How to Make chatbots productive – A user-oriented implementation framework

Antje Janssen [a,b,*], Davinia Rodríguez Cardona [a], Jens Passlick [a,c], Michael H. Breitner [a]

[a] *Leibniz Universität Hannover, Faculty of Economics and Management, Information Systems Institute, Königsworther Platz 1, 30167 Hannover, Germany*
[b] *bhn Dienstleistungs GmbH & Co. KG, Hans-Lenze-Straße 1, 31855 Aerzen, Germany*
[c] *VHV Vereinigte Hannoversche Versicherung a.G., VHV-Platz 1, 30177 Hannover, Germany*

ARTICLE INFO

ABSTRACT

Many organizations are pursuing the implementation of chatbots to enable automation of service processes. However, previous research has highlighted the existence of practical setbacks in the implementation of chatbots in corporate environments. To gain practical insights on the issues related to the implementation processes from several perspectives and stages of deployment, we conducted semi-structured interviews with developers and experts of chatbot development. Using qualitative content analysis and based on a review of literature on human computer interaction (HCI), information systems (IS), and chatbots, we present an implementation framework that supports the successful deployment of chatbots and discuss the implementation of chatbots through a user-oriented lens. The proposed framework contains 101 guiding questions to support chatbot implementation in an eight-step process. The questions are structured according to the people, activity, context, and technology (PACT) framework. The adapted PACT framework is evaluated through expert interviews and a focus group discussion (FGD) and is further applied in a case study. The framework can be seen as a bridge between science and practice that serves as a notional structure for practitioners to introduce a chatbot in a structured and user-oriented manner.

1. Introduction

In recent years, chatbots have become increasingly popular (Brandtzaeg and Følstad, 2018; Benner et al., 2021) due to major developments in machine learning (ML), and natural language processing (NLP), which have enabled new forms of chatbots (Seeger et al., 2018). The hype surrounding chatbots has led many companies in different fields to introduce them to show their technological prowess or to just offer a new channel for client interaction. Chatbot consultancies promote the ease of developing chatbots within a short timeframe (e.g., *"FAQ Chatbot In a Day"*[1] and *"Learn how to build a Facebook chatbot – in just one day!"*[2]), typically focusing on the use of a specific development tool rather than the full development process. In this regard, previous research has focused primarily on specific aspects of chatbot implementation, e.g., design elements such as "technical, situational and knowledge features" (Janssen et al., 2020, p.213), design principles

(Lewandowski et al., 2022) and tasks within chatbot introduction (Lewandowski et al., 2022; Meyer von Wolff et al., 2022) or specific prototypes developed through case studies (Laumer et al., 2019; Seeger et al., 2018).

While chatbots are increasingly being improved, they may be prone to functional failure. This has been noted as a concern in research and practice as it may lead to a loss of credibility and frustration among users (Benner et al., 2021; Brandtzaeg and Følstad, 2018; Janssen et al., 2021a). One of the main reasons why chatbots currently fail, is that user expectations do not match their functionalities (Janssen et al., 2021a). In a rapidly changing and increasingly digitized world, where people are constantly confronted with new technological changes, a key success factor within the design, implementation, and evaluation phases is to gain a deep understanding of how people interact with the technology being developed (Adam et al., 2021). Zierau et al. (2020) emphasize that task context and user characteristics have barely been studied in the

---

chatbot research so far, even though individual studies (e.g., Brandtzaeg and Følstad, 2017; Van der Goot et al., 2021) have already focused intensively on how users use chatbots. This is despite their impact on HCI just as system characteristics and task (Li and Zhang, 2005; Zierau et al., 2020). Therefore, not only technical factors, but the (potential) users, their activities, and the respective context must also be considered during implementation (Adam et al., 2021; Benyon, 2014) which can be done by using the PACT framework of Benyon (2005, 2014). A holistic framework that provides guiding questions for practitioners and researchers is helpful, as it can ensure a user-, context-, activity-, and technology-oriented alignment in each step. By putting the user at the center, the framework helps practitioners and researchers get an overview of the issues that need to be asked during the implementation process and to develop chatbots independent of the provider. With the objective of developing an artefact in the form of a user-oriented chatbot implementation framework containing guiding questions that need to be considered during implementation, we aim to address the following research question:

RQ: What questions need to be considered in a user-oriented chatbot implementation and how can these questions be structured in an implementation framework?

We apply the HCI design science research steps of Vaishnavi and Kuechler (2015) to address the implementation process through a user-oriented lens (Adam et al., 2021). We conduct 15 semi-structured interviews with practitioners who have already implemented chatbots. To gain an understanding of how the chatbot development process takes place and to use our findings for describing the most relevant aspects of chatbot implementation, we analyze these interviews using a qualitative content analysis. Furthermore, we develop an implementation framework containing 101 questions and classify them under the four PACT elements of Benyon et al. (2005, 2014). The framework questions, and the PACT allocation are evaluated through interviews, a focus group discussion (FGD), and a case study demonstration. Our results are discussed following which, the implications and limitations of the research are highlighted. Our paper ends with conclusions and an outlook for further research.

2. Research design and methodology

2.1. Design science research

Our research aims to determine the user-oriented conditions that should be considered when developing and implementing a chatbot. According to Vaishnavi and Kuechler (2015, p. 20), *"Design science*
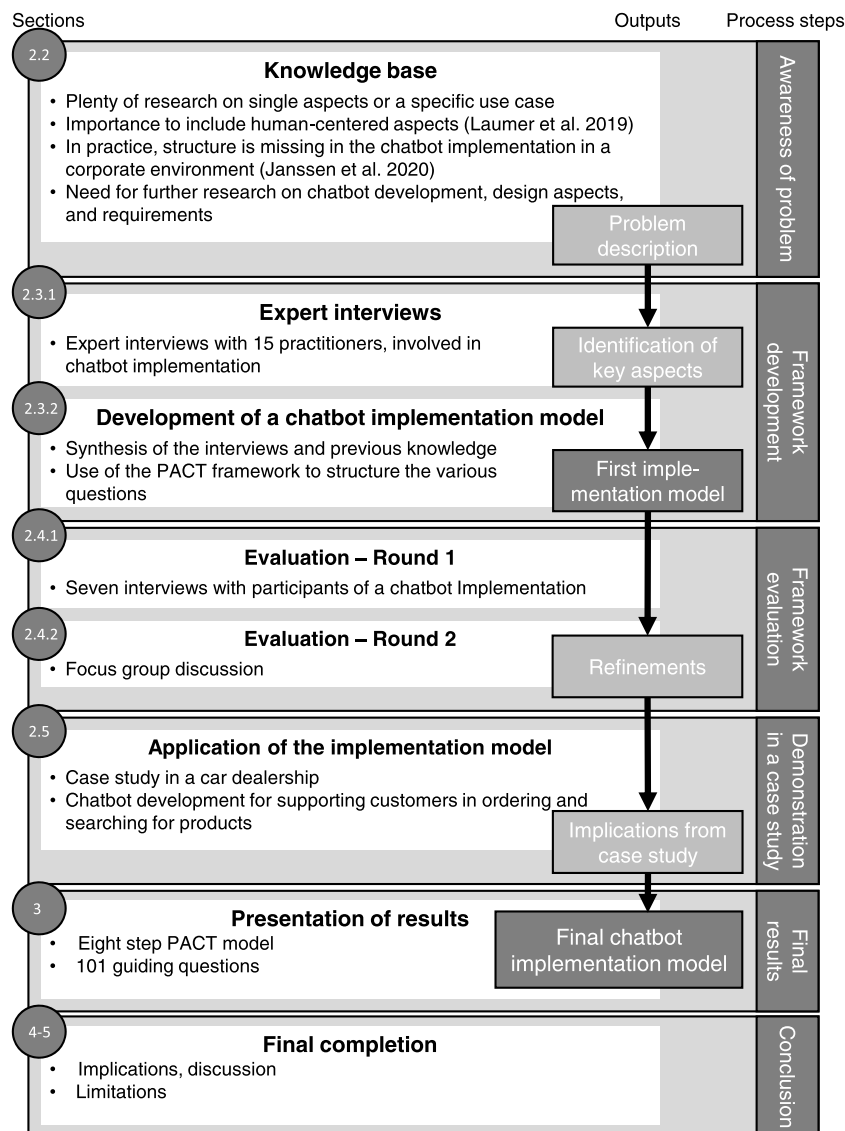

Fig. 1. Research Design based on Vaishnavi and Kuechler (2015).

knowledge can be in the form of artifacts — constructs, frameworks, architectures, design principles, methods, and/or instantiations — and design theories." Frameworks are defined as *"real or conceptual guides to serve as support or guide"* (Vaishnavi and Kuechler, 2015, p. 20). In line with this definition, we developed a framework as a human-centered HCI artifact in the context of computational design science research (Rai, 2017). In HCI, a distinction is made between the three design science research (DSR) modes. While the exterior mode focuses on the observational analysis of human-computer interactions and user behavior, and gestalt mode investigates a balance between IT systems and human behavior through a combination of technical and observational studies, in this study we focused on interior mode, which is a technical study of an IT system's design that focuses on human-computer interfaces (Adam et al., 2021). We structured our research project in several phases in accordance of the design science research approach of Vaishnavi and Kuechler (2015). The phases and research procedures are illustrated in Fig. 1.

2.2. Problem awareness based on the related literature

Chatbots are interactive application systems that can conduct a text-based conversation about a specific topic with a human while using NLP and ML techniques (Diederich et al., 2019a; Følstad et al., 2019a; Janssen et al., 2020; Meyer von Wolff et al., 2019b). Given their interactive capabilities, chatbots are used in countless private and commercial areas (e.g., education, health, daily life, collaborative work, and customer support). Each of which has widely varying requirements that are determined by their capabilities and tasks (Følstad et al., 2019a; Janssen et al., 2020), making their design all the more important.

Chatbots are seen as an ideal example of an HCI artefact (Adam et al., 2021), as the success of the technology lies directly in the interaction with the user. But, according to several researchers, chatbots differ decisively from other HCI systems in their interaction and intelligence capabilities (Maedche et al., 2019; Zierau et al., 2020). They use a natural language interface with visual elements, conversational design, emotional components (Araujo, 2018; Janssen et al., 2021a; Meyer von Wolff et al., 2022) and anthropomorphism features (Gnewuch et al., 2020). The ability of some chatbots to handoff to a human agent also distinguishes them from other IS systems (Janssen et al., 2021). Furthermore, the continuous training to understand users' input to provide appropriate information is unique to chatbots, compared to other existing frameworks (Meyer von Wolff et al., 2022).

Nevertheless, the selection of an applicable set of chatbot design elements is possible only after a clear delimitation of the context and application focus (e.g., business problem). Consequently, the determination of the preliminary socio-technical conditions for chatbot deployment is crucial for deriving design frameworks (Schuetzler et al., 2021). With regards to this, previous scientific research has largely concentrated on identifying chatbot design elements as architectural aspects of the human-chatbot interaction regarding distinct application domains (Zierau et al., 2020). These design elements constitute the *"distinctive technical, situational and knowledge features that frame the structure of chatbots and act as delimiting factors of the extent to which domain-specific chatbots can maintain a human-like interactive communication process with awareness for and understanding of the discussed topic"* (Janssen et al., 2020, p. 213). From thereon, researchers developed design principles and frameworks to facilitate and support the design and deployment of chatbots accordingly (Liu et al., 2017). For example, Di Prospero (2017) presented an architecture framework enhancing chatbot development, where chatbots are divided into four central components: user interface, application core, external services and sources, as well as personality processing, without considering domain-specific requirements. Gnewuch et al. (2017) used the social response theory to derive four design principles and 12 meta-requirements for steering the development of chatbots in customer service. On the other hand, Ma and Ho (2018) presented a "flow-based

chatbot framework" for deploying five human-chatbot dialogue patterns in three different application scenarios. Meanwhile, Meyer Von Wolff et al. (2022) developed a structured procedure model for chatbot projects in companies by focusing on design knowledge from scientific literature and their prior experiences. The procedure model includes 41 tasks within the four steps "planning," "developing," "testing," and "operating". Furthermore, Schuetzler et al. (2021, p. 3) derived three guiding implementation questions, (i.e., *"should we build a chatbot?"*, *"what technology should we use?"*, and *"how humanlike should the chatbot be?"*) based on their experiences in chatbot research and development. Regarding the first question, the article is distinct from the other literature which predominantly sees the chatbot technology as set in stone. Lewandowski et al. (2022) identified meta-requirements and design principles to manage the lifecycle of chatbots by designing four steps: "initiation," "development and training," "implementation," and "operation". Caldarini et al. (2022) gave an overview of chatbot implementation methods while distinguishing between rule-based and AI-based chatbots. Within their review, they focused on presenting appropriate algorithms and datasets that could be useful. In general, several articles can be found in the scientific literature, from which clues and design steps for chatbot implementation can be derived. However, Zierau et al. (2020), after analyzing 107 scientific papers related to chatbot design, determined, that task context and user-oriented requirements engineering have been studied little in the chatbot research so far, although these components have a fundamental impact on HCI (Li and Zhang, 2005; Zierau et al., 2020). To sum it up, several frameworks exist in terms of chatbot implementation presenting meta-requirements, general steps, and central questions, but a holistic framework with integrating HCI elements and focusing on the future user is still missing.

In Information Systems (IS), technology implementation has a long tradition, beginning with associating an implementation process as a bridge between designing and using a technology. However, today, the two research strings of technology acceptance and IS success are central elements (Lauterbach and Mueller, 2014). Therefore, the researchers developed a process framework to adopt technology in the organizational context by distinguishing between organizational and individual level as well as adaption and post-adaptive behavior. Beside general IS implementation models, numerous process models exist for specific application areas such as game development software engineering for which, Aleem et al. (2016) identified 21 elements ordered into the steps "pre-production phase," "production-phase," and "post-production phase". These models also exist for business intelligence systems, such as the lifecycle BI system by Gangadharan and Swami (2004). Even if the described stages are very similar in the basic aspects, the descriptions may not fit the chatbot's area of specialization. In the field of HCI, diverse user-oriented design frameworks and methods with different degrees of user involvement have been employed to inform the development of HCI artifacts using collaborative design approaches with users as active design partners (e.g., participatory design) or through design approaches with users as reactive informers (Salinas et al., 2020; Scaife, 1997; Wallisch et al., 2019). These include, for example, the "DIN EN ISO 9241-210:2011-01: Human-centered design for interactive systems" which presents a framework for developing an interactive system, the "Natural Conversation Framework for Conversational UX Design" focusing on conversational design features for establishing a natural conversation (Moore, 2018), and the PACT framework by Benyon et al. (2005, 2014) that presents four HCI elements (i.e., people, activities, context and technology) which should be considered within human-centered design. According to Benyon et al. (2005), *"People use technologies to undertake activities in contexts."* The quote outlines the dependencies of the four elements (people, activities, context, and technology), forming interactive systems design structured as a complex entity. In other words, PACT is an evaluation framework that assists organizations in capturing the requirements for designing interactive systems while focusing on people (Benyon, 2005; Liao et al., 2019;

Sarbazhosseini et al., 2019). To develop interactive technologies, it is essential to comprehend the diversity of the four elements (Benyon, 2005). Chatbots interact differently from other interactive systems (Maedche et al., 2019) which, in turn, is determined by the system, task, and context (besides the user) (Zierau et al., 2020). Due to this, PACT provides an integrative frame of reference to cover the fundamental technical and social HCI elements that must both be taken into account, guiding the design of expeditious user-centered chatbots (Seeger et al., 2021) Nonetheless, a holistic chatbot implementation framework that integrates all HCI elements while putting the user at the center of its focus is still missing. To address this gap, a 101 questions containing user-oriented chatbot implementation framework geared towards structuring the decision implementation process around the four fundamental HCI elements is hereby developed.

