AIS Transactions on Human-Computer Interaction

Volume 14 | Issue 4 Article 6

12-31-2022

Designing Process-based Chatbots in Enterprises: The Case of Business Travel Organization Considering the Users' Perspective and Business Value

Raphael Meyer von Wolff
University of Goettingen, raphael.meyervonwolff@gmx.de

Sebastian Hobert
University of Goettingen, shobert@uni-goettingen.de

Matthias Schumann
University of Goettingen, mschuma1@uni-goettingen.de

Recommended Citation

von Wolff, R. M., Hobert, S., & Schumann, M. (2022). Designing Process-based Chatbots in Enterprises: The Case of Business Travel Organization Considering the Users' Perspective and Business Value. *AIS Transactions on Human-Computer Interaction, 14*(4), 578-623. https://doi.org/10.17705/1thci.00180
DOI: 10.17705/1thci.00180

Transactions on Human-Computer Interaction

Volume 14 Issue 4

12-2022

Designing Process-based Chatbots in Enterprises: The Case of Business Travel Organization Considering the Users' Perspective and Business Value

Raphael Meyer von Wolff

Chair of Application Systems and E-Business, University of Goettingen, raphael.meyervonwolff@gmx.de

Sebastian Hobert

Chair of Application Systems and E-Business, University of Goettingen, shobert@uni-goettingen.de

Matthias Schumann

Chair of Application Systems and E-Business, University of Goettingen, mschuma1@uni-goettingen.de

Recommended Citation

Meyer von Wolff, R., Hobert, S., & Schumann, M. (2022). Designing process-based chatbots in enterprises: The case of business travel organization considering the users' perspective and business value. *AIS Transactions on Human-Computer Interaction*, *14*(4), pp. 578-623.

DOI: 10.17705/1thci.00180

Available at http://aisel.aisnet.org/thci/vol14/iss4/6

Transactions on Human-Computer Interaction

Research Paper DOI: 10.17705/1thci.00180 ISSN: 1944-3900

Designing Process-based Chatbots in Enterprises: The Case of Business Travel Organization Considering the Users' Perspective and Business Value

Raphael Meyer von Wolff
Chair of Application Systems and E-Business
University of Goettingen
raphael.meyervonwolff@gmx.de

Sebastian Hobert
Chair of Application Systems and E-Business
University of Goettingen

Matthias Schumann
Chair of Application Systems and E-Business
University of Goettingen

Abstract:

Chatbots have attracted much research attention in recent years, and organizations have increasingly begun applying them in everyday working life. However, researchers have rarely investigated how chatbots can support everyday tasks in enterprises. As such, we lack design knowledge for chatbots that support internal business processes since research has mostly examined customer-facing use cases. Notably, researchers have rarely considered chatbots' economic and user-related effects, which, thus, remain unknown. To address this gap, we conducted a design science research study to survey a process-based chatbot application for business processes. From examining the scenario, we deduced design principles and implemented a software artifact. We evaluated the concept with 69 participants and surveyed the users' perspective in terms of design and acceptance and the organizational perspective in terms of process efficiency and quality. In doing so, 1) we derived six design principles for process-based chatbots and implemented a respective chatbot, which enabled a user-adapted process and provided situational-dependent input options and support; 2) we found that users had a positive attitude towards using chatbots for business processes in terms of user experience and acceptance; and 3) the process performed at an economically efficient level that compared well with existing solutions and that IT affinity and prior experience had no influence on performance. Furthermore, our solution improved the process quality compared to the existing solution.

Keywords: Chatbot, Business Process, Design Science Research, Design Principles, User Perspective, Usability, Acceptance, Organizational Perspective, Efficiency, Error Probability

Miguel I. Aguirre-Urreta was the accepting senior editor for this paper.

1 Introduction

1.1 Motivation and Research Relevance

As the economy continues to digitalize, enterprises have similarly begun changing and the way employees work has become more digital and based on information systems. By leveraging new capabilities, organizations have automated many tasks and, in doing so, relieved employees, supported them in their daily work, and improved work quality. In this transformation, organizations have digitalized many established work practices, and new and particularly digital forms of work and collaboration that affect almost every employee have emerged (Byström et al., 2017; Köffer, 2015; Lacity & Willcocks, 2021; White, 2012). Notably, companies today increasingly rely on applying human-centered approaches to design IS such as natural language-based user interfaces (e.g., chatbots) (Følstad & Brandtzæg, 2017). Chatbots represent a promising technology for supporting employees' wellbeing since employees can operate them to perform business tasks without prior training due to their natural and humanlike capabilities. Rather than needing to learn extensive and complex user interfaces (UI), users can simply write or speak their needs and the chatbot can execute the corresponding business functions or provide the desired information (Aquino, 2012; Carayannopoulos, 2018; Følstad & Brandtzæg, 2017). Prior research has shown that chatbots constitute a suitable technology to implement a user-centered design in information systems by adapting to users' needs and assisting them in their daily work (Carayannopoulos, 2018; Følstad & Brandtzæg, 2017; Richter et al., 2018). Chatbots automate processes via answering questions on their own, especially in customer support scenarios, which can result in work relief and improvements in the work quality (Gnewuch et al., 2017; Meyer von Wolff et al., 2020b). Notably, researchers also expect their humanlike and natural design to contribute to positive user perceptions and service experiences and to create a feeling of actually interacting with a real person (Diederich et al., 2019). These effects coincide with overall goals for process automation or enhanced work and service quality, better employee satisfaction, and more appealing application systems that encourage use (Engel et al., 2021; Wambsganss et al., 2020).

By now, one can find chatbots in many corporate application landscapes. Organizations mainly use chatbots as a tool to support and relieve employees in daily work scenarios or as a means to acquire and provide information (Følstad & Brandtzæg, 2017; Maedche et al., 2019). Research has already investigated several primary customer-oriented business applications such as information provision, customer service, sales, or financial advice (e.g., Bavaresco et al., 2020; Diederich et al., 2022; Feng & Buxmann, 2020; Gnewuch et al., 2017; Lewandowski et al., 2021). However, while chatbots often target customers or external users, deploying them in the workplace to support employees primarily requires enabling rather time-consuming and sometimes monotonous non-value adding activities or processes, such as changing personnel data, scheduling meetings, or organizing business trips. However, little research has focused on applying chatbots to business processes in everyday office life (e.g., Diederich et al., 2022; Lewandowski et al., 2021; Meyer von Wolff et al., 2019). The studies on work task execution have focused mainly on simple tasks (e.g., resetting a password) and less on intelligent workflows or whole processes (e.g., Feine et al., 2020; Lechler et al., 2019; Winkler et al., 2019). As such, the potential of chatbots for supporting internal workflows has not yet been sufficiently explored. Recent research on this topic has shown that chatbots can (at least in principle) execute processes and transactions (e.g., Chakrabarti & Luger, 2015; Meyer von Wolff et al., 2020a; Winkler et al., 2019). However, to the best of our knowledge, no study has investigated whether one can apply chatbots at digital workplaces based on a real case.

Hence, while the literature offers preliminary design contributions for simple, short-term, and mostly customer-focused ways to use chatbots and presumes that chatbots have a potential for business processes, a research gap exists for business chatbots (i.e., applying and designing chatbots for business processes in working environments to support employees). This research gap becomes apparent in several aspects. First, we lack knowledge in research and practice about how to design and use chatbots for complex internal business processes. Second, current chatbot research does not reflect actual daily working situations in terms of business processes and does not relate or compare its results with an existing situation (e.g., comparing cycle times or successful task completion). Hence, it currently remains unclear for companies whether they should use the technology for internal business processes as the effects of using chatbots for internal processes remain unclear. Third, technology acceptance research has shown that the success of chatbot projects largely depends on users' willingness to interact with the software in a specific use case (Chen et al., 2020; Davis, 1989). So far, it has not been fully investigated whether employees are willing to interact with chatbots when working on complex internal business processes. Thus, it could also be the case that workers prefer to interact with classic user interfaces in complex internal workflows.

1.2 Research Aim and Research Questions

As we note above, chatbots have the potential to provide employees with individualized and intuitive access to resources and business processes while offering unique assistance. However, up to now, scientific research has mostly exclusively focused on external or customer-focused application areas, such as frequently asked questions (FAQs), helpdesks, or customer complaints (e.g., Diederich et al., 2022; Meyer von Wolff et al., 2019; Nißen et al., 2022). In those customer-focused scenarios, users (i.e., customers) commonly interact with a chatbot only once or only during a short time frame. The commonly short interaction times in these customer-focuses scenarios substantially differs from how companies use chatbots to support internal business processes. When a company uses chatbots to support internal business processes, workers will need to interact with the systems frequently and for longer interaction times. Up to now, to the best of our knowledge, no research has addressed chatbot adoption for business processes or provided design-oriented guidelines for developing process-based business chatbots. As such, we can see a need to examine the topic from a design science perspective to develop a corresponding process-based chatbot. Accordingly, we conducted a design science research (DSR) study based on Hevner et al.'s (2004) and Hevner's (2007) approach to develop and evaluate an exemplary process-based business chatbot to draw conclusions not only about an individual business process use case but also business processes in general. We derived three research questions to guide our research project.

First, according to the DSR approach, we 1) surveyed requirements for process-based chatbots and 2) designed a process-based chatbot based on these requirements. In particular, we assessed how well the chatbot could support an exemplary business process that exists in almost every company, that employees use repeatedly, and that employees can carry out without specific knowledge to ensure greater generalizability. To do so, we selected an existing business travel organization process as the exemplary basis. We derived requirements for process-based chatbots by 1) incorporating current scientific findings on chatbot design, 2) incorporating user stories, and 3) analyzing the selected business process. Based on the derived requirements, we formulated generalized design principles that we used to design and implement a process-based chatbot DSR artifact.

Thus, we address the following research question (RQ):

> **RQ1:** How should one design enterprise process-based chatbots to execute and support business processes?

Second, after designing an artifact, the DSR approach also necessitates that one evaluates its possible effects and outcomes. Notably, for real case applications, users should be willing to use the system, and it should have business value. Otherwise, organizations may not use the artifact and/or they could waste their investment in it. An artifact that users did not accept or that lacked business value would also indicate that an artifact lacked applicability for business processes. Consequently, the scientific community would likely reject it. Therefore, we evaluated the process-based chatbot from 1) an individual perspective (i.e., we assessed whether users accepted the artifact and their user experience) and 2) an organizational perspective (i.e., process efficiency and quality). To address both perspectives, we conducted a laboratory experiment with individuals who could potentially use the chatbot in the future. We further compared our developed process-based chatbot with the business travel organization process that we used to derive our design recommendations to benchmark our results with the status quo.

Since usage, a key factor in determining success, depends on users, we necessarily considered the user perspectives as well. Thus, we examined 1) the chatbot's usability (i.e., the user experience) and 2) users' satisfaction with its interface. In addition, we examined 3) users' acceptance.

Furthermore, we concluded that, even if users are willing to interact with a chatbot for internal business processes, we needed to investigate whether chatbots actually have business value. To investigate this organizational perspective on the chatbot usage, we examined whether our chatbot could execute business processes and how it compared to existing solutions. To do so, we measured both 1) the process lead times and 2) the chatbot's error probability and compared it with the existing baseline. In addition, we interviewed business experts (i.e., managers or project leaders) to identify their 3) opinions about the chatbot's possible business value.

In doing so, we investigated whether companies can successfully use chatbots to conduct business processes. We further outline how chatbots affect process efficiency and process quality of the investigated internal business process.

Therefore, we address two more research questions:

RQ2: How do users assess an exemplary process-based chatbot for internal business processes?

RQ3: What business value do process-based chatbots for business processes have from an organizational perspective?

We structure the paper as follows: In Section 2, we highlight the theoretical foundations on chatbots in business contexts and work that has recently evaluated chatbots. In Section 3, we describe our applied DSR approach. In Section 4, we present our process-based chatbot artifact and its corresponding requirements and design principles. In Section 5, we show the evaluation results based on user and organizational perspectives. In Section 6, we discuss our results and the study's limitations. Finally, in Section 7, we conclude the paper.

2 Related Research

2.1 Current State of Research on Chatbots in Business Contexts

Chatbots are information systems that use artificial intelligence and machine learning (and, more specifically, natural language processing) to provide a dialog-based user interface (e.g., Carayannopoulos, 2018; Diederich et al., 2022; Elshan & Ebel, 2020; Lewandowski et al., 2021; Maedche et al., 2019; Sarikaya, 2017; Winkler et al., 2019). Users can communicate either by audio or text with these systems to gather or store information and execute business processes or work tasks (e.g., Berg, 2014; Carayannopoulos, 2018; Mallios & Bourbakis, 2016; Sarikaya, 2017). From a technical point of view, chatbots process input to extract patterns and identify users' intent. Based on their intent, chatbots provide information, answer questions, or execute functions and processes. Therefore, besides the chatbots' knowledge base, one must integrate them with databases and (enterprise) systems (e.g., Berg, 2014; Mallios & Bourbakis, 2016; Meyer von Wolff et al., 2021).

As we mention in Section 1, nowadays, some researchers seek to transfer the promising results regarding chatbots to business contexts and, more specifically, workplace applications to support employees or enable them to execute business processes (e.g., Elshan & Ebel, 2020; Feine et al., 2020; Meyer von Wolff et al., 2021). These digital workplaces generally extend beyond physical places and combine work tasks, business processes, enterprise systems or databases, technologies, employees, and customers (Dery et al., 2017; White, 2012).

To apply chatbots in business contexts, chatbot research has mostly focused on customer-focused topics such as customer support or service (e.g., Corea et al., 2020; Gnewuch et al., 2017; Johannsen et al., 2018; Liebrecht & van Hooijdonk, 2020; Zierau et al., 2020b). Also, researchers have examined information acquisition or provision with chatbots (e.g., Al-Zubaide & Issa, 2011; Carayannopoulos, 2018; Chai et al., 2001; Radlinski & Craswell, 2017; Ranoliya et al., 2017). Notably, one study has highlighted professional workplace-related information acquisition (e.g., from ERP and CRM systems) (Reshmi & Balakrishnan, 2016). Furthermore, researchers have examined actual workplace applications for employees. For instance, Lechler et al. (2019) showed how chatbots may support feedback exchange. Other studies have also examined chatbots as teammates (Elshan & Ebel, 2020) as a tool for problem-solving tasks (Winkler et al., 2019) or for reducing friction in collaborative teamwork (Gyton & Jeffsry, 2017). Some researchers have attempted to enhance chatbots with more intelligence to better support processes and not only respond to questions. For example, Tavanapour et al. (2019) followed a process-like approach in which a chatbot supported the idea-generation process by asking questions and acting like a facilitator. A different study used a goal-fulfillment map similar to a finite-state machine to map a dialog and allow longer and more dynamic interactions in customer support settings (Chakrabarti & Luger, 2015). A similar study implemented a finite-state machine chatbot to provide support for complex tasks such as e-learning and education (Hobert, 2019b). In this study, the chatbot continuously adjusted its current dialog state based on users' intent and triggered corresponding actions. Likewise, another study applied an approach that dynamically adapted the dialog based on the current interaction (Winkler et al. 2020). Although the last two studies focused on e-learning and teaching, they demonstrate that chatbots can map processes as they occur in the workplace.

Apart from the research on individual application areas, many studies have analyzed chatbots' features or design aspects at a more general level. Significantly, research in one domain focuses on humanizing chatbots and their response behavior (e.g., Diederich et al., 2020). One such study examined human response behavior as a requirement for more humanlike chatbots by deriving possible linguistic elements and investigating their contributions (Liebrecht & van Hooijdonk, 2020). The authors found anthropomorphic design features to have a high impact on perceived usefulness. Another study surveyed the impact that implemented anthropomorphic and functional features had on the extent to which users accepted chatbots (Rietz et al., 2019). In contrast to the previous two studies, the results indicate that, for a workplace application in particular, usefulness has more importance than humaneness. Schuetzler et al. (2018) examined the influence that conversational relevance had on the perception of chatbots. They found that individuals perceived chatbots that provided relevant responses as more humanlike and social. Lastly, Adam and Klumpe (2019) investigated how aspects such as message interactivity and self-disclosure affected users' disclosure propensity. In addition, most design-oriented contributions in the literature have made generalized design principles. To list just a few, Gnewuch et al. (2017) presented generalized design principles for designing chatbots for customer service, and Tavanapour et al. (2019) derived design principles for a chatbot to support idea generation. Notably, Siemon and Jusmann (2021) surveyed users to identify their appearance preferences for chatbots for knowledge management tasks, Diederich et al. (2020) provided design results on anthropomorphic chatbots for enterprises, and Feine et al. (2020) outlined design recommendations for enterprise chatbots in general. In addition, Stoeckli et al. (2018) showed functional affordances and possible outcomes from using a chatbot. Følstad and Skjuve (2019) focused on the chatbot user experience and users' motivation to use a chatbot. Wuenderlich and Paluch (2017) examined how users perceived chatbots. Lastly, Nißen et al. (2022) created a taxonomy to characterize user-chatbot relationships in different time horizons to derive three time-dependent design archetypes.

In our study, we comprehensively overview whether one can apply chatbots in business contexts. Notably, some studies have pointed out generalized use cases for chatbots in business contexts (e.g., Feng & Buxmann, 2020; Laumer et al., 2019; Meyer von Wolff et al., 2020b; Stoeckli et al., 2018). They have examined whether organizations can viably use chatbots to provide information and execute business processes such as self-service tasks, which we focus on in this study. Stieglitz et al. (2018) define enterprise bots based on a literature review and further derive a research model for the field. Lastly, other research has summarized the existing knowledge based on literature reviews (like Feng & Buxmann, 2020; Lewandowski et al., 2021; Meyer von Wolff et al., 2019) or taxonomies (Janssen et al., 2020).

Overall, research on chatbots has mostly focused on simple use cases (e.g., information provision in FAQs). Only few studies have attempted to extend this capability by adjusting dialog or implementing low-level processes as they occur at the workplace. Consequently, to the best of our knowledge, researchers have not yet sufficiently applied chatbots for processes in general and business processes in particular. Accordingly, they have not yet investigated whether one can successfully use chatbots to execute and support complex internal business processes (e.g., user-centric information systems, intuitive use, individual support). Thus, we cannot simply transfer the existing design knowledge from other — simpler — use cases (like information provision) to design chatbots supporting complex internal business processes. However, a large scientific knowledge base exists for designing general chatbots and specific chatbot features, which partly applies to enterprise process applications as well. Therefore, we built on this knowledge base when deriving design principles for our process-based enterprise chatbot.

