eisingerich & Bell, 2008). Based on these arguments, we propose,

H7: *Perceived service quality is positively related to their trust (trusting beliefs) in AICSD.*

3.8 Perceived anthropomorphism and trust

Anthropomorphism is an inductive inference mechanism whereby individuals attribute distinct human characteristics to nonhuman entities, especially the ability for critical thinking (agency) and conscious feeling (Epley et al., 2007). Prior research has argued that the more technology appears to have human-like cognitive abilities, the more individuals will trust it to fulfill its intended purpose competently (Epley et al., 2006; Waytz et al., 2014). This prediction is based on the commonly understood connection between an individual’s perceptions of the mental states of others and competent action performed by them (Waytz et al., 2014). When people see an agent perform a competent action, they tend to believe that there is an intelligent mind inside that is performing the action with awareness and foresight. Attributing a human-like mind to a nonhuman entity would make the agent appear more capable of managing its own actions and, thus, more able to execute its planned functions adeptly (Waytz et al., 2014). Thus, as the agent becomes more similar to a human, users are more prone to engage in the correspondence bias, attributing human motivations, reasoning powers, and capacities to this non-human machine (Culley & Madhavan, 2013). Therefore, as the agent (AICSD) displays human-like cognitive abilities, users start attributing a human-like intelligent mind to it (perceived anthropomorphism) and consequently would feel more confident in the agent’s abilities (Moussawi et al., 2020) to perform the medical diagnosis effectively (trust). Thus, high perceived anthropomorphism would lead to high trust in AICSD. Additionally, recent research in autonomous cars (Waytz et al., 2014) and intelligent personal agents (Hu et al., 2021; Moussawi et al., 2020) have demonstrated a positive link between perceived anthropomorphism and users’ trust in technology. On this premise, we suggest the following hypothesis,

H8: *Perceived anthropomorphism will positively influence trust (trusting beliefs) in AICSD.*

3.9 Trusting beliefs and risk beliefs

In general, perceived risk refers to the likelihood of a loss as well as the perception or belief of unfavorable consequences (Cunningham, 1967). In the specific context of AICSD use, the two specific risk beliefs that can emerge are related to loss of privacy (perceived privacy risk) and related to the health and safety of the person who follows the chatbot recommendations. The research agrees that whether it is trust in people or trust in technology, both involve risk (McKnight et al., 2011). When a person trusts a cloud-based service, such as Dropbox, to safely store data, he or she exposes themselves to the risk and uncertainty associated with transferring data over the Internet and storing sensitive data on a server. Similarly, when an individual trusts an AICSD and decides to use it for self-diagnosis, he or she exposes themselves to the risk of their sensitive private data on their health and medical condition getting stolen and/or misused for target advertising (Laumer et al., 2019; Prakash & Das, 2020). Additionally, it exposes them to a risk of potential misdiagnosis (Wang & Siau, 2018) and the adverse health consequence resulting from it. In this way, risk, uncertainty, and lack of total user control are the inherent contextual conditions present in all trusting scenarios (McKnight et al., 2011). So when an individual trusts another individual or technology artifact, it is because they have confidence in the trustee to perform as intended under the conditions of risk (McKnight et al., 2011). That is, despite the risk, the person takes a leap of faith based on their assessment of the attributes of the technology (Gefen et al., 2003b; Holmes, 1991; Nicolaou & McKnight, 2006).

While research remains divided on whether trusting beliefs predict perceived risk or vice versa (Koller 1988, Pavlou & Gefen 2004), the majority of evidence shows that trust impacts risk perceptions (Gefen et al., 2003b; McKnight et al., 2017; Nicolaou & McKnight, 2006). Placing trust as a predictor of perceived risk is consistent with psychological descriptions of how trusting—as a leap of faith—provides a sense of security even when the outcomes are uncertain (Gefen et al., 2003b; Holmes, 1991; Nicolaou & McKnight, 2006). Accordingly, we propose the following hypotheses,

H9a: *Trust (trusting beliefs) in AICSD will negatively influence perceived privacy risk.*

H9b: *Trust (trusting beliefs) in AICSD will negatively influence perceived health risk.*

3.10 Impact of trust signs on risk beliefs

The trust signs include self-regulating policies such as statements about privacy security, third-party seals that guarantee that the health infomediary follows basic ethical codes (in recommending information), and reviews that reflect users' feedback. Tan and Theon (2003) noted that trustworthy cues are control mechanisms for reassuring online users. Further, according to Actor-Network Theory, trust signs are the inscription of the need to increase users' trust (Song & Zahedi, 2007). Trust signs are part of the input for assessing the risk connected with the trust context as well as the integrity of the health infomediary in the chain of trust formation (Song & Zahedi, 2007). This point of view is strengthened by "signaling theory," which divides trust-related features into observable and non-observable categories (Bacharach & Gambetta, 2001). Information quality could be viewed as an observable property in the context of a health information provider. In contrast, trust signs are non-observable properties that are meant to reassure web users that there is no risk involved in using the health infomediary and reinforce their beliefs in its integrity (Pennington et al., 2003).

Thus, trust signs are subtle measures intended to alleviate concerns associated with obtaining and using information given by AICSD. Trust cues signal to users that AICSD follows certain ethical guidelines in recommending information and adheres to values and standards of conduct. Therefore, users are likely to believe that using AICSD is less likely to cause harm to them. An earlier study conducted in the context of health infomediaries has confirmed the role of trust cues in toning down the risk perceptions of the user (Song & Zahedi, 2007). Accordingly, we believe user perceptions about trust signs will reduce privacy and health risk perceptions associated with the use of AICSD. Hence, we propose the following hypotheses,

H10a: *Trust signs will negatively influence perceived privacy risk.*

H10b: *Trust signs will negatively influence perceived health risk.*

3.11 Risk beliefs and behavioral intentions to use

3.11.1 Perceived privacy risk and behavioral intentions to use

Dinev and Hart (2006) defined perceived privacy risk as "the degree to which people believe that there is a potential loss involved with the disclosure of personal information." In this case, it refers to the perceived risk associated with disclosing sensitive private information on their health and medical condition while using the services provided by the chatbot. The uncertainty lies in the possibility of private information being sold or used for targeted advertising. Given the risk, consumers will weigh risk against benefits while considering using the services, and this "privacy calculus" will shape their attitude and behavior towards using the service (Dinev & Hart, 2006). So, if the consumers perceived that the benefits derived from using the diagnostic services of AICSD are outweighed by the risk associated with the loss of valuable personal information, they would refrain from using it. Furthermore, in various similar settings, empirical evidence from prior research has revealed a negative association between perceived privacy risk and propensity to use technology. For example, wearables (Adebesin & Mwalugha, 2020; Li et al., 2016), nutrition recommender systems (Berezowska et al., 2015; Wendel et al., 2013), health apps (Zhang et al., 2019; Zhao et al., 2018). Based on the evidence mentioned above, we suggest the following hypothesis,

H11a: *Perceived privacy risk will negatively influence intention to use AICSD.*

3.11.2 Perceived health risk and behavioral intentions to use

Relying on the Internet for health advice and self-diagnosis poses severe risks to the lives of patients (Robertson et al., 2014). Using fully autonomous chatbots for health advice can potentially aggravate these risks (Prakash & Das, 2020). Following the advice of chatbots which has unpredictable performance and that too without any further review by health professionals, poses a severe health risk to the patient (Wang & Siau, 2018). Many risk scenarios are plausible; for example, a misdiagnosis by the bot could result in 1) self-medication by patients that could cause adverse health consequences or 2) unnecessary anxiety among the patients about a disease that they do not have (Bickmore et al., 2018). There are also chances that patients may delay seeking professional care following the advice of the chatbot, which could have dire health consequences (Bickmore et al., 2018). According to the risk calculus perspective, consumers are highly likely to weigh these risks against benefits while making technology use decisions (Cocosila et al., 2007; Cocosila & Turel, 2016). If the perceived risk of adverse health consequences resulting from the use is high and exceeds the benefits of using AICSD, users' willingness to depend on the AICSD will be low, and hence, they may refrain from using it. Thus, we posit the following hypothesis,

H11b: *Perceived health risk will negatively influence intention to use AICSD.*

4 Research methodology

4.1 Study design, procedure, and participants

We used a "free simulation experiment" (FSE) (Fromkin & Streufert, 1976; Söllner et al., 2016) to test our theoretical model. Unlike the traditional lab experiment, which relies on treatments to change one or more predictor variables, FSE exposes subjects to a range of real-world conditions, such as performing a specific set of activities within a predetermined time frame (Fromkin & Streufert, 1976). The technique ensures that the subject not only completes the assigned activities but also organically explores the system to generate relevant impressions before answering the associated questions (Söllner et al., 2016). Furthermore, unlike in a field environment, this form of experiment allows us to control for various aspects, such as ruling out effects produced by different mobile devices or familiarity with an existing system (which would be difficult for assessing initial trust).

A significant number of participants (see Appendix 1) were regular students enrolled in MBA courses at two business schools in India. Apart from students, a small group of non-student participants (n = 25) were selected based on convenience sampling. Regarding demographics, the participants are tech-savvy and likely to seek health information online. A large share of the participants (85.64 percent) declared that they have previously accessed the Internet to obtain health information. It was also confirmed that none of them had previously used an AICSD (although 78.71% had previously used automated AI-based chatbots for other uses, such as e-commerce).

The experiment was conducted in different batches via online meetings, and all the participants had an internet-connected computer. We selected a popular web-based AICSD, namely, Your.MD (recently rebranded as Healthily) (Ćirković, 2020; Your.MD, 2021) for the experiment (see Figure 2). London-based Your.MD (Healthily) is a chatbot that provides customers with personalized and trustworthy healthcare assistance and information (Your.MD, 2021). It is a CE-accredited medical device that is registered with the "Medicines and Healthcare Products Regulatory Agency" (MHRA) as a Class I medical device (Your.MD, 2021). It is available on multiple platforms, including the website and smartphone apps for Android and iOS. It currently has a user rating of 4.1/5 from 14,055 users on the Google Play store and a rating of 4.4/5 from 907 users on the App Store. Its self-assessment feature allows users to ask free text questions using its chatbot. It then asks a series of questions to compile an "assessment report," which contains a list of likely conditions based on reported symptoms (with details such as reasons for suggesting the particular condition, symptoms, treatment methods etc.), the seriousness of likely conditions (e.g., red – emergency condition, orange – urgent condition, blue – routine condition, and green – no need to see a health professional) and recommendation on next steps to be taken (Ćirković, 2020; Your.MD, 2021).