2.3. Chatbot implementation framework development

2.3.1. Expert interviews

To incorporate practical experience into the development of a user-oriented chatbot implementation framework, in-depth semi-structured expert interviews were conducted with 15 experts. The selection of the experts was conducted by a sampling process in which individuals were contacted via e-mail and through career-oriented social networking sites (LinkedIn and Xing). These experts comprised those who had already implemented chatbots or were implementing it, as announced in press releases. The company size and employee strength varied from fewer than 1,000 employees (n = 5) and up to ones with 650,000 employees (n = 10). We conducted in-depth semi-structured expert interviews with (i) IT directors, product owners, and IT project managers responsible for planning or monitoring the implementation of a chatbot within companies headquartered in Germany as well as (ii) IT executives external to chatbot development firms in Germany responsible for or directly involved in the development process of a chatbot, such as design engineering and prototyping. All experts had between two and five years of experience in the relevant areas. The chatbots that the experts developed included both the private user domain (e.g., TV show media chatbot (Exp07), chatbot for financial (Exp08) or telecommunication contract advices (Exp14)) and the business domain, e.g., product FAQ chatbot in automation industry (Exp02). Depending on the availability of the interviewed experts, qualitative interviews were conducted either face-to-face (n = 4) or via telephone (n = 11). An overview of the interviewed experts is provided in Table 1.

To guide the interview process of collecting qualitative information on the approach followed by various companies that undertook chatbot implementation, a guideline for semi-structured interviews was developed first (see A4 interview guide in Appendix). This offered the advantage that a predetermined spectrum of questions is asked but with a flexible sequence. The guideline questions were open-ended. When creating the questionnaire, the wording of the interview questions was adapted such that a discussion between the expert and interviewer was created (Bogner and Menz, 2009). In addition, main questions were created and assigned to these sub-questions (Bogner and Menz, 2009). This allowed a main question to be asked first and enabled the interviewee to answer it freely, while the sub-questions could be inserted if necessary. The interview guide was sent to the experts in advance making a preparation possible. The consent of the interviewees was taken before recording the interviews. Interview lengths varied from 25:38 minutes and 44:36 minutes. All the interviews were conducted in German and translated accordingly. They were transcribed verbatim and subsequently codified using the MAXQDA software for qualitative data analysis.

The interviews were conducted in two iterations. The first iteration took place from July to August 2019. In iteration 1, Exp01 to Exp08 were interviewed about their previous experiences with the introduction of chatbots. Based on the current literature and the results of the first interviews, a prototype of the PACT implementation framework was

Table 1
Expert Descriptions.

	ID	Position	Industry sector	Number of employees
Interview iteration 1	Exp01	IT director	Insurance provider	12,000
	Exp02	Product owner	Automation industry	4,000
	Exp03	IT project manager	Automotive industry	650,000
	Exp04	Chatbot developer	Chatbot developing startup/IT consultancy	5
	Exp05	IT project manager	Banking industry	50,000
	Exp06	Chatbot developer	Chatbot developing firm/IT consultancy	350,000
	Exp07	IT project manager	Media industry	16,000
	Exp08	Product owner	Financial consultancy	2,000
Interview iteration 2	Exp09	Product owner	Transportation industry	135,000
	Exp10	Product manager	Transportation industry	9,600
	Exp11	Head of sales/ Business development	Chatbot developing firm/IT consultancy	26
	Exp12	Consultant	Chatbot developing firm/IT consultancy	160
	Exp13	Product owner/ Consultant	Healthcare industry	4
	Exp14	Chief marketing officer	Telecommunications industry	65
	Exp15	Product manager	Healthcare industry	1,400
FGD	Exp16	IT project manager	Automation industry	4,000
	Exp17	Digital analytics consultant	Automation industry	4,000
	Exp18	Business driver/ Innovation project manager	Automation industry	4,000
	Exp19	IT consultant	Automation industry	4,000

Note: Exp = Expert

developed. The second iteration took place from February to March 2020. In the second iteration, Exp09–15 were interviewed with slightly expanded questions from the guide to gain more insights on the PACT chatbot implementation framework. While the first two question blocks (entrance questions and key questions) remained identical, the questionnaire for iteration 2 was expanded with evaluation questions (see A4 interview guide in Appendix). These evaluation questions were initially asked within the third block of interviews in iteration 2 to initially evaluate the implementation framework. The procedure and results of the evaluation section are described in Section 2.4.1.

2.3.2. Development of a chatbot implementation framework

To develop a chatbot implementation framework and identify relevant questions based on the expert opinions within the interviews, we coded the interview transcripts with the help of a qualitative content analysis approach. The coding scheme was developed based on iteration 1's transcripts which was subsequently applied in iteration 2's transcripts. Appendix A1 provides a detailed procedure description and an illustrative overview of the category formation. A detailed representation of the categories formed in each stage is provided in Appendix A2. The summarized results of the coding process can be seen in Appendix A3.

In line with the artefact to be developed within this article, we grouped the statements of the interviewees under eight sequential implementation steps: (I) preliminary considerations (i.e., identification of redundant processes along with potential information, computing, communication, and/or connectivity technologies that can help to optimize them), (II) use case determination (i.e., detection of potential use cases within the organization, identification of project stakeholders and development of a stakeholder engagement plan), (III) definition of

the chatbot's characteristics (i.e., determination of the intelligence, interaction, and technical features of the chatbot), (IV) dialogue tree construction (i.e., process mapping and digitalization of the relevant technical documents), (V) prototype development (i.e., development of a proof of concept and further enhancement of the dialogue tree through training and testing), (VI) acceptance testing (i.e., in-house and target group-specific acceptance testing), (VII) performance measurement (i. e., determination and monitoring of key performance indicators), and (VIII) post-implementation (i.e., chatbot revision). These form the basis for the final chatbot implementation framework. Two authors classified the identified issues under these eight steps, reformulating them as questions wherever necessary. Furthermore, these questions were structured within the PACT elements. After conducting iteration 1 as well as a literature analysis, 63 (people = 16; activity = 21; context = 11; technology = 17) questions were identified and assigned to each element. 36 (people = 8; activity = 5; context = 12; technology = 11) questions were further formulated on the basis of the opinions from the interviews of iteration 2. The final artefact in form of the implementation framework contains 101 questions classified into the identified eight steps and PACT elements. In addition, the literature was consulted for each step to support the interview statements with scientific evidence.

2.4. Framework evaluation

2.4.1. Expert interviews

According to the DSR procedure of Vaishnavi and Kuechler (2015), a central step is to evaluate the artefact. This was done by conducting seven expert interviews (interview iteration 2) and an FGD. The first evaluation of the developed PACT framework was conducted using an interview guide (see Appendix A4). This evaluation had the objective of evaluating the general structure of the framework and took place in the second iteration of the interviews with Exp09–15. First, a representative excerpt of the framework was sent to the experts (see Appendix A5). The experts' first impressions of the framework were positive. Exp11 and Exp15 described the general structure of the framework as *"comprehensible,"* Exp14 stated to be *"relatively consistent and constant [looking],"* while Exp12 said *"There are thought-provoking impulses in it, which one must definitely be taken along away."* Suggestions for improvement were also mentioned in the interview evaluation, which were subsequently implemented as described below. With regard to the development of the first prototype, the experts pointed out that it is important to rely primarily on a minimum solution and not waste too much time on a design that will not be accepted at the end. Instead, an iterative approach should be adopted, i.e., gather results after the first "go-live" and modify the chatbot accordingly (Exp11, Exp13, Exp14). According to Exp14, this is more of a cycle, a *"permanent build-measure-learn-build-measure-learn-build-measure-learn model,"* where you can *"jump back in [the] four [steps]."* Exp11 agreed with this and said that *"the measurement of added value, [...] is a permanent cycle"* and that *"the dialogue tree construction, [...] alternates with prototype, sharpening dialogs, testing, prototype, sharpening dialogs. This always goes in this [a] loop until you get to the point where you say, this is what we have minimal."*

Another area that several experts believe should be included in the framework is "the legal aspect. Not so relevant for many companies, but for us [it] may have become an issue in the meantime due to GDPR [General Data Protection Regulation] and Co [...]" (Exp15). This could then create the questions "Which data can we use? What do we learn about the user?" (Exp12). Based on the EU-GDPR, it is possible "to obtain a large amount of data but not to be allowed to do so" (Exp12). The latter aspect was added subsequently. From the discussions, it can be concluded that issues that are relevant for one company or use case are not important for other chatbot implementations.

2.4.2. Focus group discussion

After an initial and general evaluation of the framework using expert interviews, the revised framework was discussed in an FGD in April 2020 by following the requirements of Rosemann and Vessey (2008). The focus here was on the content analysis of the questions assigned under the eight steps (see Appendix A6). Therefore, a difference to the interviews is that the participants had the entire and already expanded framework available for evaluation and could take more time to familiarize themselves with it. Since Rosemann and Vessey (2008) mandated that the participants relevant to the research area must be selected, the FGD was conducted with five participants from an industrial company who are experienced in chatbot implementation. One participant (Exp02) had already taken part in iteration 1. The participants were divided into experts with chatbot implementation experience (Exp02, Exp16, and Exp17) and IT consultants (Exp18 and Exp19) with experience in the introduction of other IT tools. Since the chatbot implementation framework to be evaluated is especially aimed at people who want to introduce a chatbot for the first time, Exp18 and Exp19 can check how well the framework helps in starting with chatbots and chatbot implementations.

The duration of the FGD was 90 minutes. Four participants took part in the discussion on site, while one was connected via Skype. The FGD began with a presentation of the implementation framework, a delivery of the framework printed on a sheet of paper (see Appendix A5) and a worksheet containing a focus group questionnaire (see Appendix A6) to familiarize the participants with the research object (Rosemann and Vessey, 2008). Each participant was asked to answer questions regarding the division into eight steps as well as the listing of relevant questions within these steps. The participants were further asked on the potential areas for using the framework and on the application of the guide for individual chatbot implementation. The focus was placed on the comprehensibility, logic, and completeness of the steps. The analysis of the FGD was performed by summarizing all the data available in the form of field notes and a tape recording (Rosemann and Vessey, 2008). One of the recurrent discussion points concerned steps I to IV. These steps were initially not comprehensive to all participants. After the first considerations (step I "preliminary considerations"), the step II "use case" was determined. However, the question of what is counted for the use case determination and for the determination of the "chatbot characteristics" in step III was raised. Exp16 specified that the step II "use case" should answer the question *"What do I want to achieve with the chatbot?"* He would then be able to define the "characteristics" more precisely by asking *"What should the chatbot have and what not?"*. Step IV would then, according to all the experts, focus more on the technical implementation of a dialogue tree. Overall, the participants agreed that the first four steps should not be combined but considered separately. Furthermore, Exp02 believed that a loop in steps II, III, and IV would represent the real sequence of these steps. This loop was based on the participants' own experience that the attempt to construct a dialogue tree showed that what was planned as an area of application and was further elaborated in the third step could not be converted into a dialogue (Exp02). From this, it follows that after the fourth step, it may be necessary to go back to the third step and see how the properties of the chatbot in the application area can be changed in such a way that the chatbot can later be realized in dialogue construction. If no profitable change can be found within the application area, the area itself will be reconsidered (step II) (Exp02). From the expert's point of view, the development is not conducted sequentially from step I to prototype development; it may be necessary to go back one or two steps from step IV. With regard to the application possibilities and the added value of the framework, the experts assumed that the basic structure can make a positive contribution to chatbot introduction by helping as a guideline.

2.5. Case study demonstration

According to vom Brocke et al. (2020), a DSR artefact should be

applied in an appropriate environment. To test the applicability of the framework, a single case study was conducted in July and August 2020. Here, a chatbot prototype was developed for a car dealership. A chatbot developer who was not previously involved in the framework development process was introduced to the framework so that one could undertake the development in a structured way. In doing so, the user experience was focused on during the development. The chatbot was designed for supporting customers in ordering and searching for products. So far, the first six steps of the framework have been completed. After a productive use of the chatbot, the added value was determined so that the post implementation phase could be run through.

At the beginning, preliminary considerations took place regarding the opportunities and future prospects for the chatbot in the car dealership. The second implementation step dealt with the definition of the use case which should generate added value in the car dealership process for both the employees and the customers, given that the latter who either want to sell or buy a new car should also be supported along the process. Since the car dealership for which the chatbot was to be developed already conducted some marketing campaigns using Facebook Messenger, this channel was chosen as a platform for the chatbot. Based on this, the important properties and characteristics of the chatbot were determined in the third step. These properties and characteristics formed the basis for the potential conversation paths to be constructed. Here, for example, the decision was made to develop the chatbot in German, since Facebook analyses showed that almost all visitors of the fan page come from Germany. In a fourth step, it was decided to provide the user with answer buttons for the communication, since this allows the conversation to proceed in fixed frames with fewer chances for comprehensibility issues. Finally, the actual development of the prototype took place.

The user acceptance of the prototype was analyzed with a survey among 20 selected users. The survey contained both closed questions, which were answered with a 7-point scale, and open questions. It was voluntary and anonymous. Around two thirds (65%) of the respondents had no previous experience with chatbots. However, they did not rate the chatbot very differently either. There was a high level (around 95% each) of approval for the clarity, ease of use and the way in which the user was addressed. The naturalness of the chatbot was rated lower (83%). However, 25% conveyed that they would not use the chatbot again, which was justified in the free text fields by a lack of functionality. The managing director of the car dealership was also satisfied with the result and saw great potential in supporting customers. Further development of the chatbot is to be carried out primarily to increase customer satisfaction.

The chatbot developer found the use of the framework useful since it structured the design and development of the chatbot and prioritized the user experience. He outlined: "*The model served as a very good roadmap on the steps of the structured approach from preliminary considerations to development and evaluation. The adaptation of the PACT framework helped into understanding the use case on the related car dealership and to analyze each of the four elements for the use of an interaction with a chatbot.*" However, the developer added that not every technology question was applicable in this use case because the chatbot development tool (i.e. ManyChat) did not offer technical possibilities concerning interaction options every time. In sum, the chatbot developer pointed out: "[The model] *facilitated the overall implementation steps and served as an overview of which aspects impact the user experience.*"

3. Results and analysis

3.1. PACT-adapted chatbot implementation framework

To answer our RQ, our final PACT-adapted chatbot implementation framework contains eight sequential implementation steps that should be performed within chatbot development. These eight steps are depicted in Fig. 2. In the following sections, we describe these sequential implementation steps from different perspectives of the PACT framework. The final framework with the full list of 101 questions can be seen in Table 2 and is also available for download at https://bit.ly/Implementation_Framework. The entire list of questions can be inserted by practitioners and researchers within chatbot implementation processes. In the following, the four elements of the PACT framework will be outlined from the perspective of chatbot implementation to show which relevant user-related questions should be considered during the implementation. Thereby, the framework helps understand people's rationale for using a system, the related activities they want to perform with the system, the context, and the activities taking place. Through this, we get an overview of the functions of the technology (Adamu, 2019). Afterwards, every implementation step will be described. The detailed step description may help to get background information from research and to find starting points for deeper understanding. A detailed overview of the practical and theoretical foundations underlying the PACT-adapted implementation framework is provided in Appendix A7.

