On the Design of and Interaction with Conversational Agents: An Organizing and Assessing Review of Human-Computer Interaction Research

Article *in* Journal of the Association for Information Systems · January 2022

4 authors:

Stephan Diederich
Georg-August-Universität Göttingen

Alfred Benedikt Brendel
Technische Universität Dresden

Stefan Morana
Universität des Saarlandes

Lutz Kolbe
Georg-August-Universität Göttingen

Accepted for publication in the Journal of the Association for Information Systems (JAIS)

On the Design of and Interaction with Conversational Agents: An Organizing and Assessing Review of Human-Computer Interaction Research

Stephan Diederich
University of Göttingen
stephan.diederich@stud.uni-goettingen.de

Alfred Benedikt Brendel
TU Dresden
alfred_benedikt.brendel@tu-dresden.de

Stefan Morana
Saarland University
stefan.morana@uni-saarland.de

Lutz Kolbe
University of Göttingen
lkolbe@uni-goettingen.de

Abstract

Conversational agents (CAs), described as software with which humans interact through natural language, have increasingly attracted interest in both academia and practice, due to improved capabilities driven by advances in artificial intelligence and, specifically, natural language processing. CAs are used in contexts like people's private life, education, and healthcare, as well as in organizations, to innovate and automate tasks, for example in marketing and sales or customer service. In addition to these application contexts, such agents take on different forms concerning their embodiment, the communication mode, and their (often human-like) design. Despite their popularity, many CAs are not able to fulfill expectations and to foster a positive user experience is a challenging endeavor. To better understand how CAs can be designed to fulfill their intended purpose, and how humans interact with them, a multitude of studies focusing on human-computer interaction have been carried out. These have contributed to our understanding of this technology. However, currently a structured overview of this research is missing, which impedes the systematic identification of research gaps and knowledge on which to build on in future studies. To address this issue, we have conducted an organizing and assessing review of 262 studies, applying a socio-technical lens to analyze CA research regarding the user interaction, context, agent design, as well as perception and outcome. We contribute an overview of the status quo of CA research, identify four research streams through a cluster analysis, and propose a research agenda comprising six avenues and sixteen directions to move the field forward.

Keywords: Conversational Agent; Chatbot; Digital Assistant; Virtual Human; Robot; Organizing Review; Assessing Review; Human-Computer Interaction.

1 Introduction

Technological advances continue to drive the digital transformation and change the way we live, work, and interact with one another (Davenport & Kirby, 2016; McAfee & Brynjolfsson, 2017). Advances in artificial intelligence (AI), such as machine learning and natural language processing, are essential drivers in this development, making machines seemingly intelligent and capable of conversing in natural language, creating meaning in written or spoken words (Brynjolfsson & McAfee, 2016). Conversational agents (CAs), benefiting from these advances, increasingly attract interest in research as well as in practice (McTear, 2017). Interacting with a system through natural language promises to increase the ease of use and to ensure faster completion of user requests, while creating the feeling of a human-like interaction (Følstad & Brandtzæg, 2017). For users, CAs can function in various contexts, ranging from digital personal assistants in mobile devices, such as Apple's Siri or Google Assistant, to specific purposes like in-car assistance (Laumer et al. 2019). For organizations, CAs offer the possibility of automating and innovating processes in areas such as human resources (Liao et al., 2018), customer service (Ashktorab et al., 2019), and sales (Vaccaro et al., 2018).

Recent examples developed by Facebook and Google underline CAs' popularity and potential. After launching its new Messenger platform, more than 100,000 bots appeared on Facebook within the first year (Johnson, 2018). Further, Google demonstrated the future potential of CAs at its 2018 developer conference by having their assistant autonomously make a hairdresser's appointment via a telephone conversation with a real person on the other end (Welch, 2018). Gartner predicts that 70 percent of white-collar workers will interact with systems using conversational interfaces in their daily work by 2022 (Goasduff, 2019). Despite their potential, many CAs do not meet expectations and are discontinued because of flaws related to their design, such as unappealing appearance, lacking conversational abilities, or unrealizable user expectations (Ashktorab et al., 2019; Ben Mimoun et al., 2012; Lahoual & Fréjus, 2019; Luger & Sellen, 2016). The complexity of designing such agents is particularly driven by the human's social responses to the cues incorporated in these artifacts, such as interaction via natural language, human names, or these agents' social roles (Feine et al., 2019; Seeger et al., 2018). Such social responses affect the individual's perception of these agents, and they foster high user expectations, which are often not in line with the agents' actual capabilities (Ben Mimoun et al., 2012; Luger & Sellen, 2016).

In summary, designing CAs and understanding how they interact with humans remain substantial challenges in research as well as practice (Schuetz & Venkatesh, 2020). In research, a multitude of studies, particularly in the field of human-computer interaction (HCI), have contributed to addressing this challenge. In particular, a renewed interest in research on CAs emerged in the information systems (IS) and computer science (CS) communities (McTear, 2017; Rzepka & Berger, 2018). Researchers from both disciplines investigate CAs with different technological properties, such as the different communication modes as in voice or text (Cho, 2019; Schroeder & Schroeder, 2018), or embodiment which can be virtual or physical (Araujo, 2018; Seymour et al., 2018), explore different contexts, such as interactive tutoring (Fryer et al., 2017) or customer service (Xu et al., 2017). Moreover, they focus on different aspects related to the perception of such agents and interaction outcomes, for example, anthropomorphism (Araujo, 2018; Seeger et al., 2018), trust (Benlian et al., 2019; Elson et al., 2018; Schuetzler et al., 2014), or number of digital products sold (Kim et al., 2018). In short, the design of and interaction with present-day CAs offer very many research opportunities, and a variety of studies are available in this research area.

However, this variety of studies also comes with the challenge of gaining an overview of the topic. Existing reviews on CAs focus on selected aspects of the interaction with and design of CAs, such as trust (Zierau et al., 2020) or social cues (Feine et al., 2019), they investigate CAs in specific contexts like the digital workplace (Wolff et al., 2019), or, more abstractly, review AI-based applications (Rzepka & Berger, 2018). An overview of CA research covering different contexts, types of CAs, and users' perceptions is not available. Without such an overview, researchers face difficulties in systematically identifying and addressing research gaps and knowledge upon which to build. To address this issue, in his project we organize existing studies on CAs, assess the status quo, and contribute avenues as well as directions for future work in this area.

The remainder of this article is organized as follows: first, we provide a brief overview of CAs and introduce the framework for our analysis. Next, we describe our research approach, present overarching observations gained in our literature review, and identify four research streams based on a cluster analysis. Building on these results, we derive six avenues containing sixteen specific directions for future research in this area.

2 Research Background

CAs are based on the idea of interacting with users through natural language as in human-to-human conversations (Dale, 2016; McTear, Callejas, & Griol, 2016). CAs are variously and interchangeably termed as digital assistant, chatbot, interactive agent, etc. (Maedche et al., 2019; Stieglitz et al., 2018). Different definitions are given for a CA, such as an agent that "interacts with users, turn by turn by using natural language" (Comendador et al., 2015, p. 137), or "computer programs designed to respond to users in natural language, thereby mimicking conversations between people" (Miner et al., 2016, p. 619), or a concept to "achieve some result by conversing with a machine in a dialogic fashion, using natural language" (Dale, 2016, page 811). While these definitions each highlight different characteristics of CAs, such as turn-taking (Comendador et al., 2015) or mimicking conversations (Miner et al., 2016), they all share the idea of natural language interaction.

Thus, for this research, we consider CAs to be technological artifacts with which users interact through natural language, both in written and spoken form.

2.1 An Overview of Conversational Agents

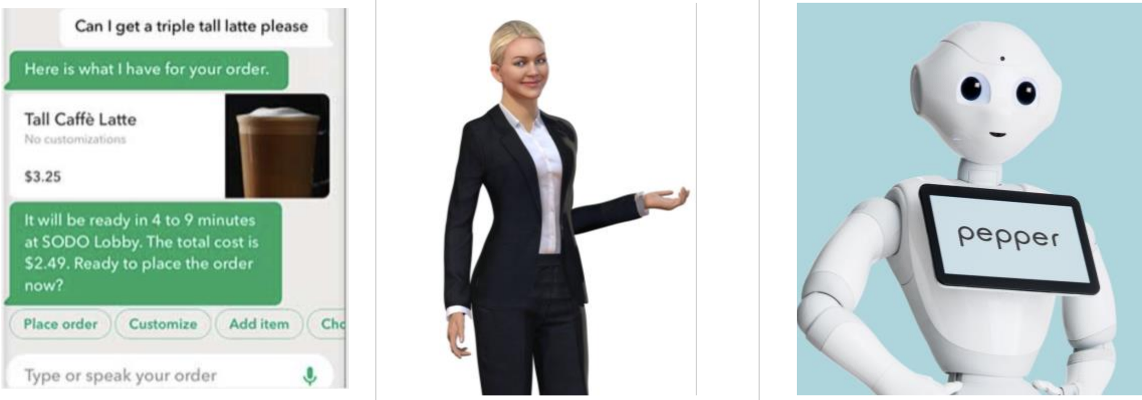
The basic idea to interact with technological artifacts through natural language already emerged in the 1960s when Joseph Weizenbaum (1966) developed ELIZA. While the fundamental idea of natural language interaction is the same for all CAs, these agents take on different forms that are distinguished by communication mode, embodiment, and the context in which they are used (Cassell et al., 1999; Cowell & Stanney, 2005; Gnewuch et al., 2017):

- **Communication mode:** CAs can communicate via voice (Cowan et al., 2015), text (Schroeder & Schroeder, 2018), or both (Cho, 2019)
- **Embodiment:** CAs can be disembodied (Araujo, 2018), virtually embodied (Diederich et al., 2019), or physically embodied (Nunamaker et al., 2011)
- **Context:** CAs can be used for general-purpose conversations or be domain-specific, for example, for a specific task or business function (Gnewuch et al., 2017).

Multiple studies have addressed the interaction with CAs regarding user responses, such as user trust (Elson et al., 2018; Seeger et al., 2017), authenticity (Wünderlich & Paluch, 2017), or empathy (Leite et al., 2013; McQuiggan & Lester, 2007). CAs have been studied in different contexts, such as in legal research (Sugumaran & Davis, 2001), lie detection (Nunamaker et al., 2011), financial advisory (Morana et al., 2020), or data analytics (Matsushita et al., 2004). Concerning their application in organizations, recent research has attended to different business functions, such as human resources (Diederich, Brendel, & Kolbe, 2020; Liao et al., 2018) or marketing and sales (Qiu & Benbasat, 2009; Vaccaro et al., 2018). In practice, different CAs have emerged over time, such as IKEA's Anna (Wakefield, 2016).

In the past, due to CAs reliance on simple pattern-matching, they were limited regarding both understanding a user's intent (i.e., the meaning behind a message) and providing purposeful feedback (Berg, 2015; Knijnenburg & Willemsen, 2016). This often leads to CAs being discontinued (Ben Mimoun et al., 2012). However, with advances in natural language processing, as well as machine learning, CA capabilities have greatly improved in recent years and led to a renewed interest in both research and practice (McTear, 2017; Oracle, 2016). Today, most smartphones are equipped out-of-the-box with voice-based CAs, such as Google Assistant or Siri (Burton & Gaskin, 2019), and devices such as Amazon's Alexa are used in private households (Purington et al., 2017) or team collaboration (Winkler et al., 2019). Further, companies are exploring the potential of physically embodied CAs, such as SoftBank's humanoid robot "Pepper," to provide services (Stock & Merkle, 2018a).

Similarly, text-based CAs, often referred to as chatbots, are increasingly available in different contexts. For example, the Dutch airline KLM introduced a text-based CA that helps users to find and book flights (Vogel-Meijer, 2018), the coffee shop chain Starbucks designed a CA called “Barista” to support ordering products (Perez, 2016), and the American railroad company Amtrak offers customer services that answer five million requests per year through “Julie” (NextIT, 2018). Regarding CAs with a virtual interactive embodiment IPSoft developed the agent known as “Amelia,” that, for example, automates information technology (IT) service desk tasks for a medical supplies manufacturer and offers customer services for a telecom provider (Ipsoft, 2020) (see Figure 1).



Text-based agent “Barista” by Starbucks	Virtual interactive agent “Amelia” by IPsoft	Service robot “Pepper” by SoftBank Robotics

Figure 1. Examples of Conversational Agents

2.2 Human-Computer Interaction

Human-Computer Interaction (HCI) research as a research domain dates back to the early 1980s (Card et al., 1980; Carroll, 2020). More recent HCI research investigated “the ways humans interact with information, technologies, and tasks, especially in business, managerial, organizational, and cultural contexts” (Zhang et al., 2002, p. 333). The first ACM Computer-Human Interaction (CHI) conference, the premier conference on HCI, was established in 1982 (SIG CHI, 2020). Since then, this research field expanded, also into other disciplines such as CS, (cognitive) psychology, human factors, and IS. In the IS discipline, Gerlach and Kuo (1991) recognized the interdisciplinary nature of HCI, after which numerous conceptual and empirical publications on HCI appeared in various disciplines (Carroll, 2020; Olson & Olson, 2003; Zhang & Li, 2004).

The interaction between humans and computers can be related to the concept of socio-technical systems. A socio-technical system relies on the interplay of three key elements (Goodhue & Thompson, 1995; Heinrich et al., 2011): the human that wants to achieve a specific goal; the task that the user has to accomplish in order to achieve the goal; the technology (i.e., software, hardware, and data) the user utilizes to complete the task. Taking a socio-technical perspective, Zhang and Li (2005) assessed the intellectual development of HCI research in the IS discipline, proposing an extended framework that consists of humans interacting with technology in a specific context, ultimately leading to a set of outcomes and perceptions. This framework adds the dimension of perception change through technology to our understanding of HCI.

In CA research, where the anthropomorphic design, related human-like perception, and application context are key research objects, the socio-technical framework provides a valuable lens to investigate the continuously growing body of knowledge on the interaction between humans and CAs. Therefore, drawing on the framework by Zhang and Li (2005), we will derive a framework to organize existing research, conduct a literature review, and analyze the identified studies in order to determine current trends as well as outline an agenda for future work.

3 Method: A Review of HCI Research on CAs

In this section, we outline our study's research framework as well as the method we used to identify and analyze research on CAs in IS and CS research.

3.1 A Framework for Human-Computer Interaction via Natural Language

As a lens for our review of CA studies, we draw on and adapt the research framework Zhang and Li (2005) proposed. Drawing on this framework, we review four dimensions of CA studies: context, human, agent, and perception and outcome (see Figure 2).

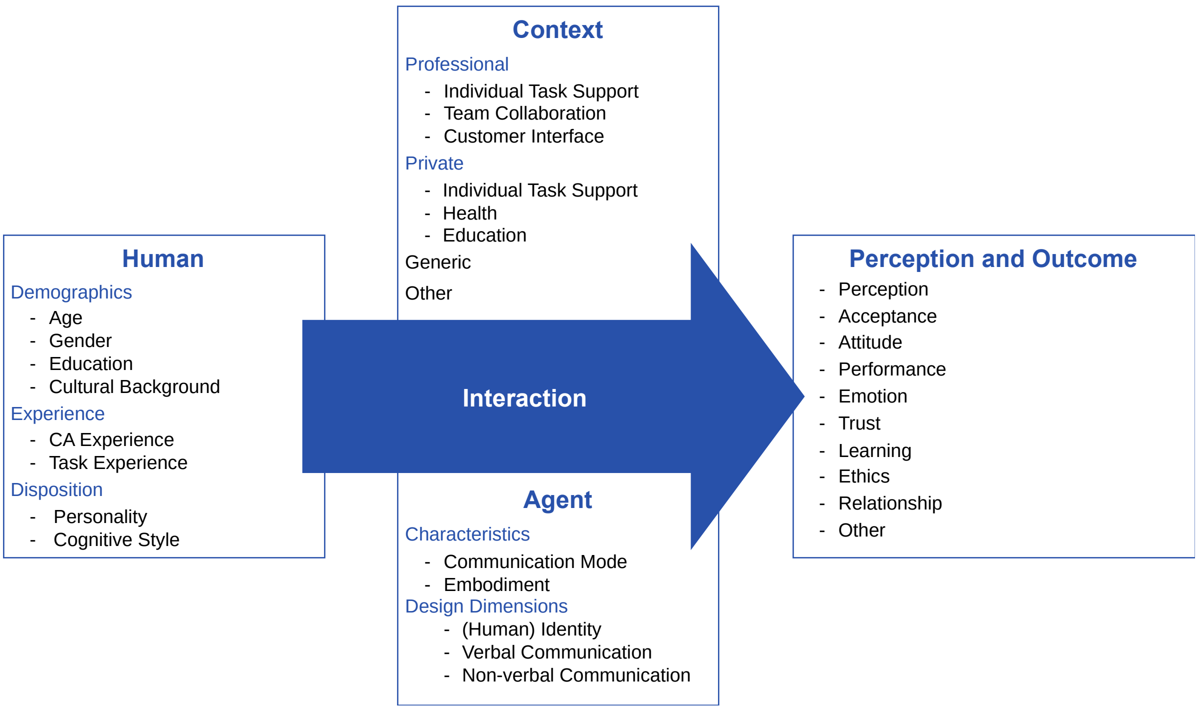


Figure 2. Research Framework

First, we consider the **context** in which the CA is applied. On an abstract level, we can distinguish professional and private contexts. Professional contexts comprise, for example, the internal use of CAs for individual task support (e.g., scheduling appointments) (Bittner & Shoury, 2019; Fast et al., 2017), CAs in team collaboration (e.g., managing tasks within a team) (Bittner et al., 2019; Seeber et al., 2019), and the customer interface (e.g., providing services) (Diederich et al., 2021; Vaccaro et al., 2018; Wünderlich & Paluch, 2017). In private life, CAs are primarily used for individual task support (e.g., searching for information online) (Porcheron et al., 2018), education (Graesse et al., 2017), or personal health (Yokotani et al., 2018). Further, studies that do not fit any of these two contexts are classified as "others" (e.g., CAs for legal research (Sugumaran & Davis, 2001)), and articles that do not fit a specific context are considered "generic" (e.g., Candello et al., 2019).

Second, the dimension **human** refers to the user interacting with the CA. In general, users' characteristics include demographic aspects such as age, gender, cultural background, experience both with CAs and with the task at hand, as well as aspects related to individual dispositions, including a user's personality or cognitive style (Zhang & Li, 2005)

The third dimension, **agent**, includes characteristics of the CA itself. These characteristics comprise the primary communication mode, i.e., the agent adjusted for interaction via speech, text, or both (Gnewuch et al., 2017), as well as the agent's embodiment, which can be physical, as with service robots (Stock & Merkle, 2018b), virtual with an interactive avatar (Seymour et al., 2018), virtual with a static avatar (Wünderlich & Paluch, 2017), or non-existent, for example, a CA without any visual embodiment (Abul et al., 2018). Further, CAs exhibit different design components comprising an identity (e.g., name), verbal communication (e.g., expression of emotions), and non-verbal communication (e.g., response delays to indicate thinking and typing) (Seeger et al., 2018).