The participants received a 10-minute briefing on the health chatbot about its functionality, features, and usage. Subsequently, they were assigned two predetermined activities that addressed the core functionality of the chatbot under evaluation.

Activity 1: Participants had to log on to the chatbot website and spend 10 minutes exploring the website and reading reviews on the Play Store.

Activity 2: Participants were required to look for symptoms of (1) an illness they experienced recently and (2) a recent illness a friend/relative experienced (15 mins).

After this, participants were asked to complete an online questionnaire based on their experience. Informed consent was obtained from all the participants, and participation in the experiment was entirely voluntary. In all, 254 people took part in the experiment, and 202 valid responses were received. The demographic profile of the participants is given in Appendix 1.

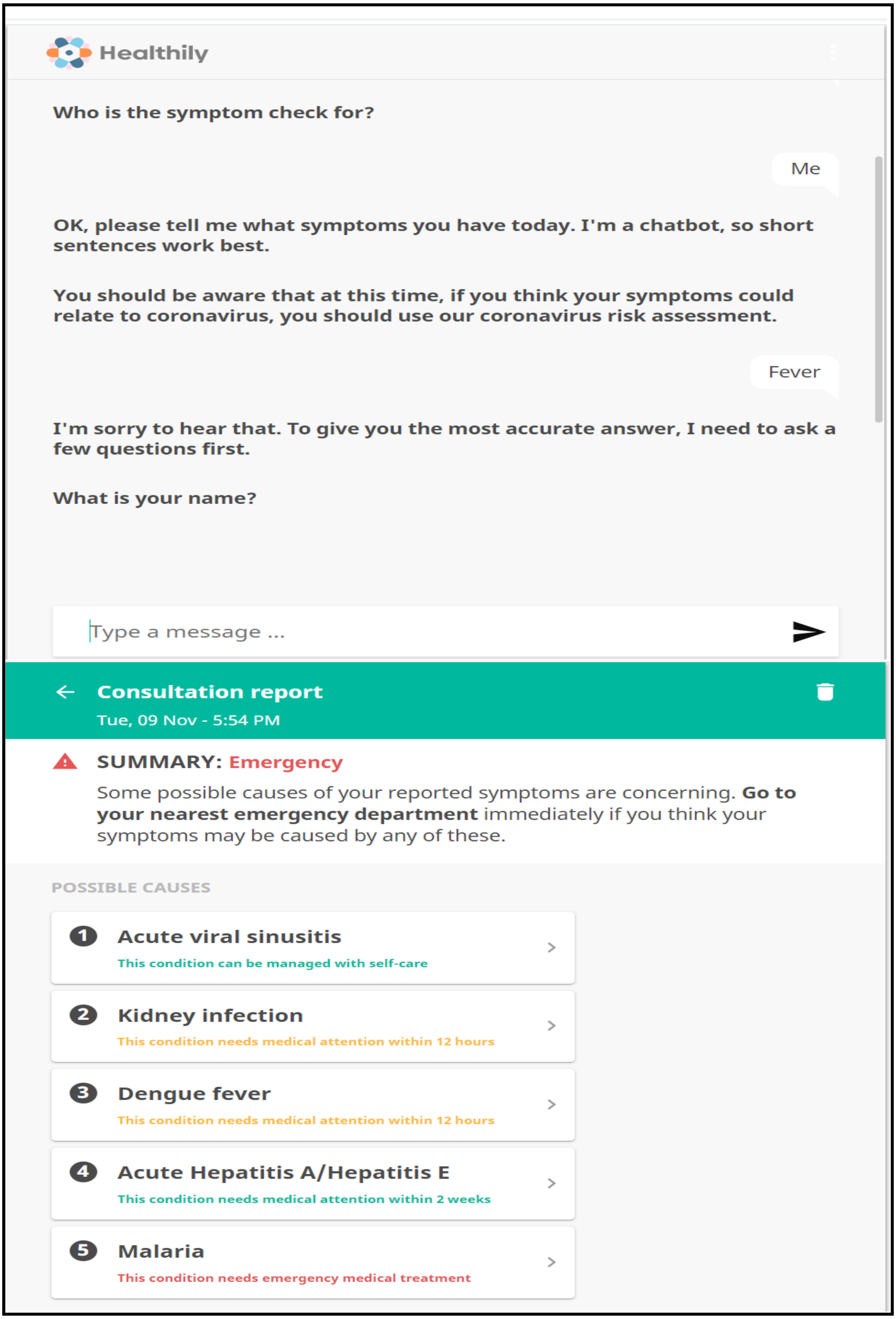

Figure 2. The user interface of the Your.MD chatbot (Source: Your.MD (2021)

4.2 Survey instrument

We collected the data using a structured questionnaire. All the items/indicators in the questionnaire were adapted versions of preexisting scales. Except for the constructs perceived anthropomorphism and perceived health risk, all indicators were measured on a seven-point Likert scale (1 = "strongly disagree" to 7 = "strongly agree"). For measuring the construct perceived anthropomorphism, we used a popular five-point semantic differential scale (Bartneck et al., 2009; Sheehan et al., 2020). A seven-point Likert scale ranging from 1 = "extremely low" to 7 = "extremely high" was used to measure perceived health risk. The measures and their sources are given in Appendix 2. In addition to the main constructs, we used certain control variables to control for the possible variation in the outcome variable due to variations in the respondent's characteristics. The control variables are age, gender, and prior experience with automated chatbots (general) (see Appendix 2).

5 Results

5.1 Structural equations modeling analysis

We employed the partial least square structural equation modeling (PLS-SEM) method to test the propositions. PLS-SEM is a "prediction-oriented" method, making it an appropriate choice for the current investigation (Hair et al., 2016). Furthermore, it performs effectively with small data sets and even in the case of non-normal data (Hair et al., 2016). Hence, this method was found to be appropriate for this study.

5.1.1 Measurement model assessment

Following the rules set by Hair et al. (2019), the validity and reliability of the empirical model were checked. First, the indicator loadings and composite reliability (CR) of constructs were used to assess the scale reliability of the constructs in the framework. As per the guidelines, indicator loadings must be greater than 0.708 to meet the acceptable item reliability (Hair et al., 2019). All the items, apart from SQ1 (0.553), had an indicator loading greater than the cut-off value. The item SQ1 was hence removed from further analysis (Hair et al., 2019). The CR values of the constructs were examined to determine the internal consistency reliability of the constructs. As shown in Appendix 3, all the CR values were greater than the cut-off value of 0.7 (Hair et al., 2019), thus confirming the internal consistency reliability. After this, the construct validity was assessed by evaluating the convergent and discriminant validity of the constructs. Convergent validity was proved when all of the extracted average variance (AVE) was larger than the cut-off of 0.5 (Hair et al., 2019). We used HTMT criteria to assess the discriminant validity (Hair et al., 2019). All relevant HTMT values (see Appendix 4) were observed to be less than the 0.85 criterion, confirming discriminant validity (Hair et al., 2019). Thus, based on the above results, we can infer that our constructs demonstrated adequate internal consistency, indicator reliability, and convergent and discriminant validity.

5.1.2 Assessing potential common method bias

In research employing a cross-sectional survey design, common method bias (CMB) might be an issue (Podsakoff et al., 2003). We used some ex-ante measures recommended by Podsakoff et al. (2012) to reduce the CMB. First, the measures of the predicted and predictor variables were proximally separated (using sections in the online survey form). Second, a panel consisting of a Professor from the IS domain and two senior doctoral students reviewed the scale items for ambiguous terms, complexity, and other discrepancies. Third, different scale types (Likert scale, Semantic differential scale) were used to measure constructs to eliminate common scale properties. We used two distinct ways to conduct post hoc statistical analysis to determine the degree of CMB. First, Harman's single-factor test (Podsakoff et al., 2003) revealed that no single factor explained most of the variance. The largest single factor that emerged from the test accounted for 37.832% of the variance, which is less than the cut-off value of 50% (Podsakoff et al., 2003). Second, we applied an advanced method known as the "marker variable" method (Lindell & Whitney 2001) by including a conceptually irrelevant marker variable in the research model (Lindell &Whitney 2001). Fashion consciousness (Malhotra et al., 2006), a popular marker variable theoretically unrelated to the constructs, was chosen for this study. The shared variance of the marker variable with other constructs in the model was found to be very low. Hence, it was concluded that no substantial CMB exists in the data (Johnson et al., 2011).

5.1.3 Structural model assessment

The latent constructs' variance inflation factor (VIF) values were used to examine the model for multi-collinearity issues (Hair et al., 2019). The VIF values of all constructs were observed to be lower than the cut-off level of 3.3, indicating that there were no multi-collinearity issues in the data (Hair et al., 2019). The VIF values are shown in Appendix 5.

The validity and significance of the path coefficients in the empirical model were assessed using the PLS bootstrapping method with 5,000 subsamples (Hair et al., 2016). Table 2 and Figure 3 summarize the findings of path analysis. First, the influence of the control variables on the dependent variable intention to use (IU) was examined. The results suggest that none of the control variables except gender (β = -0.176, $p < 0.05$) had a statistically significant influence on the outcome variable IU (see Table 2).

All proposed hypotheses, except H3, H6, H10a, and H10b, were found significant at $p < 0.05$. The factors, namely perceived anthropomorphism, perceived information quality, perceived explainability, disposition to trust technology, and perceived service quality (in the order of the magnitude of their path coefficients), were found as statistically significant predictors of trusting beliefs supporting the hypotheses H8, H5, H4, H2, and H7 respectively. However, the data did not support the effects of factors, namely perceived system quality (H6) and trust signs (H3), on trusting beliefs.