3.1.1. People

People differ physically in terms of appearance, weight, and height, as well as psychologically in terms of their personality, preferences, and cognitive abilities uttered in their needs, abilities, and mental frameworks (Benyon, 2005, 2014). This in turn implies a design for groups that are most heterogeneous. For this purpose, we identified 24 questions that should be considered from the "people" element's point of view to better classify the future users of a chatbot. Through this, the target groups as well as the ones which need to be addressed should be identified. To further comprehend the "people" element, identifying the goals, needs, and motivations that lead to the use of the technology can be helpful (Benyon, 2014; Johansson et al., 2015).
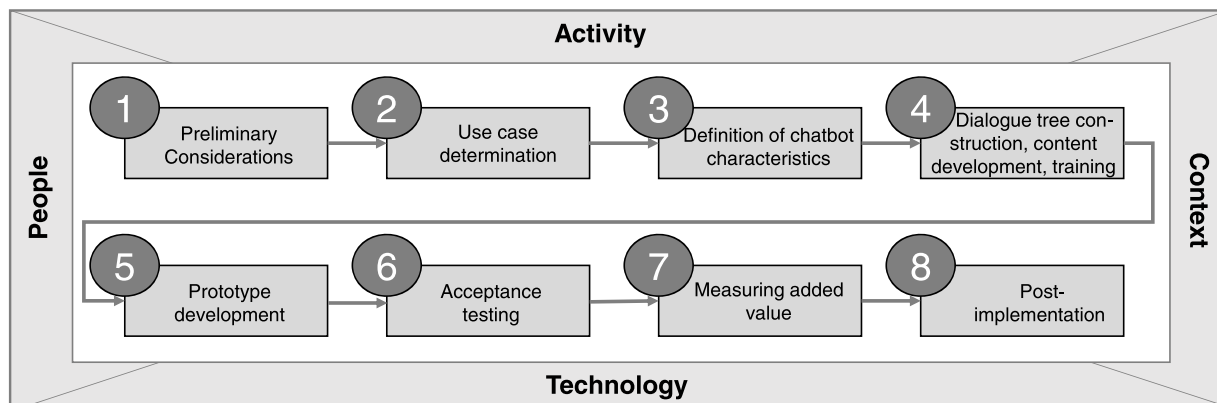

Fig. 2. Eight-step Chatbot Implementation Framework.

Table 2
Final Chatbot Implementation Framework with Guiding Questions.

People (P)	Activity (A)	Context (C)	Technology (T)
(I) Preliminary considerations			
(IP1) What are business processes in which (internal or external) users desire (need) to receive more (better) support to improve the customer/user value perception? **(IP2)** What type of communication technologies do users employ on a regular basis?	**(IA1)** What are the most repetitive/ monotonous activities from users' viewpoint? **(IA2)** What are characteristics of the previously identified activities? **(IA3)** What type of activities should be handled by a human employee to achieve the best outcome for the users?	**(IC1)** In which area or business context do users have most (special) difficulties/problems (e.g., customer service context)? **(IC2)** In which task fields can a chatbot add value to the company? **(IC3)** In which cases can a chatbot relieve employees? **(IC4)** Do employees need to be trained in handling chatbots?	**(IT1)** Considering the value proposition of the organization, is a chatbot the appropriate technology to improve the customer/user value perception (e.g., by overcoming previously identified difficulties/problems)? **(IT2)** Which technology concerns should be considered (i.e., regulations and ethical and security issues)?
(II) Use case determination			
(IIP1) Who are the end-users (i.e., target group)? **(IIP2)** How are the target groups segmented? **(IIP3)** What type of communication technologies do target groups use on a regular basis? **(IIP4)** What would be the end-users' main extrinsic motivation for using a chatbot? **(IIP5)** Which target group segments perceive added value in the potential use of a chatbot? **(IIP6)** What availability does the target groups look for (i.e., 24/7 service chatbot)?	**(IIA1)** What are collaborative requirements of the activity to be digitalized? **(IIA2)** What is the users' desired outcome? **(IIA3)** Do users need (desire) to receive additional human support to accomplish their activity? (Handover) **(IIA4)** Does the activity require historical user information to be accomplished?	**(IIC1)** On which communication platforms is the target group active? **(IIC2)** What is the application domain? **(IIC3)** Is the chatbot intended to be used for an internal or external context? **(IIC4)** Is customer data necessary to optimally support the user (i.e., login, 2-factor authentication)? **(IIC5)** Which device does the target group use (i.e., smartphone or tablet)? **(IIC6)** Should the method of communication (i.e., e-mail, web interface) also attract potential customers? **(IIC7)** Where are possible or existing touch points with customers?	**(IIT1)** How is the data situation (i.e., quality of the process/technical documentation)? **(IIT2)** Through which communication channels have users been reached so far? **(IIT3)** What type of platform integration is needed? **(IIT4)** How does a typical chatbot interface look like in the application domain? **(IIT5)** Which server fulfils the requirements (cloud or on-premises)? **(IIT6)** In-house development or outsourcing? **(IIT7)** Which provider fulfils the technical requirements?
(III) Definition of chatbot characteristics			
(IIIP1) How many users can be reached through the chatbot? **(IIIP2)** Self-evolution: What features should the chatbot have to produce the users' desired outcome? **(IIIP3)** To what degree is the behavior of using the chatbot self-motivated? **(IIIP4)** Do the users need a tutorial on how to use the chatbot? **(IIP5)** How can a chatbot measure user satisfaction? **(IIP6)** Is the user experience improved by integrating gimmicks?	**(IIIA1)** How do the users formulate their requests? **(IIIA2)** Is a chatbot-driven or user-driven dialogue preferred? **(IIIA3)** What type of objectives do the users attempt to meet by using the chatbot? **(IIIA4)** Is the intent to use the chatbot more goal-oriented or non-goal-oriented? **(IIIA5)** How did a typical conversation between a customer and an employee look like before the chatbot? **(IIIA6)** What should the chatbot be able to do? What should the chatbot be unable to do for now? (core function) **(IIIA7)** What activities are measurable after implementation?	**(IIIC1)** In what way (text/speech/ video) do users wish to communicate? **(IIIC2)** What type of context-awareness is needed by the chatbot? **(IIIC3)** How should the chatbot react if it cannot respond? **(IIIC4)** Is the explicit emotional context of the users handled properly? (i.e., stressed or frustrated users)	**(IIIT1)** Does the company have any already existing chat interfaces that can be adapted or should the company start from scratch? **(IIIT2)** To what extent is it desired for the chatbot to present human-like features (e.g., avatar, personality)? **(IIIT3)** Which interfaces to further knowledge bases are required to provide the information requested by the users? **(IIIT4)** How should the UI look from a users' viewpoint? **(IIIT5)** Are the users' desired chatbot features within the approved company budget? **(IIIT6)** Is the chatbot expected to have good speech recognition/NLP? **(IIIT7)** Does the chatbot need an interface for pictures? **(IIIT8)** Are any licenses/ permissions for access required? **(IIIT9)** Are there any data protection restrictions? **(IIIT10)** Does the chatbot need artificial intelligence?
(IV) Dialogue tree construction, content development, training			
(IVP1) In which language specifications do the users wish to communicate with? **(IVP2)** What type of characteristics should the chatbot's responses have from the user perspective (e.g., long/short answers)? **(IVP3)** Does the target group use multiple languages? Should the chatbot work with translation tools? **(IVP4)** Do answers include emojis, visualizations, and/or text? **(IVP5)** Will it be a B2B or B2C chatbot (technical or colloquial)?	**(IVA1)** Do the users prefer to use a pre-configured selection menu or would they prefer to formulate their own questions/requests? **(IVA2)** What do sample texts look like? **(IVA3)** What answers do users expect? **(IVA4)** Are there previous dialogue trees that can be used as a base? **(IVA5)** Do multiple formulations lead to the same result?	**(IVC1)** Does the chatbot match the intended context use and user's perceptions? (Exp15) **(IVC2)** How should the conversation start from the user's perspective for it to sound more human-like? **(IVC3)** What chatbot personality traits do the users expect? **(IVC4)** How should the chatbot react if it is asked something out of context (i.e., marriage proposal)?	**(IVT1)** Which data are usable? **(IVT2)** Do these data still need to be strongly classified? **(IVT3)** Are there enough data or should data be purchased? **(IVT4)** How much training does a chatbot need to obtain enough data without overloading?

Table 2 (*continued*)

People (P)	Activity (A)	Context (C)	Technology (T)
(V) Prototype development			
(VI) Acceptance testing			
(VIP1) Are the expectations of the end-users fulfilled in the test phase? **(VIP2)** Do users perceive the chatbot as a serious communicator?	**(VIA1)** What questions do the users have? **(VIA2)** Which questions can the chatbot not answer yet?		**(VIT1)** From an NLP perspective, does the chatbot interact as the users expected?
(VII) Measuring added value			
(VIIP1) What are the usage criteria for the users in the end?/What perceived value does the chatbot have to the user? **(VIIP2)** How often do the users leave the chatbot or stop writing and why?	**(VIIA1)** What is the average duration of a chat? **(VIIA2)** How profound is the response to the inquiry? **(VIIA3)** How often is the conversation surrendered to a human?	**(VIIC1)** Does the chatbot accomplish its primary task?	**(VIIT1)** How often is the chatbot used as an offer? **(VIIT2)** Does the chatbot do what it is supposed to do?
(VIII) Post-implementation			
(VIIIP1) Is the target group still reached through the chatbot?	**(VIIIA1)** Does the chatbot still represent the activity requested by the users? **(VIIIA2)** Are there any conversational flows that led to a failure because the flow was not modelled?	**(VIIIC1)** Does the context in which the chatbot is used still fit the chatbot? **(VIIIC2)** Does the chatbot fit the company? **(VIIIC3)** Is the chatbot affected by legal changes?	**(VIIIT1)** How can the answer given by a human to a question that the chatbot cannot solve be built into the chatbot? **(VIIIT2)** What newfound technologies can be included (e.g., updates)?

3.1.2. Activity

To learn more about the activities, related purpose, and target of the chatbot, we formulated 26 questions. The analysis of the activity element involves finding out the intended task the chatbot is used for, what is expressed in it, what type of objectives the users attempt to meet through using the chatbot, and what answers these users expect. The limits of context-specific chatbots are often given both technically and financially (Jain et al., 2018). To prevent a gap in user expectations, a chatbot must be familiar with all possible plot strands (Gnewuch et al., 2017; Tavanapour and Bittner, 2018).

3.1.3. Context

Activities are always embedded in a context representing the natural environment (i.e., where the user is located), indicating that these elements should be considered together (Benyon, 2005). The 24 context-related questions take the context of a chatbot's activity into account, allowing us to offer the correct solution depending on the situation and to minimize the mismatch between a chatbot's real context and the users' perception of the chatbot's context (Jain et al., 2018). Here, the context refers to the domain (e.g., daily life, work support, e-learning) where the chatbot serves specific business tasks or functions (Diederich et al., 2019a; Knote et al., 2018); it should be used for internal or external purposes (Meyer von Wolff et al. 2019b). This determines how users prefer to communicate (e.g., text/speech/video) and what their preferred language is.

3.1.4. Technology

This term refers to all the hardware and software components in interactive systems design that ideally work together to carry out the user's activities (Johansson et al., 2015). The purpose of technology is to support different people who carry out different activities in different contexts (Benyon, 2005). For this term, we identified 28 technology questions. We also incorporated questions from other elements, such as whether the chatbot understands the user's request correctly (Bittner et al., 2019; Knote et al., 2018). A chatbot can answer a message satisfactorily, independent of the message formulation, only if it is able to understand the message and analyze the content correctly (Gnewuch et al., 2017).

3.2. PACT-adapted guiding steps

3.2.1. I: Preliminary considerations

The first step in the process of implementing a chatbot to digitally redesign or integrate internal or external business processes is to identify specific business process activities **(IP1)** within an area or business context with potential for optimization from a service-oriented perspective. Exp05 pointed out the question **(IP1)**, *"In which processes can we better support our customers?"* This allows an innovative agenda, comprising the general problems experienced by an organization, to be set (Kee, 2017). The primary step of the chatbot implementation process is intricately connected with the digital business strategy defined at the organization level. Therefore, the deployment decisions must be aligned with it. A digital business strategy is a merger of the IT and business strategies that delimits the goal-oriented approach where new digital technologies are to be enforced according to the core value proposition of the organization (Bharadwaj et al., 2013). Ross et al. (2016) found that depending on the organization's strategic approach, they generally conceptualize their digital business strategy either in the form of a "digital customer engagement" or "digitized solutions" strategy. The strategic goal of the "digital customer engagement" perspective focuses on building customer loyalty and trust by reengineering the customer experience through the integration of seamless digital interactions, omnichannel capabilities, and customer-centered digital platforms. On the other hand, "digitized solutions" perspective is centered on the digital servitization of products and the reformulation of the value proposition of products and services through data and customer analytics (Ross et al., 2016; Sebastian et al., 2017). Although Sebastian et al. (2017) found that a certain degree of synergy exists between both strategies, the specific relevance of the potentially applicable information, computing, communication, and/or connectivity technologies will primarily depend on their capability to generate added value by effectively achieving the chosen digital business strategy. Consequently, the organizational adoption of a new technology should only be made when a technology constitutes a suitable solution for some of the identified issues of the organization's innovation agenda (Kee, 2017). Therefore, it is helpful to ask what kind of communication technologies the end user utilizes on a regular basis **(IP2)**.

To identify uses cases for implementing chatbots, Exp04 recommended to consider **IA1**, *"Are there any redundant processes?"* In doing so, it would be advisable to identify the characteristics and conditions of

these processes and activities **(IA2)**, as well as whether it would still be necessary that a human agent **(IA3)** performs these activities in the future.

Chatbot technology, depending on the case-specific system design and application domains, can contribute at the organizational level by accomplishing a digital business strategy oriented on a digital customer experience in terms of value creation for the company **(IC2)** and to solve difficulties in the business context previously experienced by the user **(IC1)**. Other strategic factors to be considered are both the level of chatbot-human integration, i.e., the level of workload that is intended to be shifted to the chatbot (Castro et al., 2018; Nili et al., 2019) **(IC3)**. As well as the need for trained personnel to provide assistance with complex requests beyond chatbot capabilities **(IC4),** along with skilled IT personnel to train the chatbot (Følstad and Brandtzæg, 2017; Nili et al., 2019).