Finally, **perception and outcome** refer to the investigated topic relating to the use and impact of the technology. This dimension shows how users perceive the CA during their interaction, as well as the interaction's impact. Following Zhang & Li's (2005) suggestions, these topics can be divided into eight distinct categories, as shown in Table 1.

Categories	Exemplary Constructs
Perception	Humanness, similarity, reciprocity, social distance, social presence
Acceptance	Usefulness, ease of use
Attitude	Attitude, satisfaction, preference
Performance	Productivity, effectiveness, efficiency
Emotion	Affect, hedonic quality, enjoyment, humor, intrinsic motivation
Trust	Trust, risk, loyalty, security, privacy
Learning	Learning models, learning processes, general training
Ethics	Ethical belief, ethical behavior, ethics
Relationship	Influence, interdependence, interference, agreement/disagreement

Table 1. Categories and Exemplary Constructs in the Dimension Perception and Outcome

To complement the dimensions of the research framework, we further included the research method, unit of analysis, and the theoretical grounding of the study in our review. Thereby, we sought to better understand the methodological focus of current CA research and theories used to inform CA design, and to improve insight on the user's interaction with the CA. Following Bariff & Ginzberg's (1982) explanations, we differentiate between studies on the individual level (e.g., user reactions toward CAs), on the group level (e.g., CAs as team members), on the organizational level (e.g., use cases for CAs in enterprises), and on the inter-organizational level. To analyze the methods used in the studies, we drew on the research methods Banker & Kauffman (2004) described and used in their assessment of IS research.

3.2 Identification, Coding, and Analysis of CA Literature

To collect and analyze existing studies on CAs in IS and CS research, we followed a process based on the combination of the systematic literature review guidelines by Webster and Watson (2002), vom Brocke et al. (2009), and Bandara et al. (2015). As our goal is to provide an organizing and assessing review of research on CAs (Leidner, 2018), we focused on research outcomes of the studies in the scope of this review, choosing to organize them conceptually (Cooper, 1988) using the HCI research framework adapted from Zhang & Li's (2005) seminal work. We took a three step approach (Table 2) which meant collecting literature for the review (step 1), coding the studies qualitatively using the research framework (step 2), and conducting analyses (step 3) by creating a concept matrix, investigating CA research over time and clustering the studies to identify research streams.

	Step 1: Collect Literature	**Step 2: Code Studies**	**Step 3: Conduct Analyses**
Input	Search query	Literature database	Coded literature database
Method	Literature search	Closed coding	Concept matrix, cluster analysis
Steps	Conduct database search and identify relevant CA studies	Code CA studies using dimensions from the research framework	Create concept matrix and descriptive statistics, conduct clustering
Results	Literature database	Coded literature database	Concept matrix and four research streams (cluster)

Table 2. Research Approach

To initiate our literature review, we identified the relevant outlets for our search process in step 1. The IS and CS communities conducted extensive work on CAs as part of HCI research, thus we purposefully selected journals from these fields. For the IS discipline, we focused on the Basket of Eight, and for CS, we selected four well-regarded journals that focus on HCI (i.e., Advances in Human-Computer Interaction, ACM Transactions on Computer-Human Interaction, Computers in Human Behavior, and the International Journal of Human-Computer Studies) for our search. We extended the data search by adding high-quality conference presentations to take more recent work into account, as renewed interest in CA research emerged only a few years ago (Pfeuffer, 2019; Rzepka & Berger, 2018). Thus, we complemented our review with proceedings from major IS conferences (ICIS, ECIS, HICSS, AMCIS, and PACIS) and CS conferences (ACM CHI Conference on Human Factors in Computing Systems). To collect appropriate studies, we used the Web of Science, AISeL, ACM Digital Library, and the websites of the respective outlets. We conducted the search in January 2019 and updated it in November 2019, using the following search query:

((Conversational OR Interactive OR Virtual) AND Agent) OR Chatbot OR Digital Assistant)

The query returned 8,768 results in all, for which we scanned titles, abstracts, and content to identify studies that focus on CAs. Further, we conducted a forward and backward search to identify additional studies. After this search and filtering process, 262 studies remained in our database (Table 3, outlets in alphabetical order).

Outlet	Search Results	Relevant
ACM CHI Conference on Human Factors in Computing Systems	3,661	102
ACM Transactions on Computer-Human Interaction	144	4
Advances in Human-Computer Interaction	78	6
Americas Conference on Information Systems	1,250	9
Computers in Human Behavior	1,072	42
European Conference on Information Systems	334	7
Hawaii International Conference on System Sciences	287	10
International Conference on Information Systems	731	26
International Journal of Human-Computer Studies	421	47
Journal of Management Information Systems	1	2
Journal of the Association for Information Systems	246	3
Pacific Asia Conference on Information Systems	543	4
Total	**8,768**	**262**

Note: Some studies were identified through backward and forward search

Table 3. Literature Search Results

In the second step of our research, we coded the 262 studies using the dimensions of our research framework (see Figure 2). Complementing the framework dimensions, we coded the research approach (empirical or conceptual), method (e.g., laboratory experiment or survey), unit of analysis (e.g., technology or individual), and noted the study's theoretical grounding (e.g., Similarity-Attraction theory (Byrne, 1971; Byrne & Griffitt, 1969) or Computers Are Social Actors Paradigm (Nass & Moon, 2000; Reeves & Nass, 1996)).

To ensure the reliability of the coding, three of the authors coded a random set of twenty studies independently in a pre-test using a preliminary coding guideline. The authors then discussed the coded studies to identify discrepancies and shortcomings in the codes. Based on the results of this pre-test, we adjusted the codes and the coding guidelines (e.g., adding "technology" as a unit of analysis for studies exclusively containing artifact descriptions or selection of multiple codes in the dimensions "perception and outcome") (see Appendix A). After this pre-test, one author coded the remaining studies and, when required, discussed uncertainties with the other authors. Based on the coded 262 studies, we carried out three analyses in step 3 to assess the state of research on CAs. First, we created a concept matrix to foster a conceptual understanding of the studies, going beyond descriptive content summaries, and we viewed the distribution of characteristics in the coding dimensions

(Webster & Watson, 2002). Second, we conducted a cluster analysis to identify research streams in the extant CA literature empirically. Following Punj & Stewart's (1983) recommendations, we first identified a suitable number of clusters using a hierarchical clustering approach, Ward's method with squared Euclidean distance, and afterwards used k-means as an iterative partitioning technique. We selected a solution with four clusters after reviewing the scree plot dendrogram and coefficient delta (see Appendix B). The k-means procedure then computed seven iterations until we achieved no further significant enhancements. Finally, we created graphical representations to show how CA research had developed over time, both by discipline (IS, CS) and by research stream.

4 Results

Our organizing and assessing review offers insight regarding the state of CA research and allows us to derive recommendations to advance our understanding of the design of and interaction with CAs. In this section, based on our review, we first describe our general observations on CA research. Next, we outline four research streams identified through a cluster analysis of the studies in our sample.

4.1 Overarching Observations

In accordance with our review framework (see Figure 2), we relate our findings to the technical characteristics of the studied CAs ("agent"), the application context, the user ("human"), and the perception and outcomes emerging from the interaction.

4.1.1 Agent Dimension

Regarding the agents we investigated (see Table 4), we found that around half of the studies focused on communication with the agent via written text (e.g., chatbots (Adam & Klumpe, 2019; Vaccaro et al., 2018)), around 40 percent explored speech-based CAs (e.g., digital assistants like Siri or Alexa (e.g., Burton & Gaskin, 2019; Winkler & Roos, 2019)), and nearly 10 percent studied both communication modes, mostly to identify commonalities or differences in human interaction with such agents via text or speech (e.g., Schroeder & Schroeder, 2018). Concerning the representation or embodiment of the CA, nearly half of the studies investigated agents without any form of embodiment (i.e., without a static/interactive digital avatar or physical appearance (e.g., Gnewuch et al., 2018; Schuetzler et al., 2018)). One-fourth of these studies focused on virtual interactive representations (i.e., with a virtual human (e.g., Cafaro et al., 2016; Krämer et al., 2013)). Further, 17 percent of the studies address CAs with virtual static avatars (e.g., images (Seeger et al., 2018)), 8 percent explore agents with physical embodiment (e.g., service robots (Stock & Merkle, 2018a, 2018b)), and a handful investigate and compare multiple embodiments (e.g., Gong, 2008; Seymour et al., 2017).

Around half of the studies focus on verbal communication (e.g., dialogue repair strategies (Corti & Gillespie, 2016) or expressing emotions (e.g., Beale & Creed, 2009; Niewiadomski & Pelachaud, 2010), around one-third study the agent's (human) identity (e.g., the impact of agent representations on user perception (e.g., Vugt et al., 2010)), and one-fourth explores non-verbal communication (e.g., response times (Gnewuch et al., 2018) or facial expressions (De Rosis et al., 2003). Table 4 summarizes the types of CAs investigated in the studies.

Communication Mode	
Text-based	51.9%
Speech-based	38.9%
Both	9.2%

Design Dimensions[1]	
Verbal communication	49.2%
(Human) Identity	31.7%
Non-verbal communication	24.4%

Representation / Embodiment	
None	46.2%
Virtual interactive	24.8%
Virtual static	16.8%
Physical	8.0%
Multiple	4.2%

1) Multiple selections possible

Table 4. Types of Agents investigated in the Studies

4.1.2 Context Dimension

Concerning the context in which CAs are applied (Table 5), 45 percent of the studies do not explicitly specify a context, investigating only the interaction of humans and CAs generally (e.g., Banks, 2018; Chaves & Gerosa, 2018). Around 21 percent of the studies addressed professional contexts (e.g., customer service (Baier et al., 2018; Xu et al., 2017) or marketing and sales (Kim et al., 2018; Vaccaro et al., 2018), 12 percent addressed education (e.g., interactive tutoring systems (Hobert & Wolff, 2019; Winkler & Roos, 2019), 9 percent dealt with health applications (e.g., digital health advisors (Gambino et al., 2019; Powers & Kiesler, 2006), or behavior change agents (Sebastian & Richards, 2017). The remaining studies focused on private individual task support (6.9%) (Porcheron et al., 2018; Purington et al., 2017), multiple contexts (e.g., Meyer von Wolff et al., 2019), specific contexts that do not fit the aforementioned categories (e.g., legal research (Sugumaran & Davis, 2001) or CAs as role models (Rosenberg-Kima et al., 2008)).

Generic	45.0%	Other	5.0%
Customer interface	13.4%	Professional task support	4.6%
Education	11.8%	Team collaboration	3.1%
Health	8.8%	Multiple	1.5%
Private task support	6.9%		

Table 5. CAs' Application Contexts in the Studies

4.1.3 Human Dimension

Concerning the user, the vast majority of the studies did not distinguish between users according to their characteristics (Table 6). For example, most experimental studies show that there are no significant differences between control and treatment groups concerning demographics, nor do they investigate the potential effects such characteristics have as control variables. The influence of demographic characteristics such as users' age (Chattaraman et al., 2018; Kowalski et al., 2019) or gender (Braun & Alt, 2019; Meier et al., 2019) on the perception of and interaction with CAs is discussed or controlled for in only a few of the identified studies (10-12%). Similarly, only a fraction of the studies explores the impact of individual experience with CAs (6%) (Fadhil & Villafiorita, 2017; Schroeder & Schroeder, 2018) or the task at hand (2%) (Ashktorab et al., 2019; Laumer et al., 2019). Further user characteristics, such as personality (5%) (Mou & Xu, 2017; Straßmann et al., 2018), cultural features (2%) (Duan et al., 2018; Schlesinger et al., 2018), cognitive style (1%) (Crockett et al., 2017), or the user's education (0%) are very rarely studied.

Gender	11.8%	Cultural Background	2.3%
Age	10.3%	Task Experience	1.5%
CA Experience	6.1%	Cognitive Style	1.1%
Personality	4.6%	Individual Education	0.0%

Table 6. User Characteristics investigated in the Studies

4.1.4 Perception and Outcome Dimension

Regarding the agents' perception and the interaction's outcomes, we find that many studies focus on constructs related to users' attitudes toward the CA (24%), the agent's perception (23%), system or task performance (20%), and acceptance (17%). Concerning attitude, Burgoon et al. (2016), for example, measure the user's perceived sense of connectedness with the agent after the interaction, and Vugt et al. (2010) assess perceived facial similarity with an interactive CA.

Regarding perception, many of the reviewed studies focus on comparing different design options (regarding identity, verbal communication, and non-verbal communication) have on perceived anthropomorphism. For example, Seeger et al. (2018) varied an agent's verbal and non-verbal communication (use of self-references and emoticons) and its human-like identity (a human person's name and image), to study the impact on perceived anthropomorphism. Similarly, Gnewuch et al. (2018) and Diederich et al. (2019) studied the impact of response times and preset answer options on users' perception of humanness in a customer service encounter with a chatbot. Araujo (2018) explored different language styles, framings to introduce the CA, and names in relation to anthropomorphism.

Concerning performance, CA researchers proposed different system-related measures, such as response success rate (e.g., Liao et al., 2018) or perceived conversational ability (e.g., Shah et al., 2016). Additionally, some reviewed studies used performance measures related to the interaction's outcomes, considering, for example, purchase amounts for digital content Alexa provided (Son & Wonseok, 2018), or completion time in microtask crowdsourcing (Mavridis et al., 2019).

Other authors draw on constructs in acceptance models, such as the Technology Acceptance Model (TAM) (Davis, 1989) or the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al., 2003). Examples of studies measuring acceptance-related constructs are Qiu & Benbasat (2010), who assessed the perceived usefulness of a product recommendation agent depending on different demographic embodiments, and Hobert (2019) who drew on ease-of-use and intention to use in order to evaluate a CA intended to teach programming to IS and CS students.

Further studies focus on constructs related to emotion (12%), such as empathy (Leite et al., 2013; McQuiggan & Lester, 2007), trust (Saffarizadeh et al. (2017) or Sohn (2019)). Others focus on the relationship between the user and the agent (11%), such as social roles (Seering et al., 2019) and social distance (Kim & Mutlu, 2014). Constructs less frequently studied, are the perceptions of CAs, and outcomes of the interaction include rather specific aspects, coded as other (5%), which include audience effects (Candello et al., 2019), and constructs related to learning (3%), for example, retention of learnt content (Van Der Meij, 2013). Finally, a single study focuses on ethics, developing a moral agency scale (Banks, 2018). Table 7 summarizes the constructs and shows their distribution in the studies.

Category[1]	Exemplary Constructs	%
Attitude	Connectedness with the CA (Burgoon et al., 2016) Rapport (Krämer et al., 2018) Similarity (Vugt et al., 2010)	24.0%
Perception	Anthropomorphism (Seeger et al., 2018) Humanness (Gnewuch et al., 2018) Uncanniness (Tinwell & Sloan, 2014)	22.9%
Performance	Response success rate (Liao et al., 2018) Conversational ability (Shah et al., 2016) Purchase amount (Son & Wonseok, 2018)	20.2%
Acceptance	Usefulness (Qiu & Benbasat, 2010) Performance expectancy (Laumer et al.,2019) Ease of use and intention to use (Hobert, 2019)	17.2%
Emotion	Empathy (Leite et al., 2013; McQuiggan & Lester, 2007) Enjoyment (Bell, Sarkar, & Wood, 2019) Compassion (Looije et al., 2010)	12.2%
Trust	Cognitive and emotional trust (Saffarizadeh et al., 2017) Perception of trust (Elson et al., 2018) Privacy Concern (Sohn, 2019)	12.2%
Relationship	Social distance (Kim & Mutlu, 2014) Social roles (Seering et al., 2019) Relationship building (T. W. Bickmore & Picard, 2005)	10.7%
Other	Audience effects (Candello et al., 2019) Shorthand language (Hill et al., 2015) Speech portions in dialogue (Bittner & Shoury, 2019)	4.6%
Learning	Learning outcomes (Wang et al., 2008) Retention of learned content (Van Der Meij, 2013) Facilitation of learning (Zhang et al., 2019)	3.4%
Ethics	Perceived moral agency (Banks, 2018)	0.4%

1) Multiple selections possible

Table 7. Constructs related to Perception and Outcome investigated in the Studies

4.1.5 Complementary Findings

There are different observations regarding the discipline (IS, CS), research approach (empirical or conceptual), research methods, and unit of analysis. Overall, we identified 262 studies on CAs, of which two-thirds had been published in CS journals and conference proceedings, and one-third from the IS discipline. We noted a substantially increased interest in CAs since 2016, in both the IS and CS disciplines (see Figure 4 in Appendix C). While about half of the studies in our CS sample were published before 2016 (48%), the vast majority of IS studies (85%) were published after 2016, indicating that the IS discipline recently increased the interest in CAs as a research phenomenon.

The studies in our sample comprise empirical (84%) as well as conceptual work (16%). Regarding the research methods used in the reviewed studies, we found a focus on laboratory experiments (57%) and design science research (23%), predominantly with context-specific architectural descriptions of the designed CA artifacts (e.g., Anabuki et al. (2000), Hsu et al. (2017), and Jain et al. (2018)). Less common research methods in the reviewed studies include qualitative research, such as conceptual literature reviews (e.g., Li (2015)), interviews (4%), for example concerning use cases of CAs (Laumer et al., 2019), the use of secondary data (4%) like user reviews for natural language applications (Nguyen & Sidorova, 2017), and surveys (3%). Research in the field using field studies (2%) or experiments (2%) only account for a fraction of the used methods. Concerning the unit of analysis, most strongly focus on the individual level (73%), considering the topic from a purely technological perspective (26%).

Regarding the theoretical grounding of the studies in our sample, we found that about one-third of the studies (30%) explicitly draw on different theories or paradigms to understand human interaction with CAs or guide their design. In particular, researchers have referred to the Computers Are Social Actors (CASA) paradigm that Nass & Moon (2000) formulated. Based on a review of three sets of experimental studies, they have shown that individuals mindlessly apply social rules and expectations to computers exhibiting human characteristics or behavior, such as communication via natural language. CA researchers draw on the CASA paradigm for different kinds of study: for example to investigate the extent to which knowledge gained through human-to-human interaction can be applied to human-CA interaction, as in gender stereotyping (e.g., Cowell and Stanney (2005), Pfeuffer et al. (2019), Qiu and Benbasat (2010)) or to explore how users perceive different anthropomorphic agent designs (e.g., Gong, 2008; Kim et al., 2013; Lee and Choi, 2017)).