2.2 On the Evaluation of Chatbots

Besides the design-oriented contributions, the scientific literature has already surveyed and evaluated the use of chatbots in different use cases. In this section, we summarize existing evaluation approaches from prior research. That effort guides our evaluation approach in Section 5. We identified some overviews on chatbot evaluations in scientific contributions for educational settings (Hobert, 2019a) or in general (Maroengsit et al., 2019). Also, Shawar and Atwell (2007) noted measurement metrics to evaluate chatbots, and Zierau et al. (2020a) outlined dependent and independent variables that one can use to evaluate chatbots.

In specific, many existing DSR studies have evaluated a specific chatbot's design principles and requirements to identify improvement potentials or generalize their findings (e.g., Feine et al., 2020; Gnewuch et al., 2017; Hobert, 2019b; Tavanapour et al., 2019; Winkler et al., 2020). Also, Carayannopoulos (2018) evaluated how chatbots' design elements and capabilities can help users navigate complex new situations and quickly provide them with necessary information. Furthermore, a design study investigated the difference between a human and a chatbot based on a "Wizard of Oz" study and reported that the chatbot could facilitate idea elaboration in a human-chatbot interaction in a similar way as a human would do in a human-human interaction (Bittner et al., 2019). In addition, Winkler et al. (2019) investigated the influence that chatbot applications had on group performance in problem-solving scenarios, while Winkler et al. (2020) investigated the influence that scaffolding and voice-based chatbots had on learning performance (i.e., information retention and transferability). Both studies reported positive results and confirmed that chatbots could conduct the given scenario.

Besides these design evaluations, research has also evaluated dialog designs and decisions. Chakrabarti and Luger (2015) surveyed how well chatbots conduct conversations. They showed that their finite-state machine-based chatbot could maintain the context of a conversation. In doing so, the chatbot performed beyond exchanging simply utterances, and one could virtually not distinguish the resulting dialogs from a natural one. Other studies analyzed the influence that typing delays had on users' perception (Gnewuch et al., 2018) and the influence that chatbots' conversational relevance had on how users perceive and engage with chatbots (Schuetzler et al., 2018). Gnewuch et al. (2018) showed that dynamic typing delays positively affected users' perceptions and that they perceived the chatbots that used dynamic delays as more human. Thus, users seemingly had the same expectation for the chatbot as they had for human-human interactions. The latter study concluded that users perceived conversational agents that gave conversationally relevant responses as more humanlike and social (Schuetzler et al., 2018).

Many prior research studies have focused on designing chatbots that users perceive as humanlike. Diederich et al. (2019) surveyed chatbots' empathetic behavior and their influence on customer service and found that, when chatbots can detect users' sentiment to provide empathic responses, users perceive them more favorably and have more satisfaction overall with them even if the chatbots cannot fully complete a task. Adam and Klumpe (2019) explored the influence that different features (e.g., high vs. low message interactivity and platform self-disclosure) had on the user's disclosure propensity, while Lichtenberg et al. (202) investigated the relationship between persuasive and anthropomorphic conversational agent design and performance. The latter found that just adding more anthropomorphic features did not necessarily improve perception. Manseau (2020) surveyed possible outcomes of using chatbot applications in the workplace and stressed the important role that anthropomorphism plays in increasing the extent to which users accept chatbots. Likewise, another study targeted the influence that anthropomorphic elements and usability have on system acceptance and reported that a stronger focus on humanizing the chatbot did not necessarily result in higher user enjoyment (Rietz et al., 2019). Rather, they showed that users were more likely to accept a chatbot based on utility rather than hedonism or joy. Thus, in work environments, the most important element for chatbot acceptance seems to be the benefit it provides.

Furthermore, some studies have explored chatbots' usability. For instance, Meyer von Wolff et al. (2020a) measured a chatbot's usability after users used it for three scenarios and reported initial insights into task feasibility, language comprehension, and visualization. They found that users provided a high usability rating for the chatbot. In particular, the study emphasized that users could easily learn and understand the applied chatbot and that it quickly and efficiently solved tasks without unnecessary effort. Holmes et al. (2019) conducted a similar study in comparing three usability measures (system usability scale, user experience questionnaire, and chatbot usability questionnaire) while also determining the respective task completion time for the evaluation setting. Besides highlighting rather high usability scores that were independent of the respective usability measures, the results identified all three metrics as suitable in principle.

Regardless, existing research has mostly targeted design aspects or closely related topics. Thus, to the best of our knowledge, no studies have examined or measured the effects that chatbots supporting internal business processes have on how employees perceive using chatbots as part of their daily work and the business value more comprehensively (i.e., in a way that considers both the individual perspective and the organizational perspective). In particular, the literature neglects factors such as whether internal business processes can be implemented in chatbots in a reasonable way and how well chatbots compare to existing enterprise processes and systems (e.g., in terms of usability). In particular, existing evaluations predominantly refer to individual aspects, which suggests that future evaluations should examine several aspects together to comprehensively explain the relevant effects and influences. Particularly for the enterprise context, one cannot simply consider only users' perspective because organizations need to generate business value. Otherwise, organizations will not even consider chatbots as a possible solution. Thus, we build on existing scientific findings and holistically evaluate a chatbot we developed from both the user and organizational perspectives.

3 Applied Design Science Research Approach

We conducted a DSR study to investigate whether one can use chatbots to conduct business processes. We structure our paper based on Gregor and Hevner's (2013) recommendation to ensure that we present the results clearly in a high-quality manner. In doing so, we contribute 1) a prototypical process-based chatbot artifact (RQ1), 2) provide insights into users' attitudes towards using chatbots to conduct business processes (RQ2), 3) survey the business value that process-based chatbots for business processes have (RQ3), and 4) generate first generalizable implications for process-based chatbots to extend the scientific knowledge base. To conduct our DSR study, we followed Hevner et al. (2004) and Hevner (2007) and applied the rigor, relevance, and design cycle (see Figure 1).

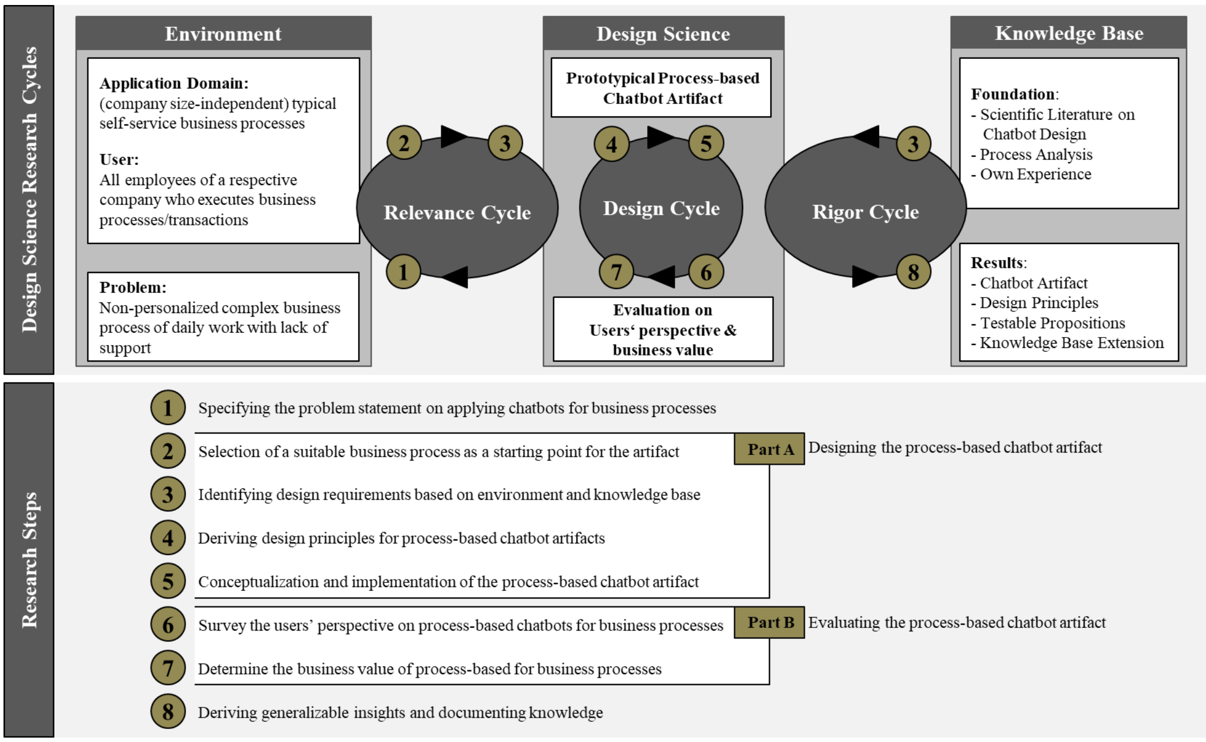

Figure 1. Applied Design Science Research Approach

To address the relevance, rigor, and design cycle, we divided our research project into eight steps (see Figure 1):

In the first step, as we note in Section 1, we focused on whether one chatbot can perform complex tasks (i.e., process-oriented, multiple steps) and, in particular, business processes. We focused on extending current knowledge by identifying new areas in which one may apply chatbots beyond the short-term and simple tasks that organizations have often used them for (e.g., FAQs or helpdesks). To do so, we derived our research objectives to guide our study (see Section 1).

Next, we designed our process-based chatbot artifact (part A) (see Section 4). Thus, in the second step, we selected a suitable business process as a basis to conceptualize and develop a chatbot artifact; namely, the process to organize business travel. Accordingly, in the third step, we surveyed the environment and knowledge base to identify the design requirements for process-based chatbots. For this purpose, we considered 1) users' experience with current processes, 2) the selected exemplary process itself, and 3) the current scientific literature on chatbot design. In the fourth step, based on these design requirements, we derived design principles for process-based chatbots for conducting business processes according to Gregor et al. (2020). In the fifth step, based on these design principles, we developed our process-based chatbot DSR-artifact. Therefore, we selected an existing business travel organization process as an example and implemented a fully functional prototype.

Subsequently, we evaluated our implemented artifact in a real-case scenario based on an experiment with 69 participants out of three groups (part B) (see Section 5). To do so, in the sixth step, we first surveyed users' perspective in terms of user experience, system design, and acceptance. In the seventh step, since we found promising results, we further analyzed the process-based chatbot artifact's business value in terms of its process efficiency and quality. Based on the evaluation, besides some minor UI changes, we could not identify any major issues or necessary improvements that we needed to make so that the chatbot could execute the example business process. Thus, we conducted no further design cycles.

Finally, in the eighth step, we documented the results in this contribution and derived generalizable findings (see Section 6).

4 Designing the Process-based Chatbot Artifact

In this section, we outline the results from the first part of our DSR study in which we designed a chatbot artifact that could execute and support business processes (RQ1). We later used this artifact for the user- and business value-focused evaluations.

4.1 Selecting an Exemplary Business Process

We conducted our study to support employees at the digital workplace in their daily work by developing a process-based chatbot and, thereby, identifying whether chatbots can execute and support business processes. Thus, this study focuses on process automation. As part of this automation, one typically introduces a chatbot to automate monotonous non-value adding tasks as far as possible. Employees should be involved to work on value adding tasks. Also, such a process automation can address aspects such as operational efficiencies, improved service quality, 24/7 availability, or employee satisfaction (Lacity & Willcocks, 2021). However, researchers recommend process automation only for rule-based processes with high standardization, maturity, and transaction volume, which means that selecting the right use case constitutes an essential step in such endeavors (Engel et al., 2021). Thus, based on Engel et al.'s (2021) and Lacity and Willcocks' (2021) work and to examine process-based chatbots in general (i.e., beyond a specific business process), we looked for a suitable exemplary process. In particular, we sought a typical company-independent process that employees would regularly or frequently conduct to allow scalability and generalizability. Therefore, we required a somewhat standardized process so that one could define single tasks or steps before conducting it and could somewhat individualize the process based on one's input (which would allow the chatbot to do so as well). If possible, the process needed to involve multiple stakeholders and require mandatory accurate information to reflect a realistic corporate situation. Furthermore, we needed a process for which a single individual had primary responsibility since that person would then have to find solutions if they became stuck. However, we needed a process that someone could do with little experience.

4.2 Scenario Description: The Business Travel Organization Process

Considering the criteria, we selected the process to organize business travels as the exemplary business process for this study. This standardized and formalized process has defined parameters yet simultaneously allows for dynamic behavior to some degree (e.g., the input order and whether one uses necessary or unnecessary inputs based on the individual travel setting). It is also a non-value adding organizational process that all business travel requires. In addition, the process exists in almost every company in the same or a comparable form, which further fosters the generalizability of our results in general and our design principles in particular. Thus, the process seems viable as an exemplary starting base to examine whether process-based chatbots can conduct and support business processes in general and to extend the scientific knowledge. It also exclusively focuses on short-term and one-time usage as we describe in Section 1. Furthermore, the process resides in the self-service domain for which chatbots seem suitable (Meyer von Wolff et al., 2020b).

To develop our DSR chatbot artifact, we used the existing process to organize business travel at a German university as the exemplary process. The university employed around 5,000 employees who could all potentially use the chatbot-based instantiation. To derive the current procedure, we analyzed the process by considering the necessary tasks (e.g., travel request or travel accounting) and reviewed the corresponding business documents and forms (i.e., business travel request and business travel accounting) to derive the process (see Figure 2, left). When we conducted the study, one could execute the process by filling out two PDF forms (one for application and one for accounting) or using a Web application that digitally represented the form. Usually, the process began when an employee planned an upcoming trip. Then, the employee created a travel request using the PDF form or the Web application. If employees had a question, they had to contact the responsible support person or derive the corresponding information from the official support documents on their own. The person then forwarded the completed application to the administrative assistant for review. Based on the results, the administrative assistant returned the form if it contained errors or, otherwise, forwarded it to the responsible supervisor. Notably, different supervisors approved the application depending on the travel location (e.g., national or international destinations) and duration. After the journey, the traveler needed to account for the business trip according to the same principle and attach receipts for the costs incurred. The existing process had many areas that one could make improvements in:

1) As the user filled out the PDF, the user could not individualize it and the PDF did not adapt to the user; thus, the user could query unnecessary data.
2) Only the user checked information's validity when entering, while a supervisor checked it later in the process, which could lead to delays to complete the process when the user entered incorrect information.
3) When the user used the PDF solution, the user needed to print out both forms before forwarding it on.
4) The user needed to always submit paper-based receipts.
5) No automatic forwarding existed. The user needed to first identify the responsible person and then manually pass on the form.
6) If the user encountered problems or needs further information, the user had to handle it on their own. However, different systems often contained the necessary information or the user needed to request it from other employees, which delayed the process or prevented others from doing their work.

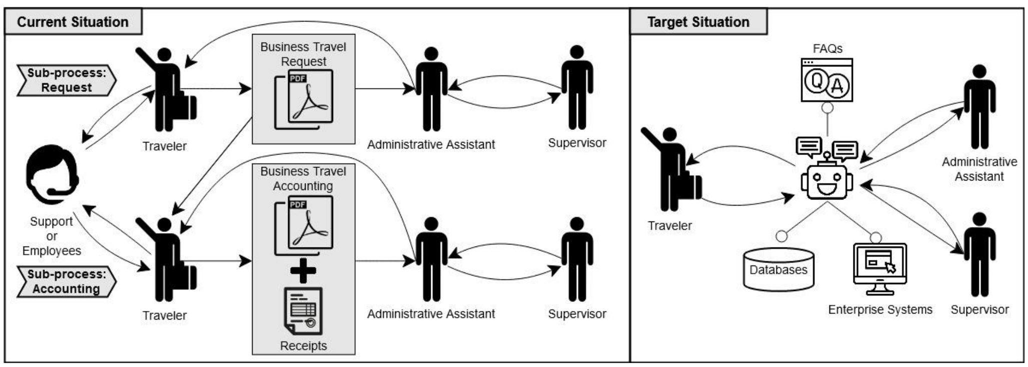

Figure 2. Exemplary Scenario: Existing Situation (Left); Target Situation (Right)

To overcome the obstacles and adapt the process to the user, we designed and implemented a chatbot for the process. As we note in Sections 1 and 2, we focused on providing a system that would require no training to use and could directly assist users if questions arose. In our target situation, the user should only need to communicate with the chatbot to perform both subprocesses (see Figure 2, right). The chatbot itself needed to adapt the necessary input fields according to each user's data and provide information in case the user had questions. When the user filled out all fields, the system needed to forward the process to the responsible persons (e.g., the supervisor or the administrative assistant) and forward the information to the other relevant databases and enterprise systems.

4.3 Requirements Analysis to Derive Design Principles

To derive design principles for our process-based chatbot, we took both inputs from employees who carried out the business travel processes and the selected exemplary business travel process itself into consideration as environmental aspects (relevance). To consider the knowledge base (rigor), we used the current scientific knowledge base. Finally, we used user stories, a process analysis, and scientific requirements to deduce generalized design principles for developing a process-based chatbot according to Hevner et al. (2004) and Hevner (2007) (see Figure 1).

Table 1. Requirements and Design Principles for Process-based Chatbots

			Design principles for process-based chatbot artifacts					
			Natural language user interface	Process guidance including progress overview	Individualized process adaptation to the user	Context-dependent input options	Integrated help function via dialog	Automatic error handling
			DP_1	DP_2	DP_3	DP_4	DP_5	DP_6
Environment: user stories	**User stories in business processes**							
	Individual understanding of the business process	US_1		●				
	Error-free input of all (necessary) information	US_2				●		●
	Processes depend on the user or their input	US_3			●			
	Individual help search, if process/inputs unclear	US_4					●	
Environment: process analysis	**Process characteristics**							
	Standardized process	P_1		●	●			
	Different types of user inputs (mandatory, optional, conditional)	P_2				●		
	Manual review of the stated information	P_3					●	●
	Submission of supporting documents in paper form	P_4		●		●		
	Clarification of questions/uncertainties outside the process	P_5					●	
	Supervisors approve or deny procedure depending on destination	P_6		●				●
	Multiple actors involved	P_7			●			
Knowledge base: scientific contributions on chatbot design	**Scientific requirements for process-based chatbots**							
	Adaptive guidance of users towards a goal	R_1		●	●	●		
	Support and Q&A in the dialog	R_2					●	
	User-friendly natural language-based user interface	R_3	●			●		
	Social cues to generate humanness	R_4	●					
	Verification of user inputs and error handling	R_5						●
	Pro- and reactive conversation mode	R_6	●					
	Transparency about the available functions	R_7		●				
	(Automatic) Feedback on the inputs	R_8		●				●
	Inclusion of trust-enhancing elements	R_9		●				
	On-demand handoff to employees	R_{10}			●		●	
	Saving the histories and user specifics	R_{11}			●			

Note: ● relation between requirement and design principle

4.3.1 User Stories for Conducting Business Processes

As a starting point to involve users or employees in designing the process-based chatbot, we briefly discussed the topic and our intended solution with different university employees to examine their experience or problems and challenges with the existing process to organize business travels. Those employees have already organized business travels in the past and have used the previous solution for this purpose. We used the collected responses to derive four relevant user-based requirements for the chatbot artifact (see Table 1) that directly relate to organizing business travel.