Further, the impact (negative) of trusting beliefs on perceived health risk (H9b) and perceived privacy risk (H9a) were significant. The impact of trusting beliefs on perceived health risk (β = -0.697, $p < 0.05$) was higher than that on perceived privacy risk (β = -0.455, $p < 0.05$). However, the impact of trust signs on perceived privacy risk (H10a) and perceived health risk (H10b) turned out to be statistically insignificant. Similarly, trusting beliefs, perceived privacy risk and perceived health risk were statistically significant predictors of intention to use AICSD, supporting hypotheses H1, H11a, and H11b.

Hypothesis	Path	*B*	t Statistics	p-value	Inference
H1	TB → IU	0.578	8.510	0.000	Supported
H2	DT → TB	0.153	3.553	0.000	Supported
H3	TS → TB	0.050	1.070	0.285	Not Supported
H4	EX → TB	0.226	3.523	0.000	Supported
H5	IQ → TB	0.236	3.476	0.001	Supported
H6	SY → TB	-0.091	1.384	0.166	Not Supported
H7	SQ → TB	0.150	1.996	0.046	Supported

H8	PA → TB	0.319	6.521	0.000	Supported
H9a	TB → PR	-0.455	6.479	0.000	Supported
H9b	TB → HR	-0.697	14.154	0.000	Supported
H10a	TS → PR	0.106	1.278	0.201	Not Supported
H10b	TS → HR	0.069	1.048	0.295	Not Supported
H11a	PR → IU	-0.194	3.503	0.000	Supported
H11b	HR → IU	-0.168	2.583	0.010	Supported
Control Variables	**Path**	***B***	**t Statistics**	**p-value**	**Inference**
Age	Age → IU	-0.060	1.064	0.287	Not significant
Gender	Gender → IU	-0.176	2.005	0.045	Significant
Experience	Exp → IU	0.007	0.080	0.937	Not significant

Table 2. Path analysis

Note. β, path coefficient; t, two-tailed t-test values; p-value, the significance level.

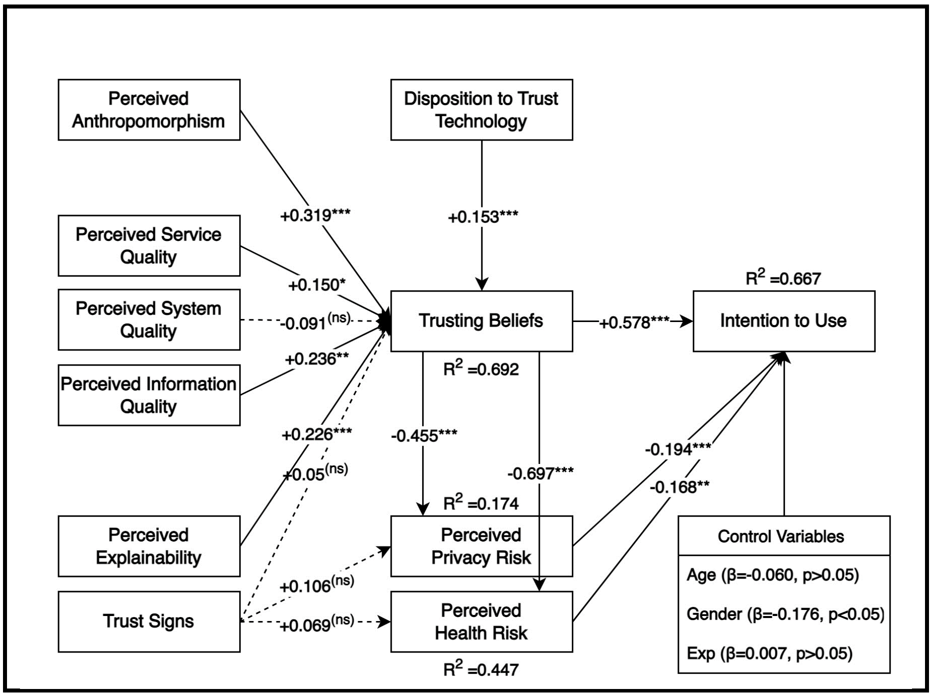

Figure 3. Structural model

Further, the R^2 values of the endogenous constructs were checked to ascertain the explanatory power of the model. The R^2 values are reported in Figure 3. The R^2 values of the constructs reveal the percentage of variance explained by its predictors in the model (Hair et al., 2019). For example, the R^2 values of the outcome variable intention to use (R^2 = 0.667) indicate that approximately 66.7% of its variance is explained by the model. Similarly, our model explains 69.2% of the variance of the key construct trusting beliefs.

Furthermore, the "predictive sample reuse technique" (Stone-Geisser's Q^2) was used to examine the predictive relevance of the model (Hair et al., 2016). It displays how successfully empirically collected data can be recreated using the model and PLS parameters (Hair et al., 2019). The Q^2 values were obtained through a "blindfolding procedure" by keeping the omission distance at 7 (Hair et al. 2019). All of the endogenous constructs had Q^2 values larger than zero, indicating substantial predictive relevance for the empirical model (Hair et al., 2016).

6 Discussion

The purpose of this research was to develop a contextualized research model to understand the determinants of consumers' trust in AICSD and its behavioral consequences. We extended the TTM with relevant contextual factors as per the guidelines proposed by Hong et al. (2014). The results supported all the hypotheses except H3, H6, and H10a. The discussions in this section synthesize the results of our study with the existing literature by rationalizing the similarities and differences with the literature.

Perceived anthropomorphism has the most prominent (positive) impact among the antecedents of trusting beliefs. This finding lends credence to the theorized link between assessments of others' mental capacities and assessments of competence, trust, and accountability (Waytz et al., 2014). It also provides evidence for the proposition made by Troshani et al. (2020) regarding anthropomorphism's role in forming trust in AI-based systems. The prominence of anthropomorphism among the other determinants also points to the argument that anthropomorphism could lead to undue trust in HCI agents (Culley & Madhavan, 2013). That is, the feelings and perceptions about an anthropomorphic agent may be used in the construction of a mental model of the system, which may result in incorrect calibrations of trust based on an emotional connection with the anthropomorphic agent rather than actual system performance (Culley & Madhavan, 2013). Prior research has also argued that incorporating human characteristics may have other unintended consequences; for example, using apologies to boost perceptions of a machine's humanness may create unreasonable expectations of changed behavior. If these expectations are not met, the agent may be perceived as deceptive or lacking in integrity, which can be extremely damaging to trust (De Visser et al., 2016). Further, past research has reported concerns that over-anthropomorphism may lead to an overestimation of the AI's capabilities, thus placing the stakeholder at risk (Culley & Madhavan, 2013), as well as leading to a host of ethical and psychological issues, such as manipulation (Salles et al., 2020). Consequently, anthropomorphism may not always lead to higher trust (Chui et al., 2019; Moussawi, 2021), although our study and some other studies in different contexts (Verberne et al., 2015; Waytz et al., 2014) report a positive impact.

The second most prominent determinant of trusting beliefs is perceived information quality. This finding is very much in line with the propositions made by the earlier studies in e-health services (Sillence et al., 2007; Song & Zahedi, 2007). Our results thus support and validate the argument that users may regard a health infomediary's high-quality information as a proxy for its goal alignment and appropriate actions as the online users' agent (Song & Zahedi, 2007). Here, such a sense of the agent's goal alignment could result in higher trust in the agent on the user's part. As is the case with the literature on interpersonal trust, where people trust a speaker who provides truthful or credible information (Giffin 1967), our study argues that users are likely to trust an AICSD that provides current, accurate, and reliable information.

Perceived explainability is the third most significant predictor of trusting beliefs. We observed that higher perceived explainability leads to higher trust. Our research adds to the specific stream of explanation-driven trust in chatbots (Zierau et al., 2020) and to the broader ongoing investigations on whether AI explanations impact people's trust in AI systems. While some studies have found that user interfaces that offer explanations are effective at increasing users' trust in AI systems (McGuinness et al., 2006; Pu & Chen, 2006), other studies have found contradictory results—providing explanations may not increase satisfaction or even erode users' trust in a system (Cramer et al., 2008; Kizilcec, 2016; Zang et al., 2021). Among the very few studies reported in the context of patient-facing AI systems, the positive impact of explainability on the trusting beliefs observed by our research is in stark contrast, for example, to the result reported by Zang et al. (2021).

As expected, the user-specific factor, i.e., disposition to trust technology, registered a significant positive impact as in the original TTM (Lankton et al., 2015; McKnight et al., 2011). Users' disposition to trust could be significant in the early usage stages, as McKnight et al. (2011) observed. Finally, the positive effect of perceived service quality was also supported, like in the case of m-health services (Akter et al., 2013). Consumers gradually build trust by evaluating many explicit and implicit indicators about the service. Among those signs, perceived service quality represents evaluations of direct experience (Chiou & Droge, 2006). If it is perceived favorably, adverse selection and moral hazard problems will be addressed, and customers will have greater confidence in the system; this, in turn, will strengthen their trust in the system. Nevertheless, we note that the role of perceived service quality in determining trust has rarely been tested in the context of AI-delivered services.

However, the insignificance of the influence of perceived system quality and trust signs on trusting beliefs is noteworthy. The finding related to perceived system quality deviates from the earlier research on health infomediaries (Song & Zahedi, 2007). The possible reason could be that the dimensions of perceived system quality, such as usability, access, and availability of the system, did not appear critical to the tech-savvy participants in this study. The users possibly considered these aspects a hygiene factor rather than critical antecedents to trusting beliefs. Similarly, trust signs, i.e., cues about assurance of performance and privacy, third-party endorsements/certifications, and customer reviews, do not seem to influence trusting beliefs as in earlier research on health infomediaries (Song & Zahedi, 2007). The likely explanation could be that these trust signals may aid in developing an initial trust before the consumer uses the product/service. However, once they use it, trust (trusting beliefs) is shaped by their actual experience of using it rather than these cues. An earlier study (McKnight et al., 2004), which observed similar results in the context of e-commerce, remarked that because of the uncertainties surrounding the web, third-party seals are insufficient to nurture trust; rather, 'seeing (or interacting) is believing.' Similar conclusions could be made in this case as well.