Additionally, several researchers, such as Seeger et al. (2018) or Tinwell & Sloan (2014), drew on the so-called "Uncanny Valley" (Mori, 1970; Mori et al., 2012) to understand adverse emotional reactions to anthropomorphic CAs and to propose how they can be overcome. The Uncanny Valley, in short, postulates that there is no monotonically increasing relationship between the human-likeness of a technological artifact and a human affinity toward it, but that there is a point (i.e., the Uncanny Valley) where "a person's response to a human-like robot would abruptly shift from empathy to revulsion as it approached, but failed to attain, a lifelike appearance" (Mori et al. 2012, p. 98). On this theme, for example, Seeger et al. (2018) found empirical evidence for an adverse emotional reaction to a CA that exhibits only partial human-like characteristics, and Seymour et al. (2017) suggested that interactivity, realized through an interactive 3D avatar that matches common human non-verbal cues, can contribute to overcoming the Uncanny Valley sensation. In addition to the CASA paradigm and the Uncanny Valley theory, we found that researchers draw on the Three-Factor Theory of Anthropomorphism (Epley et al., 2007), which explains psychological determinants of humans anthropomorphizing inanimate objects, or not. For example, Wagner and Schramm-Klein (2019) rely on this theory in discussing the anthropomorphism perception in interaction with digital assistants like Alexa. Furthermore, several researchers have adapted theories and concepts originally from human-to-human interaction for human-CA interaction. For example, Qiu and Benbasat (2010) and Vugt et al. (2010) draw on Similarity-Attraction Theory (Byrne, 1971; Byrne et al., 1967) to investigate the impact of demographic similarity and facial similarity between a user and an agent. Similarly, Kim et al. (2013) and Kim and Mutlu (2014) draw on the concept of social distance (Bogardus, 1947), manifested in physical proximity, organizational status, and task structure, to understand user responses to humanoid robots. Additionally, Gnewuch et al. (2017) draw on Grice's Maxims (Grice, 1975) for effective conversations to derive design principles for CAs in a customer service context.

Further theories and concepts adapted to human-CA interaction include social presence (e.g., Schuetzler et al. (2018) or Sohn (2019)), rapport (e.g., Krämer et al., 2018), reciprocity (e.g,. Chattaraman et al., 2018), and self-determination theory (e.g., Lechler et al., 2019).

Finally, researchers have drawn on established models to investigate the acceptance of CAs, such as the TAM (Davis, 1989) or the UTAUT (Venkatesh et al., 2003). For example, Wang & Benbasat (2005) adapt the TAM to study trust in recommendation agents, and Laumer et al. (2019) draw on the UTAUT to understand CA acceptance in healthcare.

Table 8 summarizes the theoretical grounding of the studies.

Theory	Seminal work	Exemplary studies
Computers Are Social Actors (CASA)	Nass and Moon (2000)	Cowell and Stanney (2005), Gong (2008)
Uncanny Valley	Mori (1970), Mori et al. (2012)	Strait et al. (2015), Tinwell and Sloan (2014)
Three-Factor Theory of Anthropomorphism	Epley et al. (2007)	Seeger et al. (2018), Wagner and Schramm-Klein (2019)
Similarity-Attraction Theory	Byrne (1971), Byrne et al. (1967)	Qiu and Benbasat (2010), Vugt et al. (2010)
Social Distance	Bogardus (1947)	Kim & Mutlu (2014)
Grice's Maxims	Grice (1975)	Gnewuch et al. (2017)
Social Presence	Gefen and Straub (2003)	Schuetzler et al. (2018)
Rapport	Tickle-Degnen and Rosenthal (1990)	Krämer et al. (2018)
Reciprocity	Moon (2000)	Chattaraman et al. (2018)
Self-Determination Theory	Ryan and Deci (2000)	Quynh and Sidorova (2018), Lechler et al. (2019)
Technology Acceptance Model	Davis (1989)	Wang & Benbasat (2005)
Unified Theory of Acceptance and Use of Technology	Venkatesh et al. (2003)	Laumer et al. (2019)

Table 8. Theoretical Grounding used in CA Research

4.2 Research Streams

By means of a cluster analysis we identified four streams in IS and CS research. Specifically, these streams differ in the type of investigated CA regarding communication mode and embodiment; they also differ in the design dimensions on which the studies focus, attending to the agent's identity, and to verbal and non-verbal communication. As shown in Table 9, we term the streams "Text-based CAs," "Virtual CAs," "Speech-based CAs," and "Physical CAs."

#	Research Stream	Description
1	Text-based CAs	Research on CAs with interaction via written text and no embodiment or virtual static representation.
2	Virtual CAs	Research on CAs with interaction via written text and spoken natural language, and with embodiment through virtual interactive avatars.
3	Speech-based CAs	Research on CAs with interaction via spoken natural language and without any virtual or physical embodiment.
4	Physical CAs	Research on CAs with interaction via spoken natural language and with physical embodiment.

Table 9. Research Streams for Conversational Agents

Table 10 shows the characteristics' distribution in the framework for each research stream. Note that multiple selections were used for coding human characteristics (e.g., studies that investigate the effects of age users' CA perception), design dimensions (e.g., studies that focus on CAs' verbal and non-verbal communication), and perception and outcome (e.g., studies measuring CA performance and perception), which lead to sum rows exceeding 100%.

	Text-based CAs	Virtual CAs	Speech-based CAs	Physical CAs
Number of studies	112	69	58	23
Percentage	42.7%	26.3%	22.1%	8.8%
Human				
Gender	14.3%	10.1%	8.6%	13.0%
Age	8.0%	13.0%	12.1%	8.7%
CA experience	4.5%	2.9%	12.1%	8.7%
Personality	3.6%	8.7%	1.7%	4.3%
Cultural background	2.7%	1.4%	3.4%	0.0%
Task experience	1.8%	0.0%	1.7%	4.3%
Cognitive style	0.9%	2.9%	0.0%	0.0%
Individual education	0.0%	0.0%	0.0%	0.0%
Context				
Generic	34.8%	66.7%	43.1%	34.8%
Customer interface	20.5%	2.9%	10.3%	17.4%
Education	15.2%	7.2%	6.9%	21.7%
Health	9.8%	10.1%	3.4%	13.0%
Private task support	6.3%	0.0%	17.2%	4.3%

	Text-based CAs	Virtual CAs	Speech-based CAs	Physical CAs
Other	6.3%	4.3%	5.2%	0.0%
Professional task support	3.6%	4.3%	6.9%	4.3%
Team collaboration	1.8%	1.4%	6.9%	4.3%
Multiple	1.8%	2.9%	0.0%	0.0%
Agent				
Communication mode				
Text-based	98.2%	37.7%	0.0%	0.0%
Speech-based	0.0%	43.5%	84.5%	100.0%
Both	1.8%	18.8%	15.5%	0.0%
Embodiment				
None	55.4%	1.4%	100.0%	0.0%
Virtual interactive	4.5%	85.5%	0.0%	4.3%
Virtual static	36.6%	2.9%	0.0%	4.3%
Physical	0.0%	0.0%	0.0%	91.3%
Multiple	3.6%	10.1%	0.0%	0.0%
Design Dimensions				
Verbal communication	55.4%	33.3%	62.1%	34.8%
(Human) identity	30.4%	59.4%	6.9%	17.4%
Non-verbal communication	5.4%	65.2%	8.6%	34.8%
Perception and Outcome				
Attitude	18.8%	40.6%	19.0%	13.0%
Perception	22.3%	33.3%	12.1%	21.7%
Performance	22.3%	10.1%	25.9%	26.1%
Acceptance	17.0%	13.0%	22.4%	17.4%
Emotion	7.1%	23.2%	8.6%	13.0%
Trust	13.4%	13.0%	6.9%	17.4%
Relationship	6.3%	14.5%	8.6%	26.1%
Other	4.5%	1.4%	8.6%	4.3%
Learning	3.6%	1.4%	5.2%	4.3%
Ethics	0.0%	1.4%	0.0%	0.0%

Table 10. Cross Tab Analysis

4.2.1 Research Stream 1: Text-based CAs

Studies in the first research stream focus on CAs with which users interact via written text and that either have no embodiment or only a virtual static representation (e.g., with an image). Thus, we refer to this stream as "Text-based CAs," which as a type of CA has attracted the research community's interest, particularly since 2016, which saw more CA studies published than virtual CAs, speech-based CAs, or physical CAs (see Figure 5 in Appendix C).

Research in this stream particularly investigates questions related to verbal communication (55%), such as how to repair conversation breakdown (Ashktorab et al., 2019; Corti & Gillespie, 2016), how to design effective agent communication and its user impact (Adler et al., 2016; Derrick & Ligon, 2014; Hu et al., 2018; Van Der Meij, 2013), or how to gauge message interactivity's influence on users' perception of an agent (Adam & Klumpe, 2019; Go & Sundar, 2019).

About one-third of the studies in this stream explore aspects related to the (human) identity of an agent (30%), such as how the agent's gender influences the user (Pfeuffer et al., 2019; Qiu & Benbasat, 2010) or how the degree of human-likeness in the agent's representation (e.g., a comic image of a robot or photo of an actual human person) impacts user perception (Gong, 2008). Further studies do not focus on specific design dimensions but address overarching topics, such as chatbot use cases (Laumer et al., 2019; Meyer von Wolff et al., 2019), or affordances (Stoeckli et al., 2018).

Concerning the contexts, many text-based CA studies do not specify a particular application area, or they focus on the customer interface of organizations (35%), such as online customer services (e.g., Gnewuch et al. (2018), Hu et al., (2018), Xu et al. (2017)) or marketing and sales (e.g., Al-Natour et al. (2009), Vaccaro et al., (2018), Van den Broeck et al. (2019)). Further frequently studied contexts include education (15%) and health (10%). In the context of education, text-based agents have been investigated for learning languages (Fryer et al., 2017), collaborative problem solving (Hayashi, 2013; Herborn et al., 2018), and programming education (Hobert, 2019). In the area of health, researchers have focused on, for example, chatbots for therapy (Bell et al., 2019; Constantin et al., 2019), for raising individual health awareness (Meier et al., 2019), or supporting people with allergies (Hsu et al., 2017).Research on text-based CAs focuses on constructs related to perception (e.g., humanness), attitude (e.g., attractiveness), performance (e.g., responsiveness), and acceptance (e.g., perceived usefulness). Regarding text-based agents, studies mainly investigate how design variations concerning the agent's identity (e.g., Araujo (2018), Go & Sundar (2019)) and verbal communication (e.g., Hu et al. (2018), Schuetzler et al. (2014)), influence how the user perceives the agent's anthropomorphism or humanness.

Complementary to the question of how to design human-like text-based agents, researchers have discussed how an anthropomorphic design might trigger perceptions of uncanniness in the interaction (e.g., Gnewuch et al. (2018), Wünderlich and Paluch (2017)). Finally, the outcomes of anthropomorphism perception in chatbots have been studied, regarding for example, service encounter satisfaction (Gnewuch et al., 2018), brand perception (Araujo, 2018), and learning results (Jin, 2010).

4.2.2 Research Stream 2: Virtual CAs

Research in the second stream, termed “Virtual CAs,” comprises work on agents represented by virtual animated avatars with which users interact via written and spoken natural language. The research interest in both disciplines regarding virtual agents as a type of CA has been steady, particularly in comparison to studies on text-based agents and speech-based CAs (see Figure 5 in Appendix C). In contrast to research focused on verbal communication in the first stream, studies on virtual CAs, in particular, investigate non-verbal communication (65%) and the (human) identity of an agent (59%). For example, Krämer et al. (2013) explored how an agent’s smile impacts user perception and behavior, finding that humans reciprocate an agent’s smile. Similarly, several studies in this stream explored the impact of an agent’s gaze, such as catching the user’s eye, on the user (Vertegaal et al., 2000, 2001), the effects of facial similarity between the agent and the user (Vugt et al., 2010), or the agent’s digital gesturing behavior (Biancardi et al., 2017). Finally, Seymour et al. (2018) presented emerging natural face technology for creating a realistic visual presence of agents and proposed a research agenda which includes fundamental philosophical questions arising from such agent representations becoming increasingly realistic.

Further, virtual agents are mostly studied without a specific context being described (67%). However, where context was given, the studies mostly focused on their application in health (10%) or education (7%). Regarding health, research, for example, explored virtual CAs to change stigmatizing attitudes toward mental health (Sebastian & Richards, 2017) or potential advantages such agents could have over clinical psychologists in mental health interviews (Yokotani et al., 2018). In educational context, Carlotto and Jaques (2016), for example, studied animated pedagogical agents and suggested that an agent’s movement and gestures contribute less to learning outcomes than interaction via speech. Similarly, Gulz and Haake (2006) explore the design of agent animations and student motivation to learn.

Lastly, studies on virtual CAs mainly concerned the user’s attitudes toward them, user perception, and emotions, often related to the agent’s non-verbal communication behavior. For example, Burgoon et al. (2016) studied how a match between user expectations and the agent’s behavior influenced the individual attitude toward the agent, finding that in particular positive deviations from user expectations have a substantial positive effect on perceived connectedness, receptivity, and dependability. Regarding emotions, others investigated interactive designs, for example, expressing politeness (Niewiadomski & Pelachaud, 2010) or displaying empathy (McQuiggan & Lester, 2007; Yang et al., 2017) which are direct responses to the user’s inputs. More recent studies further explored how an agent’s interactive representation and non-verbal communication behavior impact such agents’ persuasiveness (Harjunen et al., 2018; Hyde et al., 2015; Looije et al., 2010; Rosenthal-von der Pütten et al., 2018), as in decision-making tasks.

4.2.3 Research Stream 3: Speech-based CAs

The third research stream, termed "Speech-based CAs," includes primarily speech-based CAs (85%) without any physical or virtual embodiment. Similar to research on text-based CAs, the number of publications that focus on speech-based agents has substantially increased since 2016 (see Figure 5 in Appendix C). Research in this stream concentrates on verbal communication with the agent (62%) or topics not directly related to technological aspects, such as use cases (Baier et al., 2018), team settings with agents for collaborative work (Bittner et al., 2019), or the psychological impact of interacting with such assistants in a commanding voice (Burton & Gaskin, 2019). Researchers in this stream often draw on widely distributed assistants, such as Amazon's Alexa (Son & Wonseok, 2018; Winkler et al., 2019; Winkler & Roos, 2019), Apple's Siri (Burton & Gaskin, 2019), or Google Assistant (Cho, 2019). Concerning verbal communication, research on speech-based CAs is concerned with conversation design, such as for useful task guidance (Vtyurina & Fourney, 2018), for fostering a positive user experience (Burmester et al., 2019), or for combining social and functional ways of communication (Clark et al., 2019).

Similar to research on virtual CAs, studies in this stream often do not specify a particular context (43%). However, in contrast to the two previously mentioned streams, several studies on speech-based agents focus on individual task support in a private context, for example, using Alexa and Google as smart home components (Kowalski et al., 2019; Porcheron et al., 2018; Purington et al., 2017). Further studied contexts include digital assistants at the customer interface (10%), such as for advisory services (Dolata et al., 2019) or advertising (Kim et al., 2018), in education (7%), such as for learning languages (Morton et al., 2012) or team collaboration (Winkler et al., 2019a).

The aspects this research stream investigates relate particularly to performance (26%), acceptance (22%), and attitude (19%). Concerning performance, research on speech-based CAs has, for example, explored the number of tasks completed in teams (Winkler et al., 2019a; Winkler et al., 2019b), conversation turns and time required to retrieve information from the assistant (Le Bigot et al., 2006), or the agents' social dialogue capabilities (Ward & Tsukahara, 2003). With regard to acceptance, researchers have, for example, studied the influence of an agent's answers to user questions on perceived usefulness (Jung et al., 2019) or the impact of the agent's personification and social interaction capabilities on user satisfaction (Purington et al., 2017). Finally, related to attitude, studies on speech-based agents have investigated aspects such as gender stereotypes arising from an agent's female voice (Hwang et al., 2019) or the agent's answers' impact on perceived politeness and pleasantness (Jucks et al., 2018).

4.2.4 Research Stream 4: Physical CAs

The fourth research stream comprises research on CAs with a physical embodiment (91%), which interact via spoken natural language. We refer to this stream as "Physical CAs." The research interest in CAs with a physical embodiment has been comparatively steady, similar to the research stream of virtual agents (see Figure 5 in Appendix C). Studies in this stream emphasize verbal (35%) and non-verbal communication (35%). Concerning verbal communication, researchers, for example, investigated how speech style, including calling users by their names, influenced the perceived social distance toward the physically embodied agent (Kim et al., 2013) or how different sales strategies impacted the number of goods conversational robots sold in a department store (Watanabe et al., 2015). Research focusing on non-verbal communication has focused on designing different types of behaviors, such as nodding or maintaining eye contact, and evaluating such behaviors' impact on the user and her or his perception of the physical CA (Lee et al., 2004; Saerbeck et al., 2010; Sakamoto et al., 2005; Szafir & Mutlu, 2012; Yamada et al., 2013). Research on physical CAs rarely specifies a particular context (35%), but if so, investigates them in education (22%), at the customer interface (17%), or in health contexts (13%). In the context of education, these studies have explored physical CAs as learning partners that support social and cognitive aspects in the learning process (Huang, 2012; Saerbeck et al., 2010; Zhang et al., 2019). At the customer interface, such physical CAs were investigated to support product sales (Bertacchini et al., 2017; Watanabe et al., 2015) or to innovatively provide customer service (Stock & Merkle, 2018b). Finally, physical CAs have been studied in health, for example, as caregivers (Kim et al., 2013) or to promote regular physical exercise (Kanaoka & Mutlu, 2015). About the perception as well as outcomes of the interaction with physical CAs, studies have focused on the relationship between user and agent (26%), the CA's performance (26%), or the user's perception of the physical agent (21%). Concerning the relationship between the physically embodied agent and the user, Sangseok and Lionel (2019), for example, investigated how team members identify with a physically embodied CA while completing a collaborative task, and how this impacts subgroup formation. Lee et al.'s (2012) work provides a further example, discussing how users assign social roles to physical CAs and how users' relationship to the agent with a physical embodiment influences their behavior toward other users (e.g., protecting the agent or jealousy). Concerning performance, scholars, for example, have researched the number of products a sales agent sold (Watanabe et al., 2015) or learning results associated with a physical CA's support (Szafir & Mutlu, 2012). Regarding the perception of a physical agent, researchers have studied aspects such as social presence (Pereira et al., 2014), humanness (Kim et al., 2013), or social behavior (Xu et al., 2013).

5 Discussion

To date, researchers interested in CAs have lacked an overarching framework, classification, or suitable organizing device to compare and analyze existing CA research. Identifying relevant future research topics or areas has been hindered by the variety and numbers of extant studies and findings, so that an organizing structure is missing. Our framework, adapted from established research by Zhang and Li (2005), supports researchers in the classification and organization of existing and future CA research. In the following section, we discuss the observations we made during the review process, and we suggest avenues and directions for future CA research.

5.1 An Agenda for Conversational Agent Research in IS

Based on our analysis of the identified CA studies, we propose six avenues for future research that address research gaps that have not been investigated yet and need to be studied from an IS perspective. For each avenue we motivate and formulate specific directions for future research activities to address the identified gaps in the body of IS knowledge. Table 11 summarizes the avenues and gives directions for further research that will advance our understanding of the interaction between humans and CAs.