On the technical level, it is critical to question whether chatbot technology is appropriate to improve value perception of the user, considering the organization's value proposition **(IT1)** (Brown and Brown, 2019; Kane et al., 2015; Ross et al., 2016). The preliminary identification of regulatory, ethical, and security issues related to a potential chatbot implementation **(IT2)** is also of major strategic importance because higher user acceptance can also be achieved by enforcing standards and regulations that ensure the safety of users (e.g., data security and privacy) and increase user trust in the chatbot (Laumer et al., 2019; Nili et al., 2019; Rodríguez Cardona et al., 2019).

3.2.2. II: Use case determination

Exp02 recommended asking the question, "Which use case do we already see?". As indicated by Brandtzaeg and Følstad (2018 p. 2), "we are currently witnessing a rush of businesses and organizations vying to be the first to deploy chatbots in their particular service domain. In this early phase of chatbot deployment, chatbot initiatives too often aim for poor use-cases, ignoring user needs and user experiences." On these grounds, the second step of the chatbot implementation process concentrates on the identification of the personal characteristics of the target users **(IIP1)**, such as their demographic segmentation **(IIP2)**, technological preferences and habits **(IIP3)**, and motivations for using a chatbot **(IIP4)**. In particular, it is important to identify the target groups that are most likely to benefit from chatbots **(IIP5)** and determine their expectations concerning service availability **(IIP6)**. The aforementioned factors are crucial to the practical success of chatbot implementation and should therefore be kept in the foreground during the selection of a chatbot use case (De Vries et al., 2018). The results from various empirical studies (e.g., Brandtzaeg and Følstad, 2017, 2018; Følstad and Brandtzaeg, 2020; Følstad and Skjuve, 2019b; Zamora, 2017) have shown that most people use chatbots based on motivational factors in the form of gratifications or social and psychological needs. Based on the "uses and gratifications theory" as a baseline for research, Brandtzaeg and Følstad (2017) identified productivity (i.e., ease, speed, convenience, and information) as the main motivational factor underlying chatbot use, followed by other factors, such as entertainment, social interaction, and curiosity **(IIA2)**. Similarly, subsequent studies have not only reasserted the overriding importance of productivity as a motivational factor for use but have also identified "effectiveness and efficiency" as the most important productivity aspects from a user's perspective (Brandtzaeg and Følstad, 2018; Følstad and Skjuve, 2019b). In this regard, Følstad and Brandtzaeg (2020) emphasized task-oriented chatbots. In this light, it is important to consider productivity aspects of the user activities to be supported. Based on our findings, such aspects may include the collaborative requirements of the activity to be digitalized **(IIA1)**, the users desired outcome of their interaction **(IIA2)**, the possible need to handover to a human agent to achieve the users' goals **(IIA3)** and whether the user tasks and activities require knowledge of historical user information **(IIA4).**

A wide range of tasks in diverse application domains can be performed or supported by chatbots (Følstad et al., 2019a). The term application domain embodies *"the primary application purpose for which the chatbot has been designed"* (Janssen et al., 2020 p. 8). A recent systematic analysis of 103 real-world chatbots identified e-customer service, e-commerce, e-learning, finance, daily life, and work and career support as the six prevailing chatbot application domains **(IIC2)**. From a user-oriented perspective, diverse scientific studies, such as those by Zamora (2017), Brandtzaeg and Følstad (2017), Piccolo et al. (2018), Rodríguez Cardona et al. (2019), and Følstad and Skjuve (2019a), offered insights into the debate over what are the most appropriate tasks to be assigned to a chatbot. Through an analysis of 131 user-centered scientific publications on chatbot design and an evaluation published between 1975 and 2018, Piccolo et al. (2018) identified that the previous scientific knowledge suggests that chatbots are not only considered by users to be mostly appropriate for the execution of simple, non-risk-related tasks, such as rapid provisioning of information and assistance, but also can be useful for handling topics that are personal or embarrassing to ask a human agent about. Similarly, Zamora (2017) indicated that *"common tasks, such as information seeking or other administrative needs, are objective and can be fulfilled by a chatbot. Some chatbots are also designed to attempt to build relationships between human and AI"* (Zamora, 2017, p. 254). In this context, customer service is one of the most widespread use cases for chatbots, particularly with regard to simple text-based chatbots using simple pattern-matching techniques (Janssen et al., 2020; Laumer et al., 2019). In addition to information retrieval and customer support use cases, through a user survey, Laumer et al. (2019) identified a total of seven categories (smart home control, goods and services shopping, car and navigation, music and entertainment, work and office, and others, such as support for the elderly) and 33 sub-categories of chatbot use cases that users perceived as having a particular utility, especially for speech-based chatbots using more advanced NLP techniques.

In addition to chatbot use cases for external application, their implementation within the enterprise context can lead to productivity and efficiency gains as they can help automate work and other organizational processes (Nawaz and Gomes 2019) and digitalize work environments (Frommert et al., 2018) **(IIC3)**. However, the scientific literature on chatbot use in enterprise contexts is still in its early stages (Stöckli et al., 2019). Most chatbot research at the organizational and industrial levels has tended to focus on business use and acceptance of chatbots for customer engagement (e.g., Castro et al., 2018; Johannsen et al., 2018; Nuruzzaman and Hussain, 2018; Rodríguez Cardona et al., 2019). To address this gap in the customer service context due to trust and privacy issues, utilizing a hybrid interaction design where chatbots can act as transfer agents between the users and human customer service agents has been found to be particularly advisable in complex use cases associated with risks (e.g., financial, psychological, and privacy) (Piccolo et al., 2018; Rodríguez Cardona et al., 2019). In this regard, previously collected customer data may be used for login functions or two-factor authentications within the dialogue to optimally support the user in a secure manner **(IIC4)**. However, regardless of the selected chatbot use case, five organizational capabilities have been identified by Tarafdar et al. (2019) as decisive for the implementation of AI-based innovations: i) data science competence (i.e., the possession of big data and extensive data analytics capabilities) , ii) business domain proficiency (i.e., comprehensive business process know-how), iii) enterprise architecture expertise (i.e., competence for executing technology-driven transformations), iv) an operational IT backbone (i. e., adequate levels of existing operational technology, high-quality data and IT staff), and v) digital inquisitiveness (i.e., ability to question and improve the outcomes of AI algorithms).

Considering the context in which the target audience is currently addressed, it is relevant to identify the communication platforms **(IIC1)**

and the devices **(IIC5)** that the actual target audiences prefer for this purpose and if these communication channels are also appropriate to address potential new customers **(IIC6)**. In this direction, it may be worthwhile to identify possible already existing touchpoints between the company and customers **(IIC7)** and to identify the platforms and technologies through which the company **(IIT2)** is currently reaching the target groups. Questions about which platform is necessary to integrate the chatbot into existing processes belonging to the use case **(IIT3)** and which servers and technologies provide the prerequisites for data storage and processing **(IIT5)** are also aimed in this direction. Based on a strategic assessment of the development level of the aforementioned organizational capabilities, the implementing organization should consider whether an in-house or an outsourced chatbot development would be more appropriate **(IIT6)** and, if outsourcing seems the appropriate choice, consider relevant technical requirements to be fulfilled by the provider **(IIT7)**. Further questions which should be answered are how a typical interface should look like **(IIT4)** and which kind of internal data can be used **(IIT1)**.

3.2.3. III: Definition of chatbot characteristics

The next step after defining a suitable use case is to determine the set of chatbot characteristics needed to ensure that the end-user can achieve their desired outcome **(IIIP2)**. As mentioned by Janssen et al. (2020), the design decisions related to the vital chatbot characteristics (e.g., socio-emotional skills, personality, and anthropomorphic features) must be aligned with the domain application, characteristics and preferences of the end users, and platform (e.g., social media, website, app, collaboration tools) where the chatbot is expected to be utilized. The extensive body of literature on chatbot design provides a diverse classification of the structures of various design elements (e.g., Braun and Matthes, 2019; Janssen et al., 2020; Knote et al., 2018) and chatbot development frameworks (e.g., Jain et al., 2018; Power et al., 2019; Suta et al., 2020; Wei et al., 2018) that provide potential chatbot implementers with archetypal patterns to support chatbot deployment. Knote et al. (2018) classified chatbots based on the functionality principles of self-evolution, anthropomorphism, multimodality, context-awareness, platform integration, and extensibility. According to the former classification, chatbots can employ self-learning, simple reflex, model-based, goal-based, or utility-based self-evolution mechanisms to achieve a specific task **(IIIP2)**. The empirical taxonomy paper of Janssen et al. (2020, p. 7) defines the afore-mentioned mechanisms of intelligence as "*the underlying cognitive system design delimiting the technical principles under which a chatbot communicates, processes information, and/or selects an action or response*" and provides a detailed description of their architecture in the supplementary material of the article. To navigate this wide range of design and characteristic options, it helps to first consider the extent to which the user is self-motivated to consult a chatbot **(IIIP3)** (Nguyen and Sidorova, 2018), when and how users are satisfied with the content and how this can be measured within the chatbot dialogue **(IIIP5)**, as well as to what extent tutorials explaining how to use the chatbot **(IIIP4)** are valuable. Also, it should be discussed how built-in gimmicks could make the conversation more interesting and variable **(IIIP6)**. To determine to what extent investments in complex features are worthwhile, it might also be advisable to consider how many users can be reached by the chatbot **(IIIP1)**.

Several studies, such as that by Rietz et al. (2019), provide additional insights into the impact of anthropomorphic and functional chatbot design features on the user acceptance of chatbots in enterprise collaboration contexts. According to Janssen et al. (2020, p. 8), the chatbot collaboration goal "*determines whether or not the chatbot helps the user to accomplish a common goal or task*" **(IIIA3), (IIIA4)**. As a rule, the dialogue design of the locus of control to perform a common goal or task can be internal (chatbot driven) of external (user driven) (Følstad et al., 2019a; Knote et al., 2018). Depending on their application purpose and

interaction design (i.e., user or chatbot-driven locus of control and long or short length interactions), Følstad et al. (2019a) proposed a typology for four forms of chatbots: i) customer support, ii) content curation, iii) personal assistance, and iv) coaching **(IIIA3)**. Here, the locus of control of customer support and personal assistant chatbots is commonly user-driven, while for content curation and coaching chatbots, it is mostly chatbot-driven **(IIIA2)**. From a technological point of view, the design of a chatbot-driven dialogue is more complex than a user-driven dialogue (Følstad et al., 2019a). Therefore, it is recommendable to analyze how a typical analog dialogue between the user and the employee is structured in the respective situation **(IIIA5)** and which core topics of this conversation should be taken over by the chatbot in the future **(IIIA6)**. This may also be a starting point before finding out how users formulate their requests within a human-to-chatbot dialogue **(IIA1)**. To be able to determine the added value in the future, it is advisable to consider from the start how the success of these activities is intended to be measured **(IIIA7)**.

To enhance chatbot's user interaction and engagement, the scientific literature provides theoretical context and practical procedures that help adopt suitable interactive design features, such as anthropomorphic elements by which a chatbot is able to simulate unique human and mental abilities, (e.g., consciousness, intentionality, and emotions) (Feine et al., 2019; Knote et al., 2018; Muresan and Pohl, 2019; Virkar et al., 2019; Seeger et al., 2018). For instance, Feine et al. (2019) provided a configuration system of 48 social cues for chatbots (e.g., degree of human-likeness, small talk behavior, gender, age, clothing, ethnicity, interaction order), 18 influencing factors, and 192 possible user reactions toward them **(IIIC4)** to guide the decision-making process of chatbot developers. However, it is important to consider that the platform selected for chatbot deployment (e.g., Chatfuel, ManyChat, Microsoft Bot Builder SDK, Dialogflow, IBM Watson Conversation) and the delivery channel platform (e.g., Facebook Messenger, Skype, Telegram, Slack, Microsoft Teams, Amazon Alexa, Cortana, Google Assistant) will delimit the characteristics of the chatbot architectural elements and ultimately their feasible set of functionalities (Kostelník et al., 2019; Sousa et al., 2019; Suta et al., 2020). Therefore, before selecting the appropriate chatbot technology and platform, it is important to ask whether the chatbot, e.g., should be able to process text, voice or even video input **(IIIC1)** and to what extent the chatbot also should consider the context in which the dialog takes place **(IIIC2)**. Since it is not possible to predict with comprehensive reliability what questions will be asked, it is also important to consider the form in which the chatbot should respond if it is unable to help **(IIIC3)**. Kostelník et al. (2019) distinguished between two types of chatbot platforms: one-purpose only chatbot platforms (also referred to as What You See Is What You Get "WYSIWYG" platforms or high-level chatbot platforms) and all-purpose chatbot platforms. The first type (e.g., Chatfuel, ManyChat) is broadly a cloud computing platform that applies keyword matching, pre-trained datasets, and pre-defined templates to deploy chatbots, while the second type (e.g., Dialogflow, IBM Watson Conversation, Microsoft Bot Builder SDK) is an AI platform that enables users to utilize additional capabilities (e.g., image recognition, NLP analysis) through the integration of application programing interfaces (APIs) and the use of pre-built client libraries in multiple programing languages, such as Python and JavaScript (Kostelník et al., 2019). Based on their characteristics, one-purpose only chatbot platforms are the most appropriate solution for use cases on a limited budget or for implementation teams with limited technical skills. Conversely, all-purpose chatbot platforms are appropriate for complex use cases that require a higher level of NLP maturity, API options, and additional chatbot capabilities **(IIIT5)**. In terms of the features offered by the different platforms, such as the ability to behave in a human-like manner **(IIIT2)**, the use of artificial intelligence **(IIIT10)**, or the ability to upload images **(IIIT7)**, differences can be observed among chatbot platforms (Knote

et al., 2018). Suta et al. (2020) identified the features (i.e., text messages, carousels, buttons, quick reply, web view, group chatbot, list, audio, video, GIF, image, and document/file) provided by the messaging platforms of Facebook, Skype, Slack, Telegram, Microsoft Teams, and Viber. The results of their research showed that Facebook, Telegram, and Skype were the messaging platforms that enable the integration of all the analyzed features into the chatbot architecture to a larger extent. In the context of selecting the appropriate platform, it should also be asked whether the company may already have an interface, such as a Facebook profile, which can be adapted and expanded **(IIIT1)**, whether data protection restrictions apply on this platform **(IIIT9)**, and whether licenses must be acquired to operate on the platform **(IIIT8)**. According to Jain et al. (2018), chatbot users prefer to interact with a user interface **(IIIT4)** that may offer various design configurations and features such as a summary of the main functionalities of the chatbot, a horizontally scrolling carousel to view lists of options, and auto-suggestion buttons. Further, the functionality of speech recognition within a dialogue is also needed for some chatbot use cases **(IIIT6)** (Diederich et al., 2019b; Erekata et al., 2020). The information used by the chatbot to retrieve the response can come from structured, semi-structured, or unstructured data sources **(IIIT3)** (Di Prospero et al., 2017; Knote et al., 2018; Suta et al., 2020).