Further, the dampening effect of trust signs on privacy or health risk perceptions was not observed as envisaged. There is a growing notion among internet users that trying to protect privacy is a pointless endeavor (Xie et al., 2019), given the stories about data breaches/misuse in cyber settings. Adding to this, though the companies often display privacy seals and declare privacy policies, many of them secretly and deliberately violate these policies by sharing personal identifiers with third parties without declaring it (Brandtzaeg et al., 2019; Okoyomon et al., 2019; Yu et al., 2016). This growing fatalistic view of privacy (Xie et al., 2019), coupled with secret violations of the privacy policies, could be a reason that trust cues/signs were unable to produce a significant negative impact on privacy risk perceptions. Likewise, the trust signs also did not assuage the perception of health risks. Thus, having trust cues such as a third-party seal of approval, a significant amount of positive customer ratings, and privacy assurance was not seen to help alleviate the fear of negative health consequences of using AICSD. This result supports the observations by a previous study on health infomediaries that trust signs were insufficient on their own to reduce the users' risk perceptions (Song & Zahedi, 2007).

However, even though the trust cues/signs had no influence on the user's perception of trust or on their risk perceptions, our study makes an important observation that trust (trusting beliefs) in the AICSD formed after the initial use will reduce both the privacy risk perceptions and health risk perceptions. This supports the prevailing arguments and evidence in the literature on how trust reduces uncertainty and provides a sense of assurance when outcomes are unclear (Gefen et al., 2003b; Jarvenpaa et al., 2000; Nicolaou and McKnight, 2006; Pavlou & Gefen, 2004). When these results are read together, it implies that after the initial experience, the trust is dependent heavily on a person's individual experience with the technology, and cues such as a third-party seal of approval, positive customer ratings, and privacy assurance will not matter. According to the classification in signaling theory (Cook, 2001), the trust cues/signs will fall under the category of non-observable properties. These non-observable properties, even though effective in forming initial trust (similar to trust in unknown parties from the interpersonal trust literature), will not be relevant after the individual has had the initial experience of using the technology (McKnight et al., 2011). The knowledge-based trust shaped by the user experience will determine behavioral consequences (Siau & Wang, 2018). Here, the trust in AICSD is formed from an individual's perceptions about anthropomorphism, information quality, explainability, and service quality, which are derived based on his/her use experience and his/her predisposition to trust technology in general.

With respect to the antecedents of behavioral intention to use, trusting beliefs in AICSD were found to be the strongest predictor (positive) supporting the original hypothesis in the TTM (Lankton et al., 2015; McKnight et al., 2011). Similarly, the risk perception perceived privacy risk shows a significant effect on the intention to use, supporting the extant literature on various e-health services (Berezowska et al., 2015; Wendel et al., 2013; Zhang et al., 2019; Zhao et al., 2018). Likewise, health risk perception was also identified as a significant barrier, validating a similar claim made by (Prakash & Das, 2020). Finally, among the control variables, only gender ($\beta = -0.176$, $p < 0.05$) was significantly associated with intentions to use. The effect of age and prior experience of using automated chatbots (general) turned out to be insignificant. As per the findings, females are more likely to show a higher intention to use AICSD compared to males. The strong influence of gender on the intention to use AI-based health technology is noteworthy. Several studies in the United States have previously revealed that females are more likely than males to seek online health information and use health apps (Atkinson et al., 2009; Escoffery, 2018; Haluza & Wernhart, 2019). Similar findings, indicating a female majority in health app use and online searching for health-related material, have been reported in other countries, including China, Saudi Arabia, and Germany (AlGhamdi & Moussa, 2012; Baumann et al., 2017; Zhang et al., 2014).

With respect to the explanatory power of the model, it was found to be superior to the original base/generic model TTM (Mcknight et al., 2011; Lankton et al., 2015), where the values of variables trusting beliefs and trusting intentions were 50% and 50%, respectively. A likely reason for our model's superior explanatory power could be incorporating appropriate contextual factors into the base model (TTM). Thus, our model offers good explanatory power in explaining the trust (trusting beliefs) in AICSD and its consequences.

6.1 Theoretical contributions

From a theoretical perspective, this study makes a few significant contributions to the literature on trust in AI-based technologies for self-diagnosis in healthcare settings. Examining contextual factors and formulating context-specific theories is crucial for advancing research in IS (Hong et al. 2014; Orlikowski & Iacono 2001). The existing literature on trust in AICSD for healthcare is deficient in this aspect. The contextualization approach we applied by integrating relevant factors to the trust in technology framework contributes to a more profound comprehension of the specific factors that might impact the consumers' trust in this emerging AI-based technology. By empirically validating our model, we demonstrated that such integration can significantly improve the generic trust in the technology model's explanatory power. This is a valuable contribution to the literature on trust in technologies (Mcknight et al., 2011; Lankton et al., 2015), especially emerging AI-based technologies.

Second, the study contributes to the emerging literature on explanation-driven trust in AI technologies (Shin, 2021; Zierau et al., 2020) by demonstrating the mechanism and impact of perceived explainability on the trust formation process in AICSD. Our study provides evidence for establishing perceived explainability as a significant predictor of trusting beliefs. Because the impact of this factor on trust is rarely validated (Shin, 2020, 2021) in chatbots/personal assistants (Zierau et al., 2020) and AI systems in general (Rai, 2019), especially in healthcare technologies (O'Connor et al., 2021; Zhang et al., 2021) and ambiguity regarding the magnitude and direction of the relationship still exists in the literature (Cramer et al., 2008; Kizilcec, 2016; McGuinness et al., 2006; Pu & Chen, 2006; Rai, 2019; Zang et al., 2021), our study makes a valuable contribution by showing that such explanations might assist users in interpreting opaque AI recommendations and forming a higher trust towards the AI system. Additionally, in the context of patient-facing AI systems, the positive effect of explainability on trust observed by our research is noteworthy as it furthers the ongoing debate on the direction of the impact of explainability on trust (Zang et al., 2021).

Third, our study contributes to the literature on interaction-driven trust in AI-based chatbots/agents (De Visser et al., 2016; Troshani et al., 2020) by demonstrating the impact of perceived anthropomorphism on trust. It demonstrates that despite being a highly utilitarian context, users care about the humanness of the chatbot more than the other utilitarian factors. Thus, we add to the argument by Culley and Madhavan (2013) that undue trust could be formed due to an overestimation of the system's capabilities as the user gets carried away by the anthropomorphic features of the chatbot. It also points to the emerging concept of 'human-washing,' which is the deceptive use of anthropomorphic features to mislead organizational stakeholders and the general public about the machines' true capabilities (Scorici et al., 2022; Seele & Schultz, 2022).

Fourth, we validate the role of risk perceptions (privacy and health risk) as the key barriers to AICSD usage intentions. Additionally, our research confirms the critical role of trust (trusting beliefs) in alleviating the risk perceptions about the use in the post-usage context. While research is still equivocal on whether trusting beliefs predict perceived risk or vice versa, our findings add to the line of argument put forward by Gefen et al. (2003b) and Holmes (1991) and the supporting literature that followed (Jarvenpaa et al., 2000; Lou, 2002; Nicolaou and

McKnight, 2006; Pavlou 2003; Pavlou & Gefen, 2004) that trust reduces uncertainty and provides a sense of assurance when outcomes are unclear.

Additionally, the non-significant impact of trust signs and perceived system quality on trusting beliefs presents a deviation from the prior literature (Hoffmann et al., 2014; Song & Zahedi, 2007). Investigating why the relationships do not hold well in specific contexts could lead to the refinement of theory and a better understanding of causality (Christensen & Raynor, 2003). This might represent an opportunity to test a similar hypothesis in a different cultural or demographic context to understand the variability and underlying causes.

6.2 Implications to practice

From a practitioner's perspective, our context-specific model, which focuses on the specific D2C AI applications (chatbots) used for self-diagnosis in a healthcare setting, provides specific and actionable insights for enhancing the consumers' trust in and acceptance of these novel health technologies. First, the intention to use is determined by trusting beliefs and risk beliefs (privacy and health risk). Enhancing the trust in the technology (ability, integrity, benevolence, and reliability) is hence key to promoting the usage.

Our study identifies the factors, namely perceived anthropomorphism, perceived information quality, perceived explainability, disposition to trust technology, and perceived service quality, as the significant predictors of trust (trusting beliefs), which offers several actionable implications. Since perceived anthropomorphism is the most prominent determinant of building trust, designers should consider including human-like features as a conscious design choice. They should improve upon the anthropomorphic design elements of the chatbot that would make it appear more natural, human-like, and conscious. A few ways to achieve this are through visual embodiment (e.g., avatar), improving conversational ability (contingency and interactivity of the chatbot) (Go & Sundar, 2019), or empathic cues of the agent (Zierau et al., 2020). However, the decision to infuse the chatbot with more human-like attributes should be carefully and consciously considered to ensure that any impacts that emerge as a result are consistent with the intended design objectives. Because there is a possibility of consumers overtrusting the chatbot due to its human-like features (Culley & Madhavan, 2013) as well as invoking an eerie reaction happening due to the 'uncanny valley effect' (Ho & MacDorman, 2017) when the anthropomorphism exceeds a certain threshold.

Another actionable determinant of trusting beliefs is perceived information quality. This implies that in order to improve trust in AICSD, the service providers should continuously monitor the accuracy and currency of the medical/health-related information provided to the user. They must also monitor if the information shared is useful and relevant to their condition/query and ensure it is complete and understandable. Similarly, service quality perceptions can be improved by using a visually attractive user interface, providing reliable services, ensuring the user's safety, and improving the bot's responsiveness and empathy in conversations.

Further, the perceived explainability, i.e., the reason why the bot suggests a particular diagnosis and course of action, is an important determinant of trust (trusting beliefs). This implies that users are keen on receiving an explanation justifying the suggestion of the bot, and a satisfactory explanation can open up the otherwise opaque process/algorithm involved in the generation of recommendations and instill trust in the bot. However, what constitutes a good explanation and how the characteristics of the explanation affect trust is yet to be determined (Rai, 2019). Finally, consumers could be segmented according to their disposition to trust technology. Providers could put efforts into convincing the consumers (with a moderate or low propensity) about the trustworthiness of chatbots by clarifying their doubts and addressing their concerns through personalized marketing communication/customer service.