Research Avenue	Directions for Future CA Studies
Avenue 1: User Characteristics and Adaptive CA Designs	D1.1: Investigate the impact of user characteristics in the interaction between humans and CAs on human perception and interaction outcomes across different contexts. D1.2: Investigate how CAs can be designed to adapt themselves to individual users and their characteristics during the interaction. D1.3: Study the potential that CA configuration or CA co-creation has for the user, and the resulting impact on user perception and interaction outcomes.
Avenue 2: CAs on Group and Organizational Levels of Analysis	D2.1: Investigate which parts of collaborative group work can be fulfilled by CAs and the resulting impact on team behavior and performance. D2.2: Examine how CAs should be designed to efficiently support human team collaboration and how these designs influence team members' perception of and interaction with CAs. D2.3: Analyze which types of (organizational) tasks and processes are suitable for innovation and automation with CAs. D2.4: Investigate the positive and negative impact and the potential mitigation strategies when automating human work with (anthropomorphic) CAs.
Avenue 3: CAs with Virtual Interactive and Physical Embodiments	D3.1: Explore the impact of rich virtual interactive or physically embodied CAs on user perception combined with the currently available conversational capabilities. D3.2: Investigate designs of virtual interactive and physical CA embodiment to increase agent acceptance, adoption, and performance across different contexts.
Avenue 4: Transferability of Knowledge gained in CA Studies	D4.1: Partially replicate experimental CA studies in investigating whether existing prescriptive and descriptive knowledge can be transferred between different CA instances (i.e. combinations of human, agent, and context) and adapted accordingly.

Research Avenue	Directions for Future CA Studies
Avenue 5: Ethical Implications of Designing and Interacting with CAs	D5.1: Explore when, where, and how applying persuasive CAs is ethically justifiable. D5.2: Develop CA's design elements to address the ethical dimensions of the user interaction, and investigate how users perceive these design elements. D5.3: Investigate the unintended side-effects of CA design and study how to prevent such negative side-effects.
Avenue 6: Longitudinal CA Research in a Field Study Setting	D6.1: Conduct field studies to verify, extend, or refute existing knowledge gained from experimental research in controlled settings. D6.2: Investigate the interplay between CAs' limited conversational capabilities and the combination of multiple (anthropomorphic) design features of CAs. D6.3: Study how the relationship between users and CAs takes shape during initial use of the CA, and how it develops in multiple interactions over a longer period of time.

Table 11. Agenda for Conversational Agent Research in IS

5.1.1 User Characteristics and Adaptive CA Designs

We found that only a quarter of the reviewed studies investigated how individual human characteristics influence the perception of and interaction with CAs. Of these articles, several empirical studies found significant effects. For example, Schroeder and Schroeder (2018) observed that younger and male users are more likely to trust a CA. Rosenthal-von der Pütten et al. (2018) found that older users are more easily persuaded by a CA that shows dominance than by one with submissive behavior. Vugt et al. (2010) discovered that female users are more likely to use a CA with facial similarity, while male users are more likely to prefer a CA with facial dissimilarity. In addition to age and gender, characteristics like experience with CAs (e.g., Ashktorab et al. (2019), Otoo & Salam (2018)), personality traits (e.g., Krämer et al. (2018), Seeger et al. (2020)), or cognitive style (e.g., Hubal et al. (2008)) have been found to influence how users perceive and interact with CAs. These studies mostly considered human characteristics as complementary to the key constructs they investigate. However, for characteristics like age and gender, experience with CAs, the task at hand, and a user's personality, seemed to significantly influence how humans perceive and interact with CAs in different contexts (i.e., the usage scenario of the CA, such as health, education, private task support; see table 13 for further contexts). Based on these initial findings, we suggest there is a promising research opportunity to explicitly investigate the impact of individual user characteristics (e.g., gender, age, personality, cultural background, etc.) on the interaction between humans and CAs. This applies to different contexts and to aspects of the agent's perception, as well as to the interaction's outcomes. For example, questions on how cultural differences between users influence the perception of features regarding the agent's identity (e.g., name, social role), verbal communication (e.g., formal or informal language, expression of emotions), and non-verbal communication (e.g., use of gestures and facial expressions, emoticons) arise. Similar to the various culturally modifiable interface components of graphical user interfaces (Reinecke & Bernstein, 2013), we argue that a variety of CA features is likely to influence how users perceive such agents, depending on the user's cultural background. In short, a sound understanding of individual user characteristics' impact, considering the context as well as agent's design features, can allow us to better understand differences and commonalities in the perception of CAs, and in outcomes of the interaction. In summary, we propose the following direction (D):

D1.1*: Investigate the impact of user characteristics on the interaction between humans and CAs on human perception and interaction outcomes across different contexts.*

With our improved understanding of user characteristics and their influence on the interaction between CAs and their users, we propose investigating how CAs can be designed to adapt themselves to individual users during the interaction. In practice, CAs are typically implemented using a "one size fits all" approach in which all users receive the same agent and set of responses, regardless of characteristics like age, experience with such agents, or personality (Følstad & Brandtzæg, 2017). Therefore, investigating adaptive designs securely founded on empirical data and theory (Kocaballi et al., 2019) represents a lucrative research opportunity. We suggest that CAs with efficient adaptive designs for heterogenous user groups are likely to increase such agents' acceptance, adoption, and performance in line with studies for other types of IT artifacts, such as recommendation agents with graphical user interfaces (Al-Natour, Benbasat, & Cenfetelli, 2006). Researchers might be able to draw on established theories and constructs from human-to-human interaction, such as communication accommodation theory (Giles, Coupland, & Coupland, 2010), similarity attraction theory (Byrne, 1971; Byrne et al., 1967), matching and mirroring (Burgoon, Stern, & Dillmann, 1995), or mimicry (Kozlowski & Ilgen, 2006) to inform adaptive CA designs. For example, researchers can explore how CAs can efficiently tailor their verbal communication (e.g. selection of words or syntax) to different user groups in the same way as humans adjust their language style in interactions (Pickering & Garrod, 2004). Because moving away from 'one size fits all' approaches is likely to substantially increase the implementation's complexity, we further suggest evaluating and comparing different adaptive CA designs in order to identify the most efficient features drawing on the three design dimensions of such agents, namely (human-like) identity, verbal- and non-verbal communication).

D1.2*: Investigate how CAs should be designed in order to adapt themselves to individual users and their characteristics during the interaction.*

As an alternative approach to user-adaptive designs, researchers should investigate configurable agents' potential that allows users to co-create CAs that fit their preferences.

Studies on several types of technological artifacts, such as mobile devices (Carter, Grover, & Thatcher, 2013) or avatars (Belk, 2013), have shown that humans can identify with certain inanimate objects, thereby promoting emotional attachment. Individuals beginning to see objects as a part of their identity fosters this identification and the resulting emotional attachment, as the psychological process of self-extension also explains (Belk, 1988). Building or configuring one's own technology has been shown to trigger the process of self-extension for different types of artifacts such as robots (Groom et al., 2008; Robert & Sangseok, 2018) or avatars (Ducheneaut et al., 2009). Fostering identification with CAs through configurable designs or co-creation is likely to have a positive impact on, for example, enjoyment of the interaction (Li, Browne, & Chau, 2006) or team performance in collaborative settings (Robert & Sangseok, 2018) as found in studies on other types of technological artifacts. Thus, we propose that researchers investigate how CAs can offer the possibility of configuration or co-creation, including shaping the agent's identity (name, gender, appearance). They should also study how this impacts users' perception of the agent, as well as the interaction's outcomes. In considering such designs, we suggest researchers particularly investigate users' social responses to the agents and how this impacts the relationship between the user and the (human-like) CA.

D1.3*: Study the potential of CA configuration or CA co-creation for the user and the resulting impact this has on user perception and interaction outcomes.*

5.1.2 CAs on Group and Organizational Levels of Analysis

Considering the unit of analysis, we observed that the majority of the studies in our sample focus on individual interaction with CAs (73%) or on technological descriptions of CA designs (26%). Notable exceptions within the group and on (inter-) organizational levels of analysis are Bittner et al. (2019) who developed a taxonomy of design option combinations for CAs in collaborative work, and Cardona et al. (2019) who studied adoption and diffusion of conversational technology in the German insurance sector. Apart from these studies, the large majority of CA research focuses on the individual level, i.e. the interaction between a single user and a single CA.

In line with researchers in the area of computer-supported collaborative work, such as Seeber, Bittner, et al., (2019) and Seeber, Waizenegger, et al. (2019), we suggest investigating CAs in group settings. We argue that, driven by advances in natural language processing and machine learning, CAs with significantly improved capabilities emerged (McTear, 2017). Also, they have human-like characteristics that can alter the role of IT from one of providing tools that enhance team performance to eventually becoming artificial teammates (Malone, 2018). This technological progress has given rise to numerous questions related to the design of CAs in team settings, as well as to group interaction with such agents. For example, which roles in a team can an anthropomorphic CA assume, i.e., in what situations is the CA able to fulfill a gap where a human team member is missing? What are the advantages and limitations of a CA in this role? While a large body of collaboration research is available on team compositions and roles of team members (e.g., Belbin (2010)), we lack an understanding on which of the roles CAs can fulfill and what the implications are when technology assumes such roles. While we expect research on CA roles in team settings initially to focus on operational tasks, such as managing the task, gathering information, or scheduling meetings, we anticipate that even more capable CAs will emerge, which are able to assume roles typically associated with human team members. For example, CAs equipped with present-day sensing capabilities are already able to recognize individuals' sentiments (Bertacchini et al., 2017; Feine et al., 2019) and could be able to take measures where appropriate, such as proposing a break to improve the general mood within the team. Thus, we suggest investigating this and similar team roles by drawing on conversational technology available in practice, to better understand which parts of collaborative work CAs can fulfill and how an artificial teammate will impact team behavior and performance.

D2.1*: Investigate which parts of collaborative group work can be fulfilled by CAs and the resulting impact on team behavior and performance.*

Further, future studies should investigate the impact different agent designs have, such as text-based or speech-based communication in research streams 1 and 3, or embodiment as in research streams 2 and 4, in collaboration settings. For example, do team members perceive the same agent differently depending on the communication mode, i.e., on whether the members-CA interaction takes place via spoken or written natural language? And, what is the impact of a higher degree of agent anthropomorphism on the acceptance, adoption, and use of a CA within a team? According to extant research on the impact CA anthropomorphism has on the individual perception of such agents (e.g., Araujo (2018), Go & Sundar (2019), Rosenthal-Von Der Pütten & Krämer (2014)) and on the understanding of computers as social actors (Nass & Moon, 2000), we expect the agent's increased human-likeness to strengthen the social responses to an agent within a team. This can be both positive (e.g., regarding CA acceptance in the team or collective enjoyment of the interaction) and negative (e.g., regarding shared feelings of frustration when the agent is not able to fulfill high expectations fostered by a human-like design). While some of these social responses might be similar to individual interaction with and perception of CAs, we do not understand how group settings (i.e., multiple humans interacting with one or more CAs at the same time) influence such responses. Therefore, we propose studying how CAs should be designed to be efficient in team collaboration, and how such designs influence the way team members perceive and interact with (anthropomorphic) CAs.

D2.2: *Examine how CAs should be designed to efficiently support human team collaboration and how these designs influence team members' perception of and interaction with CAs.*

In addition to generating a better understanding of the potential and limitations of CAs in team collaboration, we propose studying such agents on the organizational level. Currently, the majority of CA studies in organizations focuses on the customer interface, investigating contexts like customer service (e.g., Hu et al. (2018), Stock & Merkle (2018b), Wünderlich & Paluch (2017)) or marketing and sales (e.g., Hanus & Fox (2015), Kim et al. (2018), Vaccaro et al. (2018)). Further, we found single studies in other, quite specific contexts, such as for onboarding new employees (Liao et al., 2018) or for assisting workshop moderation (Strohmann et al., 2018). However, we still lack an understanding of the key characteristics that determine whether introducing a CA makes sense in a specific context. At present, studies in new organizational contexts emerge in a bottom-up approach, possibly driven by practical interests. Thus, we posit research potential to identify the types of tasks for which CAs can be useful in a more abstract, context-independent way. For example, such task types can be characterized by their rather structured nature and by a high occurrence frequency. However, we do not know whether these tasks necessarily comprise an interaction that usually takes place between two humans, i.e., a task where an anthropomorphic agent can at least partially substitute the human contact. Nor do we know whether such tasks typically involve interaction with complex software where the CA is intended to increase ease-of-use. A sound understanding of the task types suitable for automation and innovation by means of CAs could assist us in moving from the practical, opportunity-driven identification of application contexts to a top-down approach in which such contexts could be determined by systematically reviewing tasks and processes within an organization.

D2.3: *Analyze which types of (organizational) tasks and processes are suitable for innovation and automation with CAs.*

Additionally, there is an opportunity to investigate the consequences that the introduction of CAs has on the human workforce. Davenport and Kirby (2016) discuss how cognitive technology (or AI technology in general) can support humans in performing various tasks. Similarly, Brynjolfsson and McAfee (2016) elaborate on how IT is facilitating automation, thus leading to task performance shifting from the human to the computer and so announcing the second machine age. They argue that with these advances, more and more tasks traditionally performed by humans are being automated, replacing workers with intelligent IT that has positive (e.g., less simple, repetitive tasks for people) and negative consequences (e.g., employees fearing job losses). Building on this, researchers should explore how human workers perceive the CA take-over of tasks they previously performed. For example, in a customer service context, more and more first level support is performed by CAs (Huang &

Rust, 2018) and human operators that answer basic service requests are no longer required. While this can be interpreted positively (e.g. no need to answer the same simple question multiple times a day) there are also potentially negative consequences (e.g., living in fear of losing one's job to a machine). There are more contexts in organizations where CAs can support or substitute human workforce members, such as in sales or invoice processing. While automating human tasks is an ongoing effort, in the context of CA, substituting and automating previously human work has a new component. CAs, especially when designed to be human-like, are perceived as social actors (Nass & Moon, 2000; Reeves & Nass, 1996). Therefore, when CAs take over human tasks, it is not just automation with some form of intelligent IT, it could be another social actor taking over a human's task. Further research should investigate the impact this ongoing change has on human work performance and investigate how the positive effects can be utilized even further, while potentially negative effects can be mitigated. Further, it is important to investigate whether or not the specific nature of the CA, i.e., that it is perceived as a social actor, has an impact on automation.

D2.4*: Investigate the positive and negative impact, as well as potential mitigation strategies when automating human work with (anthropomorphic) CAs.*

5.1.3 CAs with Virtual Interactive and Physical Embodiments

We disclosed that the recently increased research interest in CAs, in particular, comprises the research streams of text-based CAs (research stream 1) and speech-based CAs without embodiment (research stream 3). In contrast, the streams of virtual CAs (research stream 2) and physical CAs (research stream 4) remain on a steady, comparably low level of studies per year (see Figure 5 in Appendix C). We explain this observation by referring to the high availability of text-based CAs and speech-based agents for (experimental) research and the current interest in such types of CAs in practice. Nevertheless, we believe that there are substantial research opportunities to study physical and virtual CAs.

Recent studies on virtual agents highlight the strong social reactions humans show to virtual CAs. For example, Harjunen et al. (2018) found that participants in an experiment were comparatively more likely to accept unfair offers by a CA that smiled and touched them through a haptic glove. Krämer et al. (2018) discovered that socially responsive non-verbal agent behavior in terms of nodding, smiling, and posture shifts can reduce participant's need to engage in social activities after the interaction. Further, Seymour et al. (2018) highlight the versatile potential and the implications of natural face technology that creates a realistic visual presence, calling for "blue ocean" research in this area. Current interactive CAs, such as Amelia by IPsoft (2020), underline this potential of virtual CAs in practical application, particularly in organizational contexts like customer service or IT service desk automation. Similar to emerging research on virtual agents, recent studies on physically embodied CAs, such as Desideri (2018), Stock & Merkle (2018b), or Stock et al. (2019), demonstrate CAs' potential with a physical embodiment. For example, Stock & Merkle (2018b) investigated customer responses to behavioral cues during customer service encounters and found positive emotional reactions to the humanoid robot's behavior. Desideri et al. (2018) investigated whether humanoid robots can offer mental health assessment benefits that improve on clinical psychologists' achievements.

In short, new forms of realistic virtual interactive or physically embodied CAs are likely to have a substantial impact on user perception, complementing the verbal communication currently in the focus of IS research on CAs. We expect these additional and rich design features associated with such forms of embodiment to immediately attract users' attention in the interaction and to strengthen social responses as indicated by the aforementioned, early studies. Thus, we propose systematic study of the impact virtual interactive and physical embodiment have in combination with the advanced natural language interaction present-day CAs offer.

D.3.1: Explore the impact of rich virtual interactive or physically embodied CAs on user perception combined with the currently available conversational capabilities.

In addition to better understanding the impact of these CA embodiment forms on user perception, we propose research on how such embodiment should be designed for different contexts in order to increase users' acceptance, adoption, and performance of CAs. For example, we could ask how a pleasant embodiment of a virtual interactive customer service CA should look, or how a CA's physical embodiment should be designed to efficiently support product sales in stores. Currently available forms of CA embodiment, such as IPsoft's Amelia for virtual interactive CAs or SoftBank's humanoid robot Pepper, offer unprecedented design features that can be used, such as facial movement (e.g., eye blinking or smiling) or gestures (e.g., waving or moving the head toward a speaking person) in combination with advanced conversational capabilities. However, we lack a solid understanding of how these features should be designed and combined with one another in order to influence user perceptions to achieve specific goals (e.g. increase trust in the CA or foster the perception of anthropomorphism) across different contexts. Thus, we propose:

D.3.2: Investigate virtual interactive and physical CA embodiment designs to increase agent acceptance, adoption, and performance across different contexts.

5.1.4 Transferability of Knowledge gained in CA Studies

Besides new insight in the state of the art in CA research from a given ex-ante perspective (i.e., we preselected dimensions for coding and all analyses depended on this selection), different ex-post observations emerged from our analysis. We found several studies transferring and combining knowledge from different contexts (e.g., Seeger et al., (2018), Gnewuch et al. (2017), Tavanapour et al. (2019)) and different types of CAs (e.g., Araujo, (2018), Jeong et al. (2019), Wagner & Schramm-Klein (2019)). Justifying design decisions (Gregor et al., 2020) or proposing hypotheses using existing prescriptive and descriptive knowledge is the appropriate scientific methodology (Bhattacherjee, 2012). However, during our analysis, the question emerged as to when and under which circumstances prescriptive and descriptive knowledge can be adapted from one instance to another, as when a specific type of user interacts with a specific type of agent in a specific context (see Figure 2).

In our sample of 262 studies, 148 articles investigated the interaction between humans and a CA (or specific design variants of the CA) for a given context and measured specific outcomes (the type of users, CA design, and context – referenced as "instance" in the following). We found studies that successfully adapted knowledge from one instance to their research context in another instance. For example, Adam and Klumpe (2019) investigated a text-based CA in the context of human resources that facilitates onboarding processes and successfully drew on Lee and Choi's (2017) findings of reciprocal behavior and self-disclosure by a speech-based agent having a positive effect on user satisfaction in the context of marketing and sales. Their study provides arguments for the transferability of knowledge from one instance to another in the context of CAs.