3.2.4. IV: Dialogue tree construction, content development, training

In step IV, dialogue trees are constructed, the content is elaborated, and trainings are conducted. Dialogue training data can be applied to train an adaptive dialogue flow (Tavanapour and Bittner, 2018). As described in the previous steps, the focus should always remain on the user, and the linguistic properties and wording that the user uses **(IVP1)**, the need of having conversations in multiple languages **(IVP3)**, as well as the type of response characteristics **(IVP2)** or visualizations **(IVP4)** should be considered. The circumstance of whether the context is B2B or B2C can also have an influence on the type of dialogue **(IVP5)** (Janssen et al., 2021b). The application area can also determine whether users prefer to select from a preconfigured selection menu or whether they prefer to write their issues directly in a free text field **(IAV1)**. Ideally, communications from the company (e.g., emails or human-to-human chats) or dialogues from the industry sector are used to adopt dialogue trees **(IVA4)** and previously used phrases as well as to build up a suitable vocabulary **(IVA3)**. In this regard, Exp06 formulated the questions: *"Are the existing data usable? And do they still need to be strongly classified?"* The data already available can serve as a basis for the creation of sample texts **(IVA2)**, which allow verification that different formulations lead to the same result **(IVA5)**.

In addition, publicly accessible training dialogue datasets containing a collection of sample conversations labeled with the corresponding entities and intents can also be utilized (Tavanapour and Bittner, 2018). Therefore, it should be considered whether sufficient data is available and usable **(IVT1)** or should be purchased **(IVT3)** to ensure that different formulations of the classification lead to the same result and to what extent the chatbot must be trained to answer various questions **(IVT4)**. These external data can also be especially helpful when answering questions that do not fit the actual context of the chatbot. These must also be classified accordingly **(IVT2)**. The natural speech recognition unit constitutes the main element for understanding the user input within the conversational system, classifying the user's intention and extracting the intended and desired settings of that intention (Bashir et al., 2018). Many techniques have been used for text classification in recent years, such as convolutional neural networks (CNN) (Bashir et al., 2018). Bashir et al. (2018) also worked with neural networks that use numerical values to classify texts. According on this, Zschech et al. (2020) conducted a comprehensive technical investigation and evaluation of multiple word processing and classification process pipelines to create a system design artifact for selecting data mining

methods for text-based intelligent assistance systems. Loisel et al. (2009) described a procedure of data collection and processing to create a dialogue system by recording real conversations and analyzing the content of the dialogues by dividing them into sub-dialogues directly related to a task. Girol et al. (2008) used a classification procedure for user input that considered the complete course of the dialogue to select the system response. A language understanding module within a dialogue system comprises an intent classifier (classifies the user's intentions to guide the chatbot to the appropriate answer) and an entity extractor (extracts the main tags from the commands by assigning a label to each word in the sentence to identify its role) (Bashir et al., 2018).

Typically, the conversation starts with a greeting, which can be initiated by the chatbot or the user, e.g., by saying "Hello." However, the way this opening should be designed depends on the application area and target group, which is why the user perspective should be considered here. The dialogue ends with the chatbot saying, e.g., "Goodbye" (Tavanapour and Bittner, 2018). Depending on the greeting style, the chatbot can appear more human-like **(IVC2)** as well as convey certain personality traits **(IVC3)**. Replying to questions that are outside of a chatbot's actual area of use, such as a marriage proposal or a request to tell a joke, can also convey a certain personality and make the chatbot seem more human **(IVC4)**. To enable the user to assess what the chatbot can be used for, the chatbot introduces itself before describing the task and process (Tavanapour and Bittner, 2018) **(IVC1)**. Attention should also be paid to checking whether the chatbot communicates and advertises its functionalities to the user (Jain et al., 2018) because if the user is not aware of them, they will probably not be used.

3.2.5. V: Prototype development

In this step, the prototype is developed based on the decisions made in the previous steps. As mentioned in step III, chatbot engineers can use a variety of deployment platforms to design, program, and host a chatbot (Diederich et al., 2019b; Feine et al., 2019; Tavanapour and Bittner, 2018). The suitability of these platforms for prototyping a specific chatbot depends on various factors and requirements, such as the context, supported language, preferred hosting, and pricing model (Diederich et al., 2019b). If the chatbot is primarily based on rules that perform a simple pattern matching, a platform like ChatbotsBuilder would suffice (Diederich et al., 2019b; Feine et al., 2019). If the chatbot has to improve through self-learning while communicating with the user, Twyla would be more suitable (Diederich et al., 2019b). Platforms also differ according to the way developers make chatbots. There are platforms where chatbots are programed by writing code (e.g., wit.ai); other providers allow the modeling of user conversations using flowcharts (e.g., ManyChat and IBM Watson Assistant). In addition, the necessity of a preconfigured interface or an API, that allows the chatbot to access existing applications or web services during a conversation, such as a CRM system or a database, determines which chatbot platform provider is the most suitable (Diederich et al., 2019b; Meyer von Wolff et al., 2019a). A distinction can be made between different types of prototypes. Usually, functional chatbots are built in this step and evaluated in the following steps. However, some researchers have reported the development of a WoO (Wizard of Oz) as the first prototype (Bittner and Shoury, 2019; Sjöström et al., 2019; Tavanapour and Bittner, 2018). In this case, a chatbot interface is merely developed so that a respondent is assumed to communicate with an interactive system, although the reactions of the system are in reality generated by a human (Bittner and Shoury, 2019). Since this step focuses solely on prototype development of the requirements specified in previous steps through various questions, this step does not contain any questions of its own.

3.2.6. VI: Acceptance testing

Acceptance testing includes an evaluation and assessment that

considers future users. Thus, it is suitable to invite, e.g., between five (Feine et al., 2019) and 15 (Jain et al., 2018; Krisnawati et al., 2018) or 40 (Hobert, 2019) test users who will be asked to have a dialogue with the chatbot. The limited number of participants will help to provide quality feedback to evaluate the chatbot in terms of acceptance and satisfaction (Ghose and Barua, 2013; Krisnawati et al., 2018). Moreover, according to Exp03, an acceptance testing is needed to answer the question, *"What are the customers' expectations for testing and are the expectations met?"* This acceptance testing can be divided into two phases: an exploratory analysis and a task scenario analysis. In the exploratory phase, the test users should be asked to start a dialogue with the chatbot. Hereby, the participants should state their first general impression and overall opinion concerning the prototype (Hobert, 2019). Based on the first impressions, questions **VIA1** and **VIA2** can be answered. After introducing the test users to the context and purpose of the chatbot, providing them with a task scenario by defining concrete targets or achievements that they should find complete using a chatbot-user conversation (Hobert, 2019; Krisnawati et al., 2018) is helpful. This helps find out which phrases and formulations users enter to achieve a certain goal **(VIA1)** and if the chatbot can already answer the sentences in a satisfactory way **(VIA2)**. Any areas where the chatbot has room for improvement from a natural language processing perspective can also be identified through this activity within the acceptance testing step **(VIT1)** (Bittner et al., 2019; Knote et al., 2018). These phrases should then be used as training data so that the chatbot can respond more flexibly to similar utterances and questions (Tavanapour and Bittner, 2018).

Regarding the question **VIP1**, it should be verified whether the user's expectations of the chatbot are fulfilled. This is a fundamental question that depends on whether users see an added value in the consultation and whether they will decide to use the chatbot again in the future. It helps to give each test user a list of topics based on which, they ask the chatbot a specific number of questions. The answers from the chatbot are then classified by the test user into "satisfactory" and "unsatisfactory", representing how appropriate and accurate the chatbot responses are to the query asked by the user (Ghose and Barua, 2013). If a large number of test users exist, a quantitative 5-point Likert scale questionnaire can be used to test functional aspects, such as usefulness, form aspects, such as ease of use (Davis et al., 1989; Hobert, 2019). Jain et al. (2018) observed that users blamed themselves when chatbots did not perform their expected task or did not behave as expected, which was attributed to Norman's theory of "human error." To prevent frustration and a negative impression, this should be circumvented.

Ideally, a chatbot should have an apparent and consistent personality appropriate to its field of application, which may be expressed, for instance, in initial small talk ("Good morning, how are you?"), in appreciative farewells ("have a great day") or humorous replies (Jain et al., 2018). However, the perception of a chatbot's personality is highly dependent on its application area, which is why it is vital to consider whether the user perceives it as a serious conversational partner **(VIP2)**. In addition, the average character length can be analyzed to determine how users communicate and how long the responses of the chatbot should be (Jain et al., 2018). Based on the valuable feedback of the participants within the acceptance testing, the prototype should be revised before conducting another acceptance test (Hobert, 2019).

3.2.7. VII: Measuring added value

After a chatbot has been implemented and released, its performance should be measured by tracking the human–chatbot interactions (Przegalinska et al., 2019). To assess whether a chatbot is successful, evaluation metrics should be applied to quantify system performance (Krisnawati et al., 2018; Przegalinska et al., 2019). To do that, Exp06 noted that it is important to answer the question, *"In the end, what are the*

success criteria for the chatbot user?" In the scientific community, trusting a chatbot is mainly related to the users' perception of its knowledge and expertise (Przegalinska et al., 2019). From a user perspective, the target of the chatbot is to maximize user satisfaction (Krisnawati et al., 2018). To measure user satisfaction and the perceived value, user tests can be conducted, as described in step VI **(VIIP1)**. In addition, this is determined by performance measurement using metrics such as the bounce rate **(VIIP2)** and the reasons behind leaving the chat or the stop of writing. From an information gathering perspective, the system is evaluated by measuring, e.g., recall, and F-score (Krisnawati et al., 2018). In this regard, a central question formulated by Exp04 is *"How fast is the user's request answered?"* The quantitative evaluation of system performance can be done by dialogue-based metrics, such as the average conversation duration **(VIIA1)**. The number of turns is defined as the number of messages exchanged between the user and the chatbot within a dialogue (Jain et al., 2018). To determine how profound the responses to the inquiry are **(VIIA2)** and how effectively the chatbot engages with the user, the average number of turns necessary for each concept to be understood by the chatbot must be estimated (Jain et al., 2018; Krisnawati et al., 2018). Further indicators include the word error rate (WER), sentence error rate (SER), and task completion rate (TCR) (Glass et al., 2000; Jain et al., 2018; Krisnawati et al., 2018). Analyzing how often the chat was continued by a human employee due to the chatbot having reached its boundaries also helps to analyze how pronounced the added value of the chatbot is in operational use **(VIIA3)**.

Monitoring the content of real-world human-chatbot conversations helps obtain valuable contextual insights when communicating with the chatbot, such as why they visited the website and what they were looking for and does the chatbot accomplish its primary task **(VIIC1)**. This, in turn, can help the organization to update the chatbot and revise its marketing strategies and sales channels according to the user's needs. From the technological perspective, it is useful to ask how often the chatbot is used at all and preferred over other technologies offered **(VIIT1)** as well as whether the chatbot does what it promises with its functionalities **(VIIT2)** (Jain et al., 2018).

3.2.8. VIII: Post-implementation

The last step refers to the phase after the go-live. The crucial question here is whether the target group is still reached by the chatbot **(VIIIP1)**. Therefore, constantly checking whether the chatbot fulfills the functionalities and abilities expected by the user is crucial **(VIIIA1)** (Jain et al., 2018). These checks include questioning whether there are conversational flows that have been proven to be faulty or incomplete **(VIIIA2)**. This can also be an indication that the context in which the chatbot is being consulted may have changed **(VIIIC1)**. The expected functionalities can change constantly because users may expect a particular chatbot to give them the positive experience they experienced in other chatbot environments. Therefore, it is important to critically question, if the chatbot still fits the company **(VIIIC2)**. During the interview, Exp01 formulated the question, *"Does the context in which the chatbot is used still fit the chatbot?"* If not, the dialogues, expertise, and answers of service employees to customers' questions must be transferred to the chatbot to ensure its relevancy and efficiency.

For upcoming technologies, trends, and innovations (e.g., in AI), customer data processing should also be considered. Moreover, regulations and legislation on data protection which has evolved over the years, must also be considered **(VIIIC3)**. One such example is the General Data Protection Regulation (GDPR), which addresses the export of personal data outside the European Union (EU). To provide the user with the most personalized experience possible, many chatbots rely on the collection and processing of personal information, such as customer number or name. While this can be partially circumvented by login mechanisms on websites, it can be challenging when non-customer data

on public insurance sites, e.g., is relied upon (Koetter et al., 2019). The EU-GDPR also applies to chatbot applications, so the regulations must be fully complied with as soon as the personal data are collected and processed (Nuseibeh, 2018). In this context, it is crucial to communicate clear guidelines and agreements on data storage and use at the very beginning of the conversation and to obtain the consent of the chatbot users (Nuseibeh, 2018). Chatbot services should be capable of demonstrating that there are appropriate technical and administrative measures that tackle data breaches in the form of user data or conversation protocols (Nuseibeh, 2018). As described in step I, the purpose of the chatbot is ideally directly related to the digital business strategy. Since an organization's strategic approach can change over time (Kee, 2017), a chatbot's purpose should be regularly aligned with the changing business strategies. In this regard, it becomes important to also have new technologies and features in mind that can potentially be incorporated **(VIIIT2)** to e.g., expand the vocabulary and glossary of the chatbot **(VIIIT1)**.

4. Discussion

4.1. Contributions and implications for research and practice

We pursued a design science research-oriented HCI approach and developed a user-oriented framework that helps to implement chatbots using a set of 101 guiding questions. We focused on the interior mode of an IT system design (Adam et al., 2021) and presented a framework for developing efficient chatbots by considering aspects of the intended end-user, activities, context, and technology. With explicit focus on the four user-oriented PACT elements, including the context of the chatbot and the user (people), we attempt to close the research gap pertaining to task context and user characteristics, as highlighted first by Zierau et al. (2020). According to Zierau et al. (2020), the interaction between a chatbot and a user is formed by the characteristics of the system, user, task, and context but concluded that task context as well as user characteristics were hardly considered in previous literature, except as being a control variable, in which they see major research gaps that have a great impact on both chatbot design and user behavior (Zierau et al., 2020). The connection between the four PACT elements and their data-driven questions (Table 2) shows how many mutually dependent factors must be considered during the implementation process to mitigate the risk of implementation failure.

From considering the current research, it became clear that a general, user-oriented chatbot implementation framework which lists the entire decision-making process independent of the field of application, rarely exist in the scientific literature. This despite the fact that practically oriented introduction models are available in the field of HCI. An exception to this is the structured procedure model by Meyer Von Wolff et al. (2022) which sequentially showed and described the development of a chatbot through 41 tasks. However, besides the aspect on focusing specifically on the user, their context, and the activity as described previously, our model differs from previous research in that, we provide a loose question collection instead of providing fixed sequential tasks. By doing this, this list of questions allows flexibility and broad applicability for its users. The questions can be used as a reminder whether everything has been considered or, for example, can also be used in chatbot implementation workshops to serve as impulses and a basis for discussion. Moreover, we focused specifically on the user, their context, and the activity.