First, employees, especially new ones without expertise in the process, had to understand the underlying process and required inputs (**US_1**). Second, employees had to fill out the form completely and without errors (**US_2**). Otherwise, the process cannot be executed, will stop, or has to started again with the correct inputs, which results in unnecessary process cycle times and frustration. As such, individual employees had responsibility for the processes and their respective inputs (**US_3**). Even though they used standardized processes, the employees expressed that they needed to consider rules in the processes or individual paths and combinations as well. Lastly, if employees encountered problems, they had to manually search for solutions (e.g., FAQs or supporting documents), had to contact the support staff, or had to ask other employees for their advice and support (**US_4**). Such actions interrupted and delayed the process and, especially the latter in particular, distracted employees from their other work.

4.3.2 Process Characteristics

To derive the process's characteristics, we conducted a process analysis as we outline in Section 4.2. To do so, we analyzed the necessary forms (i.e., business travel request and business travel accounting) and all supporting information and documents for business travels (e.g., regulations or guidelines) to identify its specifics and characteristics. After analyzing all the available information, we derived seven process characteristics relevant for designing process-based chatbots (see Table 1).

First, the process was strictly standardized with defined inputs (**P_1**). The inputs, however, varied (**P_2**) since some fields were mandatory (e.g., destination and dates), some were optional (e.g., declaration of accompanying persons), and some were conditional (e.g., demand for discounts on transportation). Second, when we conducted the study, employees performed the process either with paper forms or an application system that digitally mapped both forms. Thus, employees needed to review their inputs manually (**P_3**). For the accounting subprocess in particular, employees had to submit paper-based receipts for the incurred costs in addition to the accounting form (**P_4**). If questions or ambiguities arose, the employee needed to contact a third party or look for solutions in the available documents and information sources (**P_5**). In addition, depending on the destination, different supervisors approved the forms (e.g., one for domestic travel and another for international travel). Thus, the process involved multiple actors, and employees needed to forward the process to the correct subsequent one (**P_6, P_7**).

4.3.3 Scientific Requirements for Process-based Chatbots

To include the existing scientific knowledge, we reviewed the AIS electronic database and included current literature on chatbot design that documented their design results based on requirements or generalized design principles. In doing so, we identified and used 12 relevant scientific proceedings (see the Appendix for the detailed distribution). We categorized the described requirements or design principles through joint discussions and derived 11 design requirements for process-based chatbots (see Table 1).

First, chatbots should adaptively guide users during a process or task toward a given goal (**R_1**). Hence, chatbots should encompass a goal-oriented behavior and actively pose questions to continue the conversation flow; use clarification and confirmation messages; or change the length, segmentation, and content of their messages based on the given situation (Feine et al., 2020; Tavanapour et al., 2019). In doing so, chatbots should assess the current state and decide which path to follow and which steps, depending on the given inputs or decisions, to conduct. If needed, a chatbot should implement step-by-step guidance (Hobert, 2019b). Thus, the systems must be able to adapt the process or itself to the process's actual needs and current state. Therefore, the chatbot needs to actively monitor the process and the given or missing inputs (Elshan & Ebel, 2020). Hence, the conversation should contain more than simple question-answering dialog (Gnewuch et al., 2017), which means the chatbot needs to map and implement a given business process with all its possible subpaths. Second, chatbots should provide direct support and question answering capabilities in dialog (**R_2**). For this purpose, they should offer different kinds of scaffolds to users when completing a task (Winkler & Roos, 2019). When users encounter problems or ambiguities, the chatbot should provide a Q&A component to enable on-demand explanations or clarify the necessary steps (Hobert, 2019b; Tavanapour et al., 2019). Hence, users can resolve errors on their own by questioning the chatbot, which may reduce the effort they need to expend to complete the task or remove the need to contact others (Corea et al., 2020; Winkler & Roos, 2019). Thus, chatbots should summarize necessary information, describe process's conditions, and offer explanations and clarifications if requested (Tavanapour et al., 2019). Furthermore, chatbots must provide a user-friendly natural language-based user interface (**R_3**). As usual for chatbots, users control the available functions using natural language inputs in dialog form and receive the answers or results in the same manner. Consequently, a chatbot must understand the user's messages and extract the respective intent (Bittner & Shoury, 2019; Diederich et al., 2020; Gnewuch et al., 2017). This type of interaction assumes that the system can handle typography errors or different languages and answer with correct grammar and pronunciation. The messages should use simple, short, and understandable language (Johannsen et al., 2018; Tavanapour et al., 2019). Also, the interface should provide visual input and output elements such as images, control elements, or buttons to increase efficiency or reduce the risk of input errors and, thus, maintain data consistency (Feine et al., 2020). Furthermore, chatbots should include anthropomorphic elements and social cues (**R_4**) such as an avatar, gender, typing delays, interjections, rhetorical elements, or emoticons (Diederich et al., 2020; Feine et al., 2019; Gnewuch et al., 2018; Liebrecht & van Hooijdonk, 2020). Likewise, chatbots should act in a friendly, neutral, and empathetic manner to foster an enjoyable conversation in a professional setting that evokes real human contact (Diederich et al., 2020; Elshan & Ebel, 2020; Tavanapour et al., 2019). Designers should ensure a balance between social cues and real capabilities, which also requires context-dependent social cues (Gnewuch et al., 2017). For this purpose, designers could include ways for the chatbot to predict user behavior (Corea et al., 2020). Additionally, chatbots must verify user input and provide error handling (**R_5**). On the one hand, this requirement includes unrecognized user requests, which the chatbot should clarify, and, on the other hand, incorrect or faulty inputs and given information (Bittner & Shoury, 2019; Feine et al., 2020; Tavanapour et al., 2019). To start a conversation, chatbots should use proactive methods in addition to their usual reactive conversations mode (**R_6**). In this way, chatbots can automatically notify users about changes (Bittner & Shoury, 2019; Feine et al., 2020). In addition, chatbots should also be transparent about their available functions and be identified as a machine (**R_7**). For this reason, developers should set up an adequate introduction during which the chatbot introduces itself and explains its available functions (Bittner & Shoury, 2019; Zierau et al., 2020b). But also during the conversation or process, the chatbot should always clearly and transparently communicate its available functions (Feine et al., 2020). Also, a chatbot should provide continuous feedback on the given inputs based on statically and dynamically analyzing the statements and information that users provide (Hobert, 2019b; Lechler et al., 2019) (**R_8**). Furthermore, as Winkler and Roos (2019) propose, chatbots should include trust-enhancing elements (**R_9**). Also, chatbots should offer the option to on-demand and conveniently get in touch with a human employee (**R_{10}**). Hence, in case the chatbot breaks down or users feel dissatisfied with it, it should be able to contact a human for assistance or to continue the process (Corea et al., 2020; Diederich et al., 2020; Zierau et al., 2020b). The chatbot should provide a human option instead even for when users do not want to go through the process with the chatbot (Johannsen et al., 2018). Lastly, chatbots should save histories and user specifics (**R_{11}**). This requirement allows chatbots to learn from previous conversations and provide personalized suggestions (Feine et al., 2020; Winkler & Roos, 2019)

4.3.4 Deriving Design Principles for Process-based Chatbots

To derive our design principles for a process-based chatbot artifact, we used the user stories (relevance), the process analysis (relevance), and the scientific requirements (rigor) (Hevner et al., 2004; Hevner, 2007). In doing so, we deduced six design principles for process-based chatbot artifacts (see Table 1). Afterwards, we applied the formalization method from Gregor et al. (2020) to describe the design principle based on the constructs actor, aim, mechanism, and rationale (see Table 2).

First, chatbots should enable user interaction with a natural language user interface (**DP_1** based on R_3, R_4, R_6). Hence, chatbots require a messenger-like dialog-based form with which users can control the available functions. Depending on the input, the chatbot must process and interpret the user's messages to allow the user to control the system and its underlying business process. Thus, due to the natural communication behavior, the user can control the system intuitively without previous training. To ensure users perceive the chatbot as a real person and help increase the likelihood that they accept it, chatbots should include social cues or anthropomorphic elements. Additionally, chatbots must map the respective business process and guide users through it (**DP_2**; based on US_1, P_1, P_4, P_6, R_1, R_7, R_8, R_9). Therefore, they must implement the complete business process, all subprocesses or paths, and the relevant conditions. By doing so, chatbots can query the (necessary) information from users and decide on the next steps. To help users find their way through the process and assess their current status, chatbots should offer feedback options, such as (sub-)process/task summaries or progress overviews. In a related manner, chatbots must individually adapt the process to each user (**DP_3**; based on US_3, P_1, P_7, R_1, R_{10}, R_{11}). Hence, to support flexibility, chatbots should not enforce strict process sequences. Rather, they should exhibit a goal-oriented behavior along all possible process subpaths or paths. Depending on the input, users should only go through the necessary steps or only have to share strictly necessary information, which speeds up the process and prevents unnecessary activities. To further increase efficiency, the individualized adaption should also encompass personalized suggestions based on previous interactions or recognized patterns. Chatbots should also offer context-dependent input options (**DP_4**; based on US_2, P_2, P_4, R_1, R_3). As usual in application systems for business processes or tasks, chatbots should allow interaction using control elements (e.g., buttons, selection options, file uploads). Users can work with their familiar elements, do not need to learn new techniques, and do not need to write all their inputs. Also, from a usability perspective, users know immediately what they must do when using a chatbot. For instance, a date picker shows the user that they must enter a date. Likewise, a multiple-choice list directs the user to select the suitable option. Additionally, this design principle further enhances data quality and consistency because the chatbot already correctly formats information beforehand, which ensures that the relevant actors can further process the data without errors. Furthermore, chatbots should encompass an integrated help function in their dialog (**DP_5**; based on US_4, P_3, P_5, R_2, R_{10}). If the user encounters problems or ambiguities, the user can directly ask the chatbot for help. Accordingly, chatbots need Q&A components in which they address typical questions. In this way, they do not interrupt the dialog, and users do not need to search for solutions manually or in another system. Also, this component prevents other employees from being distracted from their work and reduces how many questions users may need to ask others. Thus, users do not depend (or depend less) on third parties, which accelerates the process and reduces the potential for incorrect entries. Coincidently, chatbots must encourage users to use correct information (**DP_6**; based on US_2, P_3, P_6, R_5, R_8). Especially in business processes, users must provide error-free and complete information as subsequent tasks or processes depend on it. Otherwise, further errors or aborts can occur, which can result in delays or process restarts for users.

Table 2. Design Principles for Process-based Chatbots in Digital Workplace Settings (Gregor et al., 2020)

Design principle	Description
DP_1: Natural language user interface	For chatbots to provide user-friendly humanized user interfaces that can be used responsively and device independently with the feeling of a personal contact for employees in digital workplace settings, employ a natural language user interface with social cues (R_3, R_4, R_6).
DP_2: Process guidance including progress overview	For chatbots to enable individualized processes that can be carried out without prior knowledge by employees in digital workplace settings, employ a natural language-based step-by-step process guidance that encompasses the entire process and enables successful process execution as well as a progress overview that indicates current status (US_1, P_1, P_4, P_6, R_1, R_7, R_8, R_9).
DP_3: Individualized adaptation of the process to the user	For chatbots to allow flexibility in process execution and support a user-centered design for employees while executing business processes in digital workplace settings, employ the corresponding business process in a goal-oriented behavior with all possible tasks and enable individualized pathing based on users' input while also using previous inputs as suggestions (US_3, P_1, P_7, R_1, R_{10}, R_{11}).
DP_4: Context-dependent input options	For chatbots to offer a range of functions adapted to dialogs and comparable to classic enterprise systems as well as to enhance data quality due to a preformatted structure for employees in digital workplace settings, employ various suitable context-dependent input options (US_2, P_2, P_4, R_1, R_3).
DP_5: Integrated help function via dialog	For chatbots to provide support during the task and provide solutions for ambiguities and misunderstandings directly at the time of emergence for employees while conducting business processes in digital workplace settings, employ an integrated help function or Q&A component in the dialog where users can ask questions and receive support from the chatbots (US_4, P_3, P_5, R_2, R_{10}).
DP_6: Automatic error handling	For chatbots to ensure that all necessary entries/information are made and correct so that the process is not interrupted or delayed by employees in business processes in digital workplace settings, employ an automatic error handling to verify input and check for completeness (US_2, P_3, P_6, R_5, R_8).

4.4 Description of the Process-based Chatbot Artifact

Based on the design principles (see Table 2) and state-of-the-art chatbot architecture (Berg, 2014; Mallios & Bourbakis, 2016), we developed our process-based chatbot artifact for the process to organize business travel (Meyer von Wolff et al., 2021) (see Figure 3). We developed the user interface with common Web technologies (HTML5, JavaScript, and jQuery) to ensure it had a responsive layout. We based the chatbot itself on Node.js, TypeScript, and Rest-APIS, while we used the NLP.js-framework for natural language processing (DP_1). To provide the process functionalities, we used a finite state machine per subprocess: three general states, 69 states for the request process, and 80 states for the accounting process. Based on the results from the natural language processing, the system determines the correct state and the corresponding necessary inputs, decides the next steps, and/or adapts the dialog to the individual user and the previous inputs (DP_2, DP_3).

We structured the main user interface based on typical messenger-like systems, and it supported mobile use through responsive design (see Figure 3) (DP_1). As with other such systems, users could access a menu to log out and some control elements on the top left (1). The system also showed the progress on the current subprocess (2) at the top and at the right in the desktop view (DP_2). Thus, the system listed all necessary input categories, which turned green if the user made an input. We placed the typical chatbot input bar at the bottom of the window (3). Here, users could enter the required input and control the process or system (e.g., the user could type “return” or “back” to jump back one input or “cancel” to stop the current (sub)process (DP_1)). After a user logged in, the chatbot introduced itself, explained its purpose (4), and listed current notifications (DP_1). After the user entered an input, the chatbot queried the subsequent steps or necessary inputs (5). To do so, the chatbot passed the finite state machine and adapted the process based on the user’s inputs (DP_2, DP_3). Figure 4 displays a dialog snippet (left) and its underlying finite state machine (right) for some steps in the business travel request subprocess. Also, if the user did not make an entry for a certain time, the chatbot asked if everything was okay to generate attention (6). Users could also ask questions at any time if they found anything unclear or needed further information (DP_5).

Furthermore, besides the natural language input, we implemented several alternative input options (DP_4) to give users as much flexibility and efficiency as possible and to ensure they did not have to write everything. In implementing these options, we needed to ensure they reduced the likelihood that users would make errors as they entered information in a pre-formatted form. Thus, depending on the necessary information, users could use the following alternative input options (see Figure 5):

1) Input text by writing commands or necessary input in a messenger-like environment
2) Use a quick-reply element or type an answer when choosing between options
3) Choose from multiple selection list when selecting specified elements
4) Use a date picker to select dates and times
5) Provide receipts by attaching a file directly in the dialog window or, when using a mobile phone, by taking a photo, and
6) Confirm and sign the request and accounting forms in the dialog.