However, there is also research on CAs that investigated similar aspects, but found mixed results, which makes it difficult to adapt knowledge from one instance to another instance. Nass and Moon (2000) found that users perceived a CA's female voice as less friendly and less competent than the corresponding male voice, thus supporting their assumption that "individuals would mindlessly gender stereotype computers" (Nass & Moon, 2000, p. 85). In contrast, Forlizzi et al. (2007) found that CAs with a female-looking avatars were preferred to male-looking avatars. Moreover, female-looking avatars received higher satisfaction ratings than male avatars. Although the two studies do not assess exactly the same outcomes, they illustrate the problem regarding the transferability of existing prescriptive and descriptive CA knowledge. While Nass and Moon's (2000) study supports selecting a male gendered CA to receive positive outcomes, Forlizzi et al. (2007) provide arguments for selecting a female gendered on to create similar positive outcomes.

Against this background, the practice of transferring implications from observations between instances (e.g., from a physical CA with embodiment to an agent with a virtual embodiment) is questionable and potentially risky. Currently, researchers should only cautiously adopt knowledge from a single instance because there is insufficient literature on the transferability and adaptability of CA knowledge. Considering existing research, we were unable to identify any study empirically investigating how or how well knowledge can be transferred between different CA instances, and what the prerequisites for a successful transfer might be. Similarly, we found no research that addresses this aspect from a conceptual or theoretical point of view.

Therefore, it is still unknown how well, for example, observations regarding the effect of designing the CA as gendered (e.g., a CA having a male or female name or an avatar representing a specific gender) can be transferred between a text-based chatbot (research stream 1) and a physical CA (research stream 4). The two CA types offer a different set of potential design elements to be used in communicating and portraying the CA's gender or, more generally, the agent's degree of humanness. Thus, one can expect differences in the users' perception of the resulting CAs. The same is true for other CA design features (e.g., non-verbal communication of positive emotions through emoticons in a text-based agent, compared to a smiling interactive agent) and for different contexts (e.g., from a CA in education to a CA in customer services). In essence, the question remains regarding the circumstances under which existing prescriptive and descriptive knowledge can be transferred and adapted between CA instances. Our research framework can be helpful in systematically analyzing the commonalities and differences between user groups (i.e., human users and their characteristics), contexts (e.g., customer service or education), CA types (i.e., agent), and resulting perceptions and outcomes across different instances. We assume that some effects are observable across several instances, while others might be specific to a particular group of users, application context, or CA type. In short, we propose that experimental CA studies be partially replicated by deliberately varying one dimension from the original study (human, context, agent) and then investigate the effect this has on user perception of the agent and outcome(s) of the interaction.

D4.1*: Partially replicate experimental CA studies in investigating whether existing prescriptive and descriptive knowledge can be transferred between different CA instances (i.e. combinations of human, agent, and context) and adapted accordingly.*

Eventually, by conducting such studies where one or two dimensions are deliberately changed from the original work, our knowledge of CA design and interaction will mature from a plethora of seemingly related observations into a systematic framework, consisting of generalized statements on overarching phenomena, such as CA gender stereotyping.

5.1.5 Ethical Implications of Designing and Interacting with CAs

We found only a single study investigating ethical aspects related to human interaction with CAs. Banks (2018) proposes and validates a scale to measure the agent's behavior regarding morality and dependency on its implementation. Considering this substantial lack of research on ethical implications of the design of and interaction with CAs, we formulate three directions for future research in the following section.

Clearly, CA can be a tool for persuasion (Lehto & Oinas-Kukkonen, 2017), also because they have been able to influence user's cognition, emotions, and behavior. They offer developers options to achieve other goals besides supporting users in making decisions or completing tasks. For example, researchers, such as Adler et al. (2016), Derrick and Ligon (2014), or

Harjunen et al. (2018) adapt established approaches for persuasion known from human-to-human interaction for designing CAs, such as emotional persuasion strategies (Adler et al., 2016). However, subconsciously manipulating users comes with great ethical implications. Referring to the literature on digital nudging (Lembcke et al., 2019), various aspects have to be considered and weighted against one another before interference with users' free will can be considered as ethically justified: (1) the individual's freedom of choice should be preserved, (2) the intention behind the design should be transparent, and (3) the goal-oriented intention of the interference should be justified.

Similar considerations are presented in other seminal models and frameworks, such as the principles of ethical and persuasive technology design. Berdichevsky and Neuenschwander (2002) proposed or the rules for building trustworthy AI as Floridi (2019) put forward. Nonetheless, to the best of our knowledge, we do not have a dedicated discourse on the ethical design of CAs. Extending from the general considerations on the ethics of persuasion in other disciplines, future research should systematically identify, analyze, and discuss CA design's unique aspects. For instance, the CA's capabilities have drastically improved in recent years (McTear, 2017), leading to the ability to display empathy via sentiment-analysis (Diederich et al., 2019), which falls in an area of unique ethical challenges (e.g., justifying chatbots that dynamically adapt to the emotional state of the user, making it increasingly more difficult for the user to make a free decision). Similarly, we need to investigate other new means CAs have to take on a human-like appearance and interact with users.

D5.1*: Explore when, where, and how applying persuasive CAs is ethically justifiable.*

Against this background, there is also a need for research on design elements that are specifically intended to make the CA "more ethical." For instance, a common feature is for CAs to self-disclose that they are not human (Grudin & Jacques, 2019; O'Leary, 2019) because users can find it difficult to distinguish increasingly human-like CAs from actual humans (Welch, 2018). Similar features could be developed for specific aspects of the communication between humans and CAs. For instance, a CA could let the user know that its social cues (e.g., having a name, avatar, using self-references) are intended to change the user's perception of the CA in a certain way (e.g., letting the CA appear more trustworthy) to achieve a certain goal (e.g., riding a bike more often). Understanding how users perceive such features is an important new area of research. For instance, referencing the previous example of trustworthiness, a CA explaining its persuasive design to the user might not lead to the user becoming more aware of such an intention, but could instead increase the perception of trustworthiness (e.g., "the CA takes care of the ethical dimensions of our interaction, therefore, I can trust it").

D5.2*: Develop CA's design elements to address the ethical dimensions of the user interaction, and investigate how users perceive these design elements.*

Besides the CA’s clear design intention to be persuasive (or in the future, ethically aware), there are also unintended effects that need to be considered. For example, children interacting with voice-based CAs, such as Amazon Alexa, can by design be encouraged to say “please” when issuing voice commands (BBC 2019), a behavior that is then potentially adapted to human-to-human interaction as well. However, frequently formulating commands that are fulfilled instantly can also lead to similar behavior of children in human-to-human interactions (Truong, 2016). Further, a recent UNESCO (2019) report outlined that voice-based CAs can reinforce gender stereotypes as users continuously interact with mostly female CA voices in a commanding tone. Similarly, researchers established gender bias in the design of text-based CAs, finding that developers show a clear preference for implementing CAs with obvious female-gender traits (e.g., having a traditional female name) (Feine et al., 2019). Therefore, expanding current research clearly focused on the expected and intended effects of certain CA designs, we need research on the unintended side-effects of CAs and also on the developers’ biases and assumptions, leading to these side-effects.

D5.3*: Investigate the unintended side-effects of CA design and study how to prevent such negative side-effects.*

5.1.6 Longitudinal CA Research in a Field Study Setting

Most studies in our sample (84%) consist of empirical research, preferably in a controlled setting (57%), both through on-site laboratory and online experiments. In contrast, only a few studies (4%) were carried out in a field study setting, despite the high availability of CAs in different contexts in practice. For example, Adam and Klumpe (2019), conducted a randomized field experiment with 2,095 visitors of an e-commerce website to investigate the impact of message interactivity and an agent’s information self-disclosure in onboarding new customers. In agreement with studies in controlled experimental settings, on information disclosure (e.g. Pickard et al. (2016), Saffarizadeh et al. (2017)) and on message interactivity (e.g. Schuetzler et al. (2014)), scholars find a positive effect of both constructs on actual user behavior regarding information disclosure. Thus, the work by Adam and Klumpe (2019) provides empirical support for insights from different laboratory experiments with field data. Thereby, the study validated existing findings with actual user information sharing behavior in a real-world scenario. Toxtli et al. (2018) provided a further example of initial fieldwork in CA research by implementing and deploying a text-based CA for task management. Based on data from information workers’ interaction with the agent, and from a survey, the authors investigated the human nature of interactions and discuss issues related to response failure or the handling of multi-threaded conversations.

These exemplary studies underline CA research potential in the field. First, field studies can help to overcome shortcomings related to external validity of the rich body of knowledge on CAs gained in controlled experimental settings (Dennis & Valacich, 2001; Karahanna et al., 2018). As shown in Adam and Klumpe's (2019) study, knowledge gained in single or multiple laboratory experiments can be applied in field study settings. By conducting field studies on the design of and human interaction with CAs, existing knowledge can be validated (see Adam and Klumpe (2019)), extended (see Toxtli et al. (2018) who identified new aspects for study), or (partially) refuted. Thus, we formulate the following direction for future research studies:

D6.1*: Conduct field studies to verify, extend, or refute existing knowledge gained from experimental research in controlled settings.*

While assuming that many insights gained from laboratory experiments remain valid in the field, it will be particularly interesting to study how human perception of CAs changes when insights from different kinds of studies are combined. For example, various studies, such as Gnewuch et al. (2018), Go and Sundar (2019), and Araujo (2018), identify positive effects of different social cues on user perception (e.g. response delays to simulate a CA typing, message interactivity, or a human-like avatar), intended to make agents appear human-like. A question now is what happens if these rich social cues, identified in separate experiments, are combined in a single CA in the field. Will it be able to induce a high level of perceived anthropomorphism and, if so, will users perceive this agent as appealing or uncanny? Additionally, as indicated in Toxtli et al.'s (2018) study, we lack a solid understanding of the complex interplay of different design aspects in the field. For example, does existing research help us to know how the limited conversational capabilities of present-day CAs, often manifested in the agent’s inability to provide a purposeful response, practically influence users’ perception of agents with a human-like design. Will they swiftly diminish the identified positive effects of the social cues the various laboratory experiments disclosed as good in avoiding failure? In short, current CA research offers rich knowledge gathered in controlled experimental settings. However, we do not know whether these insights hold true in the field where various aspects, both related to the design of the agent itself and to the application context, influence users’ CA interaction and perception. To advance our overall understanding of CA’s design and human interaction with them, it would be good to move from investigating single specific aspects in controlled settings to combining existing knowledge from laboratory experiments and applying it in the field. As initial field studies such as Toxtli et al. (2018), indicate, several issues that substantially impact user perception are likely to remain in practice. Also, a complex interplay of different design aspects, so far investigated separately in different studies, could influence human-CA interaction in, as yet, unforeseeable ways.

D6.2*: Investigate the interplay between CAs’ limited conversational capabilities and the combination of multiple (anthropomorphic) design features of CAs.*

Further, we propose that researchers conduct CA studies over an extended period of time by means of longitudinal research. Many of the studies in our sample investigate research questions in experimental settings where participants engage in single or a few interactions with an agent. While this approach certainly has benefits for investigating the isolated impact different design alterations have on human perception of an agent, we argue that observing users' perception of CAs over the course of multiple interactions can help us understand the emerging relationships between human users and (human-like) agents. As CAs exhibit various social cues, first and foremost interacting via natural language as opposed to graphical user interfaces, they trigger social responses as shown in many empirical studies in our sample (e.g. Hong & Williams (2019), Lee & Choi (2017), Xu & Lombard (2017)). Users are likely to form relationships with CAs over the course of multiple interactions where such agents can assume a variety of roles ranging from simple digital assistants to companions or even friends. As Bickmore and Picard's (2005) early study shows, CAs can be designed to leverage human relationship building approaches which support establishing a social connection. Such features can induce higher levels of trust and likability. Similarly, adverse experiences, such as misunderstandings or an agent's failure to appropriately complete a user's request, are likely to negatively impact user perception of CAs. This will also influence the relationship between the user and the CA. We need information on how users will react if their befriended agent is suddenly unable to adequately answer a simple question. Will the reaction be similar to a computer's failure or will it trigger a stronger social reaction similar to being disappointed by a good friend?

Further, as the specific capabilities of a CA are at first, to some extent hidden from the user, they can require deeper investigation and learning compared to graphical user interfaces (Følstad & Brandtzæg, 2017). Then, it will be interesting to see how the adoption and use of CAs changes over a longer period of time. Thus, we propose that researchers investigate the emerging relationship between users and CAs over time, and we suggest applying an interaction-centric approach (Al-Natour & Benbasat, 2009) that recognizes the different roles an agent can assume for a specific user and in a specific context.

D6.3*: Study how the relationship between users and CAs takes shape during initial use of the CA, and how it develops in multiple interactions over a longer period of time.*

5.2 A Springboard for CA Studies in the IS Discipline

Our review analyzes and discusses the findings of 262 publications on the CAs' design. However, given the selection criteria of our review approach, we were only able to investigate a subset of the available CA research that has been published in IS and CS as well as other disciplines, such as didactics and pedagogy, medicine, ethics, and psychology.

In the following table, we provide a non-exhaustive list of CA research that we consider as important to serve researchers and practitioners alike, as a starting point for addressing the proposed avenues.

Authors	Title	Discipline
Bickmore & Cassell (2005)	Social dialogue with embodied conversational agents	CS
Cassell (2000)	Embodied conversational interface agents	CS
Gulz & Haake (2006)	Design of animated pedagogical agents - A look at their look	D
Graesser, Hu, & Person (2001)	Teaching with the help of talking heads	D
Graesser, Li, & Forsyth (2014)	Learning by communicating in natural language with conversational agents	P & D
Graesser & McNamara (2010)	Self-regulated learning in learning environments with pedagogical agents that interact in natural language	D
Johnson, Rickel, & Lester (2000)	Animated pedagogical agents: Face-to-face interaction in interactive learning environments	D
Laranjo et al. (2018)	Conversational agents in healthcare: a systematic review	M
Louwerse, Graesser, McNamara, & Lu (2009)	Embodied conversational agents as conversational partners	P
Massaro, Cohen, Daniel, & Cole (1999)	Developing and evaluating conversational agents	P
Montenegro, da Costa, & da Rosa Righi (2019)	Survey of conversational agents in health	M
Moreno (2012)	Multimedia Learning with Animated Pedagogical Agents	D
Provoost, Lau, Ruwaard, & Riper (2017)	Embodied conversational agents in clinical psychology: a scoping review	M & P
Vaidyam, Wisniewski, Halamka, Kashavan, & Torous (2019)	Chatbots and conversational agents in mental health: a review of the psychiatric landscape	M & P
Veletsianos & Russell (2014)	Pedagogical agents	D
Wik & Hjalmarsson (2009)	Embodied conversational agents in computer assisted language learning	D
Luxton (2020)	Ethical implications of conversational agents in global public health	M & E
McGreevey, Hanson, & Koppel (2020)	Clinical, Legal, and Ethical Aspects of Artificial Intelligence–Assisted Conversational Agents in Health Care	M & E
European Commission (2019)	Ethics Guidelines for Trustworthy AI	E
Floridi (2019)	Establishing the rules for building trustworthy AI	E

CS = computer science, D = didactics and pedagogy, M = medicine, P = psychology, E = ethics

Table 12. Springboard to Further CA Research

5.3 Practical Implications

Our overview of CA research's status quo and of avenues to advance the field is primarily targeted at (HCI) scholars in the IS and CS disciplines with an interest in this technological phenomenon. However, this paper also offers three key implications for practitioners.

First, the adapted research framework (Figure 2) can support CA conceptualization and implementation in practice by highlighting aspects that should be considered during the design process. For example, the framework can enable designers to reflect the relevant characteristics of the user group for whom the CA is intended (e.g., concerning age, gender, cultural background, or experience with CAs and the task at hand); or it can reflect the key aspects related to how users perceive the agent and the interaction outcomes (e.g., measures to evaluate the agent).

Second, the coded literature can help practitioners to identify empirically grounded design approaches to influence certain design aspects related to how CAs are perceived. For example, designers can draw on the literature database in identifying conversational strategies that build trust in a CA (e.g., Schroeder & Schroeder, 2018; Stock et al., 2019) or foster enjoyment in the interaction (e.g., Beale & Creed, 2009; Liao et al., 2018). Similarly, the literature database can be useful in identifying undesirable effects on user perception due to an agent's design. Such effects could relate to privacy concerns (Sohn, 2019) or feelings of uncanniness (Seeger et al., 2018; Strait et al., 2015; Tinwell & Sloan, 2014).

Third and finally, our research contributes ideas for field studies through a collaboration between CA researchers and practitioners. Particularly, longitudinal studies could be beneficial in validating findings from controlled experimental research in the field, and helping to identify and overcome CAs' shortcomings in practice. In our opinion, the substantial potential for both research and practice lies in such collaborative studies, because in practice CAs often do not meet (high) user expectations (Luger & Sellen, 2016) and much of the insights gained in research have not been transferred to their application in the field.

5.4 Limitations

Although we conducted our organizing and assessing review and subsequent analysis according to established guidelines, potential limitations should be considered.

First, in our review and analysis, we have a restricted view of the available literature because of the applied review methodology. We followed established guidelines to be as rigorous as possible, given our self-chosen review constraints (i.e., selecting keywords, timeframe, and outlets). We believe our selection of keywords and outlets is representative and sufficient for the scope of this review article. In total, we found 262 publications in various outlets that allowed us to draw a holistic picture of state-of-the-art CA research. However, our selection of the outlets could be open to criticism. We decided to focus on an IS perspective and selected the Basket of Eight, as well as leading conferences as sources for our review. Also, we selected a representative subset of CS outlets to identify research published outside of IS research. Future research could include further non-IS outlets to broaden our rather IS-centered impression of research on the human-CA interaction.

Second, a limitation of all review articles is the ongoing availability of new publications. We conducted the search process in January 2019 and updated it in November 2019. Considering the publications over time (see Figure 4 in Appendix C), we can assume that even more publications have become available while writing this article; in the future such work could enrich our CA research presentation. One way to tackle this problem could be implementing an online database that enables authors to submit their studies and each's classification following our framework. This would allow the community to have a reasonably recent overview of the available research, and to discover further trends and opportunities for future research. Similar online databases exist, for example, to search for variables and items[1] (Larsen & Bong, 2016) or conversational agents' social cues[2] (Feine et al., 2019).

Third, our analysis and discussion depend on our sample. As outlined, we followed a rigorous process in our review and assumed our set of publications to be representative for discussing the status quo of HCI-related CA research from an IS perspective. Moreover, including research published in non-IS outlets provides an even more holistic picture; however, a different sample of publications might result in different findings. Nevertheless, we perceive our identified research streams as insightful, relatively stable, and consistent independent of potential changes in the underlying data by, e.g., new publications appearing.