When considering implementation models from other domains in IS and HCI, it has been frequently noted that while the steps (in general) are present in the chatbot introduction, they are absent in the description. This is quite concerning as it can hinder chatbot development. One example is "DIN EN ISO 9241-210:2011-01: Human-centered design for

interactive systems", which presents a framework on how to develop highly usable human-centered systems and products. Although the basic characteristics and steps of this model are very close to those of our developed framework (e.g., "understand and specify the context of use" and "produce design solutions to meet user requirements" [DIN EN ISO 9241-210:2011-01]), our developed framework differs from other IS and HCI-based models as it was especially designed for chatbots. These chatbot specialties can be observed, for example, in step IV which is not relevant for other interactive systems as well as within individual questions (e.g., IIIC2; VIA2; VIT1). In addition, although the steps are intended to provide structure and orientation, the added value lies in the listing of questions within the steps for the development of text-based and domain-specific chatbots. Several questions focus on the interaction and intelligence capabilities which differentiate chatbots from other HCI technologies. This can be observed in the questions as to what extent the chatbot should show human-like features (IIIT2) (Knote et al., 2018; Exp12), if the answers should include emojis (IVP4) (Exp12; Exp14) as well as what personality traits do the users expect (IVC3), (Jain et al., 2018; Exp 11).

However, not all the questions presented are necessarily relevant for every chatbot development project and the involved stakeholders. Rather, the purpose is to provide a broad spectrum of potentially important questions to maintain an overview throughout the entire process. The framework can be seen as a bridge between science and practice, where both sides benefit from the extensive list. For research, we present "a big picture" (Vaishnavi and Kuechler, 2015, p. 205) about issues currently addressed in research and practice with identifying several gaps for future research.

For practice, we see our contribution by offering project stakeholders a certain independence with our guiding questions so that they can focus on the future user instead of looking at what the chatbot platform provider offers. It reveals many issues that are highly relevant in practice but have rarely been considered in academia. Even though most of the steps are mentioned by both researchers in literature and practitioners in interviews, it is noticeable that step I, in particular, is almost exclusively addressed by the practitioners. While the chatbot literature is more concerned with its development, e.g., (IIIA4), (Ghose and Barua, 2013) or specific design aspects, e.g., (IIIP2), (Virkar et al., 2019), the practitioners emphasize the need to take a step back and consider where a chatbot actually makes sense, e.g., (IA3, IT1). This highlights the need to first question whether a chatbot is the appropriate communication tool before starting with the development. A chatbot, no matter how well designed, is superfluous if the use case does not fit because of the user not utilizing it (Schuetzler et al., 2021). With the questions in step I, we provide a basis for researchers reconsidering the preliminary considerations to find clues for future research, e.g., comparison of different communication tools with chatbots. Furthermore, the interviews have shown that practitioners are often very focused on their own use cases and their environment. Here the summarized results of literature and practice within our framework ensure decisions are made based on scientific studies and draw attention to design considerations, which otherwise would not have been considered at all. Moreover, by providing the sources, it gives starting points for practical purposes, especially to read deeper into individual topics. From a research point of view, the use case determination is more descriptive which has not been in the research focus so far.

4.2. Limitations and research directions

Our theoretical and practical contributions as well as limitations give rise to 11 research directions (RD), which can be addressed by HCI and IS researchers in the future. To identify the relevant questions involved in chatbot design and deployment, our focus was on involving 19

experts. Even though we also considered the literature on chatbot implementation, a fully comprehensive and structured review was not conducted. When interviewing the experts, we reached a saturation point. Even though the goal was to get insights from as many deployment areas as possible, we might have missed out certain questions due to our focus on introducing chatbots with a holistic perspective. Besides this, future research could also focus on each of the eight steps identified in this work (RD1).

The collection gives future researchers the opportunity to identify thematic areas that receive broad attention in practice but are rarely addressed in research (RD2). They can use the eight identified steps as well as the questions as a basis to develop design principles (RD3). Another example is the post-implementation step. Through the interviews, we learned that post-implementation is just as essential as the implementation steps for a chatbot's long-term success. This is crucial to ensure that chatbots evolve according to the needs of their users. Nevertheless, the scientific literature often focuses on chatbot introduction, ignoring the post-implementation phases. Our research makes a first contribution to this, which can be expanded in the future. The results of the interviews reflect the subjective opinions of the interviewees and, therefore, may be self-biased. To generalize the individual experiences of the interviewees, we consolidated the results with the scientific literature. Future research could focus on assessing each step as well as its questions which could be represented by color coding (RD3). This would make it more visible and measurable for future research and practice. We have assigned the identified questions to the four PACT elements according to Benyon (2005, 2014). Even if these questions have different perspectives, as described in Section 3.1, it may be possible that certain questions can also be assigned to one of the other elements or other steps. For the sake of clarity, we have decided to assign the questions to the element that fits best, which we then had confirmed in the evaluation.

In this article, the focus was to find out which issues should be considered when implementing chatbots. Even though this analysis was carried out in a user-centric manner based on the four PACT elements, chatbot end-users were not surveyed as part of this study. The reason is that while they can give feedback on the performance of the chatbot, they will not be able to exactly explain how chatbot developers can make the necessary changes so as to ensure better performance. In the future, it may be useful to draw conclusions about the relevance and correctness of the questions based on the decisions made in the development process (RD4).

Although we have applied three-stage evaluation process, we can only partially generalize the success of using the framework within chatbot development. While all questions of the respective levels were discussed in the FGD, within the expert interviews, we concentrated on an excerpt due to a more conceptual focus in the evaluation interviews (see framework excerpt in Appendix A5). It is possible that the randomly selected questions may distort the overall impression of the framework. However, since the focus was on questions about the general framework (see interview guide, Appendix A4) and the questions in the framework were only for illustration purposes, this can be relativized. In this regard, the study's scope is limited as well; it only illustrates how a concrete application of the developed framework can look like in practice and how it can enhance the user acceptance towards chatbots. Future research needs to deeply investigate how the presented framework influences this acceptance. While the case study suggests a positive influence, a larger study could expound on its practical efficiency (RD5). Therefore, it would be useful in the future for different chatbot development teams to apply this framework in analyzing their practical application as done within the case study demonstration (see Section 2.5) (RD6). In this context, discovering and evaluating additional methods is necessary, such as workshops, exercises, questionnaires, or experiments, which could be used to apply the framework in organizational settings (RD7).

Our questions also provide a basis for the future development of critical success factors (Williams and Ramaprasad, 1996). Further research could focus on the question of how crucial the presented questions are for the success of an implementation (RD8). Here, certain platforms for the realization of a chatbot can make an adjustment necessary for part of the questions; for example, the choice of platform can determine the possible communication channels. To develop the PACT framework, we interviewed different stakeholders, such as chatbot developers, IT project managers, and product owners, who had already been involved in a chatbot implementation process. It became apparent that the participants had different perspectives depending on their area of responsibility. Further research can systematically examine the role of these stakeholders at different steps of the process and broaden our questions by specifically considering each stakeholder (RD9). This would allow us to formulate further stakeholder-related questions to obtain a holistic view of the implementation. The chatbot environment has evolved dramatically through developments in areas such as NLP and AI, and it will continue to evolve at least as rapidly in the future. Hence, the framework should be regularly reviewed and updated (RD10). Within this study, we focused on text-based and domain-specific chatbots. Future research could focus on how these questions and stages can be applied to the speech-based virtual assistant context (RD11).

5. Conclusions

To fill the research gap on the requirements and implementation of chatbots, we identified aspects which require consideration while developing a chatbot. Therefore, 15 experts in this field, who had already been involved in chatbot implementations, shared their expertise in semi-structured interviews. Contributing to the knowledge on chatbot implementation, we developed the user-oriented PACT implementation framework by Benyon (2005, 2014). Our framework comprises 101 questions for the development of a user-oriented chatbot implementation using the results from conducted interviews as well as from literature. We evaluated this framework in a three-step evaluation process by conducting interviews as well as an FGD. The findings from our research provide a comprehensive understanding of how the successful introduction of chatbots can take place. Our results can help practitioners to keep track of the relevant issues throughout the chatbot implementation process as well as guide academic researchers in gathering design knowledge as a basis for further research.

Declaration of Competing Interest

None.

Acknowledgements

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijhcs.2022.102921.

Appendix

A1. Illustrative Category System Formation Procedure

As described by Mayring (2015), qualitative content analysis is a systematic method used by researchers to generate a conceptual understanding by analyzing the content of the text components embedded in the data material derived from, for example, narrative or semi-structure interviews. The target of this qualitative analysis method is segmenting the text components of the primary data into analysis units (i.e., coding, context and evaluation units) as well as the allocation of these analysis units into categories (Mayring, 2015). The coding units are the smallest text components that may be evaluated and classified under a category. The systematic process used to analyze the primary data obtained through our semi-structured expert interviews can be divided into three feedback loops or stages where the data material is assigned to categories. To gain an initial understanding of the data material's content, we collected primary data through inductive analysis using extraction via the open coding of an initial set of 54 coding units related to more than 1000 segment quotes of the verbatim interview transcripts (Myers, 2020) resulting in the first stage. The open coding units were induced by words used by the interviewees and therefore reflect the substantive nature of the interviewees' statements (e.g., response speed, standard features, limits to technology, human-like performance). In the second stage, to increase the level of abstraction, the previously identified codes were sub-grouped into 18 second-order inductive categories (context units), e.g., environmental context conditions, organizational context conditions, strategical causal conditions, operational causal conditions, positive effects, and negative outcomes (Myers, 2020). In the third stage, to reduce the data material into essential content and achieve a deeper insight into the pattern regularities of the primary data, the sub-categories identified through inductive category formation were subsumed into the five main deductive categories (evaluation units): (i) context conditions, (ii) causal conditions, (iii) intervening conditions (e.g., variables limiting the causal conditions), (iv) routine or strategic actions and/or interactions, and (v) consequences (Corbin and Strauss, 2015). Within the text analysis, the categories developed through deduction were assigned to the text passages. After performing the category formation stages, the code list was extended to include a total of 77 codes. With the help of this category system, we have determined when a text passage can be assigned to each category. Appendix A1 provides an illustrative overview of the category formation. A detailed representation of the procedure and the categories formed in each stage is provided in Appendix A2. The summarized results of the coding process can be seen in Appendix A3.

| Exemplary interview statements	Stage 1: Coding units	State 2: Context units (subcategories)	Stage 3: Evaluation units (main categories)
"(…) There must be no delays in providing the information. The response must not take too long[1]. The speed must be approximately the same as the human speed, which should be seen as the lower limit[2]. Under no circumstances should it be slower[2]. The chatbot should be able to use the amount of information necessary to make qualitative judgements. It must be a stable channel and must not suddenly break off[3]. These are the same requirements I would have for a human being. A person must be competent, understand as quickly as possible what the customer wants, make adequate suggestions, be able to react to these suggestions, be able to provide the necessary information on them or find the knowledge from databases very quickly and not break off[4] (…)"	Response speed Standard features Limits to technology Human-like performance	Technology-related intervening conditions	Intervening conditions

A2. Category Formation Procedure

Stage 1: Coding units	State 2: Context units (subcategories)	Stage 3: Evaluation units (main categories)
Competitive pressure	Environmental context conditions	Context conditions
Data protection regulations	Organizational context conditions	
Diffusion of new technologies		
Available expert knowledge		
Budget constraints		
Customer-oriented improvements		
Top-level manager's attitude		
Business competition	Strategical causal conditions	Causal conditions
Digital transformation process	Operational causal conditions	
Potential benefits		
Technology acceptance		
Suitable application areas		
Use case driven design		
Intern chatbot user	People-related intervening conditions	Intervening conditions
Stakeholder support level	Activity-related intervening conditions	
Target group	Context-related intervening conditions	
User acceptance	Technology-related intervening conditions	
Deciding factors		
Economic viability		

(*continued*)

Technical feasibility		
Ethical implications		
Legal aspects		
Cost		
Economic viability		
Human-like performance		
Implementation boundary conditions (e.g., level of technical documentation)		
Implementation time		
Limits of technology		
Pace of technology change		
Prototype tools		
Response speed		
Standard features		
Transparency in HCI		
Employee/user participation	Preliminary considerations	Routine or strategic actions/interactions
HCI expert support	Use case determination	
IT security	Definition of chatbot characteristics	
Pace of technology change	Dialogue tree construction, content development, training	
Business case (cost-benefit analysis)		
Definition of application context	Prototype development	
Stakeholders	Acceptance testing	
Chatbot characteristics	Measuring added value	
Hybrid model	Post-implementation	
Platform determination		
Dialogue tree		
Flowchart generation		
Practical/performance test (proof of concept)		
Group-specific acceptance test		
In-house testing		
Measuring added value (KPIs)		
Chatbot maintenance		
Customer value added	Positive effects	Consequences
Socially acceptable alternative	Negative outcomes	
Chatbot failure		
Internal resistance		

A3. Chatbot Implementation Process Conditions
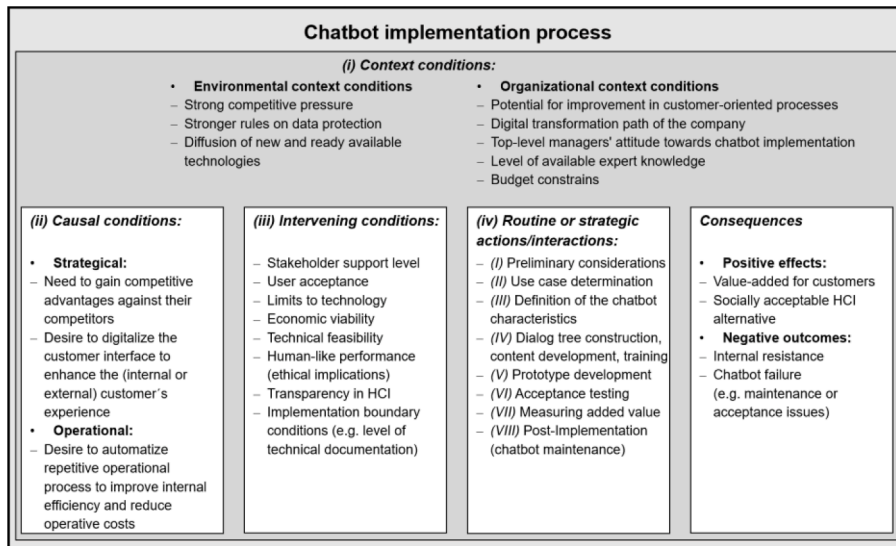
A4. Interview Guide
Entrance questions.

- Can you please present yourself (age, profession, industrial sector, number of employees)?

- How long have you been dealing with chatbots?

- What is your role (job title and responsibilities) in the chatbot implementation?

- What was your last chatbot implementation project? What type of chatbot is it? (name, purpose, target group, type of development, platform)

- Are your customers more likely to have a use case first by wanting a chatbot as a solution or did you want to install a chatbot as a new communication technology in your company and then searched for a use case?

Key questions.

- What do you think is the first question to consider before implementing a chatbot?

- How do the requirements analysis and definition of a chatbot work?
  - To what extent does the target group play a role? How are future users involved?
  - How do you define the tasks and purpose of the chatbot?
  - To what extent do you consider the environment in which the chatbot will be used?
  - How are the technical functionalities determined? Are there any choices? Are the target group and content determined first or are the definitions of the functions considered in isolation?
  - In which order are the different aspects in the decision-making process considered?
  - Which areas are also considered?
  - Which challenges arise?

- How do you proceed after the requirements analysis and definition?
  - Is there a previously defined procedure?

- How do you measure the success of the chatbot?
  - Which key performance indicators (KPIs) do you use?

- When do you think the introduction phase is finished?
  - In your opinion, what are the three biggest challenges facing the introduction of a chatbot?