Further, reducing the risk belief is critical to improving the usage, as per our study. Accordingly, user perception of privacy risk could be assuaged by safeguarding user privacy by implementing privacy-preserving settings, privacy seals, and certifications issued by reputable third-party organizations such as TRUSTe/VeriSign and communicating the same to the public. However, this could become very challenging as new research indicates that online users are becoming increasingly fatalistic about privacy protection (Xie et al., 2019) due to deliberate/secret violations of privacy policies by the service providers (Brandtzaeg et al., 2019; Okoyomon et al., 2019; Yu et al., 2016). In this study itself, the impact of privacy-related trust signs was found to be ineffective in reducing privacy risk perceptions. Strict regulatory oversight by government agencies could help in rebuilding consumer trust in privacy protection measures/settings. Similarly, for managing health risks, endorsements from doctors representing professional medical bodies, such as the Indian Medical Association, etc., could be added to their marketing communications. Consumers could also be given the option to consult with a human medical practitioner in case of doubts regarding the chatbot recommendations.

Even though, according to our results, the trust cues are ineffective at reducing the risk perceptions, our results indicate that these risk perceptions can be assuaged by building trust in the AICSD. Given the evidence that this 'trust in AICSD' depends on the user's perception of attributes of the technology, i.e., anthropomorphism, information quality, explainability, and service quality, it allows enhancing/manipulating some of these attributes to create an enhanced sense of security (even a false perception), which would reduce the risk perception and thereby promote the adoption/use. For example, as previously indicated, perceived anthropomorphism can be enhanced by manipulating design characteristics such as voice, text, or physical embodiment or improvements in conversational ability, speech recognition, or adaptive behavior (Go & Sundar, 2019; Moussawi, 2021), which could build trust in AICSD and, in turn, will help reduce the risk perceptions.

6.3 Limitations & future research direction

Withstanding its findings, this work has some limitations, some of which may provide avenues for further research. First, by selecting India as the research setting, our findings may primarily apply to comparable developing nations. We acknowledge the influence of culture in the formation of trust and acceptance of emerging technology and suggest that cross-cultural research might help to confirm our findings. Second, participants of this study were predominantly post-graduate students (although there was a small group of non-student respondents), which could restrict the generalizability of the findings. However, we contend that our samples sufficiently represent the population since our participants are representative of the target consumers for health apps in general. Nevertheless, substantiating our findings with a larger representative sample would add credibility to our conclusions. Lastly, our research was cross-sectional in nature, and it did not examine actual behavior; instead, it focused on "use intentions." A longitudinal approach is advised for further empirical evaluation of the framework in similar scenarios with other contextual factors. It would also be worthwhile to investigate the link between stated intention and actual usage.

Declaration

The authors report there are no competing interests to declare. The first author was partially supported by the University Grants Commission, Government of India [grant number F.15-6(DEC. 2015)/2016(NET)].

References

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE Access*, *6*, 52138-52160.

Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, *31*(2), 427–445. https://doi.org/10.1007/s12525-020-00414-7

Adebesin, F., & Mwalugha, R. (2020). The Mediating Role of Organizational Reputation and Trust in the Intention to Use Wearable Health Devices: Cross-Country Study. *JMIR Mhealth Uhealth*, *8*(6), e16721. https://doi.org/10.2196/16721

Akter, S., Ray, P., & Ambra, J. D. (2013). Continuance of mHealth services at the bottom of the pyramid : the roles of service quality and trust. *Electronic Markets*, *23*(1), 29–47. https://doi.org/10.1007/s12525-012-0091-5

AlGhamdi, K. M., & Moussa, N. A. (2012). Internet use by the public to search for health-related information. *International Journal of Medical Informatics*, *81*(6), 363-373.

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, *85*, 183–189. https://doi.org/10.1016/j.chb.2018.03.051

Asan, O., & Choudhury, A. (2021). Research trends in artificial intelligence applications in human factors health care: mapping review. *JMIR Human Factors*, *8*(2), e28236.

Atkinson, N., Saperstein, S., & Pleis, J. (2009). Using the internet for health-related activities: findings from a national probability sample. *Journal of Medical Internet Research*, *11*(1), e1035.

Bach, T. A., Khan, A., Hallock, H., Beltrão, G., & Sousa, S. (2022). A Systematic Literature Review of User Trust in AI-Enabled Systems: An HCI Perspective. *International Journal of Human–Computer Interaction*, 1-16.

Bacharach, M., & Gambetta, D. (2001). Trust in signs. In K. Cook (ed.), *Trust in Society* (pp. 148–184). New York, NY, USA:: Russell Sage Foundation.

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics*, *1*(1), 71–81. https://doi.org/10.1007/s12369-008-0001-3

Baumann, E., Czerwinski, F., & Reifegerste, D. (2017). Gender-specific determinants and patterns of online health information seeking: results from a representative German health survey. *Journal of Medical Internet Research*, *19*(4), e92.

Bedué, P., & Fritzsche, A. (2022). Can we trust AI? An empirical investigation of trust requirements and guide to successful AI adoption. *Journal of Enterprise Information Management*, *35*(2), 530-549.

Benbasat, I., & Wang, W. (2005). Trust in and adoption of online recommendation agents. *Journal of the Association for Information Systems*, *6*(3), 4.

Berezowska, A., Fischer, A. R. H., Ronteltap, A., van der Lans, I. A., & van Trijp, H. C. M. (2015). Consumer adoption of personalised nutrition services from the perspective of a risk–benefit trade-off. *Genes & Nutrition*, *10*(6), 42. https://doi.org/10.1007/s12263-015-0478-y

Bickmore, T. W., Trinh, H., Olafsson, S., O'Leary, T. K., Asadi, R., Rickles, N. M., & Cruz, R. (2018). Patient and Consumer Safety Risks When Using Conversational Assistants for Medical Information: An Observational Study of Siri, Alexa, and Google Assistant. *Journal of Medical Internet Research*, *20*(9), e11510. https://doi.org/10.2196/11510

Brandtzaeg, P. B., Pultier, A., & Moen, G. M. (2019). Losing control to data-hungry apps: A mixed-methods approach to mobile app privacy. *Social Science Computer Review*, *37*(4), 466-488. https://doi.org/10.1177/0894439318777706

Breward, M., Hassanein, K., & Head, M. (2017). Understanding consumers' attitudes toward controversial information technologies: A contextualization approach. *Information Systems Research*, *28*(4), 760-774.

Chi, O. H., Jia, S., Li, Y., & Gursoy, D. (2021). Developing a formative scale to measure consumers' trust toward interaction with artificially intelligent (AI) social robots in service delivery. *Computers in Human Behavior*, *118*, 106700. https://doi.org/10.1016/j.chb.2021.106700

Chiles, T. H., & McMackin, J. F. (1996). Integrating Variable Risk Preferences, Trust, and Transaction Cost Economics. *Academy of Management Review*, *21*(1), 73–99. https://doi.org/10.5465/amr.1996.9602161566

Chiou, J.-S., & Droge, C. (2006). Service quality, trust, specific asset investment, and expertise: Direct and indirect effects in a satisfaction-loyalty framework. *Journal of the Academy of Marketing Science*, *34*(4), 613. https://doi.org/10.1177/0092070306286934

Christensen, C. M., & Raynor, M. E. (2003). Why hard-nosed executives should care about management theory. *Harvard Business Review, 81*(9), 66-75.

Chui, K. T., Liu, R. W., Lytras, M. D., & Zhao, M. (2019). Big data and IoT solution for patient behaviour monitoring. *Behaviour & Information Technology*, *38*(9), 940-949.

Ćirković, A. (2020). Evaluation of Four Artificial Intelligence–Assisted Self-Diagnosis Apps on Three Diagnoses: Two-Year Follow-Up Study. *Journal of Medical Internet Research*, *22*(12), e18097. https://doi.org/10.2196/18097

Cocosila, M., & Turel, O. (2016). A dual-risk model of user adoption of mobile-based smoking cessation support services. *Behaviour & Information Technology*, *35*(7), 526-535.

Cocosila, M., Turel, O., Archer, N., & Yuan, Y. (2007). Perceived health risks of 3G cell phones: do users care?. *Communications of the ACM*, *50*(6), 89-92.

Cook, K. (ed.). (2001). *Trust in society*. New York, NY, USA: Russell Sage Foundation.

Cramer, H., Evers, V., Ramlal, S., Van Someren, M., Rutledge, L., Stash, N., ... & Wielinga, B. (2008). The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-adapted Adapted Interaction*, *18*, 455-496.

Cronin, J. J., Brady, M. K., & Hult, G. T. M. (2000). Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments. *Journal of Retailing*, *76*(2), 193–218. https://doi.org/10.1016/S0022-4359(00)00028-2

Culley, K. E., & Madhavan, P. (2013). A note of caution regarding anthropomorphism in HCI agents. *Computers in Human Behavior*, *29*(3), 577–579. https://doi.org/10.1016/j.chb.2012.11.023

Cunningham, S.M. (1967). The major dimensions of perceived risk. In D. F. Cox (Ed.), *Risk taking and information handling in consumer behavior* (pp. 82–108). Harvard University Press.

Dagger, T. S., Sweeney, J. C., & Johnson, L. W. (2007). A Hierarchical Model of Health Service Quality: Scale Development and Investigation of an Integrated Model. *Journal of Service Research*, *10*(2), 123–142. https://doi.org/10.1177/1094670507309594

De Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A. B., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016). Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied*, *22*(3), 331.

Delone, W. H., & McLean, E. R. (2003). The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *Journal of Management Information Systems*, *19*(4), 9–30. https://doi.org/10.1080/07421222.2003.11045748

Denecke, K., Gabarron, E., Grainger, R., Konstantinidis, S. T., Lau, A., Rivera-Romero, O., ... & Merolli, M. (2019). Artificial intelligence for participatory health: applications, impact, and future implications. *Yearbook of Medical Informatics*, *28*(01), 165-173.