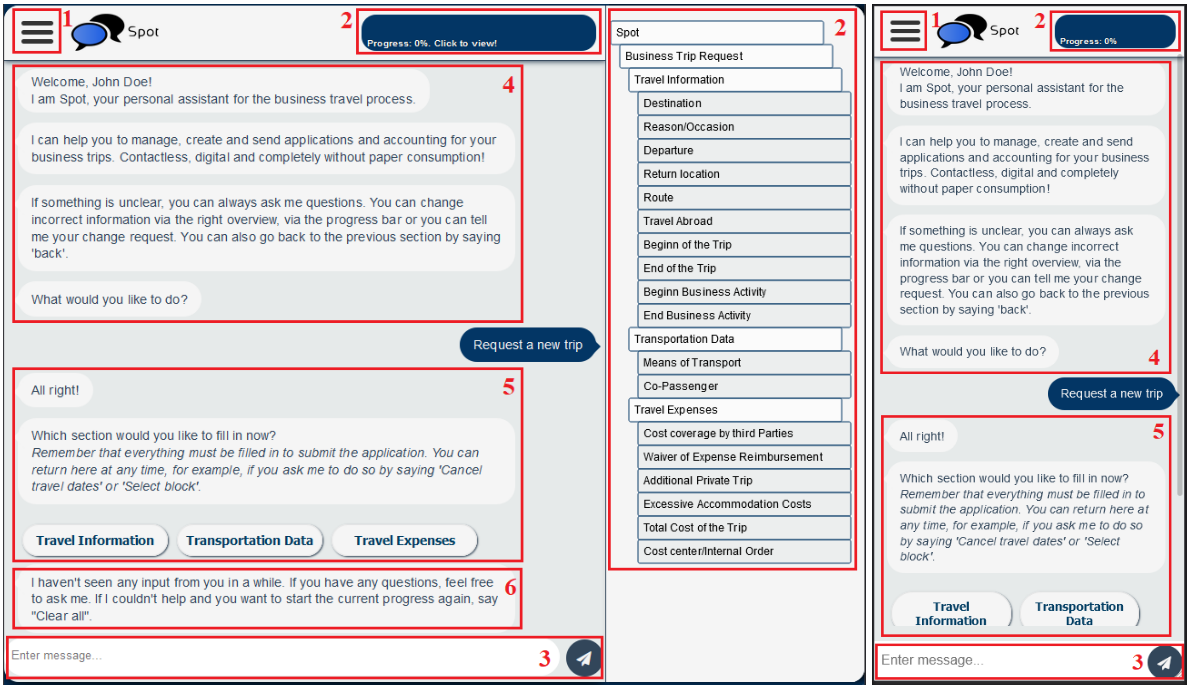

Figure 3. User Interface: Desktop View (Left); Mobile View (Right)

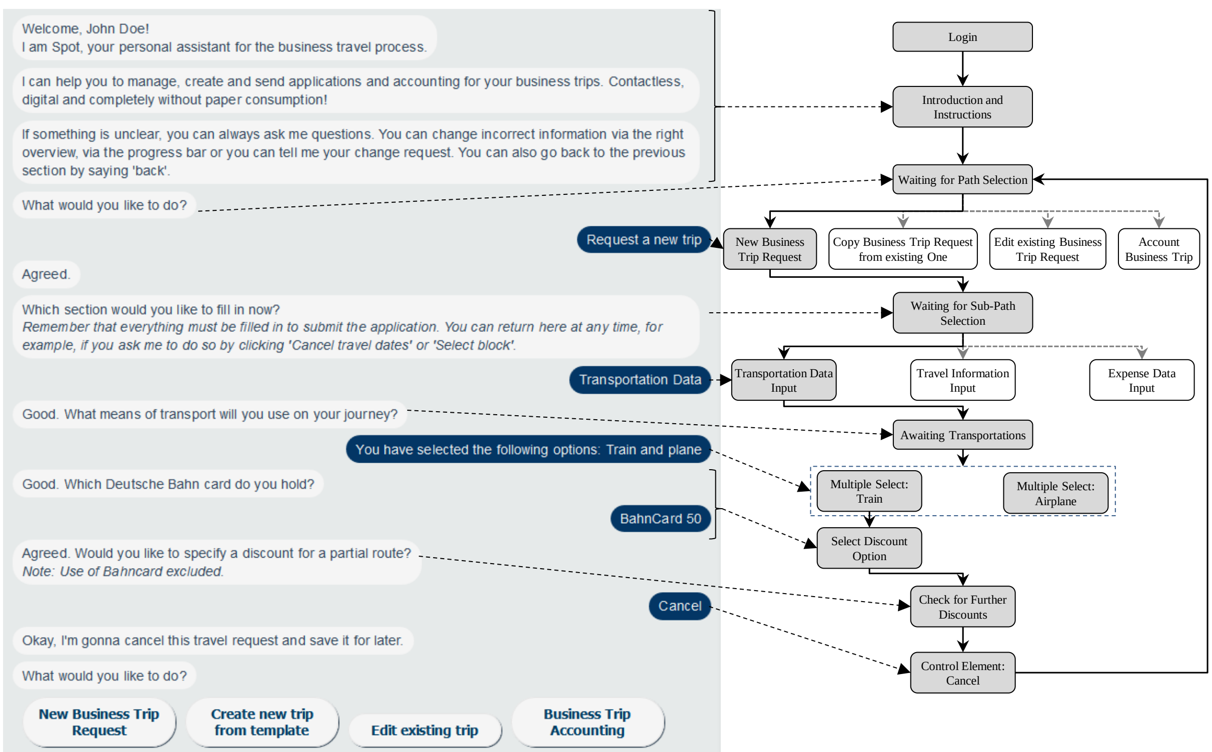

Figure 4. Exemplary Dialog Flow (Left); Corresponding Finite State Machine Excerpt (Right)

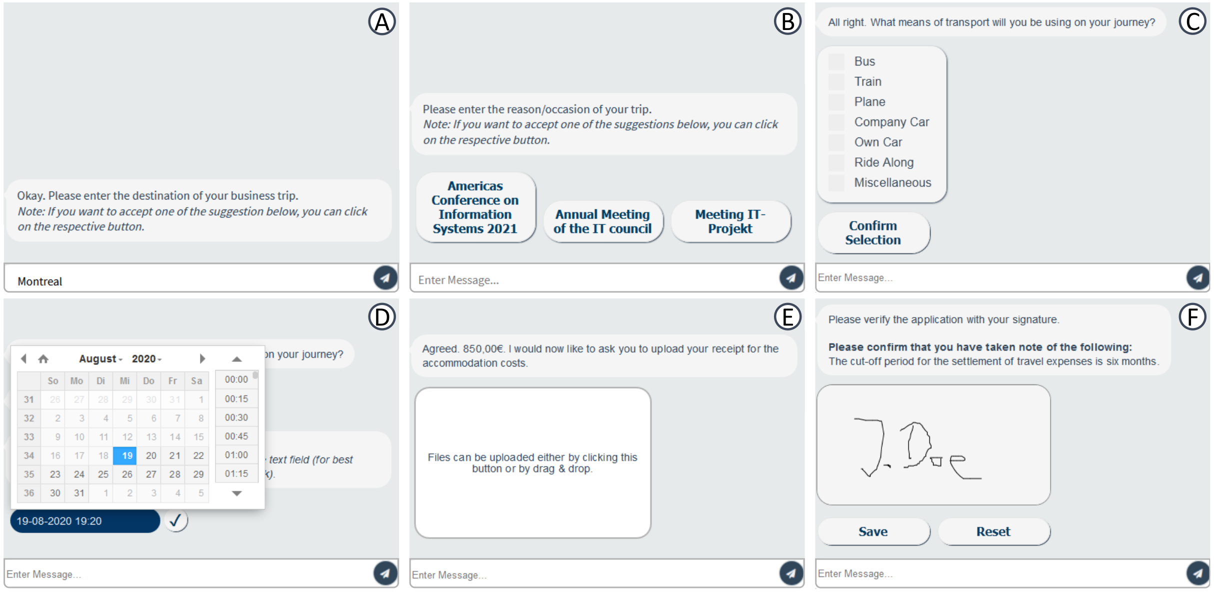

Figure 5. Input Options: (A) Text; (B) Quick Replies; (C) Multiple Selection; (D) Date Picker; (E) File Upload; (F) Signature Box

5 Evaluating the Process-based Chatbot Artifact

After we successfully developed our process-based chatbot, we needed to evaluate the concept to determine its effects and how users would perceive it. Therefore, we looked at the situation from two perspectives. On the one hand, we needed to check how chatbots' users assessed the concept and whether they accepted it, a crucial factor for the success of information systems (RQ2). On the other hand, we needed to determine a potential business value to justify using the chatbot at the company level. Otherwise, if the chatbot did not add any value, companies would not even consider using it to execute and support business processes (RQ3). Consequently, process-based chatbots can only be successful if they meet both users' and businesses' concerns.

5.1 Experimental Evaluation Approach

To survey how individuals (RQ2) and organizations (RQ3) perceived our process-based chatbot, we conducted a two-stage experimental evaluation approach with participants from three distinct user groups (see Figure 6).

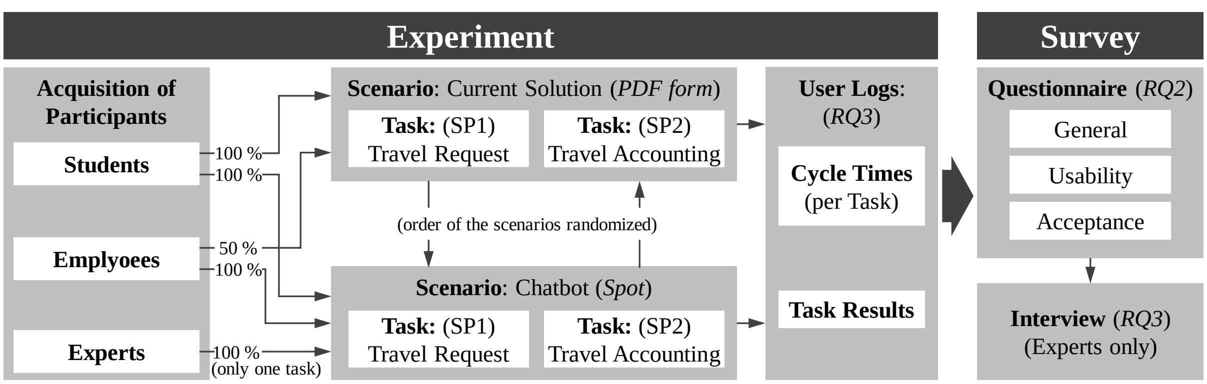

Figure 6. Research Design

5.1.1 Experiment Setting

To reflect a comprehensive most-likely cross-section participants group that corresponded to an actual company, we considered participants in three user groups (see Figure 6):

1) Local students who had little to no experience in using the sample process. This group represents novice or new employees who would enter the business world in the future.
2) Experienced employees who had experience with the sample process or similar processes and would soon face a possible system change. This group represents typical employees who organize business travel or perform similar business processes.
3) External experts (i.e., manager or project lead) who dealt with (re-)designing or practically applying chatbots in the workplace. This group represents the management level and can give practice-oriented insights from projects or make decisions in their company. Thus, the group could assess the process-based chatbot concept from an organizational level (e.g., its potential business value and cost benefits).

We conducted the experiment as a within-subject experiment (Charness et al., 2012) so we could compare the PDF form and the chatbot with each other. For this purpose, the participants used both the PDF form and the chatbot. We gave each participant an evaluation scenario for each setting (i.e., the PDF form and the chatbot). In each scenario, we asked the participants to complete a scenario in which they had to request a business trip (SP1) and do the accounting (SP2) for it. The tasks differed in content between the settings but followed the same structure. We controlled the experiment using the groups. All students used both the PDF form and the chatbot. In order not to unnecessarily disturb the employees in their actual work and since we did not focus on evaluating the chatbot rather than the process, half of the employees used both the PDF form and the chatbot and the other half used the chatbot only. For the same reason and because the experts were external partners, the experts used the chatbot only based on one of the two subprocesses (i.e., either travel request or travel accounting). We randomly assigned the employees and the experts and randomly ordered all scenarios (i.e., either the PDF form first or the chatbot first) across the groups to prevent biased results.

Lastly, we further logged the subprocess results and the cycle times for each participant to compare the existing form method and the chatbot in terms of their efficiency (i.e., process cycle times) and quality (i.e., error rates).

5.1.2 Survey Approach

Following the experiments, we conducted an evaluation survey (see Figure 6). To do so, we analyzed usage logs and conducted an online questionnaire and a short interview study.

Each participant filled out a questionnaire after each setting in the experiment (i.e., one to the chatbot) (see Table 3 and the Appendix) and one to the PDF form (the questionnaire for the PDF form only included the usability part in terms of the user experience questionnaire). The questionnaire had three parts:

1) We used some general and demographic questions to classify the participants and their results.
2) To measure the process-based chatbot's usability, we included the user experience questionnaire (UEQ) (Laugwitz et al., 2008; Schrepp et al., 2017) – a standardized and simple measure for user experience based on 26 items grouped into six scales: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty. We measured the items on a six-point Likert scale in pairs. In addition, we assessed our design principles based on one item for each design principle to verify them.
3) To address the users' perspective, we included the technology acceptance model (TAM) (Davis, 1993) and the information systems success model (ISSM) (DeLone & McLean, 2003) to use established and accepted measurement items. In doing so, we measured the potential degree to which participants would accept process-based chatbots and their satisfaction to determine the usage probability. Based on comparing available TAM and ISSM questionnaires and analyzing whether the items fit with our research goal, we included three TAM constructs (i.e., perceived usefulness, perceived ease of use, and behavioral intention to use) and three ISSM constructs (i.e., information quality, service quality, and user satisfaction). We measured the constructs with three to five items based on a seven-point Likert scale.

Lastly, we briefly interviewed only the experts given that they had a management or corporate perspective and could or needed to make the decision for actual deployment. In doing so, we discussed subjects such as the chatbot's fit for practice applications, system features, and challenges to practically assess the chatbot concept and look at it more critically.

Table 3. Evaluation Questionnaire

Part	Construct	Items	Type	Reference
(A) General	Age	1	Free text	
	Gender	1	Single choice	
	Chatbot experience	1	Five-point Likert scale	
	IT affinity	9	Six-point Likert scale	Franke et al. (2019)
	Process experience	1	Single choice	
(B) Usability	Design Principles	6	Seven-point Likert scale	
	User experience	26		Laugwitz et al. (2008), Schrepp et al. (2017)
(C) Acceptance	Information quality	4		Freeze et al. (2010), Yu & Qian (2018)
	Service quality	4		Alshibly (2014), Ojo (2017)
	Perceived usefulness	5		Davis (1989), Venkatesh & Bala (2008)
	Perceived ease of use	5		Venkatesh & Bala (2008), Venkatesh & Davis (2000)
	Behavioral intention to use	3		Constantinides et al. (2013), Venkatesh & Davis (2000)
	User Satisfaction	3		Alshibly (2014), Freeze et al. (2010), Yu & Qian (2018)

5.2 Sample Distribution

Based on the evaluation design, we recruited 69 participants (40% students, 42 percent employees, and 19 percent experts) (see Figure 7). Thus, 46 participants represented actual future users, while 13 participants represented the management's perspective (see Appendix for the industry that the experts came from). Most participants were male (60%), but across the student and employee groups, the ratio was equal (28 female, 28 male). On average, the participants were roughly 30 years old, which may have been due to the high proportion of students and younger employees. Nevertheless, we were able to attract at least one participant from all age groups. Also, our participants had a rather high IT affinity (Franke et al., 2019) (mean: 4.38) possible due to today's increasing digitalization and the growing preoccupation with digital and IT technologies. Regardless, the sample covered the complete range of IT affinity. Furthermore, 50 percent of our participants knew the existing business process, while 23 percent knew similar processes. Only 28 percent had no previous experience with the process. Lastly, most participants (57%) had used chatbots on an occasional basis, while 28 percent had less to no experience with chatbots. Around 14 percent had frequent to regular usage experience. Thus, our participants represented a good cross-section of the targeted population.

5.3 Users' Perspectives on Using Chatbots for Business Processes

Even if a system has possible business value, whether it succeeds depends above all on future users (RQ2). Therefore, we needed to consider this perspective as a precondition. One can assume that users will use a system only if it matches their expectations and they accept it. Consequently, we assessed our chatbot's usability and design from users' perspective and determined their acceptance of this new solution for the process to organize business travel.

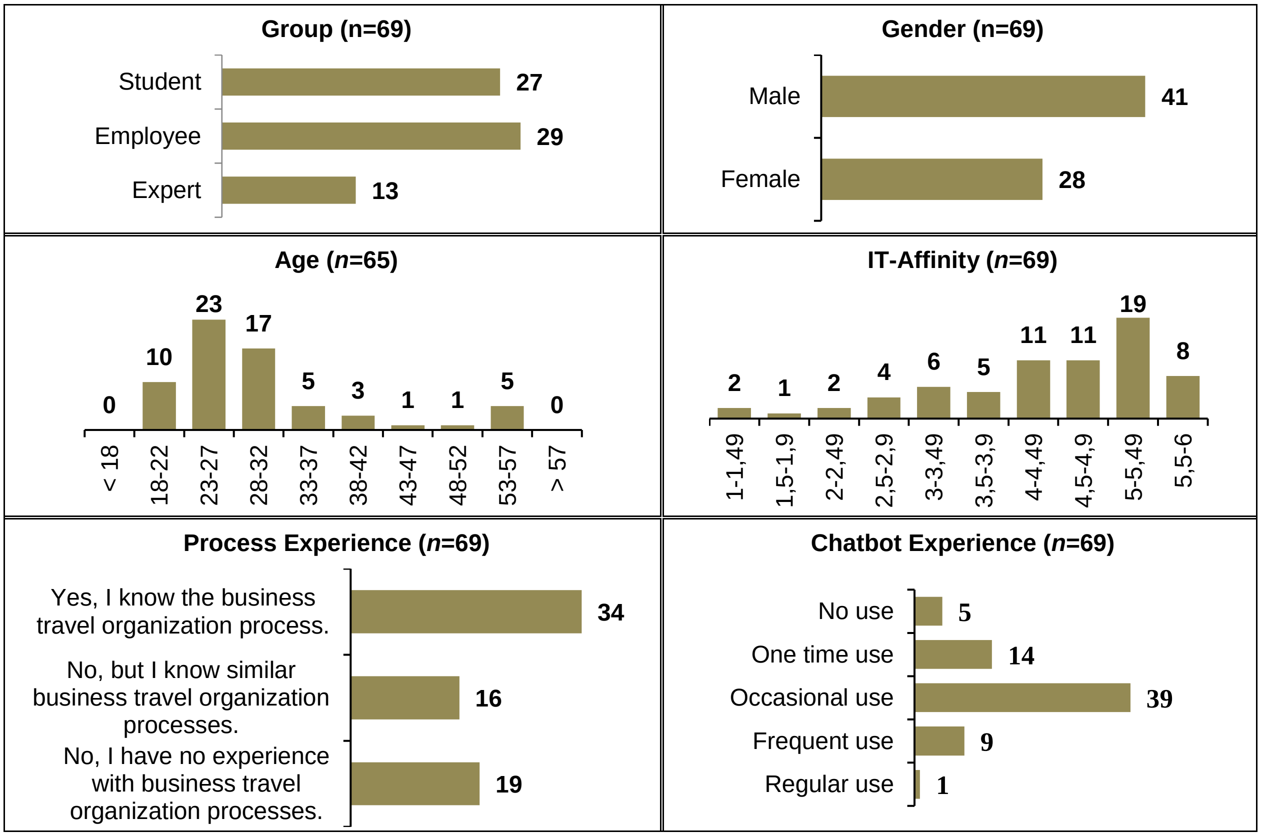

Figure 7. Sample Distribution

5.3.1 Usability of Chatbots for Business Processes

To evaluate the chatbot, we surveyed users' experience with the user interface in detail first. Specifically, we assessed the user experience in terms of attractiveness, pragmatic quality (perspicuity, efficiency, and dependability), and hedonic quality (stimulation and novelty) based on the user experience questionnaire (UEQ), which contained 26 items measured with seven–point Likert scales (see Table 3) (Laugwitz et al., 2008; Schrepp et al., 2017). Other researchers have successfully applied this approach to examine other chatbots (e.g., Holmes et al., 2019; Meyer von Wolff et al., 2020a). Except for the missing values, we analyzed the data set with the official UEQ analysis tool that the UEQ scale's authors provided and removed suggested suspicious data sets, which we identified based on the difference between an item's best and worst evaluation. As a result, we had 66 datasets for the chatbot and 35 data sets for the existing method. Additionally, we compared our results with the official UEQ benchmark data set based on 452 UEQ studies (Schrepp et al., 2017).