  [The following questions were further asked in Interview Iteration 2:]

Evaluation.

- How do you evaluate the designed chatbot introduction model?

- How do you evaluate the outline of 8 steps?

- Are you missing steps or should something be summarized?

- What is your opinion about the formulation of questions instead of key points?

- How well can a guideline be applied to the introduction of individual chatbots?

- Are there any other areas that should be covered additionally?

A5 Model Excerpt

Steps	People	Activity	Context	Technology
I Preliminary considerations	**(IP2)** What type of communication technologies do the users use on a regular basis?	**(IA1)** What are the most repetitive/monotonous activities from a user viewpoint?	**(IC1)** In which area or business context do users present more (special) difficulties/problems? (e.g., customer service context)	
II Use case determination	**(IIP3)** Which target group segments perceive added value in the potential use of a chatbot?	**(IIA4)** Does the activity require historical user information to be accomplished?		**(IIT3)** What type of platform integration is needed?
III Definition of chatbot characteristics				**(IIIT4)** How should the user interface (UI) look from the user's viewpoint?
IV Dialogue tree construction, content development, training	**(IVA4)** Are there previous dialogues trees that can be used as a base?			**(IVT1)** Which data are usable?
V Prototype development				
VI Acceptance testing		**(VIA1)** What questions do users have?		
VII Measuring added value			**(VIIC1)** Does the context in which the chatbot is used still fit the chatbot?	**(VIIT1)** How often is the chatbot used as an offer?
VIII Post-implementation			**(VIIIC1)** Does the context in which the chatbot is used still fit the chatbot?	

A6. Focus Group Questionnaire

How do you assess the division into 8 steps?

Are there any steps missing?

Should any steps be deleted?

Should any steps be combined?

Are all relevant questions listed?

Are any relevant questions missing?

Should any questions be left out?

Can any questions be summarized?

Are the questions correctly classified?

Old position		New position	
Step	Element	Step	Element

What are the possible application areas?

How can the guide be applied to individual chatbot implementation?

_____

_____

Do you prefer the formulation of questions or key points?

_____	
Questions	Key points

What are some other issues and opportunities for improvement?

_____

_____

A7 Practical and Theoretical Foundations of the PACT Implementation Framework

Steps	PACT-Elements	PACT-Item	Practical / Theoretical Base
(I) Preliminary Considerations	People	(IP1)	Exp5; Exp11; Exp13; Exp15
		(IP2)	Exp13; Exp14
	Activity	(IA1)	Exp05; Exp06; Exp12
		(IA2)	Exp04; Exp14
		(IA3)	Exp12
	Context	(IC1)	Exp04; Exp11; Exp13; Exp15
		(IC2)	Exp11; Exp12; Exp14; Exp15
		(IC3)	Exp12
		(IC4)	Exp10; Følstad and Brandtzæg 2017; Nili et al. 2019
	Technology	(IT1)	Exp06; Exp11; Exp13; Exp14; Exp15; Brown and Brown 2019; Kane et al. 2015; Ross et al. 2016; Zumstein and Hundertmark 2018
		(IT2)	Laumer et al. 2019; Castro et al. 2018; Nili et al. 2019; Rodríguez Cardona et al. 2019
(II) Use case determination	People	(IIP1)	Exp02; Exp06; Exp12; Exp15
		(IIP2)	Exp06; Exp11; Exp15
		(IIP3)	Exp13; Exp14
		(IIP4)	Brandtzaeg and Følstad 2018
		(IIP5)	Exp06; Exp11; Exp15
		(IIP6)	Exp12
	Activity	(IIA1)	Exp11
		(IIA2)	Exp04; Exp11; Brandtzaeg and Følstad 2017; 2018; Følstad and Brandtzaeg 2020; Følstad and Skjuve 2019b; Zamora 2017
		(IIA3)	Exp14; Exp15
		(IIA4)	Exp04; Exp12; Følstad et al. 2019a
	Context	(IIC1)	Exp7; Exp11; Exp13; Exp15; Sjöström et al. 2019
		(IIC2)	Exp13; Zamora 2017; Brandtzaeg and Følstad 2017; Knote et al. 2018; Piccolo et al. 2018; Diederich et al. 2019a; Rodríguez Cardona et al. 2019; Følstad and Skjuve 2019a; Janssen et al. 2020
		(IIC3)	Exp13; Meyer von Wolff et al. 2019b
		(IIC4)	Exp11; Exp12; Exp15
		(IIC5)	Exp11; Exp15
		(IIC6)	Exp12; Exp15
		(IIC7)	Exp13; Exp14
	Technology	(IIT1)	Exp4; Exp6; Exp13
		(IIT2)	Exp06; Exp11
		(IIT3)	Exp13; Exp15; Diederich et al. 2019b
		(IIT4)	Exp6; Exp14
		(IIT5)	Exp9; Exp15
		(IIT6)	Exp09
		(IIT7)	Exp15
(III) Definition of Chatbot Characteristics	People	(IIIP1)	Exp11
		(IIIP2)	Exp11; Knote et al. 2018
		(IIIP3)	Exp14; Nguyen and Sidorova 2018
		(IIIP4)	Exp14
		(IIIP5)	Exp10; Exp13
		(IIIP6)	Exp11
	Activity	(IIIA1)	Exp06; Exp13; Exp14
		(IIIA2)	Exp14; Følstad et al. 2019a
		(IIIA3)	Exp14; Exp10; Følstad et al. 2019a; Tavanapour and Bittner 2018
		(IIIA4)	Exp14; Exp10; Bittner et al. 2019; Jain et al. 2018; Janssen et al. 2020
		(IIIA5)	Exp14
		(IIIA6)	Exp11; Exp15
		(IIIA7)	Exp10; Exp13
	Context	(IIIC1)	Exp14

(*continued*)

Steps	PACT-Elements	PACT-Item	Practical / Theoretical Base
		(IIIC2)	Exp15
		(IIIC3)	Exp14
		(IIIC4)	Exp09
	Technology	(IIIT1)	Exp6; Exp13
		(IIIT2)	Exp12; Knote et al. 2018
		(IIIT3)	Exp11; Exp13; Exp14; Di Prospero et al. 2017; Knote et al. 2018; Suta et al. 2020
		(IIIT4)	Exp14; Diederich et al. 2019b; Erekata et al. 2020
		(IIIT5)	Exp04; Exp06; Exp11; Exp13; Exp15
		(IIIT6)	Exp11
		(IIIT7)	Exp12
		(IIIT8)	Exp13
		(IIIT9)	Exp9; Exp10; Exp11; Exp12; Exp14; Exp15
		(IIIT10)	Exp12; Exp15
(IV) Dialogue Tree Construction, Content Development, Training	People	(IVP1)	Exp11; Tavanapour and Bittner 2018
		(IVP2)	Exp14; Feine et al. 2019; Tavanapour and Bittner 2018
		(IVP3)	Exp10; Exp11; Exp13
		(IVP4)	Exp06; Exp12; Exp14
		(IVP5)	Exp13
	Activity	(IVA1)	Exp4; Exp14; Exp15
		(IVA2)	Exp12
		(IVA3)	Exp06; Exp14
		(IVA4)	Exp06
		(IVA5)	Exp13
	Context	(IVC1)	Exp15; Jain et al. 2018
		(IVC2)	Feine et al. 2019
		(IVC3)	Exp11; Jain et al. 2018
		(IVC4)	Exp12
	Technology	(IVT1)	Exp06; Exp11; Exp12; Exp14; Exp15
		(IVT2)	Exp06; Exp11; Exp14
		(IVT3)	Exp13; Exp15
		(IVT4)	Exp13; Exp15
(VI) Acceptance Testing	People	(VIP1)	Exp03; Exp11; Exp14
		(VIP2)	Exp11
	Activity	(VIA1)	Exp05; Exp12
		(VIA2)	Exp04; Exp11; Exp15
	Technology	(VIT1)	Exp11; Exp14; Bittner et al. 2019; Knote et al. 2018
(VII) Measuring Added Value	People	(VIIP1)	Exp6; Exp11; Exp15
		(VIIP2)	Exp6; Exp11; Exp13
	Activity	(VIIA1)	Exp03; Exp13; Jain et al. 2018
		(VIIA2)	Exp04; Exp15
	Context	(VIIC1)	Exp13; Exp14; Jain et al. 2018
	Technology	(VIIT1)	Exp06; Exp11; Exp12; Exp13; Exp15; Przegalinska et al. 2019
		(VIIT2)	Exp12; Exp13; Jain et al. 2018
(VIII) Post-implementation	People	(VIIIP1)	Exp15
	Activity	(VIIIA1)	Exp15; Jain et al. 2018
		(VIIIA2)	Exp12; Jain et al. 2018
	Context	(VIIIC1)	Exp1; Exp13
		(VIIIC2)	Exp11
		(VIIIC3)	Exp12
	Technology	(VIIIT1)	Exp04; Exp13; Exp14
		(VIIIT2)	Exp12

Note: Exp = Expert

References

Adam, M.T., Gregor, S., Hevner, A., Morana, S., 2021. Design science research modes in human-computer interaction projects. AIS Trans. Hum.-Comput. Interact. 13 (1), 1–11.

Adamu, M.S., 2019. Designing and evaluating learning technology: an African dilemma and approach. In: Proceedings of the International Conference on Computer Supported Education, pp. 184–191.

Aleem, S., Capretz, L.F., Ahmed, F., 2016. Game development software engineering process life cycle: a systematic review. J. Softw. Eng. Res. Dev. 41, 1–30.

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. Comput. Hum. Behav. 85, 183–189.

Bashir, A.M., Hassan, A., Rosman, B., Duma, D., Ahmed, M., 2018. Implementation of a neural natural language understanding. Procedia Comput. Sci. 142, 222–229. https://doi.org/10.1016/j.procs.2018.10.479.

Benner, D., Elshan, E., Schöbel, S., Janson, A., 2021. What do you mean? A review on recovery strategies to overcome conversational breakdowns of conversational agents. In: International Conference on Information Systems.

Benyon, D., Turner, P., Turner, S., 2005. Designing Interactive Systems: People, Activities, Contexts, Technologies. Addison-Wesley, New-York.

Benyon, D., 2014. Designing Interactive Systems: A Comprehensive Guide to HCI, UX and Interaction Design. Pearson, Harlow, England.

Bharadwaj, A., El Sawy, O.A., Pavlou, P.A., Venkatraman, N., 2013. Digital business strategy: toward a next generation of insights. MIS Q. 37 (2), 471–482.

Bittner, E., Oeste-Reiß, S., Leimeister, J.M., 2019. Where is the bot in our team? Toward a taxonomy of design option combinations for conversational agents in collaborative work. In: Proceedings of the Hawaii International Conference on System Sciences, pp. 284–293.

Bittner, E., Shoury, O., 2019. Designing automated facilitation for design thinking: A chatbot for supporting teams in the empathy map method. In: Proceedings of the Hawaii International Conference on System Sciences.

Bogner, A., Menz, W., 2009. *The Theory-Generating Expert Interview: Epistemological Interest, Forms of Knowledge, Interaction.* Interviewing Experts. Palgrave Macmillan, London, pp. 43–80.

Brandtzaeg, P.B., Følstad, A., 2017. Why people use chatbots. In: International Conference on Internet Science, DOI. Springer, Cham, pp. 377–392. https://doi.org/10.1007/978-3-319-70284-1_30.

Brandtzaeg, P.B., Følstad, A., 2018. Chatbots: changing user needs and motivations. Interactions 25 (5), 38–43. https://doi.org/10.1145/3236669.

Braun, D., Matthes, F., 2019. Towards a framework for classifying chatbots. In: Proceedings of the International Conference on Enterprise Information Systems, pp. 484–489.

Brown, N., Brown, I., 2019. From digital business strategy to digital transformation-how: a systematic literature review. In: Proceedings of the South African Institute of Computer Scientists and Information Technologists.

Caldarini, G., Jaf, S., McGarry, K., 2022. A literature survey of recent advances in chatbots. Information 13 (1), 1–22. https://doi.org/10.3390/info13010041.

Castro, F., Tedesco, F., Alves, H., Quintino, J.P., Steffen, J., Oliveira, F., Soares, R., Santos, A.L.M., da Silva, F.Q.B., 2018. Developing a corporate chatbot for a customer engagement program: a roadmap. In: Proceedings of the International Conference on Intelligent Computing, pp. 400–412. https://doi.org/10.1007/978-3-319-95930-6_37.

Corbin, J., Strauss, A., 2015. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Sage Publications, Thousand Oaks, CA.

Davis, F.D., Bagozzi, R.P., Warshaw, P.R., 1989. User acceptance of computer technology: a comparison of two theoretical models. Manag. Sci. 35 (8), 982–1003. https://doi.org/10.1287/mnsc.35.8.982.

De Vries, H., Tummers, L., Bekkers, V., 2018. The diffusion and adoption of public sector innovations: a meta-synthesis of the literature. Perspect. Public Manag. Governance 1 (3), 159–176.

DIN EN ISO 9241-210:2011-01, 2011. Ergonomics of Human-System Interaction – Part 210: Human-Centred Design for Interactive Systems (ISO 9241-210:2010).

Diederich, S., Brendel, A.B., Kolbe, L.M., 2019a. On conversational agents in information systems research: analyzing the past to guide future work. In: Proceedings of the Wirtschaftsinformatik, 2019, pp. 1550–1565.

Diederich, S., Brendel, A.B., Kolbe, L.M., 2019b. Towards a taxonomy of platforms for conversational agent design. In: Proceedings of the Wirtschaftsinformatik, 2019, pp. 1100–1114.

Di Prospero, A., Norouzi, N., Fokaefs, M., Litoiu, M., 2017. Chatbots as assistants: an architectural framework. In: Proceedings of the Annual International Conference on Computer Science and Software Engineering, pp. 76–86.

Erekata, O., Azeta, A., Misra, S., Odusami, M., Ahuja, R., 2020. Development of a text and speech enabled conversational agent for students' activities planning using dialog flow. In: Proceedings of the International Conference on Futuristic Trends in Networks and Computing Technologies.

Feine, J., Morana, S., Maedche, A., 2019. Designing a chatbot social cue configuration system. In: Proceedings of the International Conference on Information Systems.

Følstad, A., Brandtzæg, P.B., 2017. Chatbots and the new world of HCI. Interactions 24 (4), 38–42. https://doi.org/10.1145/3085558.

Følstad, A., Skjuve, M., Brandtzæg, P.B., 2019a. Different chatbots for different purposes: towards a typology of chatbots to understand interaction design. In: Proceedings of the International Conference on Internet Science, 11551. LNCS, pp. 145–156.

Følstad, A., Skjuve, M., 2019b. Chatbots for customer service: user experience and motivation. In: Proceedings of the International Conference on Conversational User Interfaces, p. 9. https://doi.org/10.1145/3342775.3342784.

Følstad, A., Brandtzaeg, P.B., 2020. Users' experiences with chatbots: findings from a questionnaire study. Q. User Exper. 5, 1–14. https://doi.org/10.1007/s41233-020-00033-2.

Frommert, C., Häfner, A., Friedrich, J., Zinke, C., 2018. Using chatbots to assist communication in collaborative networks. In: Proceedings of the Working Conference on Virtual Enterprises, pp. 257–265.