Fourth, our framework can be extended with more dimensions and more granular characteristics. The framework allowed us to classify the studies sufficiently to provide a holistic picture of CA's research status quo. Even so, future work could, for instance, extend the agent dimension with a more detailed set of characteristics that would provide a more productive overview of research on this specific aspect. To illustrate, Feine et al. (2019) differentiate CAs' design features into four categories (i.e., verbal, visual, auditory, and invisible) with multiple subcategories. Similarly, the research could enrich the human dimensions and the context with more detailed characteristics. The resulting insights of the interaction might enable us to further investigate the transferability and adaptability of knowledge.

[1] https://inn.theorizeit.org/
[2] https://design.chatbotresearch.com/

6 Conclusion

In this article, we have organized a rich body of knowledge on the interaction between humans and CAs in the IS and CS disciplines. We contribute a framework adapted from established research that allows for classifying the vast body of knowledge on CAs. The framework enables the research community to understand the interaction between humans and CAs in specific contexts and can guide future research on CAs. Based on the findings of our literature review, we have assessed the status quo of CA research and now propose six avenues with sixteen actionable directions to move CA research forward. We invite researchers to address the outlined directions for future studies and contribute valuable knowledge to this exciting research area.

Acknowledgements

We would like to thank our Senior Editor Dorothy Leidner and the anonymous reviewers for their support, encouragement and valuable feedback that helped us to substantially improve the paper throughout the review process.

7 References

Abul, M., Siddike, K., Spohrer, J., Demirkan, H., & Kohda, Y. (2018). People's Interactions with Cognitive Assistants for Enhanced Performances. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS)* (pp. 1640–1648). Waikoloa Village, Hawaii, USA.

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 European Conference on Information Systems (ECIS)* (pp. 0–17). Stockholm, Sweden.

Adler, R. F., Iacobelli, F., & Gutstein, Y. (2016). Are you convinced? A Wizard of Oz study to test emotional vs. Rational persuasion strategies in dialogues. *Computers in Human Behavior*, *57*, 75–81.

Al-Natour, S., & Benbasat, I. (2009). The Adoption and Use of IT Artifacts: A New Interaction-Centric Model for the Study of User-Artifact Relationships. *Journal of the Association for Information Systems*, *10*(9), 661–685.

Al-Natour, S., Benbasat, I., & Cenfetelli, R. T. (2006). The Role of Design Characteristics in Shaping Perceptions of Similarity: The Case of Online Shopping Assistants. *Journal of the Association for Information Systems*, *7*(12), 821–861.

Al-Natour, S., Benbasat, I., & Cenfetelli, R. T. (2009). The Antecedents of Customer Self-Disclosure to Online Virtual Advisors. In *Proceedings of the International Conference on*

Information Systems (ICIS) (pp. 1–18). Phoenix, USA.

Anabuki, M., Kakuta, H., Yamamoto, H., & Tamura, H. (2000). Welbo: An Embodied Conversational Agent Living in Mixed Reality Spaces. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 10–11). The Hague, Netherlands.

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

Ashktorab, Z., Jain, M., Liao, V. Q., & Weisz, J. D. (2019). Resilient chatbots: Repair strategy preferences for conversational breakdowns. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–12). Glasgow, Scotland.

Baier, D., Rese, A., & Röglinger, M. (2018). Conversational User Interfaces for Online Shops? A Categorization of Use Cases. In *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–17). San Francisco, USA.

Bandara, W., Gorbacheva, E., Beekhuyzen, J., Furtmueller, E., & Miskon, S. (2015). Achieving Rigor in Literature Reviews: Insights from Qualitative Data Analysis and Tool-Support. *Communications of the Association for Information Systems*, *34*(8), 154–204.

Banker, R. D., & Kauffman, R. J. (2004). 50th Anniversary Article: The Evolution of Research on Information Systems: A Fiftieth-Year Survey of the Literature in Management Science. *Management Science*, *50*(3), 281–298.

Banks, J. (2018). Perceived Moral Agency Scale: Development and Validation of a Metric for Humans and Social Machines. *Computers in Human Behavior*.

Bariff, M. L., & Ginzberg, M. J. (1982). MIS and the behavioral sciences: Research patterns and prescriptions. *ACM SIGMIS Database*.

Beale, R., & Creed, C. (2009). Affective interaction: How emotional agents affect users. *International Journal of Human Computer Studies*, *67*(9), 755–776.

Belbin, R. M. (2010). *Management Teams: Why they succeed or fail*. Routledge.

Belk, R. W. (1988). Possessions and the Extended Self. *Journal of Consumer Research*. https://doi.org/10.1086/209154

Belk, R. W. (2013). Extended Self in a Digital World. *Journal of Consumer Research*. https://doi.org/10.1086/671052

Bell, S., Sarkar, A., & Wood, C. (2019). Perceptions of chatbots in therapy. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Glasgow, Scotland.

Ben Mimoun, M. S., Poncin, I., & Garnier, M. (2012). Case study-Embodied virtual agents: An analysis on reasons for failure. *Journal of Retailing and Consumer Services*, *19*(6), 605–612.

Benlian, A., Klumpe, J., & Hinz, O. (2019). Mitigating the Intrusive Effects of Smart Home Assistants by using Anthropomorphic Design Features: A Multi-Method Investigation. *Information Systems Journal*, (forthcoming).

Berdichevsky, D., & Neuenschwander, E. (2002). Toward an Ethics of persuasive technology. *Communications of the ACM*, *42*(5), 51–58.

Berg, M. (2015). NADIA: A Simplified Approach Towards the Development of Natural Dialogue Systems. In *Natural Language Processing and Information Systems* (pp. 144–150). Springer International Publishing.

Bertacchini, F., Bilotta, E., & Pantano, P. (2017). Shopping with a robotic companion. *Computers in Human Behavior*, *77*, 382–395.

Bhattacherjee, A. (2012). *Social Science Research: Principles, Methods, and Practices*. *Book 3.* Tampa, Florida, USA.

Biancardi, B., Cafaro, A., & Pelachaud, C. (2017). Could a virtual agent be warm and competent? investigating user's impressions of agent's non-verbal behaviours. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 22–24). Denver, USA.

Bickmore, T., & Cassell, J. (2005). Social Dialongue with Embodied Conversational Agents. In *Advances in natural multimodal dialogue systems* (pp. 23–54).

Bickmore, T. W., & Picard, R. W. (2005). Establishing and Maintaining Long-term Human-computer Relationships. *ACM Transactions on Computer-Human Interaction*, *12*(2), 293–327.

Bittner, E. A. C., Oeste-Reiß, S., & Leimeister, J. M. (2019). Where is the Bot in our Team? Toward a Taxonomy of Design Option Combinations for Conversational Agents in Collaborative Work. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS)*. Grand Wailea, Maui.

Bittner, E., & Shoury, O. (2019). Designing Automated Facilitation for Design Thinking: A Chatbot for Supporting Teams in the Empathy Map Method. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS)* (pp. 227–236). Grand Wailea, Maui.

Bogardus, E. S. (1947). Measurement of Personal-Group Relations. *Sociometry*.

Braun, M., & Alt, F. (2019). Affective assistants: A matter of states and traits. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Glasgow, Scotland.

Brynjolfsson, E., & McAfee, A. (2016). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. New York, USA: Norton & Company.

Burgoon, J. K., Bonito, J. A., Lowry, P. B., Humpherys, S. L., Moody, G. D., Gaskin, J. E., & Giboney, J. S. (2016). Application of Expectancy Violations Theory to communication with and judgments about embodied agents during a decision-making task. *International Journal of Human Computer Studies*, *91*, 24–36.

Burgoon, Judee K., Stern, L. A., & Dillmann, L. (1995). *Interpersonal adaptation: Dyadic interaction patterns*. Cambridge University Press.

Burmester, M., Schippert, K., Zeiner, K. M., & Platz, A. (2019). Creating positive experiences with digital companions. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Glasgow, Scotland.

Burton, N., & Gaskin, J. (2019). “Thank You , Siri ”: Politeness and Intelligent Digital Assistants. In *Proceedings of the Americas Conference on Information Systems (AMCIS)* (pp. 1–10). Cancun, Mexico.

Byrne, D. (1971). *The Attraction Paradigm*. Michigan, USA: Academic Press.

Byrne, D., & Griffitt, W. (1969). Similarity and awareness of similarity of personality characteristics as determinants of attraction. *Journal Of Experimental Research In Personality*.

Byrne, D., Griffitt, W., & Stefaniak, D. (1967). Attraction and similarity of personality characteristics. *Journal of Personality and Social Psychology*.

Cafaro, A., Vilhjalmsson, H. H., & Bickmore, T. (2016). First impressions in human-agent virtual encounters. *ACM Transactions on Computer-Human Interaction*, *24*(4), 1–40.

Candello, H., Pinhanez, C., Pichiliani, M., Cavalin, P., Figueiredo, F., Vasconcelos, M., & Carmo, H. Do. (2019). The effect of audiences on the user experience with conversational interfaces in physical spaces. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–13). Glasgow, Scotland.

Card, S. K., Moran, T. P., & Newell, A. (1980). The Keystroke-Level Model for User Performance Time with Interactive Systems. *Communications of the ACM*.

Cardona, D. R., Schönborn, S., Werth, O., & Breitner, M. H. (2019). A Mixed Methods Analysis of the Adoption and Diffusion of Chatbot Technology in the German Insurance Sector. In *Proceedings of the Americas Conference on Information Systems (AMCIS)* (pp. 1–10). Cancun, Mexico.

Carlotto, T., & Jaques, P. A. (2016). The effects of animated pedagogical agents in an English-as-a-foreign-language learning environment. *International Journal of Human Computer Studies*, *95*, 15–26.

Carroll, J. M. (2020). Human Computer Interaction - brief intro. In *The Encyclopedia of Human-Computer Interaction* (2nd editio).

Carter, M., Grover, V., & Thatcher, J. B. (2013). Mobile Devices and the Self: Developing the Concept of Mobile Phone Identity. In *Strategy, Adoption, and Competitive Advantage of Mobile Services in the Global Economy*. https://doi.org/10.4018/978-1-4666-1939-5.ch008

Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjálmsson, H., & Yan, H. (1999). Embodiment in conversational interfaces. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 520–527). Pittsburgh, USA.

Cassell, Justine. (2000). Embodied conversational interface agents. *Communications of the ACM*, *43*(4), 70–78.

Chattaraman, V., Kwon, W.-S., Gilbert, J. E., & Ross, K. (2018). Should AI-Based, Conversational Digital Assistants Employ Social- or Task-Oriented Interaction Style? A Task-Competency and Reciprocity Perspective for Older Adults. *Computers in Human Behavior*, *90*, 315–330.

Chaves, A. P., & Gerosa, M. A. (2018). Single or multiple conversational agents? An interactional coherence comparison. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–13). Montréal, Canada.

Cho, E. (2019). Hey Google, can i ask you something in private? The effects of modality and device in sensitive health information acquisition from voice assistants. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–9). Glasgow, Scotland.

Clark, L., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., … Cowan, B. R. (2019). What makes a good conversation? Challenges in designing truly conversational agents. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–12). Glasgow, Scotland.

Comendador, B. E. V., Francisco, B. M. B., Medenilla, J. S., Nacion, S. M. T., & Serac, T. B. E. (2015). Pharmabot: A Pediatric Generic Medicine Consultant Chatbot. *Journal of Automation and Control Engineering*, *3*(2), 137–140.

Constantin, A., Lai, C., Farrow, E., Alex, B., Jeuring, J., Pel-Littel, R., & Nap, H. H. (2019). “Why is the doctor a man?” Reactions of older adults to a virtual training doctor. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Glasgow, Scotland.

Cooper, H. M. (1988). Organizing knowledge synthesis: a taxonomy of literature reviews. *Knowledge in Society*, *1*(1), 104–126.

Corti, K., & Gillespie, A. (2016). Co-constructing intersubjectivity with artificial conversational agents: People are more likely to initiate repairs of misunderstandings with agents represented as human. *Computers in Human Behavior*, *58*, 431–442.

Cowan, B. R., Branigan, H. P., Obregón, M., Bugis, E., & Beale, R. (2015). Voice anthropomorphism, interlocutor modelling and alignment effects on syntactic choices in human-computer dialogue. *International Journal of Human Computer Studies*, *83*, 27–42.

Cowell, A. J., & Stanney, K. M. (2005). Manipulation of non-verbal interaction style and demographic embodiment to increase anthropomorphic computer character credibility.

International Journal of Human Computer Studies, *62*(2), 281–306.

Crockett, K., Latham, A., & Whitton, N. (2017). On predicting learning styles in conversational intelligent tutoring systems using fuzzy decision trees. *International Journal of Human Computer Studies*, *97*, 98–115.

Dale, R. (2016). The return of the chatbots. *Natural Language Engineering*, *22*(5), 811–817.

Davenport, T. H., & Kirby, J. (2016). Just How Smart Are Smart Machines? *MIT Sloan Management Review*, *57*(3), 21–25.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *Management Information Systems Quarterly*, *13*(3), 319–340.

De Rosis, F., Pelachaud, C., Poggi, I., Carofiglio, V., & De Carolis, B. (2003). From Greta's mind to her face: Modelling the dynamics of affective states in a conversational embodied agent. *International Journal of Human Computer Studies*, *59*(1–2), 81–118.

Dennis, A. R., & Valacich, J. S. (2001). Conducting Experimental Research in Information Systems. *Communications of the Association for Information Systems*, *7*(5), 1–41.

Derrick, D. C., & Ligon, G. S. (2014). The affective outcomes of using influence tactics in embodied conversational agents. *Computers in Human Behavior*, *33*, 39–48.

Desideri, L., Ottaviani, C., Malavasi, M., di Marzio, R., & Bonifacci, P. (2018). Emotional processes in human-robot interaction during brief cognitive testing. *Computers in Human Behavior*, (August), 0–1.

Diederich, S., Brendel, A. B., & Kolbe, L. M. (2020). Designing Anthropomorphic Enterprise Conversational Agents. *Business & Information Systems Engineering*, (62), 193–209.

Diederich, S., Brendel, A. B., Lichtenberg, S., & Kolbe, L. M. (2019). Design for Fast Request Fulfillment or Natural Interaction? Insights from an Online Experiment with a Conversational Agent. In *Proceedings of the European Conference on Information Systems (ECIS)* (pp. 1–17). Stockholm, Sweden.

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 International Conference on Information Systems (ICIS)* (pp. 1–17). Munich, Germany.

Diederich, S., Lembcke, T.-B., Brendel, A. B., & Kolbe, L. M. (2021). Understanding the Impact that Response Failure has on How Users Perceive Anthropomorphic Conversational Service Agents: Insights from an Online Experiment. *AIS Transactions on Human-Computer Interaction*, *13*(1), 82–103.

Diederich, S., Lichtenberg, S., Brendel, A. B., & Trang, S. (2019). Promoting Sustainable Mobility Beliefs with Persuasive and Anthropomorphic Design: Insights from an Experiment with a Conversational Agent. In *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–17). Munich, Germany.

Dolata, M., Kilic, M., & Schwabe, G. (2019). When a computer speaks institutional talk: Exploring challenges and potentials of virtual assistants in face-to-face advisory services. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS)* (pp. 105–114). Grand Wailea, Maui.

Duan, W., Yamashita, N., Hwang, S. Y., & Fussell, S. R. (2018). " Let Me Ask Them to Clarify If You Don't Want To "— A Clarification Agent for Nonnative Speakers. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Montréal, Canada.

Ducheneaut, N., Don Wen, M. H., Yee, N., & Wadley, G. (2009). Body and mind: A study of avatar personalization in three virtual worlds. In *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/1518701.1518877

Elson, J. S., Derrick, D. C., & Ligon, G. S. (2018). Examining Trust and Reliance in Collaborations between Humans and Automated Agents. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS)* (pp. 430–439). Waikoloa Village, Hawaii, USA.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On Seeing Human: A Three-Factor Theory of Anthropomorphism. *Psychological Review*, *114*(4), 864–886.

Ethics Guidelines for Trustworthy AI. (2019).

Fadhil, A., & Villafiorita, A. (2017). An Adaptive Learning with Gamification & Conversational UIs. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 408–412). Denver, USA.

Fast, E., Chen, B., Mendelsohn, J., Bassen, J., & Bernstein, M. (2017). Iris: A Conversational Agent for Complex Tasks. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–12). Denver, USA.

Feine, Jasper;, Gnewuch, U., Morana, S., & Maedche, A. (2019). Gender Bias in Chatbot Design. In *Proceedings of CONVERSATIONS 2019* (pp. 79–93).

Feine, Jasper, Gnewuch, U., Morana, S., & Maedche, A. (2019). A Taxonomy of Social Cues for Conversational Agents. *International Journal of Human-Computer Studies*, *132*(December), 138–161.

Feine, Jasper, Morana, S., & Gnewuch, U. (2019). Measung Service Encounter Satisfaction with Customer Service Chatbots using Sentiment Analysis. In *Proceedings of the International Conference on Wirtschaftsinformatik* (pp. 0–11).

Floridi, L. (2019). Establishing the rules for building trustworthy AI. *Nature Machine Intelligence*, *1*(6), 261–262. https://doi.org/10.1038/s42256-019-0055-y

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

Forlizzi, J., Zimmerman, J., Mancuso, V., & Kwak, S. (2007). How interface agents affect interaction between humans and computers. *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces, DPPI'07*, (August), 209–221.

Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior*, *75*, 461–468.

Gambino, A., Shyam Sundar, S., & Kim, J. (2019). Digital doctors and robot receptionists: User attributes that predict acceptance of automation in healthcare facilities. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Glasgow, Scotland.

Gefen, D., & Straub, D. (2003). Managing User Trust in B2C e-Services. *E-Service Journal*, *2*(2), 7–24.

Gerlach, J. H., & Kuo, F. Y. (1991). Understanding human-computer interaction for information systems design. *Management Information Systems Quarterly*.

Giles, H., Coupland, N., & Coupland, J. (2010). Accommodation theory: Communication, context, and consequence. In *Contexts of Accommodation*. https://doi.org/10.1017/cbo9780511663673.001

Gnewuch, U., Morana, S., Adam, M. T. P., & Maedche, A. (2018). Faster Is Not Always Better: Understanding the Effect of Dynamic Response Delays in Human-Chatbot Interaction. In *Proceedings of the European Conference on Information Systems (ECIS)* (pp. 1–17). Portsmouth, United Kingdom.

Gnewuch, U., Morana, S., & Maedche, A. (2017). Towards Designing Cooperative and Social Conversational Agents for Customer Service. In *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–13). Seoul, Korea.

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

Goasduff, L. (2019). Chatbots Will Appeal to Modern Workers. Retrieved June 19, 2020, from https://www.gartner.com/smarterwithgartner/chatbots-will-appeal-to-modern-workers/

Gong, L. (2008). How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior*, *24*(4), 1494–1509.

Goodhue, D. L., & Thompson, R. L. (1995). Task-Technology Fit and Individual Performance. *Management Information Systems Quarterly*, *19*(2), 213–236.

Graesser, A. C., Cai, Z., Morgan, B., & Wang, L. (2017). Assessment with computer agents that engage in conversational dialogues and trialogues with learners. *Computers in Human Behavior*, *76*, 607–616.