Overall, participants rated our artifact substantially higher than the existing system (see Figure 8, top; axis dimensions reduced from -3/+3 to -2.5/+2.5; see the Appendix for item distributions and scale consistencies). Notably, the artifact achieved excellent values based on the official benchmark, especially for perspicuity (mean: 2.06) and novelty (mean: 1.61). Thus, participants perceived the artifact as easy to learn and understand and as creative or innovative. They also perceived its efficiency (e.g., fast and efficient in solving tasks without unnecessary effort) (mean: 1.65), and stimulation (e.g., as exciting and motivating) (mean: 1.42) as good. Therefore, the results confirmed the basic idea that a chatbot can provide users with a single answer and guide users through processes via natural dialog and, thus, avoid the need for them to search for solutions and instructions themselves and possibly experience problems such as information overload due to too many sources and systems. However, the participants rated attractiveness (mean: 1.59) and dependability (mean: 1.43) lower than all other aspects. These values only slightly exceeded the average scores from the benchmark's perspective. In contrast to the good ratings they provided to the chatbot, the participants rated the existing method as bad in all aspects. Notably, only dependability (mean: 0.35) attracted a positive value. Thus, from user experience, the chatbot did not perform worse than the existing one and, therefore, could both keep up with the previous one and even implement the process in a much more appealing way. Hence, from a user experience perspective, users would likely use the chatbot over the existing solution.

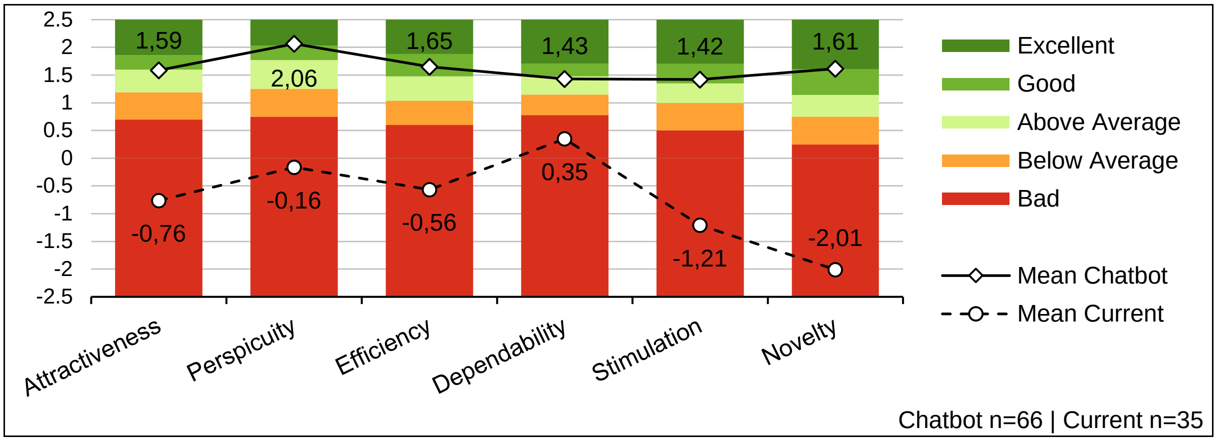

Figure 8. UEQ Distribution including the Official UEQ Benchmark (Schrepp et al., 2017)

5.3.2 System Design of Chatbots for Business Processes

Further, we evaluated our six design principles for process-based chatbots with seven-point Likert scales (1: very negative; 7: very positive) and tested whether the results differed significantly from the Likert scale's mean using a one-sided t-test (see Figure 9).

Our results showed that participants rated all design principles positively and all average values significantly differed. In particular, they assessed process guidance (DP_2) (mean: 6.45) and context-dependent input options (DP_4) (mean: 6.07) positively. Thus, the results identified process guidance combined with suitable input options, depending on the respective required information, as viable for process-based chatbots. Also, automated error handling (DP_6) (mean: 5.78), natural language user interface (DP_1) (mean: 5.74), and process adaption to the individual user (DP_3) (mean: 5.72) seemed useful for implementing the process in its existing form. However, the integrated help function (DP_5) (mean: 5.36) could be improved in the current version even though this value exceeded the mean and participants still rated it well.

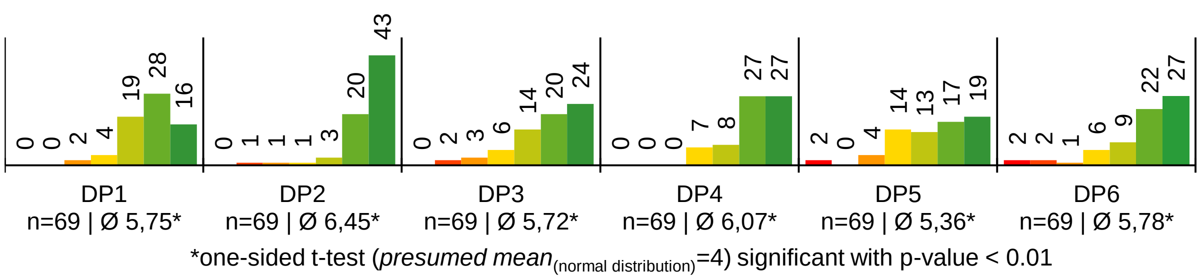

Figure 9. Design Principle Distribution

Based on the interviews with the experts, we also confirmed our deduced design principles for process-based chatbots (see Table 4). The experts considered the user interface (DP_1) to be quite useful as users could easily and naturally run it without having to deal with the system behind it. In addition, the chatbot reduced user effort and, thus, allowed the users to focus on important aspects. In particular, the experts reported that the overview function seemed appropriate for process-based chatbots (DP_2). Furthermore, they asserted that the suggestions seemed viable for individualized adaptation (DP_3) and would allow users to reuse previous input and further reducing effort. In addition, they confirmed that the input options supported users as they showed at a glance what kind of input users needed to provide (DP_4). Hence, users could perform functionalities faster and did not have to rely on textual input. Also, the experts agreed that, if questions arose, the chatbot should be able to provide an integrated help function (DP_5). In this way, users could directly ask questions and get solutions or clarifications easily. Lastly, the experts supported our belief that the system needed error handling to prevent processes from being interrupted (DP_6). Notably, speech processing can further support error handling by automatically correcting typos.

Table 4. Exemplary Interview Quotes for the Design Principles

Design principle	Exemplary quotes
DP_1: Natural language user interface	"*For a chatbot that has to be the case*" "*It is helpful because it translates what it actually does into my language. [In processes] are a lot of things and I don't really care what's behind them. But what I want is someone to ask me when did you leave, what did it cost, and so on*"
DP_2: Process guidance including progress overview	"*It takes away the problem that I forget something, that I do something wrong, and that in the end, the person who gets it is satisfied with the result, […] without me having to think about what I actually have to do. So it makes the task easier, that I get these guidelines*" "*What always annoys me about chatbots…is that I don't know where I am actually in my process.... What it did well was that I had this organizer on the right. I always knew this is the structure…and I can follow this structure and the chatbot actually just supports me in filling out these individual things*"
DP_3: Individualized process adaptation to the user	"*The chatbot already knows something about me and saves me work and effort*" "*Characteristics where patterns happen in my preference.... I think they can be recognized and suggested to me next time…as a preference*"
DP_4: Context-dependent input options	"*That's what the user expects and also takes away a bit of this cognitive load, ...it guides me much better.... I see that and I know immediately what I have to do. The task is just intuitively clear to me*" "*I also have these things with a normal web form...and if you now only offer chat, because that's cool...that's super inefficient and it has to be broken up as often as possible. Ideally, the chatbot knows exactly what I want and it only asks me questions and I say "Yes". Since this is not possible in its entirety, …support me with intelligent input options…that I don't always have to write text.*"
DP_5: Integrated help function via dialog	"*If questions arise they are answered as briefly and concisely as possible and very quickly and also in the immediate context of the process support of the tool. ...Preferably still in the dialog*" "*You are in a process right now and need this knowledge now. The knowledge is already somewhere else. I'll put it right into your process. I don't have to look in another place*"
DP_6: Automatic error handling	"*You always have a return because something is missing or because you have entered something wrong…. This is very good at least from an IT point of view, but also for the user*" "*Validations in any form help in any system, same with chatbots. It is also good if you have an NLP framework behind it, which also filters out typos and so on*"

5.3.3 Acceptance of Chatbots for Business Processes

In this section, we focus on surveying the extent to which users accepted the chatbot and considered their satisfaction with it. Measuring actual usage requires a long-term evaluation and technical and organizational changes such as installation, integration, and introduction of the chatbot in the productive operation, which one cannot achieve with a prototypical artifact. Therefore, we examined the process-based chatbot for its suitability and, thus, can only determine the probability that one would use it. To do so, we rely on applied measurement items from TAM and ISSM available in scientific research. In particular, we applied the TAM constructs – perceived usefulness (PU), perceived ease of use (PEoU), and behavioral intention to use (BI) – and the ISSM constructs – information quality (IQ), service quality (SQ), and user satisfaction (US). In comparing available TAM and ISSM studies in current research, we noticed that the constructs for perceived ease of use, perceived usefulness, and behavioral intention to use (TAM) also represented the items representing usage intention and actual use (ISSM). We found the same situation with system quality (ISSM) and perceived ease of use (TAM). Thus, to avoid ambiguity, we only used the TAM constructs and their items. We measured the resulting six constructs based on 24 items with 7-point Likert scales (see Table 3; 1: do not agree; 7: fully agree; see the Appendix for the detailed applied items). Furthermore, we aggregated the distributions for each item to the distribution of the constructs (see Figure 10 at the top; we evaluate the items in detail in the Appendix) and tested whether the results significantly differed from the Likert scale's mean using a one-sided t-test.

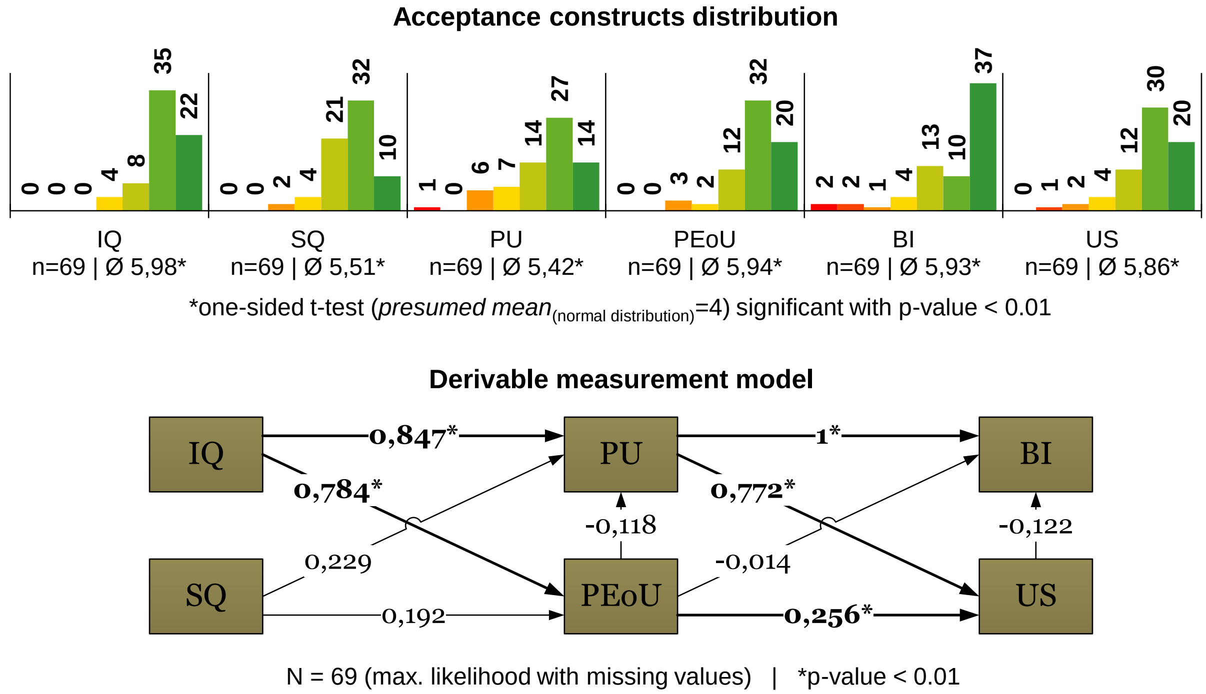

Figure 10. Evaluation of the Acceptance Constructs and the Derivable Measurement Model

In particular, we found positive ratings for IQ (mean: 5.98), PEoU (mean: 5.94), and BI (mean: 5.93). Thus, the system could provide relevant information, users could learn to handle it easily, and users would even recommend the system to others. In particular, the results for PEoU coincided with those for the UEQ part of the evaluation in which perspicuity also received quite high ratings. From the acceptance perspective, however, the constructs SQ (mean: 5.51), PU (mean: 5.42), and US (mean: 5.86) received lower but still good ratings. In particular, the participants criticized the support that the chatbot gave in case problems arose, increased productivity, and conformity to expectations. Nonetheless, our results indicated a potential actual system use as participants rated the constructs quite highly and, therefore, seemed to accept the chatbot. Notably, they rated BI highly, which indicates that they accepted the chatbot and would likely use it. Also, the overall US received quite high ratings, which confirms the positive results from the UEQ. Thus, the results indicate that the participants were rather satisfied with the process-based chatbot and reported probable future usage. Consequently, we seemed to have ensured acceptance and, thus, applicability from the user perspective. We did not identify results that indicated any risk that users would not use the chatbot and, thus, that it lacked applicability from their perspective.

Additionally, based on the applied items and their constructs, we derived a measurement model with help from the original TAM and ISSM (Davis, 1993; DeLone & McLean, 2003). Consequently, our proposed derived measurement model comprises our six applied constructs (see Figure 10 at the bottom). Based on the evaluation results, we could deduce initial insights on the respective influences using STATA for calculating the structural equation model. We identified only some significant relations: in particular, IQ had a positive influence on PU and PEoU. Moreover, the BI only depended on PU. Lastly, PU and PEoU positively influenced US, which demonstrates that a process-based chatbot can achieve user satisfaction. Since PU had had a greater influence than PEoU on US, we can conclude a chatbot's ability to conduct a given business process in a digital workplace environment has more importance than its usability aspects.

5.4 Organizational Perspectives on the Business Value of Chatbots for Business Processes

Following the positive results we achieved from an individual perspective, we next examined chatbot's business value to justify their operation from an organizational perspective (RQ3). Usability and acceptance from the user perspective do not necessarily justify deploying technology at a company level. Rather, organizations typically only adopt such technologies if they contribute economically or they perform at a similar or better level than existing solutions. Therefore, we measured our developed chatbot's process efficiency and quality based on lead time and errors that occurred and considered the experts' opinions on chatbot usage.

5.4.1 Process Efficiency: Comparison of Lead Times

To measure process efficiency, we first calculated the lead time for the two exemplary processes depending on the setting (see Figure 11). Except for missing values, we cleansed the data by removing the extreme outliers ($> 3 \times interquartil\ range$) from each data set until only mild outliers remained (Tukey, 1977), which resulted in fewer valid datasets than actual participants.

The results showed that the users could conduct the business travel request process scenario (SP1) in a similar timeframe with or without the chatbot (median: 12:45 minutes without it to 12:48 minutes with it) (see Figure 11). For the scenario process business travel accounting (SP2), however, the participants performed more quickly with the chatbot by 3:20 minutes (median: 14:40 minutes to 11:20 minutes). Thus, because each setting included accounting as the second task, we assert that the chatbot supported a high learning effect. Notably, the mean time between SP1 and SP2 decreased compared to the existing method. The total lead time confirmed these results. In total, the participants needed 24:45 minutes (median) for both subprocesses with the chatbot, while they needed 27:08 minutes (median) with the existing method. Thus, our results indicate that chatbots can keep up with and even undercut previous lead times, especially for as yet not (fully) digitalized processes as in our example.

Additionally, we compared the participants' lead times with their IT affinity (see Figure 11). We found that IT affinity seemed to have little to no influence on lead time. We found that it had a small positive influence only for SP2. Thus, our results showed that IT affinity had no effect on whether the participants could conduct chatbot-based processes, which would prove advantageous for new and older employees in particular. Also, our results suggest that one needs no specific IT affinity level to carry out work tasks with process-based chatbots.

Furthermore, we compared the participants' experience with the process with their respective lead times (see Figure 11). Regardless of their personal experience level, we identified comparable lead times were. Only participants who knew similar processes had higher lead times for the accounting task. As expected, those without previous experience had the highest lead times for SP1. Interestingly, however, they had the lowest lead times for SP2, which further supports our statement on chatbots' learning effects.

5.4.2 Process Quality: Risk of Faulty Processes

To survey the process quality, we analyzed the results for each setting and scenario in terms of correctness and errors (see Table 5; we removed incomplete evaluations). For this purpose, we checked the individual process results (i.e., submitted requests or accountings) for both the PDF form and chatbot. For each scenario, we classified the individual result as error free or faulty. We classified the result as error free if it included all entries and all critical entries matched the correct response. We classified the result as faulty if it included missing or wrong entries (i.e., missing signature, missing attached supporting documents, or non-valid dates). We further divided the errors into critical errors (that stopped or refused the process) and false positives (that one could check only with background knowledge, such as entering the destination Los Angeles rather than San Francisco).

For SP1, the chatbot achieved more correct requests than the existing method, which resulted in a difference of 17.47 percent. For SP2, this difference increased to 30.85 percent in the chatbot's favor. As for erroneous processes, the chatbot reduced the number of refusals (SP1: -47.14%, SP2: -36.10%). However, it also increased the number of false positives (e.g., forgotten optional inputs), erroneous entries (e.g., wrong start or destination), and selections due to insufficient process knowledge (i.e., incorrect transport mode) (SP1: ~30%, SP2: ~5.24%). The chatbot cannot validate all of these errors automatically. Evidently, our chatbot could not represent the entire process as humans still needed to check the proposals, but it did reduce the number of fundamentally incorrect or incomplete processes. Thus, these results suggest the need to further refine the chatbot and identify possible problems. In some cases, for example, it seems advisable to rephrase the chatbot's questions to avoid input errors in users' answers.