Gangadharan, G.R., Swami, S.N., 2004. Business intelligence systems: design and implementation strategies. In: Proceedings of the International Conference on Information Technology Interfaces, pp. 139–144.

Ghose, S., Barua, J.J., 2013. Toward the implementation of a topic specific dialogue based natural language chatbot as an undergraduate advisor. In: Proceedings of the International Conference on Informatics, Electronics and Vision. IEEE.

Girol, D., Hurtado, L.F., Segarra, E., Sanchis, E., 2008. A statistical approach to spoken dialog systems design and evaluation. Speech Commun. 50 (8-9), 666–682. https://doi.org/10.1016/j.specom.2008.04.001.

Glass, J., Polifroni, J., Seneff, S., Zue, V., 2000. Data collection and performance evaluation of spoken dialogue systems: the MIT experience. In: Proceedings of the International Conference on Spoken Language Processing, pp. 1–4.

Gnewuch, U., Morana, S., Maedche, A., 2017. Towards designing cooperative and social conversational agents for customer service. In: Proceedings of the International Conference on Information Systems.

Gnewuch, U., Yu, M., Maedche, A., 2020. The effect of perceived similarity in dominance on customer self-disclosure to chatbots in conversational commerce. In: Proceedings of the European Conference on Information Systems.

Hobert, S., 2019. Say hello to 'coding tutor'! Design and evaluation of a chatbot-based learning system supporting students to learn to program. In: Proceedings of the International Conference on Information Systems.

Jain, M., Kumar, P., Kota, R., Patel, S.N., 2018. Evaluating and informing the design of chatbots. In: Proceedings of the Designing Interactive Systems Conference, pp. 895–906. https://doi.org/10.1145/3196709.3196735.

Janssen, A., Rodríguez Cardona, D., Breitner, M.H., 2020. Virtual assistance in any context – a taxonomy of design elements for domain-specific chatbots. Bus. Inf. Syst. Eng. 62 (3), 211–225. https://doi.org/10.1007/s12599-020-00644-1.

Janssen, A., Grützner, L., Breitner, M.H., 2021a. Why do chatbots fail? A critical success factors analysis. In: Proceedings of the International Conference on Information Systems.

Janssen, A., Rodríguez Cardona, D., Breitner, M.H., 2021b. More than FAQ! Chatbot taxonomy for business-to-business customer services. In: Chatbot Research and Design. CONVERSATIONS 2020, 12604. LNCS, pp. 175–189. https://doi.org/10.1007/978-3-030-68288-0_12.

Johansson, B., Alkan, D., Carlsson, R., 2015. Self-service BI does it change the rule of the game for BI systems designers. In: Proceedings of the Business Informatics Research Workshops, pp. 48–61.

Johannsen, F., Leist, S., Konadl, D., Basche, M., de Hesselle, B., 2018. Comparison of commercial chatbot solutions for supporting customer interaction. In: Proceedings of the European Conference on Information Systems, pp. 23–28.

Kane, G.C., Palmer, D., Nguyen Phillips, A., Kiron, D., Buckley, N., 2015. Strategy, Not Technology, Drives Digital Transformation, 14. MIT Sloan Management Review and Deloitte University Press, pp. 1–25.

Kee, K.K., 2017. Adoption and diffusion. In: Scott, C., Lewis, L (Eds.), International Encyclopedia of Organizational Communication. Wiley-Blackwell, Hoboken, New York, pp. 1–14.

Knote, R., Janson, A., Eigenbrod, L., Söllner, M., 2018. The what and how of smart personal assistants: principles and application domains for IS research. In: Proceedings of the Multikonferenz Wirtschaftsinformatik, pp. 1083–1094.

Koetter, F., Blohm, M., Drawehn, J., Kochanowski, M., Goetzer, J., Graziotin, D., Wagner, S., 2019. Conversational agents for insurance companies: from theory to practice. In: Proceedings of the International Conference on Agents and Artificial Intelligence. Springer, Cham, pp. 338–362. https://doi.org/10.1007/978-3-030-37494-5_17.

Kostelník, P., Pisařovic, I., Muroň, M., Dařena, F., Procházka, D., 2019. Chatbots for enterprises: outlook. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis 67 (6), 1541–1550. https://doi.org/10.11118/actaun201967061541.

Krisnawati, L.D., Butar-Butar, B.E., Virginia, G., 2018. Prototyping a chatbot for student supervision in a pre-registration process. CommIT (Communication and Information Technology) J. 12 (2), 87–96.

Laumer, S., Gubler, F.T., Racheva, A., Maier, C., 2019. Use cases for conversational agents: an interview-based study. In: Proceedings of American Conference on Information Systems.

Lauterbach, J., Mueller, B., 2014. Adopt, adapt, enact or use?. In: Working Conference on Information Systems and Organizations.

Lewandowski, T., Heuer, M., Vogel, P., Böhmann, T., 2022. Design knowledge for the lifecycle management of conversational agents. In: Proceedings of the Wirtschaftsinformatik, 2022.

Liao, Y.W., Huang, Y.M., Huang, Y.M., Su, Z.Y., Wei, C.W., 2019. An empirical evaluation of online cooperative programming platforms based on the PACT framework and technology readiness. J. Internet Technol. 20 (2), 345–352. https://doi.org/10.3966/160792642019032002003.

Li, N.L., Zhang, P., 2005. The intellectual development of human-computer interaction research: A critical assessment of the MIS literature (1990-2002). J. Assoc. Inf. Syst. 6 (11), 227–292. https://doi.org/10.17705/1jais.00070.

Liu, X., He, S., Maedche, A., 2017. Designing an AI-based advisory platform for design techniques. In: Proceedings of the European Conference on Information Systems.

Loisel, A., Chaingnaud, N., Kotowicz, J, 2009. Modeling human interaction to design a human-computer dialog system. In: Proceedings of the International Conference on Enterprise Information Systems.

Ma, S.P., Ho, C.T., 2018. Modularized and flow-based approach to chatbot design and deployment. J. Inf. Sci. Eng. 34 (5), 1187–1201.

Mayring, P., 2015. Qualitative content analysis: theoretical background and procedures. Approaches to Qualitative Research in Mathematics Education. Springer, Dordrecht, pp. 365–380. Advances in Mathematics Education.

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., Söllner, M., 2019. AI-based digital assistants. Bus. Inf. Syst. Eng. 61 (4), 535–544. https://doi.org/10.1007/s12599-019-00600-8.

Meyer von Wolff, R., Hobert, S., Schumann, M., 2019a. How may I help you?–State of the art and open research questions for chatbots at the digital workplace. In: Proceedings of the Hawaii International Conference on System Sciences.

Meyer von Wolff, R., Masuch, K., Hobert, S., Schumann, M., 2019b. What do you need today? – an empirical systematization of application areas for chatbots at digital workplaces. In: Proceedings of the American Conference on Information Systems.

Meyer Von Wolff, R., Hobert, S., Schumann, M., 2022. Chatbot introduction and operation in enterprises–a design science research-based structured procedure model for chatbot projects. In: Proceedings of the Hawaii International Conference on System Sciences.

Moore, R.J., 2018. A natural conversation framework for conversational UX design. Studies in Conversational UX Design. Springer, Cham, pp. 181–204.

Muresan, M., Pohl, H., 2019. Chats with bots: balancing imitation and engagement. In: *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–6. DOI. https://doi.org/10.1145/3290607.3313084.

Myers, M.D., 2020. Qualitative Research in Business & Management. Sage Publications, London.

Nguyen, Q.N., Sidorova, A., 2018. Understanding user interactions with a chatbot: a self-determination theory approach. In: Proceedings of the American Conference on Information Systems.

Nili, A., Barros, A., Tate, M., 2019. The public sector can teach us a lot about digitizing customer service. MIT Sloan Management Review 60 (2), 84–87.

Nuruzzaman, M., Hussain, O.K., 2018. A survey on chatbot implementation in customer service industry through deep neural networks. In: Proceedings of the IEEE International Conference on e-Business Engineering, pp. 54–61. https://doi.org/10.1109/ICEBE.2018.00019.

Nuseibeh, R., 2018. Is my chatbot GDPR compliant? Chatbot Mag. https://chatbotsmagazine.com/is-my-chatbot-gdpr-compliant-fc18dd57f9e6?gi=7d054de061f0 Accessed: 23-07-2022.

Piccolo, L.S., Mensio, M., Alani, H., 2018. Chasing the chatbots. In: International Conference on Internet Science, DOI. Springer, Cham, pp. 157–169. https://doi.org/10.1007/978-3-030-17705-8_14.

Power, D., Kaparthi, S., Mann, A., 2019. Building decision adviser bots. In: Proceedings of the Midwest Association for Information Systems Conference.

Przegalinska, A., Ciechanowski, L., Stroz, A., Gloor, P., Mazurek, G., 2019. In bot we trust: A new methodology of chatbot performance measures. Bus. Horiz. 62 (6), 785–797. https://doi.org/10.1016/j.bushor.2019.08.005.

Rai, A., 2017. Editor's comments: diversity of design science research. MIS Q. 41 (1), iii–xviii.

Rietz, T., Benke, I., Maedche, A., 2019. The impact of anthropomorphic and functional chatbot design features in enterprise collaboration systems on user acceptance. In: Proceedings of the Wirtschaftsinformatik.

Rodríguez Cardona, D., Werth, O., Schönborn, S., Breitner, M.H., 2019. A mixed methods analysis of the adoption and diffusion of chatbot technology in the German insurance sector. In: Proceedings of the American Conference on Information Systems.

Rosemann, M., Vessey, I., 2008. Toward improving the relevance of information systems research to practice: the role of applicability checks. MIS Quarterly 1–22.

Ross, J.W., Sebastian, I., Beath, C., Mocker, M., Moloney, K., Fonstad, N., 2016. Designing and executing digital strategies. In: Proceedings of the International Conference on Information Systems.

Salinas, E., Cueva, R., Paz, F., 2020. A systematic review of user-centered design techniques. In: International Conference on Human-Computer Interaction.

Sarbazhosseini, H., Banihashemi, S., Adikari, S., 2019. Human-centered framework for managing IT project portfolio. In: Proceedings of the International Conference on Human-Computer Interaction.

Scaife, M., Rogers, Y., Aldrich, F., Davies, M., 1997. Designing for or designing with? Informant design for interactive learning environments. In: Proceedings of the CHI 97, Human Factors in Computing Systems.

Schuetzler, R., Grimes, M., Giboney, J.S., Rosser, H.K., 2021. Deciding whether and how to deploy chatbots. MIS Q. Execut. 20 (1), 1–15.

Sebastian, I., Ross, J., Beath, C., Mocker, M., Moloney, K., Fonstad, N., 2017. How big old companies navigate digital transformation. MIS Q. Execut. 16 (3), 197–213, 2017.

Seeger, A.M., Pfeiffer, J., Heinzl, A., 2018. Designing anthropomorphic conversational agents: development and empirical evaluation of a design framework. In: Proceedings of the International Conference on Information Systems.

Seeger, A.M., Pfeiffer, J., Heinzl, A., 2021. Texting with human-like conversational agents: Designing for anthropomorphism. J. Assoc. Inf. Syst. 22 (4), 931–967. https://doi.org/10.17705/1jais.00685, 2021.

Sjöström, J., Aghaee, N., Dahlin, M., Ågerfalk, P.J., 2019. Designing chatbots for higher education practice. In: Proceedings of the International Conference on Artificial Intelligence and Machine Learning.

Sousa, D.N., Brito, M.A., Argainha, M.C., 2019. Virtual customer service: building your chatbot. In: Proceedings of the International Conference on Business and Information Management, pp. 174–179. https://doi.org/10.1145/3361785.3361805.

Stöckli, E., Dremel, C., Uebernickel, F., Brenner, W., 2019. How affordances of chatbots cross the chasm between social and traditional enterprise systems. Electron. Market. 1–35. https://doi.org/10.1007/s12525-019-00359-6. DOI.

Suta, P., Lan, X., Wu, B., Mongkolnam, P., Chan, J.H., 2020. An overview of machine learning in chatbots. Int. J. Mech. Eng. Robotic. Res. 9 (4), 502–510. https://doi.org/10.18178/ijmerr.9.4.502-510. DOI.

Tarafdar, M., Beath, C.M., Ross, J.W., 2019. Using AI to enhance business operations. MIT Sloan Manag. Rev. 60 (4), 37–44.

Tavanapour, N., Bittner, E., 2018. Automated facilitation for idea platforms: design and evaluation of a chatbot prototype. In: Proceedings of the International Conference on Information Systems.

Van der Goot, M.J., Hafkamp, L., Dankfort, Z., 2021. Customer Service Chatbots: A Qualitative InterviewStudy into the Communication Journey of Customers. In: Følstad, A., et al. (Eds.), Chatbot Research and Design, Lecture Notes in Computer Science, 12604. Springer, Cham, pp. 190–204.

Virkar, M., Honmane, V., Rao, S.U., 2019. Humanizing the chatbot with semantics based natural language generation. In: Proceedings of the International Conference on Intelligent Computing and Control Systems. IEEE, pp. 891–894. https://doi.org/10.1109/ICCS45141.2019.9065723.

Vaishnavi, V.K., Kuechler, W., 2015. Design Science Research Methods and Patterns: Innovating Information and Communication Technology, 2015. Crc Press, Boca Raton.

vom Brocke, J., Hevner, A., Maedche, A., 2020. Introduction to design science research. In: vom Brocke, J., Hevner, A., Maedche, A. (Eds.), Design Science Research. Cases. Progress in IS. Springer, Cham. https://doi.org/10.1007/978-3-030-46781-4_1.

Wallisch, A., Sankowski, O., Krause, D., Paetzold, K., 2019. Overcoming fuzzy design practice: revealing potentials of user-centered design research and methodological concepts related to user involvement. In: Proceedings of the IEEE International Conference on Engineering, Technology and Innovation.

Wei, C., Yu, Z., Fong, S., 2018. How to build a chatbot: chatbot framework and its capabilities. In: Proceedings of the International Conference on Machine Learning and Computing, pp. 369–373. https://doi.org/10.1145/3195106.3195169.

Williams, J.J., Ramaprasad, A., 1996. A taxonomy of critical success factors. Eur. J. Inf. Syst. 5 (4), 250–260. https://doi.org/10.1057/ejis.1996.30.

Zamora, J., 2017. I'm sorry, Dave, I'm afraid I can't do that: chatbot perception and expectations. In: Proceedings of the International Conference on Human Agent Interaction, pp. 253–260. https://doi.org/10.1145/3125739.3125766.

Zierau, N., Elshan, E., Visini, C., Janson, A., 2020. A review of the empirical literature on conversational agents and future research directions. In: Proceedings of the International Conference on Information Systems.

Zschech, P., Horn, R., Höschele, D., Janiesch, C., Heinrich, K., 2020. Intelligent user assistance for automated data mining method selection. Bus. Inf. Syst. Eng. 62 (3), 227–247. https://doi.org/10.1007/s12599-020-00642-3.

Zumstein, D., Hundertmark, S., 2018. Chatbots: an interactive technology for personalized communication and transaction. IADIS Int. J. on www/Internet 15 (1), 96–109.