Graesser, A. C., Hu, X., & Person, N. (2001). Teaching with the help of talking heads. In

Proceedings - IEEE International Conference on Advanced Learning Technologies, ICALT 2001.

Graesser, A. C., Li, H., & Forsyth, C. (2014). Learning by Communicating in Natural Language With Conversational Agents. *Current Directions in Psychological Science*, *23*(5), 374–380.

Graesser, A., & McNamara, D. (2010). Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist*, *45*(4), 234–244. https://doi.org/10.1080/00461520.2010.515933

Gregor, S., Kruse, L. C., & Seidel, S. (2020). The Anatomy of a Design Principle. *Journal of the Association for Information Systems*, 1–49.

Grice, H. P. (1975). Logic and Conversation. *Syntax and Semantics, Vol 3*.

Groom, V., Takayama, L., Ochi, P., & Nass, C. (2008). I am my robot: The impact of robot-building and robot form on operators. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction, HRI'09*. https://doi.org/10.1145/1514095.1514104

Grudin, J., & Jacques, R. (2019). Chatbots, humbots, and the quest for artificial general intelligence. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–11). Glasgow, Scotland.

Gulz, A., & Haake, M. (2006). Design of animated pedagogical agents - A look at their look. *International Journal of Human Computer Studies*, *64*(4), 322–339.

Hanus, M. D., & Fox, J. (2015). Persuasive avatars: The effects of customizing a virtual salespersons appearance on brand liking and purchase intentions. *International Journal of Human Computer Studies*, *84*, 33–40.

Harjunen, V. J., Spapé, M., Ahmed, I., Jacucci, G., & Ravaja, N. (2018). Persuaded by the machine: The effect of virtual nonverbal cues and individual differences on compliance in economic bargaining. *Computers in Human Behavior*, *87*, 384–394.

Hayashi, Y. (2013). Pedagogical conversational agents for supporting collaborative learning. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (p. 655). Paris, France.

Heinrich, L. J., Heinzl, A., & Riedl, R. (2011). *Wirtschaftsinformatik: Einführung und Grundlegung*. *Springer*.

Herborn, K., Stadler, M., Mustafić, M., & Greiff, S. (2018). The Assessment of Collaborative Problem Solving in PISA 2015: Can Computer Agents Replace Humans? *Computers in Human Behavior*.

Hill, J., Randolph Ford, W., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations. *Computers in Human Behavior*, *49*, 245–250.

Hobert, S. (2019). Say Hello to ‘ Coding Tutor ’! Design and Evaluation of a Chatbot-based Learning System Supporting Students to Learn to Program. In *Proceedings of the International Conference on Information Systems* (pp. 1–17). Munich, Germany.

Hobert, S., & Wolff, R. M. Von. (2019). Say Hello to Your New Automated Tutor – A Structured Literature Review on Pedagogical Conversational Agents. In *Proceedings of International Conference on Wirtchaftsinformatik* (pp. 301–314). Siegen, Germany.

Hong, J. W., & Williams, D. (2019). Racism, responsibility and autonomy in HCI: Testing perceptions of an AI agent. *Computers in Human Behavior*, *100*, 79–84. https://doi.org/10.1016/j.chb.2019.06.012

Hsu, P. (Pei-T., Zhao, J., Liao, K., Liu, T., & Wang, C. (2017). AllergyBot. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 74–79). Denver, USA.

Hu, T., Xu, A., Liu, Z., You, Q., Guo, Y., Sinha, V., … Akkiraju, R. (2018). Touch Your Heart: A Tone-aware Chatbot for Customer Care on Social Media. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–12). Montréal, Canada.

Huang, C.-M. (2012). Designing effective behaviors for educational embodied agents. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (p. 931). Austin, Texas.

Huang, M. H., & Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, *21*(2), 155–172.

Hubal, R. C., Fishbein, D. H., Sheppard, M. S., Paschall, M. J., Eldreth, D. L., & Hyde, C. T. (2008). How do varied populations interact with embodied conversational agents? Findings from inner-city adolescents and prisoners. *Computers in Human Behavior*, *24*(3), 1104–1138.

Hwang, G., Oh, C. Y., Lee, J., & Lee, J. (2019). It sounds like a woman: Exploring gender stereotypes in South Korean voice assistants. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Glasgow, Scotland.

Hyde, J., Carter, E. J., Kiesler, S., & Hodgins, J. K. (2015). Using an Interactive Avatar’s Facial Expressiveness to Increase Persuasiveness and Socialness. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1719–1728). Seoul, Korea.

Interaction, A. S. I. G. on C.-H. (2020). ACM CHI Conference on Human Factors in Computing Systems. Retrieved from https://sigchi.org/conferences/conference-history/CHI/

Ipsoft. (2020). Amelia in Action: A Selection of Stories from Organizations adopting IPsoft’c cognitive agent. Retrieved from http://www.ipsoft.com/amelia/

Jain, M., Kota, R., Kumar, P., & Patel, S. (2018). Convey: Exploring the Use of a Context View for Chatbots. In *Proceedings of the ACM CHI Conference on Human Factors in*

Computing Systems (pp. 1–6). Montréal, Canada. Retrieved from http://dl.acm.org/citation.cfm?doid=3173574.3174042%0Ahttps://homes.cs.washington.edu/~mohitj/convey-chi2018.pdf

Jeong, Y., Kang, Y., & Lee, J. (2019). Exploring effects of conversational fillers on user perception of conversational agents. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Glasgow, Scotland.

Jin, S. A. A. (2010). The effects of incorporating a virtual agent in a computer-aided test designed for stress management education: The mediating role of enjoyment. *Computers in Human Behavior*, *26*(3), 443–451.

Johnson, K. (2018). Facebook Messenger hits 100,000 bots. Retrieved June 5, 2018, from https://venturebeat.com/2017/04/18/facebook-messenger-hits-100000-bots/

Johnson, W., Rickel, J., & Lester, J. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education*, *11*(1), 47–78.

Jucks, R., Linnemann, G. A., & Brummernhenrich, B. (2018). Student Evaluations of a (Rude) Spoken Dialogue System Insights from an Experimental Study. *Advances in Human-Computer Interaction*, *2018*, 1–10.

Jung, H., Hwang, G., Lee, J., Oh, C., Oh, C. Y., & Suh, B. (2019). Tell me more: Understanding user interaction of smart speaker news powered by conversational search. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Glasgow, Scotland.

Kanaoka, T., & Mutlu, B. (2015). Designing a Motivational Agent for Behavior Change in Physical Activity. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1445–1450). Seoul, Korea.

Karahanna, E., Benbasat, I., Bapna, R., & Rai, A. (2018). Opportunities and Challenges for Different Types of Online Experiments. *Management Information Systems Quarterly*, *42*(4), 3–11.

Kim, D., Park, K., & Park, Y. (2018). Alexa, Tell Me More: The Effect of Advertisements on Memory Accuracy from Smart Speakers. In *Proceedings of the Pacific Asia Conference on Information Systems (PACIS)*. Yokohama, Japan.

Kim, K. J., Park, E., & Shyam Sundar, S. (2013). Caregiving role in human-robot interaction: A study of the mediating effects of perceived benefit and social presence. *Computers in Human Behavior*, *29*(4), 1799–1806.

Kim, Y., Kwak, S. S., & Kim, M. S. (2013). Am I acceptable to you? Effect of a robot's verbal language forms on people's social distance from robots. *Computers in Human Behavior*, *29*(3), 1091–1101.

Kim, Y., & Mutlu, B. (2014). How social distance shapes human-robot interaction. *International*

Journal of Human Computer Studies, *72*(12), 783–795.

Knijnenburg, B. P., & Willemsen, M. C. (2016). Inferring Capabilities of Intelligent Agents from Their External Traits. *ACM Transactions on Interactive Intelligent Systems*, *6*(4), 1–25.

Kocaballi, A. B., Berkovsky, S., Quiroz, J. C., Laranjo, L., Tong, H. L., Rezazadegan, D., … Coiera, E. (2019). The personalization of conversational agents in health care: Systematic review. *Journal of Medical Internet Research*.

Kowalski, J., Skorupska, K., Kopeć, W., Jaskulska, A., Abramczuk, K., Biele, C., & Marasek, K. (2019). Older adults and voice interaction: A pilot study with google home. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Glasgow, Scotland.

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the Effectiveness of Work Groups and Teams: A Reflection. *Perspectives on Psychological Science*, *13*(2), 205–212.

Krämer, N. C., Lucas, G., Schmitt, L., & Gratch, J. (2018). Social snacking with a virtual agent – On the interrelation of need to belong and effects of social responsiveness when interacting with artificial entities. *International Journal of Human Computer Studies*, *109*, 112–121.

Krämer, N., Kopp, S., Becker-Asano, C., & Sommer, N. (2013). Smile and the world will smile with you - The effects of a virtual agent's smile on users' evaluation and behavior. *International Journal of Human Computer Studies*, *71*(3), 335–349.

Lahoual, D., & Fréjus, M. (2019). When users assist the voice assistants: From supervision to failure resolution. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–8). Glasgow, Scotland.

Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., … Coiera, E. (2018). Conversational Agents in Healthcare: A Systematic Review. *Journal of the American Medical Informatics Association*, *25*(9), 1248–1258.

Larsen, K. R., & Bong, C. H. (2016). A tool for addressing construct identity in literature reviews and meta-analyses. *Management Information Systems Quarterly*, *40*(3), 529–551.

Laumer, S., Maier, C., & Gubler, F. T. (2019). Chatbot Acceptance in Healthcare: Explaining User Adoption of Conversational Agents for Disease Diagnosis. In *Proceedings of the European Conference on Information Systems (ECIS)* (pp. 0–18). Stockholm, Sweden. Retrieved from https://aisel.aisnet.org/ecis2019_rp/88

Laumer, S., Racheva, A., Gubler, F., & Maier, C. (2019). Use Cases for Conversational Agents : An Interview-based Study. In *Proceedings of the Americas Conference on Information Systems (AMCIS)* (pp. 1–10). Cancun, Mexico.

Le Bigot, L., Jamet, E., Rouet, J. F., & Amiel, V. (2006). Mode and modal transfer effects on performance and discourse organization with an information retrieval dialogue system in natural language. *Computers in Human Behavior*, *22*(3), 467–500.

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 European Conference on Information Systems (ECIS)* (pp. 0–17). Stockholm, Sweden. Retrieved from https://aisel.aisnet.org/ecis2019_rp/119

Lee, C., Lesh, N., Sidner, C. L., Morency, L.-P., Kapoor, A., & Darrell, T. (2004). Nodding in conversations with a robot. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (p. 785). Vienna, Austria.

Lee, K. M., Jung, Y., Kim, J., & Kim, S. R. (2006). Are physically embodied social agents better than disembodied social agents? The effects of physical embodiment, tactile interaction, and people's loneliness in human-robot interaction. *International Journal of Human Computer Studies*, *64*(10), 962–973.

Lee, M. K., Kiesler, S., Forlizzi, J., & Rybski, P. (2012). Ripple effects of an embedded social agent. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (p. 695). Austin, Texas.

Lee, S. Y., & Choi, J. (2017). Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. *International Journal of Human Computer Studies*, *103*, 95–105.

Lehto, T., & Oinas-Kukkonen, H. (2017). Examining the Persuasive Potential of Web-based Health Behavior Change Support Systems. *AIS Transactions on Human-Computer Interaction*, *7*(3), 126–140.

Leidner, D. (2018). Review and Theory Symbiosis: An Introspective Retrospective. *Journal of the Association for Information Systems*, *19*(06), 552–567.

Leite, I., Pereira, A., Mascarenhas, S., Martinho, C., Prada, R., & Paiva, A. (2013). The influence of empathy in human-robot relations. *International Journal of Human Computer Studies*, *71*(3), 250–260.

Lembcke, T. B., Engelbrecht, N., Brendel, A. B., & Kolbe, L. M. (2019). To nudge or not to nudge: Ethical considerations of digital nudging based on its behavioral economics roots. In *Proceedings of the European Conference on Information Systems (ECIS)* (pp. 0–17). Stockholm, Sweden.

Li, D., Browne, G. J., & Chau, P. Y. K. (2006). An empirical investigation of web site use using a commitment-based model. *Decision Sciences*, *37*(3), 427–444.

Li, J. (2015). The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human Computer Studies*, *77*, 23–27.

Liao, Q. V., Hussain, M. M., Chandar, P., Davis, M., Crasso, M., Wang, D., … Geyer, W. (2018). All Work and no Play? Conversations with a Question-and-Answer Chatbot in the Wild. In *Proceedings of the ACM CHI Conference on Human Factors in Computing*

Systems (pp. 1–13). Montréal, Canada.

Looije, R., Neerincx, M. A., & Cnossen, F. (2010). Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors. *International Journal of Human Computer Studies*, *68*(6), 386–397.

Louwerse, M. M., Graesser, A. C., McNamara, D. S., & Lu, S. (2009). Embodied conversational agents as conversational partners. *Applied Cognitive Psychology*, *23*(9), 1244–1255.

Luger, E., & Sellen, A. (2016). “Like Having a Really Bad PA”: The Gulf between User Expectation and Experience of Conversational Agents. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 5286–5297). San José, USA.

Luxton, D. D. (2020). Ethical implications of conversational agents in global public health.

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., … Söllner, M. (2019). AI-Based Digital Assistants. *Business & Information Systems Engineering*, (4), 1–28.

Malone, T. W. (2018). How human-computer “superminds” are redefining the future of work. *MIT Sloan Management Review*.

Massaro, D. W., Cohen, M. M., Daniel, S., & Cole, R. A. (1999). Developing and Evaluating conversational Agents. In *Human Performance and Ergonomics* (pp. 173–194). Academic Press. https://doi.org/10.1016/b978-012322735-5/50008-7

Matsushita, M., Maeda, E., & Kato, T. (2004). An interactive visualization method of numerical data based on natural language requirements. *International Journal of Human Computer Studies*, *60*(4), 469–488.

Mavridis, P., Huang, O., Qiu, S., Gadiraju, U., & Bozzon, A. (2019). Chatterbox: Conversational interfaces for microtask crowdsourcing. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 243–251). Glasgow, Scotland.

McAfee, A., & Brynjolfsson, E. (2017). *Machine, Platform, Crowd: Harnessing Our Digital Future*. Norton & Company.

McGreevey, J. D., Hanson, W., & Koppel, R. (2020). Clinical, Legal, and Ethical Aspects of Artificial Intelligence–Assisted Conversational Agents in Health Care. *Journal of the American Medical Association*, *324*(6).

McQuiggan, S. W., & Lester, J. C. (2007). Modeling and evaluating empathy in embodied companion agents. *International Journal of Human Computer Studies*, *65*(4), 348–360.

McTear, M. (2017). The rise of the conversational interface: A new kid on the block? *Lecture Notes in Computer Science*.

McTear, M., Callejas, Z., & Griol, D. (2016). *The Conversational Interface: Talking to Smart Devices*. Basel, Switzerland: Springer Publishing Company.

Meier, P., Beinke, J. H., Fitte, C., Behne, A., & Teuteberg, F. (2019). FeelFit – Design and Evaluation of a Conversational Agent to Enhance Health Awareness. In *Proceedings of the International Conference on Information Systems* (pp. 1–17). Munich, Germany.

Meyer von Wolff, R., Hobert, S., Masuch, K., & Schumann, M. (2019). What Do You Need Today ? – An Empirical Systematization of Application Areas for Chatbots at Digital Workplaces. In *Proceedings of the Americas Conference on Information Systems (AMCIS)* (pp. 1–10). Cancun, Mexico.

Miner, A., Chow, A., Adler, S., Zaitsev, I., Tero, P., Darcy, A., & Paepcke, A. (2016). Conversational Agents and Mental Health: Theory-Informed Assessment of Language and Affect. In *Proceedings of the International Conference on Human Agent Interaction*. Singapore.

Miner, A. S., Milstein, A., Schueller, S., Hegde, R., Mangurian, C., & Linos, E. (2016). Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. *JAMA Internal Medicine*, *176*(5), 619–625.

Montenegro, J. L. Z., da Costa, C. A., & da Rosa Righi, R. (2019). Survey of conversational agents in health. *Expert Systems with Applications*, *129*, 56–67. https://doi.org/10.1016/j.eswa.2019.03.054

Moon, Y. (2000). Intimate Exchanges: Using Computers to Elicit Self-Disclosure From Consumers. *Journal of Consumer Research*, *26*(4), 323–339.

Morana, S., Gnewuch, U., Jung, D., & Granig, C. (2020). The Effect of Anthropomorphism on Investment Decision-Making with Robo-Advisor Chatbots. In *Proceedings of the European Conference on Information Systems (ECIS)* (pp. 0–18). Marrakech, Morocco.

Moreno, R. (2012). Multimedia Learning with Animated Pedagogical Agents. In *The Cambridge Handbook of Multimedia Learning* (pp. 507–523). Cambridge University Press. https://doi.org/10.1017/cbo9780511816819.032

Mori, M. (1970). The Uncanny Valley. *Energy*.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The Uncanny Valley. *IEEE Robotics and Automation Magazine*, *19*(2), 98–100.

Morton, H., Gunson, N., & Jack, M. (2012). Interactive language learning through speech-enabled virtual scenarios. *Advances in Human-Computer Interaction*.

Mou, Y., & Xu, K. (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior*, *72*, 432–440.

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

NextIT. (2018). Helping a railroad service conduct business. Retrieved November 1, 2018, from http://nextit.com/case-studies/amtrak

Nguyen, Q. N., & Sidorova, A. (2017). AI capabilities and user experiences: a comparative study of user reviews for assistant and non-assistant mobile apps. In *Proceedings of the Americas Conference on Information Systems (AMCIS)* (pp. 1–10). Boston, USA.

Niewiadomski, R., & Pelachaud, C. (2010). Affect expression in ECAs: Application to politeness displays. *International Journal of Human Computer Studies*, *68*(11), 851–871.

Nunamaker, J. F., Derrick, D. C., Elkins, A. C., Burgoon, J. K., & Patton, M. W. (2011). Embodied Conversational Agent-Based Kiosk for Automated Interviewing. *Journal of Management Information Systems*, *28*(1), 17–48.

O'Leary, D. E. (2019). GOOGLE'S Duplex: Pretending to be human. *Intelligent Systems in Accounting, Finance and Management*, *26*(1), 46–53.

Olson, G. M., & Olson, J. S. (2003). Human-Computer Interaction: Psychological Aspects of the Human Use of Computing. *Annual Review of Psychology*.

Oracle. (2016). Can Virtual Experiences Replace Reality? The future role for humans in delivering customer experience, 19.

Otoo, B. A., & Salam, A. F. (2018). Mediating Effect of Intelligent Voice Assistant (IVA), User Experience and Effective Use on Service Quality and Service Satisfaction and Loyalty. In *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–9). San Francisco, USA.

Pereira, A. T., Prada, R., & Paiva, A. (2014). Improving social presence in human-agent interaction. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1449–1458). Toronto, Canada.