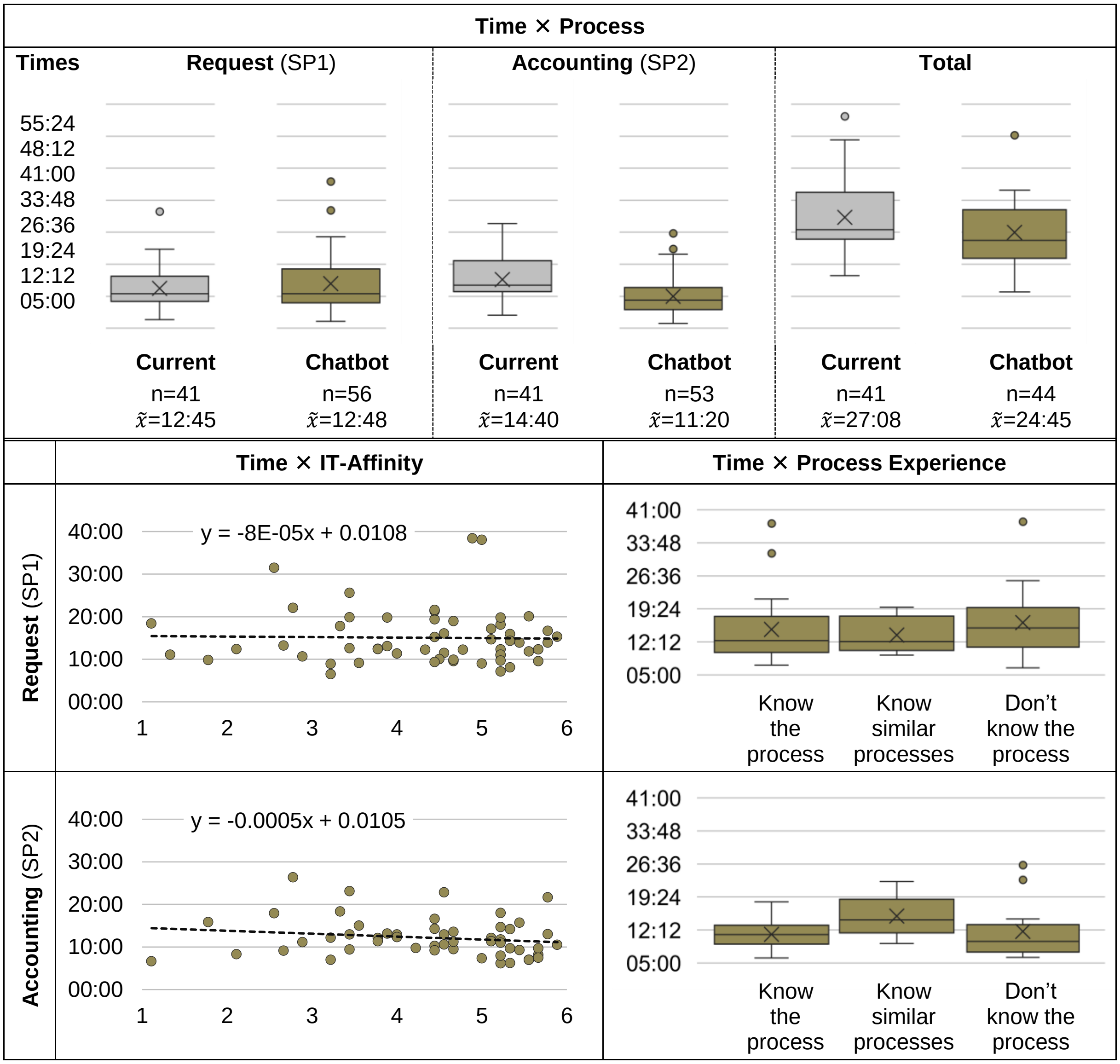

Figure 11. Lead Time Distributions

Table 5. Process Outcome Quality

		N	Correct	Erroneous	Δ%	Thereof refusals	Δ%	Thereof false positives	Δ%
SP1	Current	41	31.71%	68.29%		48.78%		19.51%	
	Chatbot	61	49.18%	50.82%	↓ 17.47%	1.64%	↓ 47.14%	49.18%	↑ 29.67%
SP2	Current	41	34.15%	65.85%		56.10%		9.76%	
	Chatbot	60	65.00%	35.00%	↓ 30.85%	20.00%	↓ 36.10%	15.00%	↑ 5.24%

5.4.3 Qualitative Reflection by the Experts

Lastly, we asked our experts (see Figure 6) to reflect on our chatbot concept and assess its suitability from an actual real-case business perspective in the interviews.

Notably, all of our participating experts (n = 13) agreed on the chatbot's general suitability. In particular, one reported that "nothing is really missing; I think that the way it is designed, it is also well suited for practice". Especially, one participant presented a viable application scenario for chatbots in digital workplaces:

> *[They should target processes] that quite a lot of people do, also do the same, but not frequently. So from each individual not frequently...that have many branches, which means many alternative paths and that you only run very rarely, that you don't get lost somewhere*.

In a related manner, some experts noted that chatbots could be a burden, especially for processes that employees perform frequently. For example:

> *I wouldn't use the chatbot if it would slow me down and that, that might be the case if you do such a process three or more times a week. Then you actually know which fields you click on, then the option would be faster for me that I have an input field or a form and do not have to be led through such a dialog again…*

In particular, one participant stated that users high in IT affinity or "power users" would rather rely on other systems than a chatbot in noting that such users "just want to click through the form, and that quickly, and not always have to wait for a response from the chatbot".

Crucially, the experts viewed the effort required to reach the desired quality level critically. For example: one said:

> *From an economic point of view, you have to think carefully about where you use it.... Let's take business trips, that's certainly cool to have, but I think a well-designed business trip form does this as well and you can probably run that with one developer. Whereas with a business trip chatbot, you're going to have to do a lot of work. But there are other processes, for example, the support process at Amazon, which always starts with a chat. I don't think there's any question at all as to whether it's worth it.*

In this context, the experts also brought up continuous care and maintenance. For example, one said:

> *Sooner or later, the thing dies if you don't put work in there reasonably, and the work you put in is much greater than what you need to set it up at the beginning*.

Hence, one expert concluded:

> *What I think is a hindrance at the moment is just the effort that chatbots take to make them good*.

Thus, despite chatbots' potential positive effects, the interviews revealed that companies must select use cases of chatbots carefully. Furthermore, companies must be aware that the development and use of chatbots involves a considerable amount of work (e.g., maintenance, NLP training).

6 Discussion

6.1 How should one Design Enterprise Process-based Chatbots to Execute and Support Business Processes?

Based on our DSR project, we identified six design principles and pointed out how one should design a chatbot to ensure it can execute and support processes to organize business travel (RQ1). In this study, we could demonstrate that a chatbot based on our deduced design principles could, from a technical perspective, execute and, thus, support the chosen process to organize business travel. Based on subsequently analyzing our design principles, we concluded that the chatbot could also execute general business processes in digital workplaces. Therefore, we recommend that chatbots for business process applications, besides using natural language-based user interaction (DP_1), should map and adapt processes based on each individual user, the current dialog, or previous conversations (DP_2, DP_3). It seems viable to enable content-related input options to foster efficiency during the process and offer as much flexibility as possible without forcing the user to type every command or make every necessary input (DP_4). Thus, the relevant processes must be mapped to enable different possibilities for each user. In addition, since processes require correct data (especially in business settings), chatbots need to automatically handle errors (DP_6). Besides these process-relevant requirements, we also propose that chatbots should provide direct individualized help to each user (DP_5). However, as we only developed a prototypical chatbot for the process to organize business travel, readers should examine the extent to which our findings generalize to other contexts critically. As we outline in Section 4.2, we selected the process because it fulfills many typical characteristics that self-service processes in particular and business processes in general possess (e.g., regular usage, defined amount individual steps/tasks, need correct inputs). Thus, we expect our results to generalize to similar processes in specific and process-based chatbots in general.

Notably, in this study, we confirm existing chatbot design recommendations as we note in Section 4.3.4 (e.g., Bittner & Shoury, 2019; Gnewuch et al., 2017; Tavanapour et al., 2019) and extend them with process-specific requirements. In particular we found similar results as existing chatbot research in terms of the user interface and general requirements concerning natural language processing. However, our results go beyond general fundamental requirements for chatbots. We also demonstrated in one case that chatbots can execute complex internal business processes. In this case, we demonstrated that process-based business chatbots can be designed similarly to chatbots in other use cases (like Q&A chatbots). However, process-based chatbots differ regarding their process characteristics and how they implement step-by-step process guidance (DP_2). In order to enable a proper process support, the chatbot must be adapted to the respective process characteristics in every use case.

Our results show that many research findings from prior chatbot studies in other use cases can be transferred to process-based chatbots as well. In particular, some general design recommendations seem valid for applying chatbots in business contexts and, in particular, business processes (e.g., from Feine et al., 2020; Diederich et al., 2020). Nevertheless, each company must examine and address the peculiarities of its respective processes individually. Additionally, companies should ensure that they provide a suitable user experience. When implementing process-based chatbots, for instance, humanizing dialog (Liebrecht & van Hooijdonk, 2020) and applying anthropomorphic features (Diederich et al., 2020) seem to be important. Overall, our study supports the assertion that one can use chatbots to execute and support business processes. In specific, we extend the current scientific knowledge base with new insights into design principles for process-based chatbots.

6.2 RQ2: How do Users Assess an Exemplary Process-based Chatbot for Internal Business Processes?

From the user perspective (RQ2), we show that many participants (~71%) had already used chatbots on at least an occasional basis, possibly from private use, and, therefore, had experience with conversational user interfaces and chatbots. Thus, organizations may not face the risks or need to train staff as they often do when introducing new systems.

In this study, we found that our chatbot exhibited rather high usability for the given business processes. Several factors exhibit usability. First, it should successfully conduct the evaluated processes in a timely manner. Therefore, if our chatbot had poor usability, we would not have achieved such good results. Additionally, as we identified with the UEQ benchmark (Schrepp et al., 2017), participants perceived our chatbot as user friendly to a substantial degree. In particular, perspicuity and novelty received the highest ratings. The ratings for novelty differed the most between the chatbot and the existing method.

Furthermore, we evaluated our design principles and received good feedback on them. The evaluation results suggest confirmed the design principles as relevant for designing process-based chatbots. Thus, we confirmed and verified our derived design principles and demonstrated their applicability to process-based chatbots. In particular, process guidance and context-dependent input options both received the highest scores. However, future studies should further validate the design principles to determine whether they generalize to other contexts or if they need further adaptation.

Based on our derived measurement model, our results indicate that users are willing to use chatbots for business processes. Since acceptance constitutes a rather critical success factor for application systems, our results further support the assumption that chatbots can be used to support business processes. . Our derived measurement model based on TAM and the ISSM (Davis, 1989; DeLone & McLean, 2003) can serve as a starting point for further studies (in research and practice). We observed that usefulness rather than ease of use mostly influenced user satisfaction. Thus, in business applications, chatbots need the ability to map and execute the given tasks or processes, and their appearance has less importance as Rietz et al. have already shown (2019).

However, as we based our DSR artifact on a rather outdated starting basis (an existing process to organize business travel), one could argue that it automatically led to better results. Nevertheless, we did not focus on replacing the existing process. Rather, we evaluated users' perspectives on a chatbot that conducts business processes. With a better baseline, the results would probably not be quite so far apart. In this context, we also mention the risk of novelty effects (Winkler & Söllner, 2018) since a process-based chatbot in the developed form does not yet exist. This fact may have positively influenced the results and, thus, suggest the need for long-term assessments. Regardless, the result remains the same: one can basically use chatbots to execute and support business processes, and users perceived the concept well.

Lastly, we deliberately gave little consideration to the human or anthropomorphic elements that the literature has frequently discussed (e.g., Diederich et al., 2020; Liebrecht & van Hooijdonk, 2020) because we focused on the applicability, usability, and resulting benefits of a chatbot for executing and supporting business processes. Indeed, Rietz et al. (2019) already highlighted such a focus as the most important factor in a working environment rather than anthropomorphic features. Nonetheless, our results show that the chatbot had good usability and levels of human behavior. Thus, our results represent a good starting point for future studies. Following the literature, researchers may achieve even better results by focusing more on humanness and anthropomorphism.

We can summarize our findings on the users' perspective (RQ2) with three propositions:

1) Users perceive process-based chatbots as more usable than traditional enterprise systems.
2) Users accept process-based chatbots when the chatbots actually support them in complex business process.
3) Users feel more satisfied with process-based chatbots than with traditional enterprise systems.

6.3 RQ3: What Business Value do Process-based Chatbots for Business Processes have from an Organizational Perspective?

From an organizational perspective, we demonstrate that our chatbot had similar or better process cycle times compared to existing enterprise systems. Therefore, we conclude that using chatbots has higher learning effects than classic systems. However, we found that existing individual IT affinity and knowledge of the evaluation process, at least in our evaluation case, only had a marginal influence on process execution. Hence, all users (no matter their personal knowledge) can conduct the given process with the chatbot. This finding further confirms our chatbot's capability to support in processes in which users have little to no experience.

We found similar results concerning the likelihood that our chatbot would fail. We observed that the chatbot reduced the likelihood that process aborts or restarts (i.e., critical errors) would occur by approximately 17 and 31 percent per task, respectively, compared to the existing method. Thus, the tasks may have been clearer with the chatbot because users only needed to answer questions, which reduced information overload, and could receive targeted assistance if they had any issues. Nonetheless, we determined that the chatbot also increased the danger that false positives could occur. Even classical systems find identifying such errors difficult, and they usually require human review. Hence, we interpret this result as a signal to adjust the prototype (e.g., conduct a second design cycle and pay more attention to the chatbot's messages and instructions).

Nonetheless, our results confirm what Manseau (2020) found, especially regarding the positive outcomes on efficiency and productivity. However, the results would probably not differ that much with a better baseline (e.g., different process implementation). From a business perspective, we also show that chatbots represent a suitable option for executing and supporting business processes and can keep up with the status quo.

Overall, we could summarize our findings on the organizational perspective (RQ3) with three propositions:

1) Users can execute business processes in a similar or faster time when using process-based chatbots compared to existing solutions.
2) Users' IT affinity or existing process knowledge has no influence on their ability to use process-based chatbots to execute the corresponding business process.
3) Users reduce the risk that they will make critical errors or experience interruptions when using process-based chatbots to execute a business process.

Nevertheless, one should view our results critically. As the experts we interviewed noted, even if chatbots achieve positive effects, the effort required to create and maintain them somewhat counters the effects. Still, companies should consider using chatbots as a means to improve work quality even though they may not make a monetary contribution or reduce costs. Therefore, our results indicate process-based chatbots might have a positive effect for users and companies. Users might benefit from better process support and better user experience while companies might benefit from added business value. Decision makers must compare users' and companies' perspectives with how much it will cost their organization to acquire and operate chatbots to assess their options and make a decision that all involved or responsible parties can stand behind. In doing so, companies can hopefully avoid unnecessary investments in developing and implementing chatbots that do not provide suitable benefits for themselves and users. In addition, as we only evaluated one point in time, the question arises as to whether the results hold with recurrent use or rely on novelty effects.

We found that users without previous knowledge can use chatbots. Compared to other information systems, chatbots might have a reduced usage barrier (i.e., reduced training time due to individual guidance). They represent an especially useful option for users who occasionally or rarely conduct the corresponding process. However, as some experts mentioned, a chatbot may hinder power users as they know the process in detail and do not need process guidance or input validation. Experienced users may also feel forced into the dialog and slowed down in the process. Nonetheless, based on our data, we could not identify substantial differences in process efficiency and quality between the participant groups (i.e., novice vs. power users). Therefore, we argue that researchers still have much work to do in this area (e.g., does one need to design chatbots differently for power users than for novice users?). It also seems necessary to clarify how one can encourage users to use chatbots and whether a chatbot represents the right system for such a use case (e.g., organizing business travel). Depending on the context, perhaps future users should be given a choice regarding which system they want to use. For instance, inexperienced users could use a chatbot, whereas experienced users could choose between a chatbot and traditional systems.

6.4 Limitations

Despite our positive results, our study has some limitations. First, we surveyed a process-based chatbot application based on only one exemplary case. To obtain more generalizable results about designing process-based chatbots, researchers must test our design principles in other application scenarios by implementing corresponding artifacts. Especially, if the processes differ significantly from our analyzed process, that would hamper our results' generalizability. Nevertheless, our results indicate that chatbots generally constitute a suitable method to execute and support processes to organize business travel in particular and, to a lesser degree, other business processes. Therefore, we assume that our results can help practitioners to implement process-based chatbots in companies.

Second, we did not consider anthropomorphic aspects for the most part as we prioritized chatbots' capability to execute processes (Rietz et al., 2019). Thus, future studies should survey the extent to which these elements also influence process deployments. However, our results already show quite good usability, and one could theoretically improve it if one further considered anthropomorphic elements (Diederich et al., 2020).

Third, even though we completed the DSR cycle, our artifact remained in a prototype state. Hence, it had limited functionality and could still be improved. In addition, depending on the chosen scenario, one could create different artifacts based on our design principles. However, our results verify that chatbots can generally execute and support business processes and present quite good value for user experience.

Fourth, our evaluation results depend on participants, their experience with chatbots or given business processes, and their willingness to participate. Therefore, we incorporated a suitable number of participants (students, employees, and external experts) who differed in age, IT affinity level, experience with the business process, and experience with chatbots. Additionally, to further prevent bias, we randomized the order of the tasks in the evaluation among the participants as we outline in Section 5.1. However, as we could only acquire 69 participants, researchers should expand our evaluation.

Fifth, we derived our measurement model from the items, constructs, and relationships in the original TAM and ISSM because we did not focus on model development. Future research could use our preliminary model as a starting point to further evaluate process-based chatbots. However, researchers need to verify the model and its relationships in a proper model development study.

Sixth, as we have measured only at one point in time, we cannot make statements about long-term or repetitive chatbot usage. The experts also brought up this subject. Thus, we need longitudinal studies to determine the effects depending on the use duration. In this context, researchers could also address possible novelty effects.

7 Conclusion

In this DSR study, we applied a process-based chatbot to a real business process and evaluated its impacts. We demonstrated that chatbots 1) can execute and support business processes and 2) have high usability. Further, we 3) derived six design principles for process-based chatbots. Thus, this study contributes to both research and practice. For the scientific community, we contribute to closing the existing research gap by deducing design principles useable for future process-based chatbots, particularly for form-based processes such as the process to organize business travels. Our propositions also provide starting points for future chatbot studies. For practice, we showed that chatbots can successfully execute business processes. For reducing errors and increasing usability in particular, using chatbots makes sense. Companies can use our results to evaluate possible chatbot projects to make a decision. Nonetheless, researchers must verify our results on a wider scale and with different business processes.