Dinev, T., & Hart, P. (2006). An Extended Privacy Calculus Model for E-Commerce Transactions. *Information Systems Research*, *17*(1), 61–80. https://doi.org/10.1287/isre.1060.0080

Eisingerich, A. B., & Bell, S. J. (2008). Perceived Service Quality and Customer Trust: Does Enhancing Customers' Service Knowledge Matter? *Journal of Service Research*, *10*(3), 256–268. https://doi.org/10.1177/1094670507310769

Epley, N., Caruso, E. M., & Bazerman, M. H. (2006). When perspective taking increases taking: reactive egoism in social interaction. *Journal of Personality and Social Psychology*, *91*(5), 872.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864.

Escoffery, C. (2018). Gender similarities and differences for e-health behaviors among US adults. *Telemedicine and e-Health*, *24*(5), 335-343.

Fan, X., Chao, D., Zhang, Z., Wang, D., Li, X., & Tian, F. (2021). Utilization of Self-Diagnosis Health Chatbots in Real-World Settings: Case Study. *Journal of Medical Internet Research*, *23*(1), e19928. https://doi.org/10.2196/19928

Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. *Interactions*, *24*(4), 38-42.

Fromkin, H. L., & Streufert, S. (1976). Laboratory experimentation. In M. D. Dunnette (ed.), *Handbook of Industrial and Organizational Psychology* (pp. 415-465). Rand McNally.

Gambetta, D., & Bacharach, M. (2001). Trust in signs. *Trust and Society, New York, NY,USA: Russell Sage Foundation*, 148-184.

Gefen, D., Karahanna, E., & Straub, D. W. (2003a). Trust and TAM in Online Shopping: An Integrated Model. *MIS Quarterly*, *27*(1), 51–90. https://doi.org/10.2307/30036519

Gefen, D., Srinivasan Rao, V., & Tractinsky, N. (2003). The conceptualization of trust, risk and their electronic commerce: The need for clarifications. In R. H. Sprague (Ed.), *Proceedings of the 36th Annual Hawaii International Conference on System Sciences* (HICSS 2003), Article 1174442. Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/HICSS.2003.1174442

Gille, F., Jobin, A., & Ienca, M. (2020). What we talk about when we talk about trust: Theory of trust for AI in healthcare. *Intelligence-Based Medicine*, *1–2*, 100001. https://doi.org/10.1016/j.ibmed.2020.100001

Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, *97*, 304-316.

Hair Jr, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. SAGE Publications.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Haluza, D., & Wernhart, A. (2019). Does gender matter? Exploring perceptions regarding health technologies among employees and students at a medical university. *International Journal of Medical Informatics*, *130*, 103948.

Hengstler, M., Enkel, E., & Duelli, S. (2016). Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices. *Technological Forecasting and Social Change*, *105*, 105-120.

Ho, C.-C., & MacDorman, K. F. (2017). Measuring the uncanny valley effect. *International Journal of Social Robotics*, *9*(1), 129–139.

Hoffmann, C. P., Lutz, C., & Meckel, M. (2014). Digital Natives or Digital Immigrants? The Impact of User Characteristics on Online Trust. *Journal of Management Information Systems*, *31*(3), 138–171. https://doi.org/10.1080/07421222.2014.995538

Holmes, J. G. (1991). Trust and the appraisal process in close relationships. In W. H. Jones & D. Perlman (Eds.), *Advances in personal relationships: A research annual,* Vol. 2, pp. 57–104). London, UK: Jessica Kingsley Publishers.

Hong, W., Chan, F. K., Thong, J. Y., Chasalow, L. C., & Dhillon, G. (2014). A framework and guidelines for context-specific theorizing in information systems research. *Information Systems Research*, *25*(1), 111-136.

Hu, P., Lu, Y., & Gong, Y. (2021). Dual humanness and trust in conversational AI: A person-centered approach. *Computers in Human Behavior*, *119*, 106727. https://doi.org/10.1016/j.chb.2021.106727

Jiang, J. J., Klein, G., & Carr, C. L. (2002). Measuring Information System Service Quality: SERVQUAL from the Other Side. *MIS Quarterly*, *26*(2), 145–166. https://doi.org/10.2307/4132324

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review*, *31*(2), 386-408.

Johnson, V. L., Kiser, A., Washington, R., & Torres, R. (2018). Limitations to the rapid adoption of M-payment services: Understanding the impact of privacy risk on M-Payment services. *Computers in Human Behavior*, *79*, 111–122.

Kizilcec, R. F. (2016). How much information? Effects of transparency on trust in an algorithmic interface. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 2390–2395). Association for Computing Machinery (ACM). https://doi.org/10.1145/2858036.2858402

Koller, M. (1988). Risk as a determinant of trust. *Basic and Applied Social Psychology*, *9*(4), 265-276.

Lankton, N. K., McKnight, D. H., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. *Journal of the Association for Information Systems, 16*(10), 880.

Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A. Y. S., & Coiera, E. (2018). Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association*, *25*(9), 1248–1258. https://doi.org/10.1093/jamia/ocy072

Laumer, S., Maier, C., & Gubler, F. (2019). Chatbot acceptance in healthcare: explaining user adoption of conversational agents for disease diagnosis. *Proceedings of the 27th European Conference on Information Systems (ECIS), 27*(1), 88. https://aisel.aisnet.org/ecis2019_rp/88

Li, H., Wu, J., Gao, Y., & Shi, Y. (2016). Examining individuals' adoption of healthcare wearable devices: An empirical study from privacy calculus perspective. *International Journal of Medical Informatics*, *88*, 8–17. https://doi.org/10.1016/j.ijmedinf.2015.12.010

Li, X., Hess, T. J., & Valacich, J. S. (2006). Using attitude and social influence to develop an extended trust model for information systems. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, *37*(2-3), 108-124.

Li, X., Hess, T. J., & Valacich, J. S. (2008). Why do we trust new technology? A study of initial trust formation with organizational information systems. *The Journal of Strategic Information Systems*, *17*(1), 39–71. https://doi.org/10.1016/j.jsis.2008.01.001

Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *The Journal of Applied Psychology*, *86*(1), 114–121. https://doi.org/10.1037/0021-9010.86.1.114

London, A. J. (2019). Artificial intelligence and black-box medical decisions: accuracy versus explainability. *Hastings Center Report*, *49*(1), 15-21.

Malhotra, N. K., Kim, S. S., & Patil, A. (2006). Common method variance in IS research: A comparison of alternative approaches and a reanalysis of past research. *Management Science*, *52*(12), 1865-1883.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709-734. https://doi.org/10.5465/amr.1995.9508080335

McGuinness, D. L., Ding, L., Glass, A., Chang, C., Zeng, H., & Furtado, V. (2006, November). Explanation interfaces for the semantic web: Issues and models. In *Proceedings of the 3rd International Semantic Web User Interaction Workshop (SWUI'06)*. https://hdl.handle.net/20.500.13015/4736

McKnight, D. H., & Chervany, N. L. (2001). What Trust Means in E-Commerce Customer Relationships: An Interdisciplinary Conceptual Typology. *International Journal of Electronic Commerce*, *6*(2), 35–59. https://doi.org/10.1080/10864415.2001.11044235

Mcknight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a Specific Technology: An Investigation of Its Components and Measures. *ACM Transactions on Management Information Systems*, *2*(2), 12:1-12:25. https://doi.org/10.1145/1985347.1985353

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research*, *13*(3), 334-359.

McKnight, D. H., Kacmar, C. J., & Choudhury, V. (2004). Shifting Factors and the Ineffectiveness of Third Party Assurance Seals: A two-stage model of initial trust in a web business. *Electronic Markets*, *14*(3), 252-266.

McKnight, D. H., Lankton, N. K., Nicolaou, A., & Price, J. (2017). Distinguishing the effects of B2B information quality, system quality, and service outcome quality on trust and distrust. *The Journal of Strategic Information Systems*, *26*(2), 118-141.

McTear, M., Callejas, Z., & Griol, D. (2016). Conversational Interfaces: Devices, Wearables, Virtual Agents, and Robots. In M. McTear, Z. Callejas, & D. Griol (Eds.), *The Conversational Interface: Talking to Smart Devices* (pp. 283–308). Cham: Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-32967-3_13

Mesbah, N., & Pumplun, L. (2020). "Hello, I'm here to help you" – Medical care where it is needed most: Seniors' acceptance of health chatbots. *Proceedings of the European Conference on Information Systems. 28*(1), 209. https://aisel.aisnet.org/ecis2020_rp/209

Miner, A. S., Shah, N., Bullock, K. D., Arnow, B. A., Bailenson, J., & Hancock, J. (2019). Key Considerations for Incorporating Conversational AI in Psychotherapy. *Frontiers in Psychiatry*, *10*. https://doi.org/10.3389/fpsyt.2019.00746

Minutolo, A., Esposito, M., & De Pietro, G. (2017). A Conversational Chatbot Based on Kowledge-Graphs for Factoid Medical Questions. In *New Trends in Intelligent Software Methodologies, Tools and Techniques* (pp. 139-152). Amsterdam, The Netherlands: IOS Press.

Moussawi, S., Koufaris, M., & Benbunan-Fich, R. (2021). How perceptions of intelligence and anthropomorphism affect adoption of personal intelligent agents. *Electronic Markets*, *31*(2), 343-364. https://doi.org/10.1007/s12525-020-00411-w

Nadarzynski, T., Miles, O., Cowie, A., & Ridge, D. (2019). Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. *Digital Health*, *5*, 2055207619871808.

Naneva, S., Sarda Gou, M., Webb, T. L., & Prescott, T. J. (2020). A systematic review of attitudes, anxiety, acceptance, and trust towards social robots. *International Journal of Social Robotics*, *12*(6), 1179-1201.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*(1), 81-103.

Nicolaou, A. I., & McKnight, D. H. (2006). Perceived information quality in data exchanges: Effects on risk, trust, and intention to use. *Information Systems Research*, *17*(4), 332–351.

Nowak, K. L., & Rauh, C. (2005). The influence of the avatar on online perceptions of anthropomorphism, androgyny, credibility, homophily, and attraction. *Journal of Computer-Mediated Communication*, *11*(1), 153-178.

Nundy, S., Montgomery, T., & Wachter, R. M. (2019). Promoting trust between patients and physicians in the era of artificial intelligence. *JAMA*, *322*(6), 497-498.