References

Adam, M., & Klumpe, J. (2019). Onboarding with a chat: The effects of message interactivity and platform self-disclosure on user disclosure propensity. In *Proceedings of the 27th European Conference on Information Systems.*

Alshibly, H. H. (2014). Evaluating E-HRM success: A validation of the information systems success model. *International Journal of Human Resource Studies*, *4*(3), 107-124.

Al-Zubaide, H., & Issa, A. A. (2011). OntBot: Ontology based chatbot. In *Proceedings of the 4th International Symposium on Innovation in Information & Communication Technology.*

Aquino, J. (2012). Virtual assistants enter the enterprise. *CRM Magazine*, *16*(12), 30-34.

Bavaresco, R., Silveira, D., Reis, E., Barbosa, J., Righi, R., Costa, C., Antunes, R., Gomes, M., Gatti, C., Vanzin, M., Junior, S. C., Silva, E., & Moreira, C. (2020). Conversational agents in business: A systematic literature review and future research directions. *Computer Science Review*, *36*(4).

Berg, M. M. (2014). *Modelling of natural dialogues in the context of speech-based information and control systems.* IOS Press.

Bittner, E., Küstermann, G. C., & Tratzky, C. (2019). The facilitator is a bot: Towards a conversational agent for facilitating idea elaboration on idea platforms. In *Proceedings of the 27th European Conference on Information Systems.*

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 52nd Hawaii International Conference on System Science.*

Byström, K., Ruthven, I., & Heinström, J. (2017). Work and information: Which workplace models still work in modern digital workplaces? In *Proceedings of the 9th International Conference on Conceptions of Library and Information Science.*

Carayannopoulos, S. (2018). Using chatbots to aid transition. *International Journal of Information and Learning Technology*, *35*(2), 118-129.

Chai, J., Lin, J., Zadrozny, W., Ye, Y., Stys-Budzikowska, M., Horvath, V., Kambhatla, N., & Wolf, C. (2001). The role of a natural language conversational interface in online sales: A case study. *International Journal of Speech Technology*, *4*(3-4), 285-295.

Chakrabarti, C., & Luger, G. F. (2015). Artificial conversations for customer service chatter bots: Architecture, algorithms, and evaluation metrics. *Expert Systems with Applications*, *42*(20), 6878-6897.

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, *81*(1), 1-8.

Chen, H.-L., Vicki Widarso, G., & Sutrisno, H. (2020). A chatbot for learning Chinese: Learning achievement and technology acceptance. *Journal of Educational Computing Research*, *58*(6), 1161-1189.

Constantinides, E., Lorenzo-Romero, C., & Alarcón-del-Amo, M.-C. (2013). Social networking sites as business tool: A study of user behavior. In M. Glykas (Ed.), *Business process management* (vol. 444, pp. 221-240). Springer.

Corea, C., Delfmann, P., & Nagel, S. (2020). Towards intelligent chatbots for customer care—practice-based requirements for a research agenda. In *Proceedings of the* 53rd Hawaii International Conference on System Sciences.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319-340.

Davis, F. D. (1993). User acceptance of information technology: System characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies*, *38*(3), 475-487.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, *19*(4), 9-30.

Dery, K., Sebastian, I. M., & van der Meulen, R. (2017). The digital workplace is key to digital innovation. *MIS Quarterly*, *16*(2), 135-152.

Diederich, S., Brendel, A. B., & Kolbe, L. M. (2020). Designing anthropomorphic enterprise conversational agents. *Business & Information Systems Engineering*, *62*(3), 193-209.

Diederich, S., Brendel, A. B., Morana, S., & Kolbe, L. M. (2022). On the design of and interaction with conversational agents: An organizing and assessing review of human-computer interaction research. *Journal of the Association for Information Systems*, *23*(1), 96-138.

Diederich, S., Janßen-Müller, M., Brendel, A. B., & Morana, S. (2019). Emulating empathetic behavior in online service encounters with sentiment-adaptive responses: Insights from an experiment with a conversational agent. In *Proceedings of the 40th International Conference on Information Systems.*

Elshan, E., & Ebel, P. (2020). Let's team up: Designing conversational agents as teammates. In *Proceedings of the 41th International Conference on Information Systems.*

Engel, C., Elshan, E., & Ebel, P. (2021). Deploying a model for assessing cognitive automation use cases: Insights from action research with a leading European manufacturing company. In *Proceedings of the 54th Hawaii International Conference on System Sciences.*

Feine, J., Adam, M., Benke, I., Maedche, A., & Benlian, A. (2020). Exploring design principles for enterprise chatbots: An analytic hierarchy process study. In *Proceedings of the 15th International Conference on Design Science Research in Information Systems and Technology.*

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A taxonomy of social cues for conversational agents. *International Journal of Human-Computer Studies*, *132*, 138-161.

Feng, S., & Buxmann, P. (2020). My virtual colleague: A state-of-the-art analysis of conversational agents for the workplace. In *Proceedings of the 53rd Hawaii International Conference on System Sciences.*

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

Følstad, A., & Skjuve, M. (2019). Chatbots for customer service: User experience and motivation. In *Proceedings of the 1st International Conference on Conversational User Interfaces.*

Franke, T., Attig, C., & Wessel, D. (2019). A personal resource for technology interaction: Development and validation of the affinity for technology interaction (ATI) scale. *International Journal of Human-Computer Interaction*, *35*(6), 456-467.

Freeze, R. D., Alshare, K. A., Lane, P. L., & Wen, H. J. (2010). IS success model in e-learning context based on students' perceptions. *Journal of Information Systems Education*, *21*(2), 173-184.

Gnewuch, U., Morana, S., Adam, M., & Maedche, A. (2018). Faster is not always better: Understanding the effect of dynamic response delays in human-chatbot interaction. In *Proceedings of* the 26th European Conference on Information Systems.

Gnewuch, U., Morana, S., & Maedche, A. (2017). Towards designing cooperative and social conversational agents for customer service. In *Proceedings of the 38th International Conference on Information Systems.*

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly*, 37(2), 337-355.

Gregor, S., Kruse, L., & Seidel, S. (2020). Research perspectives: The anatomy of a design principle. *Journal of the Association for Information Systems*, *21*(6), 1622-1652.

Gyton, G., & Jeffsry, R. (2017). These are the experts deciding the future of HR … shouldn't you know who they are? *People Management*, 24-31.

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems*, *19*(2), 1-6.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, *28*(1), 75-105.

Hobert, S. (2019a). How are you, chatbot? Evaluating chatbots in educational settings—results of a literature review. In N. Pinkwart & J. Konert (Eds.), *Fachtagung bildungstechnologien* (pp. 259-270). Gesellschaft für Informatik.

Hobert, S. (2019b). Say hello to "coding tutor"! Design and evaluation of a chatbot-based learning system supporting students to learn to program. In *Proceedings of the 40th International Conference on Information Systems.*

Holmes, S., Moorhead, A., Bond, R., Zheng, H., Coates, V., & Mctear, M. (2019). Usability testing of a healthcare chatbot: Can we use conventional methods to assess conversational user interfaces? In *Proceedings of the 31st European Conference on Cognitive Ergonomics*.

Janssen, A., Passlick, J., Rodríguez Cardona, D., & Breitner, M. H. (2020). Virtual assistance in any context. *Business & Information Systems Engineering*, *62*(3), 211-225.

Johannsen, F., Leist, S., Konadl, D., & Basche, M. (2018). Comparison of commercial chatbot solutions for supporting customer interaction. In *Proceedings of the 26th European Conference on Information Systems.*

Köffer, S. (2015). Designing the digital workplace of the future—what scholars recommend to practitioners. In *Proceedings of the 36th International Conference on Information Systems.*

Lacity, M., & Willcocks, L. (2021). Becoming strategic with intelligent automation*. MIS Quarterly Executive*, *20*(2), 1-14.

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In A. Holzinger (Ed*.), HCI and usability for education and work* (LNCS vol. 5298, pp. 63-76). Springer.

Laumer, S., Gubler, F., Racheva, A., & Maier, C. (2019). Use cases for conversational agents: An interview-based study. In *Proceedings of the 25th Americas Conference on Information Systems.*

Lechler, R., Stoeckli, E., Rietsche, R., & Uebernickel, F. (2019). Looking beneath the tip of the iceberg: The two-sided nature of chatbots and their roles for digital feedback exchange. In *Proceedings of the 27th European Conference on Information Systems.*

Lewandowski, T., Delling, J., Grotherr, C., & Böhmann, T. (2021). State-of-the-art analysis of adopting AI-based conversational agents in organizations: A systematic literature review. In *Proceedings of the 25th Pacific Asia Conference on Information Systems.*

Lichtenberg, S., Bührke, J., Brendel, A. B., Trang, S., Diederich, S., & Morana, S. (2021). "Let us work Together": Insights from an experiment with conversational agents on the relation of anthropomorphic design, dialog support, and performance. In F. Ahlemann, R. Schütte, & S. Stieglitz (Eds.), *Innovation through information systems—volume II: A collection of latest research on technology issues* (LNISO vol. 47. pp. 299-315). Springer.

Liebrecht, C., & van Hooijdonk, C. (2020). Creating humanlike chatbots: What chatbot developers could learn from webcare employees in adopting a conversational human voice. In A. Følstad, T. Araujo, S. Papadopoulos, E. L.-C. Law, O.-C. Granmo, E. Luger, & P. B. Brandtzaeg (Eds.), *Chatbot research and design* (LNICS vol. 11970., pp. 51-64). Springer.

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., & Söllner, M. (2019). AI-based digital assistants: Opportunities, threats, and research perspektives. *Business & Information Systems Engineering*, *61*(4), 535-544.

Mallios, S., & Bourbakis, N. (2016). A survey on human machine dialogue systems. In *Proceedings of the 7th International Conference on Information, Intelligence, Systems & Applications.*

Manseau, J. (2020). AI in the workplace: A qualitative analysis of intelligent employee assistants. In *Proceedings of the 26th Americas Conference on Information Systems.*

Maroengsit, W., Piyakulpinyo, T., Phonyiam, K., Pongnumkul, S., Chaovalit, P., & Theeramunkong, T. (2019). A survey on evaluation methods for chatbots. In *Proceedings of the 7th International Conference on Information and Education Technology.*

Meyer von Wolff, R., Briesemeister, M., Gobrecht, L., Horndasch, F. M., Kupfer, D., & Hobert, S. (2021). Hey spot, can you help me with my business travel organization? Design of a process-based chatbot artifact. In *Proceedings of the 27th Americas Conference on Information Systems.*

Meyer von Wolff, R., Heuzeroth, T., Hobert, S., & Schumann, M. (2020a). The students' view on IT-Support chatbots at universities: A case-based pilot study. In *Proceedings of the 26th Americas Conference on Information Systems.*

Meyer von Wolff, R., Hobert, S., Masuch, K., & Schumann, M. (2020b). Chatbots at digital workplaces: A grounded-theory approach for surveying application areas and objectives. *Pacific Asia Journal of the Association for Information Systems*, *12*(2), 63-101.

Meyer von Wolff, R., Hobert, S., & Schumann, M. (2019). How may I help you? State of the art and open research questions for chatbots at the digital workplace. In *Proceedings of the 52nd Hawaii International Conference on System Science.*

Nißen, M., Selimi, D., Janssen, A., Cardona, D. R., Breitner, M. H., Kowatsch, T., & Wangenheim, F. von (2022). See you soon again, chatbot? A design taxonomy to characterize user-chatbot relationships with different time horizons. *Computers in Human Behavior*, *127*(2), 1-15.

Ojo, A. I. (2017). Validation of the DeLone and McLean information systems success model. *Healthcare Informatics Research*, *23*(1), 60-66.

Radlinski, F., & Craswell, N. (2017). A theoretical framework for conversational search. In *Proceedings of the Conference Human Information Interaction and Retrieval.*

Ranoliya, B. R., Raghuwanshi, N., & Singh, S. (2017). Chatbot for university related FAQs. In *Proceedings of the International Conference on Advances in Computing, Communications and Informatics.*

Reshmi, S., & Balakrishnan, K. (2016). Implementation of an inquisitive chatbot for database supported knowledge bases. *Sadhana*, *41*(10), 1173-1178.

Richter, A., Heinrich, P., Stocker, A., & Schwabe, G. (2018). Digital work design. *Business & Information Systems Engineering*, *60*(3), 259-264.

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 14th International Conference on Wirtschaftsinformatik.*

Sarikaya, R. (2017). The technology behind personal digital assistants: An overview of the system architecture and key components. *IEEE Signal Processing Magazine*, *34*(1), 67-81.

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Construction of a benchmark for the user experience questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence*, *4*(4), 40-44.

Schuetzler, R. M., Grimes, G. M., & Giboney, J. S. (2018). An investigation of conversational agent relevance, presence, and engagement. In *Proceedings of the 24th Americas Conference on Information Systems.*

Shawar, B. A., & Atwell, E. (2007). Different measurements metrics to evaluate a chatbot system. In *Proceedings of the Workshop on Bridging the Gap: Academic and Industrial Research in Dialog Technologies.*

Siemon, D., & Jusmann, S. (2021). Preferred appearance of embodied conversational agents in knowledge management. In *Proceedings of the 27th Americas Conference on Information Systems.*

Stieglitz, S., Brachten, F., & Kissmer, T. (2018). Defining bots in an enterprise context. In *Proceedings of the 39th International Conference on Information Systems.*

Stoeckli, E., Uebernickel, F., & Brenner, W. (2018). Exploring affordances of slack integrations and their actualization within enterprises: Towards an understanding of how chatbots create value. In *Proceedings of the 51st Hawaii International Conference on System Sciences.*

Tavanapour, N., Poser, M., & Bittner, E. (2019). Supporting the idea generation process in citizen participation: Toward an interactive system with a conversational agent as a facilitator. In *Proceedings of the 27th European Conference on Information Systems.*

Tukey, J. W. (1977). *Exploratory data analysis* (1st ed.). Addison-Wesley.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, *39*(2), 273-315.

Venkatesh, V., & Davis, F. D. (2000). A Theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, *46*(2), 186-204.

Wambsganss, T., Winkler, R., Söllner, M., & Leimeister, J. M. (2020). A conversational agent to improve response quality in course evaluations. In *Proceedings of the CHI Conference on Human Factors in Computing Systems.*

White, M. (2012). Digital workplaces: Vision and reality. *Business Information Review*, *29*(4), 205-214.

Winkler, R., Hobert, S., Salovaara, A., Söllner, M., & Leimeister, J. M. (2020). Sara, the lecturer: Improving learning in online education with a scaffolding-based conversational agent. In *Proceedings of the CHI Conference on Human Factors in Computing Systems.*

Winkler, R., Neuweiler, M. L., Bittner, E., & Söllner, M. (2019). Hey Alexa, please help us solve this problem! How interactions with smart personal assistants improve group performance. In *Proceedings of the 40th International Conference on Information Systems*.

Winkler, R., & Söllner, M. (2018). Unleashing the potential of chatbots in education: A state-of-the-art analysis. In *Proceedings of the Academy of Management Annual Meeting.*

Winkler, R., & Roos, J. (2019). Bringing AI into the classroom: Designing smart personal assistants as learning tutors. In *Proceedings of the 40th International Conference on Information Systems.*

Wuenderlich, N. V., & Paluch, S. (2017). A nice and friendly chat with a bot: User perceptions of AI-based service agents. In *Proceedings of the 38th International Conference on Information Systems.*

Yu, P., & Qian, S. (2018). Developing a theoretical model and questionnaire survey instrument to measure the success of electronic health records in residential aged care*. PloS One*, *13*, 1-18.

Zierau, N., Elshan, E., Visini, C., & Janson, A. (2020a). A review of the empirical literature on conversational agents and future research directions. In *Proceedings of the 41th International Conference on Information Systems.*

Zierau, N., Hausch, M., Bruhin, O., & Söllner, M. (2020b). Towards developing trust-supporting design features for AI-based chatbots in customer service. In *Proceedings of the 41th International Conference on Information Systems.*

Appendix A: Requirement Analysis

Table A1. Requirements Analysis

References	Requirements										
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11
Bittner & Shoury (2019)	●	●	●		●	●	●				●
Corea et al. (2020)	●	●		●		●				●	
Diederich et al. (2020)	●		●	●			●			●	
Elshan & Ebel (2020)	●	●		●							
Feine et al. (2020)	●		●		●	●	●				●
Gnewuch et al. (2017)	●		●	●							
Hobert (2019b)	●	●	●					●			
Johannsen et al. (2018)			●	●						●	
Lechler et al. (2019)	●	●						●			
Tavanapour et al. (2019)	●	●	●	●	●						
Winkler & Roos (2019)	●	●		●		●	●		●		●
Zierau et al. (2020b)	●			●			●			●	

Appendix B: Experts' Distribution

Table B1. Experts' Distribution

Expert	Industry
1	Automotive engineering
2	Information & communication technology
3	Chemicals, pharmaceuticals & raw materials
4	Information & communication technology
5	Chemicals, pharmaceuticals & raw materials
6	Information & communication technology
7	Information & communication technology
8	Finance & insurance
9	Automotive engineering
10	Information & communication technology
11	Public sector
12	Finance & insurance
13	Finance & insurance

Appendix C: Questionnaire

Table C1. Evaluation Questionnaire

	Construct	Theory	Item	Type	Scale	Reference
General	Age		How old are you?	Free text	-	
	Gender		What is your gender?	Single choice	Male	
					Female	
					Diverse	
					Not specified	
	Chatbot experience		What is your previous chatbot experience?	Five-point Likert scale	No use <-> regular use	
	IT affinity	Affinity for technology interaction scale	I like to occupy myself in greater detail with technical systems.	Six-point Likert scale	Do not agree <-> fully agree	Franke et al. (2019)
			I like testing the functions of new technical systems.			
			I predominantly deal with technical systems because I have to.			
			When I have a new technical system in front of me, I try it out intensively.			
			I enjoy spending time becoming acquainted with a new technical system.			
			It is enough for me that a technical system works; I don't care how or why.			
			I try to understand how a technical system exactly works.			
			It is enough for me to know the basic functions of a technical system.			
			I try to make full use of the capabilities of a technical system.			
	Process experience		Do you have experience with the university's business travel process?	Single choice	Yes, I know the business trip organization process of the university	
					No, but I know similar business travel organization processes	
					No, I have no experience with business travel organization processes	

Table C1. Evaluation Questionnaire

Usability	System features (DP)		Natural language interaction for process execution	Seven-point Likert scale	Very negative <-> very positive	
			Process guidance including progress overview and forwarding of the process			
			Adaptation of the process to the user / free choice of procedure			
			Provided input options (free text, selection errors, buttons, file upload)			
			Integrated help function for the necessary entries directly in the dialog			
			Automatic error handling of the entries made			
	User experience (UEQ)	User experience questionnaire	Please rate the system based on the following items.	Seven-point Likert scale	Annoying <-> enjoyable	Laugwitz et al. (2008), Schrepp et al. (2017)
					Not understandable <-> understandable	
					Creative <-> dull	
					Easy to learn <-> difficult to learn	
					Valuable <-> inferior	
					Boring <-> exciting	
					Not interesting <-> interesting	
					Unpredictable <-> predictable	
					Fast <-> slow	
					Inventive <-> conventional	
					Obstructive <-> supportive	
					Good <-> bad	
					Complicated <-> easy	
					Unlikable <-> pleasing	
					Usual <-> leading edge	
					Unpleasant <-> pleasant	
					Secure <-> not secure	
					Motivating <-> demotivating	
					Meets expectations <-> does not meet expectations	
					Inefficient <-> efficient	
					Clear <-> confusing	
					Impractical <-> practical	
					Organized <-> cluttered	
					Attractive <-> unattractive	
					Friendly <-> unfriendly	
					Conservative <-> innovative	
Acceptance	Information quality (IQ)	Information systems success model	The chatbot presents the information/answers in a useful format.	Seven-point Likert scale	Do not agree <-> fully agree	Yu & Qian (2018)
			The outputs of the chatbot are easy to understand.		Do not agree <-> fully agree	Freeze et al. (2010)
			The chatbot provides the information I need to organize business trips.		Do not agree <-> fully agree	Freeze et al. (2010)
			The chatbot provides relevant information for the business trip organization/substeps.		Do not agree <-> fully agree	Freeze et al. (2010)
	Service quality	Information systems	When using the chatbot, I feel safe in terms of data protection and data security.	Seven-point	Do not agree <-> fully agree	Alshibly (2014)

Table C1. Evaluation Questionnaire

	(SQ)	success model	The messages and results of the chatbot are complete for the business trip organization.	Likert scale	Do not agree <-> fully agree	Ojo (2017)
			The chatbot supports me individually in organizing business trips.		Do not agree <-> fully agree	Alshibl (2014)
			If I have a problem using it, the chatbot helps me find a solution.		Do not agree <-> fully agree	Alshibly (2014)
	Perceived usefulness (PU)	Technology acceptance model	Using the chatbot allows me to do the business trip organization quickly.	Seven-point Likert scale	Do not agree <-> fully agree	Davis (1989)
			Using the Chabot makes it easier for me to organize business trips.		Do not agree <-> fully agree	Davis (1989)
			Using the chatbot for business travel organization increases my productivity.		Do not agree <-> fully agree	Davis (1989)
			I find the chatbot useful for business travel organization.		Do not agree <-> fully agree	Venkatesh & Bala (2008)
			Using the chatbot for business travel organization increases my effectiveness.		Do not agree <-> fully agree	Davis (1989)
	Perceived ease of use (PEoU)	Technology acceptance model	My interaction with the system is clear and understandable.	Seven-point Likert scale	Do not agree <-> fully agree	Venkatesh & Davis (2000)
			I find it easy to get the system to do what I want.		Do not agree <-> fully agree	Venkatesh & Bala (2008)
			Using the chatbot is easy for me to learn.		Do not agree <-> fully agree	Davis (1989)
			I find the chatbot easy to use.		Do not agree <-> fully agree	Venkatesh & Davis (2000)
			Using the chatbot for business travel organization requires little mental effort.		Do not agree <-> fully agree	Venkatesh & Davis (2000)
	Behavioral intention to use (BI)	Technology acceptance model	If I have access to the chatbot, I would probably use it for business travel organization.	Seven-point Likert scale	Do not agree <-> fully agree	Venkatesh & Davis (2000)
			I would recommend the chatbot for organizing the business trip.		Do not agree <-> fully agree	Constantinides et al. (2013)
			Provided I have access to the chatbot, I will use it for business travel organization in the future.		Do not agree <-> fully agree	Venkatesh & Davis (2000)
	User satisfaction (US)	Information systems success model	The chatbot for business travel organization has met my expectations.	Seven-point Likert scale	Do not agree <-> fully agree	Alshibly (2014)
			Overall, I am satisfied with the chatbot for business trip organization.		Do not agree <-> fully agree	Yu & Qian (2018)
			I find the chatbot very helpful for business travel organization.		Do not agree <-> fully agree	Freeze et al. (2010)

Appendix D: User Experience Questionnaire

Table D1. Items: Existing Form

Item	Left	Right	Scale	N	Mean	Std. dev.
1	Annoying	Enjoyable	Attractiveness	35	-0.629	1.215
2	Not understandable	Understandable	Perspicuity	35	0.200	1.431
3	Creative	Dull	Novelty	35	-1.571	1.399
4	Easy to learn	Difficult to learn	Perspicuity	35	0.143	1.438
5	Valuable	Inferior	Stimulation	35	-0.171	1.150
6	Boring	Exciting	Stimulation	35	-1.657	1.413
7	Not interesting	Interesting	Stimulation	35	-1.457	1.442
8	Unpredictable	Predictable	Dependability	35	0.457	1.245
9	Fast	Slow	Efficiency	35	-0.914	1.147
10	Inventive	Conventional	Novelty	35	-2.086	0.951
11	Obstructive	Supportive	Dependability	35	-0.200	1.256
12	Good	Bad	Attractiveness	35	-0.343	1.162
13	Complicated	Easy	Perspicuity	35	-0.543	1.358
14	Unlikable	Pleasing	Attractiveness	35	-0.714	1.073
15	Usual	Leading edge	Novelty	34	-2.176	1.114
16	Unpleasant	Pleasant	Attractiveness	35	-0.771	1.140
17	Secure	Not secure	Dependability	35	0.314	1.323
18	Motivating	Demotivating	Stimulation	35	-1.543	1.172
19	Meets expectations	Does not meet expectations	Dependability	35	0.829	1.465
20	Inefficient	Efficient	Efficiency	35	-0.429	1.461
21	Clear	Confusing	Perspicuity	35	-0.457	1.578
22	Impractical	Practical	Efficiency	35	-0.200	1.368
23	Organized	Cluttered	Efficiency	35	-0.714	1.467
24	Attractive	Unattractive	Attractiveness	35	-1.286	1.073
25	Friendly	Unfriendly	Attractiveness	35	-0.829	1.224
26	Conservative	Innovative	Novelty	35	-2.229	1.003

Table D2. Constructs and Scale Consistency: Existing Form

Item	Scale	N	Mean	Std. dev.	Cronbach's alpha coefficient	Guttman's Lambda-2 coefficient
1	Attractiveness	35	-0.762	0.94	0.92	0.92
2	Perspicuity	35	-0.164	1.43	0.85	0.84
3	Efficiency	35	-0.564	1.08	0.76	0.77
4	Dependability	35	0.350	0.90	0.69	0.68
5	Stimulation	35	-1.207	1.23	0.87	0.88
6	Novelty	35	-2.010	0.79	0.82	0.80

Table D3. Items: Chatbot

Item	Left	Right	Scale	N	Mean	Std. dev.
1	Annoying	Enjoyable	Attractiveness	66	1.909	0.940
2	Not understandable	Understandable	Perspicuity	66	2.106	0.879
3	Creative	Dull	Novelty	66	1.318	1.179
4	Easy to learn	Difficult to learn	Perspicuity	66	2.394	0.742
5	Valuable	Inferior	Stimulation	66	1.742	0.847
6	Boring	Exciting	Stimulation	66	1.091	0.779
7	Not interesting	Interesting	Stimulation	66	1.409	0.841
8	Unpredictable	Predictable	Dependability	66	1.045	1.221
9	Fast	Slow	Efficiency	66	1.500	1.206
10	Inventive	Conventional	Novelty	66	1.439	1.191
11	Obstructive	Supportive	Dependability	66	1.924	1.012
12	Good	Bad	Attractiveness	65	2.123	0.893
13	Complicated	Easy	Perspicuity	66	1.894	1.010
14	Unlikable	Pleasing	Attractiveness	65	0.985	1.068
15	Usual	Leading edge	Novelty	66	1.773	1.064
16	Unpleasant	Pleasant	Attractiveness	66	1.576	1.096
17	Secure	Not secure	Dependability	66	1.409	1.289
18	Motivating	Demotivating	Stimulation	66	1.439	0.994
19	Meets expectations	Does not meet expectations	Dependability	64	1.313	1.167
20	Inefficient	Efficient	Efficiency	66	1.682	1.025
21	Clear	Confusing	Perspicuity	66	1.864	1.094
22	Impractical	Practical	Efficiency	66	1.576	1.138
23	Organized	Cluttered	Efficiency	66	1.833	1.284
24	Attractive	Unattractive	Attractiveness	64	1.375	0.968
25	Friendly	Unfriendly	Attractiveness	65	1.508	1.017
26	Conservative	Innovative	Novelty	66	1.924	0.933

Table D4. Constructs and Scale Consistency: Chatbot

Item	Scale	N	Mean	Std. dev.	Cronbach's alpha coefficient	Guttman's Lambda-2 coefficient
1	Attractiveness	66	1.585	0.53	0.82	0.82
2	Perspicuity	66	2.064	0.59	0.84	0.83
3	Efficiency	66	1.648	0.91	0.83	0.83
4	Dependability	66	1.431	0.75	0.72	0.72
5	Stimulation	66	1.420	0.43	0.76	0.76
6	Novelty	66	1.614	0.77	0.82	0.81

Appendix E: Acceptance

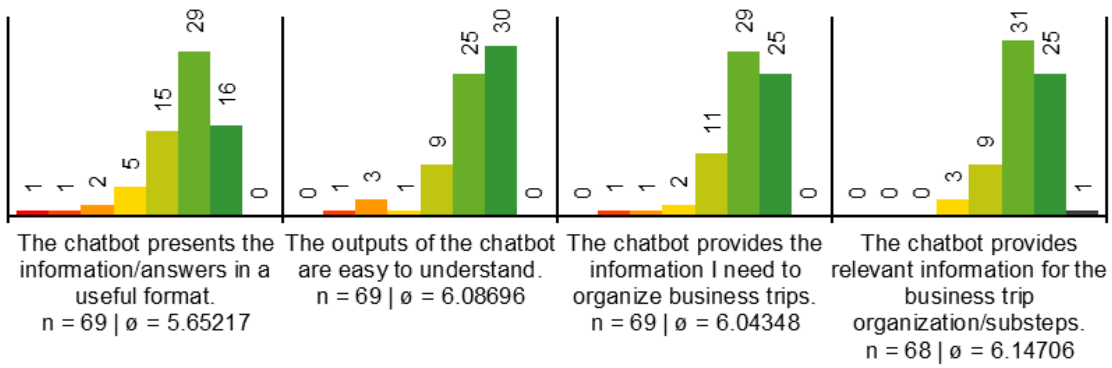

Figure E1. Information Quality

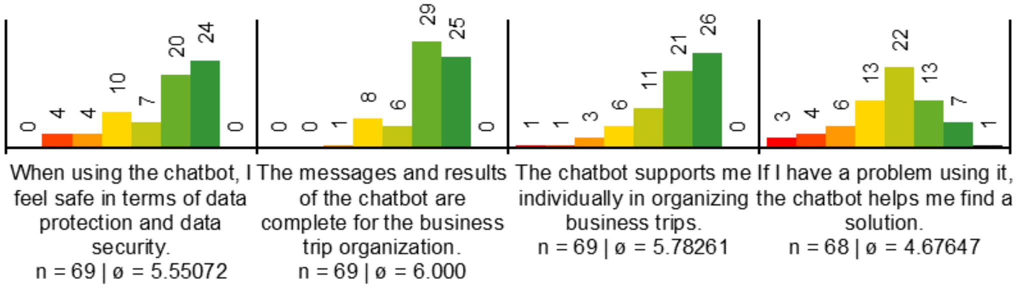

Figure E2. Service Quality

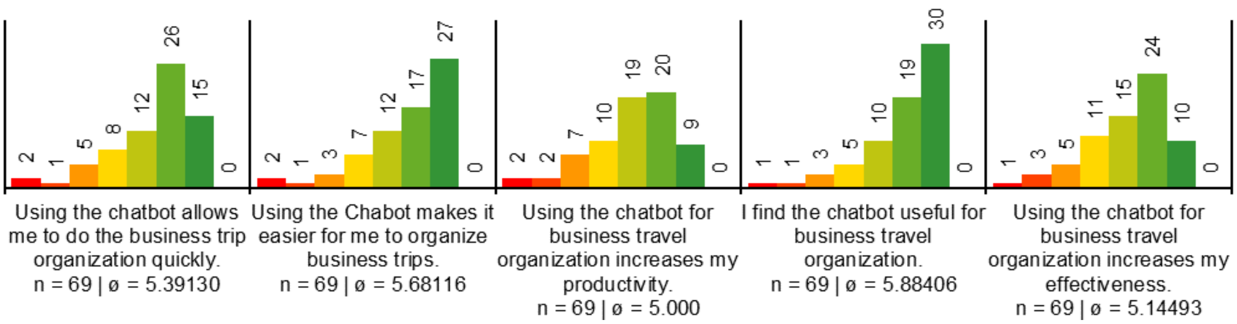

Figure E3. Perceived Usefulness

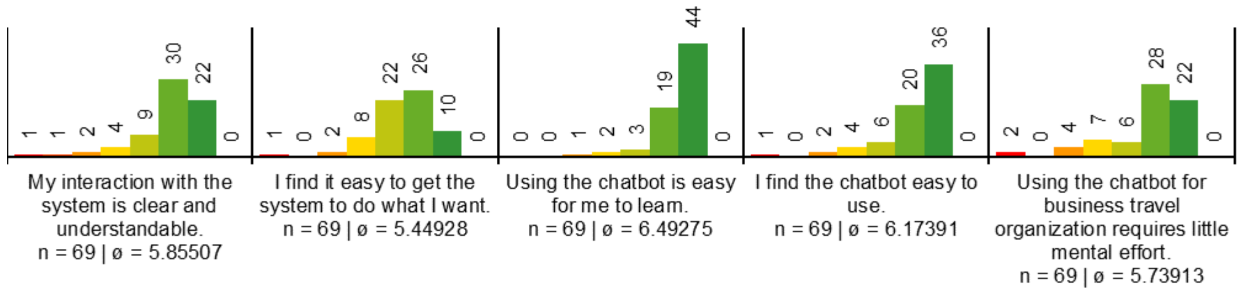

Figure E4. Perceived Ease of Use

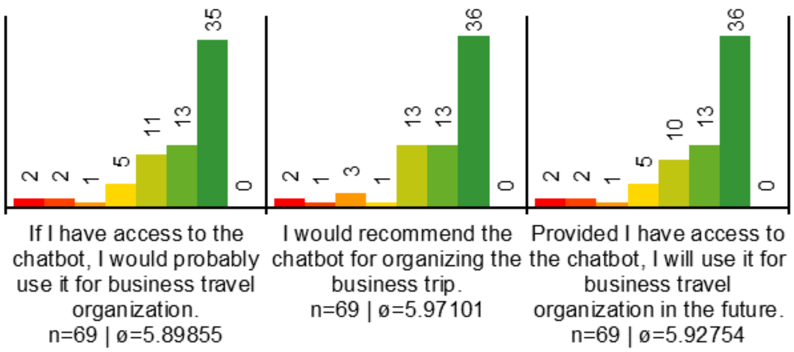

Figure E5. Behavioral Intention to Use

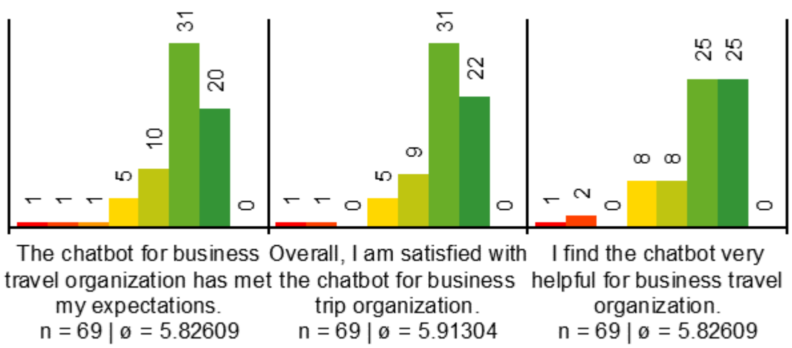

Figure E6. User Satisfaction

About the Authors

Raphael Meyer von Wolff graduated in business information systems from the University of Goettingen, Germany, in 2016 and received his PhD in 2022 in the research area of chatbots at digital workplaces. His research interests focus on applying chatbots in professional workplace domains based on multiple years of experience in the scientific and practice domain. His research focuses on human behavior, human-computer interaction, design science research, application areas, and impact on businesses. His works have been presented and published in international conferences and journals such as *Pacific Asia Journal of the Association for Information Systems*, Hawaii International Conference on System Sciences, Americas Conference on System Sciences, and *Lecture Notes on Computer Science* in which he also acted as a reviewer. He is the corresponding author and can be contacted via the indicated mail address.

Sebastian Hobert is a postdoctoral researcher at the University of Goettingen, Germany. He graduated in applied computer science in 2014 and received his PhD in 2018 in the information systems domain. Now, he works in of the computer science, human-computer interaction, and information systems domains. In his current research, he mainly focuses on design-oriented research on chatbots, technology-enhanced learning, and learning analytics. His work has been published in journals and leading information systems and computer science conference proceedings.

Matthias Schumann is a Professor of Information Systems at the University of Goettingen, Germany. He graduated in Business Administration from the University of Goettingen and received his PhD in information systems from the University of Erlangen-Nuremberg. His research interests include digitization of knowledge work, Social Media in the business context, and competence measurement and E-learning. He has published a large number of refereed publications including books, journal articles, book chapters, and conference articles.

Transactions on Human – Computer Interaction