O'Connor, Y., Kupper, M., & Heavin, C. (2021). Trusting Intentions Towards Robots in Healthcare: A Theoretical Framework. *Proceedings of the 54th Hawaii International Conference on System Sciences*, *54*(1), 586. https://doi.org/10.24251/HICSS.2021.071

Okoyomon, E., Samarin, N., Wijesekera, P., Elazari Bar On, A., Vallina-Rodriguez, N., Reyes, I., ... & Egelman, S. (2019). On the ridiculousness of notice and consent: Contradictions in app privacy policies. In *Workshop on Technology and Consumer Protection (ConPro 2019), in conjunction with the 39th IEEE Symposium on Security and Privacy* (pp. 1-7). http://hdl.handle.net/20.500.12761/690

Orlikowski, W. J., & Iacono, C. S. (2001). Research commentary: Desperately seeking the "IT" in IT research—A call to theorizing the IT artifact. *Information Systems Research*, *12*(2), 121-134.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing, 64*(1), 12–40.

Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, *7*(3), 101-134.

Pavlou, P. A., & Gefen, D. (2004). Building effective online marketplaces with institution-based trust. *Information Systems Research*, *15*(1), 37-59.

Pelaez, A., Chen, C.-W., & Chen, Y. X. (2019). Effects of Perceived Risk on Intention to Purchase: A Meta-Analysis. *Journal of Computer Information Systems*, *59*(1), 73–84. https://doi.org/10.1080/08874417.2017.1300514

Pennington, R., Wilcox, H. D., & Grover, V. (2003). The Role of System Trust in Business-to-Consumer Transactions. *Journal of Management Information Systems*, *20*(3), 197–226. https://doi.org/10.1080/07421222.2003.11045777

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of Method Bias in Social Science Research and Recommendations on How to Control It. *Annual Review of Psychology*, *63*(1), 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879. https://doi.org/10.1037/0021-9010.88.5.879.

Powell, J. (2019). Trust Me, I'm a Chatbot: How Artificial Intelligence in Health Care Fails the Turing Test. *Journal of Medical Internet Research*, *21*(10), e16222. https://doi.org/10.2196/16222

PR Newswire. (2021, March 5). Healthcare Chatbot Market to Reach US$ 967.7 Million by 2027, Globally. *PR Newswire*. https://www.prnewswire.com/in/news-releases/healthcare-chatbot-market-to-reach-us-967-7-million-by-2027-globally-cagr-21-56-univdatos-market-insights-840043251.html

Prakash, A. V., & Das, S. (2020). Would you trust a bot for healthcare advice? An empirical investigation. *Proceedings of the Pacific Asia Conference on Information Systems*, *24*(1), 62. https://aisel.aisnet.org/pacis2020/62

Pu, P., & Chen, L. (2006). Trust building with explanation interfaces. In *Proceedings of the 11th International Conference on Intelligent User Interfaces* (pp. 93–100). Association for Computing Machinery (ACM). https://doi.org/10.1145/1111449.1111475

Qiu, L., & Benbasat, I. (2009). Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems. *Journal of Management Information Systems*, *25*(4), 145–182. https://doi.org/10.2753/MIS0742-1222250405

Rai, A. (2020). Explainable AI: from black box to glass box. *Journal of the Academy of Marketing Science*, *48*(1), 137–141. https://doi.org/10.1007/s11747-019-00710-5

Razzaki, S., Baker, A., Perov, Y., Middleton, K., Baxter, J., Mullarkey, D., Sangar, D., Taliercio, M., Butt, M., Majeed, A., DoRosario, A., Mahoney, M., & Johri, S. (2018). *A comparative study of artificial intelligence and human doctors for the purpose of triage and diagnosis*. 1–15. http://arxiv.org/abs/1806.10698

Rice, S. C. (2012). Reputation and uncertainty in online markets: an experimental study. *Information Systems Research*, *23*(2), 436-452.

Robertson, N., Polonsky, M., & McQuilken, L. (2014). Are my symptoms serious Dr Google? A resource-based typology of value co-destruction in online self-diagnosis. *Australasian Marketing Journal (AMJ)*, *22*(3), 246–256. https://doi.org/10.1016/j.ausmj.2014.08.009

Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. PUBLISHER London, UK: Penguin UK.

Salles, A., Evers, K., & Farisco, M. (2020). Anthropomorphism in AI. *AJOB Neuroscience, 11*(2), 88-95.

Sbaffi, L., & Rowley, J. (2017). Trust and Credibility in Web-Based Health Information: A Review and Agenda for Future Research. *Journal of Medical Internet Research*, *19*(6), e218–e218. https://doi.org/10.2196/jmir.7579

Scorici, G., Schultz, M. D., & Seele, P. (2024). Anthropomorphization and beyond: conceptualizing humanwashing of AI-enabled machines. *AI & Society*, *39*(2), 789–795. https://doi.org/10.1007/s00146-022-01492-1

Seele, P., & Schultz, M. D. (2022). From greenwashing to machinewashing: a model and future directions derived from reasoning by analogy. *Journal of Business Ethics*, *178*(4), 1063-1089.

Shah, H., Warwick, K., Vallverdú, J., & Wu, D. (2016). Can machines talk? Comparison of Eliza with modern dialogue systems. *Computers in Human Behavior*, *58*, 278–295. https://doi.org/10.1016/j.chb.2016.01.004

Sheehan, B., Jin, H. S., & Gottlieb, U. (2020). Customer service chatbots: Anthropomorphism and adoption. *Journal of Business Research*, *115*, 14–24. https://doi.org/10.1016/j.jbusres.2020.04.030

Shin, D. (2020). User Perceptions of Algorithmic Decisions in the Personalized AI System: Perceptual Evaluation of Fairness, Accountability, Transparency, and Explainability. *Journal of Broadcasting & Electronic Media*, *64*(4), 541–565. https://doi.org/10.1080/08838151.2020.1843357

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, *146*, 102551. https://doi.org/10.1016/j.ijhcs.2020.102551

Sillence, E., Briggs, P., Harris, P., & Fishwick, L. (2007). Health websites that people can trust–the case of hypertension. *Interacting with Computers*, *19*(1), 32–42.

Siwicki, B. (2018, February 1). Special Report: AI voice assistants have officially arrived in healthcare. *Healthcare IT News*. http://www.healthcareitnews.com/news/special-report-ai-voice-assistants-have-officially-arrived-healthcare. Accessed October 22, 2021.

Söllner, M., Hoffmann, A., & Leimeister, J. M. (2016). Why different trust relationships matter for information systems users. *European Journal of Information Systems*, *25*(3), 274–287. https://doi.org/10.1057/ejis.2015.17

Song, J., & Zahedi, F. (2007). Trust in health infomediaries. *Decision Support Systems*, *43*(2), 390–407. https://doi.org/10.1016/j.dss.2006.11.011

Tan, Y. H., & Thoen, W. (2003). Electronic contract drafting based on risk and trust assessment. *International Journal of Electronic Commerce*, *7*(4), 55-71.

Troshani, I., Rao Hill, S., Sherman, C., & Arthur, D. (2020). Do We Trust in AI? Role of Anthropomorphism and Intelligence. *Journal of Computer Information Systems*, 1–11. https://doi.org/10.1080/08874417.2020.1788473

Verberne, F. M., Ham, J., & Midden, C. J. (2015). Trusting a virtual driver that looks, acts, and thinks like you. *Human Factors*, *57*(5), 895-909.

Wang, W., & Siau, K. L. (2018). Living with Artificial Intelligence: Developing a Theory on Trust in Health Chatbots-Research in Progress. *Proceedings of the SIGHCI 2018, 4.* https://aisel.aisnet.org/sighci2018/4

Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, *52*, 113–117. https://doi.org/10.1016/j.jesp.2014.01.005

Wells, J. D., Valacich, J. S., & Hess, T. J. (2011). What Signal Are You Sending? How Website Quality Influences Perceptions of Product Quality and Purchase Intentions. MIS Quarterly, 35(2), 373–396. https://doi.org/10.2307/23044048

Wendel, S., Dellaert, B. G., Ronteltap, A., & Van Trijp, H. C. (2013). Consumers' intention to use health recommendation systems to receive personalized nutrition advice. *BMC Health Services research*, *13*(1), 126.

Whetten, D. A. (2009). An examination of the interface between context and theory applied to the study of Chinese organizations. *Management and Organization Review*, *5*(1), 29-56.

Xie, H., Prybutok, G., Peng, X., & Prybutok, V. (2020). Determinants of trust in health information technology: An empirical investigation in the context of an online clinic appointment system. *International Journal of Human–Computer Interaction*, *36*(12), 1095-1109.

Xie, W., Fowler-Dawson, A., & Tvauri, A. (2019). Revealing the relationship between rational fatalism and the online privacy paradox. *Behaviour & Information Technology*, *38*(7), 742-759. https://doi.org/10.1080/0144929X.2018.1552717

Your.MD. (2021). *Healthily: Health questions, answered*. https://www.livehealthily.com/. Assessed November 9, 2021.

Yu, L., Luo, X., Liu, X., & Zhang, T. (2016). Can we trust the privacy policies of android apps?. *Proceedings of the IEEE/IFIP International Conference on Dependable Systems and Networks, 46*(1), 538-549. https://doi.org/10.1109/DSN.2016.55

Zahedi, F. & Song, J. (2008). Dynamics of Trust Revision: Using Health Infomediaries. *Journal of Management Information Systems*, *24*(4), 225–248. https://doi.org/10.2753/MIS0742-1222240409

Zhang, X., Guo, X., Lai, K. H., Guo, F., & Li, C. (2014). Understanding gender differences in m-health adoption: a modified theory of reasoned action model. *Telemedicine and e-Health*, *20*(1), 39-46.

Zhang, Y., Liu, C., Luo, S., Xie, Y., Liu, F., Li, X., & Zhou, Z. (2019). Factors influencing patients' intentions to use diabetes management apps based on an extended unified theory of acceptance and use of technology model: web-based survey. *Journal of Medical Internet Research*, *21*(8), e15023.

Zhang, Z., Genc, Y., Wang, D., Ahsen, M. E., & Fan, X. (2021). Effect of AI Explanations on Human Perceptions of Patient-Facing AI-Powered Healthcare Systems. *Journal of Medical Systems*, *45*(6), 64. https://doi.org/10.1007/s10916-021-01743-6

Zhao, Y., Ni, Q., & Zhou, R. (2018). What factors influence the mobile health service adoption? A meta-analysis and the moderating role of age. *International Journal of Information Management*, *43*, 342–350. https://doi.org/10.1016/j.ijinfomgt.2017.08.006

Zhou, T. (2011). The impact of privacy concern on user adoption of location-based services. *Industrial Management & Data Systems*, *111*(2), 212–226. https://doi.org/10.1108/02635571111115146

Zierau, N., Engel, C., Söllner, M., & Leimeister, J. M. (2020). Trust in Smart Personal Assistants: A Systematic Literature Review and Development of a Research Agenda. *Proceedings of the International Conference on Wirtschaftsinformatik (WI 2020), 15*(1), 1-15. https://doi.org/10.30844/wi_2020_a7-zierau

Appendix

Appendix 1. Respondent demographics

Item	Category	Frequency	Percentage
Gender	Female	65	32.18
	Male	137	67.82
Age	21-25	123	60.89
	26-30	73	36.14
	31-35	2	0.99
	51-55	2	0.99
	56-61	2	0.99
Educational Qualification	Postgraduate	22	10.89
	Graduate	180	89.11
Employment status	Student	177	87.62
	Employed	25	12.38
Prior experience of using automated AI-based chatbots (general)	No	43	21.29
	Yes	159	78.71

Appendix 2. Measurement items

Construct	Item Code	Item	Source
Intention to use (IU)	IU1	I intend to use this health chatbot for making health decisions when a health issue arises in the future	Adapted from Venkatesh et al. (2003)
	IU2	I predict I would use this health chatbot for making health decisions when a health issue arises in the future	
	IU3	I plan to use this health chatbot for making health decisions when a health issue arises in the future	
Perceived Privacy Risk	PR1	In general, it would be risky to give personal health information to this Health Chatbot.	(Xu et al., 2011)
	PR2	There would be a high potential for privacy loss associated with giving personal health information to this Health Chatbot.	

(PR)	PR3	Personal health information could be inappropriately used by this Health Chatbot.	
Perceived Health Risk (HR)	HR1	How risky do you feel it would be to make a decision based on the health information provided by this chatbot?	Adapted from Mun et al. (2013)
	HR2	How risky do you feel it would be to accept and apply the provided health information to your life?	
	HR3	How risky do you feel it would be to accept and apply the provided health information to the lives of others important to you?	
Perceived Information Quality (IQ)	IQ1	This health chatbot provides useful health information to the user	(Mun et al., 2013; Song & Zahedi, 2007)
	IQ2	This health chatbot provides sufficient health information regarding the symptoms of the questioner.	
	IQ3	This health chatbot provides accurate health information.	
	IQ4	This health chatbot provides health information that the questioner is seeking for.	
	IQ5	The diagnosis of the consultation is based on the newest (up to date) health information.	
	IQ6	The information provided by the health chatbot is easy to comprehend	
Perceived System Quality (SY)	SY1	This health chatbot is responsive to your request	(Aladwani & Palvia, 2002; McKinney et al., 2002; Pantano et al., 2017; Song & Zahedi, 2007)
	SY2	This health chatbot is easy to use	
	SY3	This health chatbot is always up and available	
	SY4	This health chatbot allows me to interact with it to receive tailored information about my health issue	
Perceived Service Quality (SQ)	SQ1	*This health chatbot has a modern-looking/visually attractive interface**	(Ashfaq et al., 2020; Roca et al., 2006)
	SQ2	This health chatbot provides dependable services.	
	SQ3	I feel safe in my transactions with this health chatbot	
	SQ4	This health chatbot gives me a prompt response.	
	SQ5	This health chatbot has the users' best interests at heart.	
Perceived Anthropo-morphism (PA)	PA1	Please rate your impression of the health chatbot on these scales: Fake (unnatural) 1 2 3 4 5 Natural	(Bartneck et al., 2009; Sheehan et al., 2020)
	PA2	Machine-like 1 2 3 4 5 Human-like	
	PA3	Artificial 1 2 3 4 5 Lifelike	
	PA4	Unconscious 1 2 3 4 5 Conscious	
	PA5	Communicates Inelegantly 1 2 3 4 5 Communicates Elegantly	
Perceived Explainability (EX)	EX1	This Health Chatbot clearly explained the reason why it arrived at a particular diagnosis.	Developed from (Rai, 2020)
	EX2	This Health Chatbot clearly explains the reason for suggesting a particular diagnosis.	
	EX3	This Health Chatbot provided a clear explanation as to why it is recommending a particular course of action.	
Trust Signs (TS)	TS1	The services provided by this Health Chatbot was tested by an independent organization and given a seal/certificate of approval	(Hoffmann et al., 2014)
	TS2	This Health Chatbot received a significant amount of positive ratings from other users.	
	TS3	This Health Chatbot has a third-party privacy assurance seal (e.g., ePrivacy) that guarantees privacy protection.	

	TS4	This Health Chatbot has a strong privacy policy to protect my sensitive information	(Beldad et al., 2010; Song & Zahedi, 2007)
Trusting Beliefs (TB)	TB1	I believe that this Health Chatbot would act in my best interest.	Adapted from McKnight et al.(2002)
	TB2	I believe this Health Chatbot is truthful in its dealings with me.	
	TB3	This Health Chatbot is competent and effective in identifying health conditions from symptoms (diagnosis) and recommending an appropriate course of action.	
	TB4	I believe this Health Chatbot is very reliable for symptom checking & triage	
Disposition to trust technology	DT1	My typical approach is to trust new technologies until they prove to me that I shouldn't trust them.	(Lankton et al., 2015)
	DT2	I usually trust a technology until it gives me a reason not to trust it.	
	DT3	I generally give a technology the benefit of the doubt when I first use it.	
Marker Variable: Fashion Consciousness	MV1	When I must choose between the two, I usually dress for fashion, not for comfort.	(Malhotra et al., 2006)
	MV2	An important part of my life and activities is dressing smartly.	
	MV3	A person should try to dress in style.	

Note. * item removed from the analysis due to poor loadings (<0.708).

Appendix 3. Reliability and validity statistics

Construct	Item	Loading
Disposition to Trust Technology (DT)	*CR = 0.946; AVE = 0.853; μ = 4.850; σ = 1.484*	
	DT1	0.910
	DT2	0.933
	DT3	0.928
Perceived Explainability (EX)	*CR = 0.960; AVE = 0.890; μ = 4.686; σ = 1.139*	
	EX1	0.935
	EX2	0.950
	EX3	0.946
Perceived Health Risk (HR)	CR = 0.963; AVE = 0.897; μ = 4.687; σ = 1.430	
	HR1	0.934
	HR2	0.963
	HR3	0.943
Perceived Information Quality (IQ)	*CR = 0.938; AVE = 0.717; μ = 4.872; σ = 1.001*	
	IQ1	0.833
	IQ2	0.879
	IQ3	0.898
	IQ4	0.868
	IQ5	0.857

	IQ6	0.737
Perceived Anthropomorphism (PA)	*CR = 0.873; AVE = 0.579; μ = 3.373; σ = 0.626*	
	PA1	0.768
	PA2	0.795
	PA3	0.736
	PA4	0.788
	PA5	0.716
Perceived Privacy Risk (PR)	*CR = 0.945; AVE = 0.852; μ = 4.252; σ = 1.487*	
	PR1	0.916
	PR2	0.946
	PR3	0.908
Perceived Service Quality (SQ)	*CR = 0.896; AVE = 0.683; μ = 4.979; σ = 0.969*	
	SQ2	0.788
	SQ3	0.849
	SQ4	0.777
	SQ5	0.888
Perceived System Quality (SY)	*CR = 0.924; AVE = 0.754; μ = 5.673; σ = 0.945*	
	SY1	0.839
	SY2	0.917
	SY3	0.883
	SY4	0.831
Trusting Beliefs (TB)	*CR = 0.967; AVE = 0.854; μ = 4.472; σ = 1.249*	
	TB1	0.911
	TB2	0.916
	TB3	0.933
	TB4	0.925
	TB5	0.936
Trust Signs (TS)	*CR = 0.860; AVE = 0.607; μ = 4.395; σ = 0.921*	
	TS1	0.744
	TS2	0.791
	TS3	0.740
	TS4	0.837
Intention to Use (IU)	*CR = 0.984; AVE = 0.953; μ = 4.554; σ = 1.689*	
	IU1	0.974
	IU2	0.973
	IU3	0.983

Note. CR, Composite Reliability; AVE, Average Variance Extracted; μ, Mean; σ, Standard Deviation

Appendix 4. Discriminant validity

	DT	EX	HR	IQ	PA	PR	SQ	SY	TB	IU
DT										
EX	0.401									
HR	0.349	0.432								
IQ	0.433	0.749	0.508							
PA	0.252	0.647	0.550	0.685						
PR	0.397	0.334	0.544	0.345	0.367					
SQ	0.458	0.774	0.409	0.806	0.685	0.390				
SY	0.379	0.681	0.227	0.730	0.580	0.332	0.801			
TB	0.468	0.739	0.701	0.759	0.759	0.433	0.750	0.583		
IU	0.463	0.673	0.685	0.739	0.699	0.545	0.639	0.547	0.796	
TS	0.227	0.556	0.274	0.574	0.447	0.173	0.580	0.497	0.511	0.370

Note. As per the HTMT criteria, all the HTMT values should be less than 0.85 to confirm discriminant validity (Hair et al., 2019).

Appendix 5. Variance Inflation Factor (VIF) Values

	HR	PR	TB	IU
DT			1.256	
EX			2.506	
HR				2.077
IQ			2.837	
PA			1.789	
PR				1.445
SQ			2.87	
SY			2.178	
TB	1.261	1.261		1.841
TS	1.261	1.261	1.415	

doi: https://doi.org/10.3127/ajis.v28.4235