Perez, S. (2016). Starbucks unveils a virtual assistant that takes your order via messaging or voice. Retrieved November 23, 2018, from https://techcrunch.com/2017/01/30/starbucks-unveils-a-virtual-assistant-that-takes-your-order-via-messaging-or-voice/?guccounter=1

Pfeuffer, N., Toutaoui, J., Adam, M., Hinz, O., & Benlian, A. (2019). Mr. and Mrs. Conversational Agent - Gender Stereotyping in Judge-Advisor Systems and the Role of Egocentric Bias. In *Proceedings of the International Conference on Information Systems* (pp. 1–17). Munich, Germany.

Pickard, M. D., Roster, C. A., & Chen, Y. (2016). Revealing sensitive information in personal interviews: Is self-disclosure easier with humans or avatars and under what conditions? *Computers in Human Behavior*, *65*, 23–30.

Pickering, M. J., & Garrod, S. (2004). Toward a Mechanistic Psychology of Dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–190. https://doi.org/10.1017/s0140525x04000056

Porcheron, M., Fischer, J. E., Reeves, S., & Sharples, S. (2018). Voice Interfaces in Everyday Life. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–12). Montréal, Canada.

Powers, A., & Kiesler, S. (2006). The advisor robot: Tracing people's mental model from a robot's physical attributes. In *Proceedings of the 2006 ACM Conference on Human-Robot Interaction* (pp. 1–8). Salt Lake City, USA.

Provoost, S., Lau, H. M., Ruwaard, J., & Riper, H. (2017). Embodied conversational agents in clinical psychology: A scoping review. *Journal of Medical Internet Research*, *19*(5). https://doi.org/10.2196/jmir.6553

Punj, G., & Stewart, D. W. (1983). Cluster Analysis in Marketing Research: Review and Suggestions for Application. *Journal of Marketing Research*, *20*(2), 134–148.

Purington, A., Taft, J. G., Sannon, S., Bazarova, N. N., & Taylor, S. H. (2017). "Alexa is my new BFF": Social roles, user satisfaction, and personification of the Amazon Echo. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (Vol. Part F1276, pp. 2853–2859). Denver, USA.

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

Qiu, L., & Benbasat, I. (2010). A study of demographic embodiments of product recommendation agents in electronic commerce. *International Journal of Human Computer Studies*, *68*(10), 669–688.

Quynh, N., & Sidorova, A. (2018). Understanding User Interactions with a Chatbot: A Self-determination Theory Approach. In *Proceedings of the Americas Conference on Information Systems (AMCIS)* (pp. 1–5). New Orleans, USA.

Reeves, B., & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television and New Media Like Real People and Places*. The Center for the Study of Language and Information Publications.

Reinecke, K., & Bernstein, A. (2013). Knowing what a user likes: A design science approach to interfaces that automatically adapt to culture. *MIS Quarterly: Management Information Systems*, *37*(2), 427–453. https://doi.org/10.25300/MISQ/2013/37.2.06

Robert, J. L. R., & Sangseok, Y. (2018). Emotional Attachment, Performance, and Viability in Teams Collaborating with Embodied Physical Action (EPA) Robots. *Journal of the Association for Information Systems*, *17*, 267–307.

Rosenberg-Kima, R. B., Baylor, A. L., Plant, E. A., & Doerr, C. E. (2008). Interface agents as social models for female students: The effects of agent visual presence and appearance on female students' attitudes and beliefs. *Computers in Human Behavior*, *24*(6), 2741–2756.

Rosenthal-Von Der Pütten, A. M., & Krämer, N. C. (2014). How design characteristics of robots determine evaluation and uncanny valley related responses. *Computers in Human Behavior*, *36*, 422–439.

Rosenthal-von der Pütten, A. M., Straßmann, C., & Krämer, N. C. (2018). Dominant and submissive nonverbal behavior of virtual agents and its effects on evaluation and negotiation outcome in different age groups. *Computers in Human Behavior*.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*.

Rzepka, C., & Berger, B. (2018). User Interaction with AI-enabled Systems: A Systematic Review of IS Research. In *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–16). San Francisco, USA.

Saerbeck, M., Schut, T., Bartneck, C., & Janse, M. D. (2010). Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1613–1622). Atlanta, USA.

Saffarizadeh, K., Boodraj, M., & Alashoor, T. M. (2017). Conversational Assistants: Investigating Privacy Concerns, Trust, and Self-Disclosure. In *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 0–12). Seoul, Korea.

Sakamoto, D., Kanda, T., Ono, T., Kamashima, M., Imai, M., & Ishiguro, H. (2005). Cooperative embodied communication emerged by interactive humanoid robots. *International Journal of Human Computer Studies*, *62*(2), 247–265.

Sangseok, Y., & Lionel, P. R. J. (2019). Subgroup Formation in Human – Robot Teams. In *Proceedings of the International Conference on Information Systems* (pp. 1–17). Munich, Germany.

Schlesinger, A., Hara, K. P. O., & Taylor, A. S. (2018). Let's Talk About Race: Identity, Chatbots, and AI. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–14). Montréal, Canada.

Schroeder, J., & Schroeder, M. (2018). Trusting in Machines: How Mode of Interaction Affects Willingness to Share Personal Information with Machines. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS)* (Vol. 9, pp. 472–480). Waikoloa Village, Hawaii, USA.

Schuetz, S., & Venkatesh, V. (2020). The rise of human machines: How cognitive computing systems challenge assumptions of user-system interaction. *Journal of the Association for Information Systems*, 1–42.

Schuetzler, R. M., Giboney, J. S., Grimes, G. M., & Nunamaker, J. F. (2018). The Influence of Conversational Agents on Socially Desirable Responding. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS)* (Vol. 9, pp. 283–292). Waikoloa Village, Hawaii, USA.

Schuetzler, R. M., Grimes, G. M., Giboney, J. S., & Buckman, J. (2014). Facilitating Natural Conversational Agent Interactions: Lessons from a Deception Experiment. In *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–16). Auckland, New Zealand.

Sebastian, J., & Richards, D. (2017). Changing stigmatizing attitudes to mental health via education and contact with embodied conversational agents. *Computers in Human Behavior*, *73*, 479–488.

Seeber, I., Bittner, E., Briggs, R. O., de Vreede, T., de Vreede, G.-J., Elkins, A., … Söllner, M. (2019). Machines as Teammates: A Research Agenda on AI in Team Collaboration. *Information & Management*, 103174.

Seeber, I., Waizenegger, L., Seidel, S., Morana, S., Benbasat, I., & Lowry, P. B. (2019). Collaborating with Technology-Based Autonomous Agents: Issues and Research Opportunities. *Internet Research*.

Seeger, A.-M., Pfeiffer, J., & Heinzl, A. (2017). When Do We Need a Human? Anthropomorphic Design and Trustworthiness of Conversational Agents. In *Special Interest Group on Human-Computer Interaction* (Vol. 15, pp. 1–6).

Seeger, A.-M., Pfeiffer, J., & Heinzl, A. (2018). Designing Anthropomorphic Conversational Agents: Development and Empirical Evaluation of a Design Framework. In *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–17). San Francisco, USA.

Seering, J., Luria, M., Kaufman, G., & Hammer, J. (2019). Beyond dyadic interactions: Considering chatbots as community members. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–13). Glasgow, Scotland.

Seymour, Michael, Riemer, K., & Kay, J. (2017). Interactive Realistic Digital Avatars - Revisiting the Uncanny Valley. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS)* (pp. 547–556). Waikoloa Village, Hawaii, USA.

Seymour, Mike, Riemer, K., & Kay, J. (2018). Actors, Avatars and Agents: Potentials and Implications of Natural Face Technology for the Creation of Realistic Visual Presence. *Journal of the Association for Information Systems*, *19*, 953–981.

Shah, H., Warwick, K., Vallverdú, J., & Wu, D. (2016). Can machines talk? Comparison of Eliza with modern dialogue systems. *Computers in Human Behavior*, *58*(January), 278–295.

Shamekhi, A., Liao, Q. V., Wang, D., Bellamy, R. K. E., & Erickson, T. (2018). Face Value? Exploring the Effects of Embodiment for a Group Facilitation Agent. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–13). Montréal, Canada.

Sohn, S. (2019). Can Conversational User Interfaces Be Harmful? The Undesirable Effects on Privacy Concern. In *Proceedings of the International Conference on Information Systems* (pp. 1–9). Munich, Germany.

Son, Y., & Wonseok, O. (2018). “ Alexa, Buy Me a Movie !”: How AI Speakers Reshape Digital Content Consumption and Preference. In *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–17). San Francisco, USA.

Stieglitz, S., Brachten, F., & Kissmer, T. (2018). Defining Bots in an Enterprise Context. In *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–9). San Francisco, USA.

Stock, R. M., & Merkle, M. (2018a). Can Humanoid Service Robots Perform Better Than Service Employees? A Comparison of Innovative Behavior Cues. In *Proceedings of the 51st Hawaii International Conference on System Sciences*. Hawaii International Conference on System Sciences.

Stock, R. M., & Merkle, M. (2018b). Customer Responses to Robotic Innovative Behavior Cues During the Service Encounter. In *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–17). San Francisco, USA.

Stock, R., Merkle, M., Eidens, D., Hannig, M., Heineck, P., Nguyen, M. A., & Völker, J. (2019). Understanding Employee Trust in Assistive Robots When Robots Enter Our Workplace: Understanding Employee Trust in Assistive Robots. In *Proceedings of the International Conference on Information Systems* (pp. 1–9). Munich, Germany.

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 Hawaii International Conference on System Sciences (HICSS)* (pp. 2016–2025). Waikoloa Village, Hawaii, USA.

Strait, M., Vujovic, L., Floerke, V., Scheutz, M., & Urry, H. (2015). Too Much Humanness for Human-Robot Interaction. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 3593–3602). Seoul, Korea.

Straßmann, C., Pütten, A. M. R. Der, & Krämer, N. C. (2018). With or against Each Other? The Influence of a Virtual Agent's (Non) cooperative Behavior on User's Cooperation Behavior in the Prisoners ' Dilemma. *Advances in Human-Computer Interaction*.

Strohmann, T., Fischer, S., Siemon, D., Brachten, F., Lattemann, C., Robra-Bissantz, S., & Stieglitz, S. (2018). Virtual Moderation Assistance: Creating Design Guidelines for Virtual Assistants Supporting Creative Workshops. In *Proceedings of the Pacific Asia Conference on Information Systems (PACIS)* (pp. 3580–3594). Yokohama, Japan.

Sugumaran, V., & Davis, K. (2001). A Natural Language-Based Multi-Agent System for Legal Research. In *Proceedings of the Americas Conference on Information Systems (AMCIS)*. Boston, USA.

Szafir, D., & Mutlu, B. (2012). Pay attention! Designing adaptive agents that monitor and improve user engagement. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 11–20). Austin, Texas.

Tavanapour, N., Poser, M., & Bittner, E. A. C. (2019). Supporting the Idea Generation Process in Citizen Participation - Toward an Interactive System With a Conversational Agent. In *Proceedings of the European Conference on Information Systems (ECIS)* (pp. 1–17).

Stockholm, Sweden.

Tickle-Degnen, L., & Rosenthal, R. (1990). The Nature of Rapport and Its Nonverbal Correlates. *Psychological Inquiry*.

Tinwell, A., & Sloan, R. J. S. (2014). Children's perception of uncanny human-like virtual characters. *Computers in Human Behavior*, *36*, 286–296.

Toxtli, C., Monroy-Hernández, A., & Cranshaw, J. (2018). Understanding Chatbot-mediated Task Management. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Montréal, Canada.

Truong, A. (2016). Parents Are Worried the Amazon Echo Is Conditioning Their Kids to Be Rude.

Vaccaro, K., Agarwalla, T., Shivakumar, S., & Kumar, R. (2018). Designing the Future of Personal Fashion Experiences Online. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–11). Montréal, Canada.

Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape. *Canadian Journal of Psychiatry*, *64*(7), 456–464. https://doi.org/10.1177/0706743719828977

Van den Broeck, E., Zarouali, B., & Poels, K. (2019). Chatbot advertising effectiveness: When does the message get through? *Computers in Human Behavior*, *98*, 150–157.

Van Der Meij, H. (2013). Motivating agents in software tutorials. *Computers in Human Behavior*, *29*(3), 845–857.

Veletsianos, G., & Russell, G. S. (2014). Pedagogical agents. In *Handbook of Research on Educational Communications and Technology: Fourth Edition* (pp. 759–769). Springer. https://doi.org/10.1007/978-1-4614-3185-5_61

Venkatesh, Morris, Davis, & Davis. (2003). User Acceptance of Information Technology: Toward a Unified View. *Management Information Systems Quarterly*, *27*(3), 425.

Vertegaal, R., Slagter, R., van der Veer, G., & Nijholt, A. (2000). Why conversational agents should catch the eye. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (p. 257). The Hague, Netherlands.

Vertegaal, R., Slagter, R., van der Veer, G., & Nijholt, A. (2001). Eye gaze patterns in conversations. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 301–308). Seattle, USA.

Vogel-Meijer, K. (2018). KLM: Making airline customer service soar with Messenger. Retrieved November 1, 2018, from https://www.facebook.com/business/success/klm-messenger

vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. *Proceedings of the European Conference on Information Systems*

(ECIS), *9*, 2206–2217.

Vtyurina, A., & Fourney, A. (2018). Exploring the Role of Conversational Cues in Guided Task Support with Virtual Assistants. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 208:1--208:7). Montréal, Canada.

Vugt, H. C. Van, Bailenson, J. N., Hoorn, J. F., & Konijn, E. A. (2010). Effects of facial similarity on user responses to embodied agents. *ACM Transactions on Computer-Human Interaction*, *17*(2), 1–27.

Wagner, K., & Schramm-Klein, H. (2019). Alexa, Are You Human? Investigating the Anthropomorphism of Digital Voice Assistants – A Qualitative Approach. In *Proceedings of the International Conference on Information Systems* (pp. 1–17). Munich, Germany.

Wakefield, J. (2016). Would you want to talk to a machine? Retrieved July 9, 2017, from https://www.bbc.com/news/technology-36225980

Wang, N., Johnson, W. L., Mayer, R. E., Rizzo, P., Shaw, E., & Collins, H. (2008). The politeness effect: Pedagogical agents and learning outcomes. *International Journal of Human Computer Studies*, *66*(2), 98–112.

Wang, W., & Benbasat, I. (2005). Trust in and Adoption of Online Recommendation Agents. *Journal of the Association for Information Systems*, *6*(3), 72–101.

Ward, N., & Tsukahara, W. (2003). A study in responsiveness in spoken dialog. *International Journal of Human Computer Studies*, *59*(5), 603–630.

Watanabe, M., Ogawa, K., & Ishiguro, H. (2015). Can Androids Be Salespeople in the Real World? In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 781–788). Seoul, Korea.

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *Management Information Systems Quarterly*, *26*(2), xiii–xxiii.

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45.

Welch, C. (2018). Google just gave a stunning demo of Assistant making an actual phone call. Retrieved May 18, 2018, from https://www.theverge.com/2018/5/8/17332070/google-assistant-makes-phone-call-demo-duplex-io-2018

Wik, P., & Hjalmarsson, A. (2009). Embodied conversational agents in computer assisted language learning. *Speech Communication*, *51*(10), 1024–1037. https://doi.org/10.1016/j.specom.2009.05.006

Winkler, R., 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 International Conference on Information Systems* (pp. 1–17). Munich, Germany.

Winkler, R., & Roos, J. (2019). Bringing AI into the Classroom: Designing Smart Personal

Assistants as Learning Tutors. In *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–9). Munich, Germany.

Winkler, R., Söllner, M., Neuweiler, M. L., Conti Rossini, F., & Leimeister, J. M. (2019). Alexa, Can You Help Us Solve This Problem? In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Glasgow, Scotland.

Wolff, R. M. Von, 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 Hawaii International Conference on System Sciences (HICSS)* (Vol. 6, pp. 95–104). Grand Wailea, Maui.

Wünderlich, N. V., & Paluch, S. (2017). A Nice and Friendly Chat With a Bot: User Perceptions of AI-based Service Agents. In *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–11). Seoul, Korea.

Xu, A., Liu, Z., Guo, Y., Sinha, V., & Akkiraju, R. (2017). A New Chatbot for Customer Service on Social Media. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 3506–3510). Denver, USA.

Xu, K., & Lombard, M. (2017). Persuasive computing: Feeling peer pressure from multiple computer agents. *Computers in Human Behavior*, *74*, 152–162. https://doi.org/10.1016/j.chb.2017.04.043

Xu, Q., Li, L., & Wang, G. (2013). Designing engagement-aware agents for multiparty conversations. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (p. 2233). Paris, France.

Yamada, S., Terada, K., Kobayashi, K., Komatsu, T., Funakoshi, K., & Nakano, M. (2013). Expressing a robot's confidence with motion-based artificial subtle expressions. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (p. 1023). Paris, France.

Yang, Y., Ma, X., & Fung, P. (2017). Perceived Emotional Intelligence in Virtual Agents. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 2255–2262). Denver, USA.

Yokotani, K., Takagi, G., & Wakashima, K. (2018). Advantages of virtual agents over clinical psychologists during comprehensive mental health interviews using a mixed methods design. *Computers in Human Behavior*, *85*, 135–145.

Zhang, P., Benbasat, I., Carey, J., Davis, F., Galletta, D. F., & Strong, D. (2002). Human-Computer Interaction Research in the MIS Discipline. *Communications of the Association for Information Systems*, *9*.

Zhang, P., & Li, N. (2004). An assessment of human – computer interaction research in management information systems : topics and methods. *Computers in Human Behavior*, *20*, 125–147.

Zhang, P., & Li, N. (2005). The Intellectual Development of Human-Computer Interaction Research: A Critical Assessment of the MIS Literature (1990-2002). *Journal of the Association for Information Systems*, *6*(11), 227–292.

Zhang, P., Li, N., Scialdone, M., & Carey, J. (2008). The Intellectual Advancement of Human-Computer Interaction Research: A Critical Assessment of the MIS Literature (1990-2008). *AIS Transactions on Human-Computer Interaction*, *1*(3), 55–107.

Zhang, Y., Song, W., Tan, Z., Zhu, H., Wang, Y., Lam, C. M., … Yi, L. (2019). Could social robots facilitate children with autism spectrum disorders in learning distrust and deception? *Computers in Human Behavior*, *98*, 140–149.

Zierau, N., Engel, C., Söllner, M., & Leimeister, J. M. (2020). Trust in Smart Personal Assistants: A Systematic Literature Review and Development of a Research Agenda. *Proceedings of the International Conference on Wirtschaftsinformatik*, 99–114.

Appendix A. Coding Guideline

Human	
Age	Did the study collect and analyze the users' age (was it explicitly analyzed or discussed in-depth independent of the analysis results)?
Gender	Did the study collect and analyze the users' gender (was it explicitly analyzed or discussed in-depth independent of the analysis results)?
Education	Did the study collect and analyze the users' education (was it explicitly analyzed or discussed in-depth independent of the analysis results)?
Cultural Background	Did the study collect and analyze the users' cultural background (e.g., ethnicity or nationality); (was it explicitly analyzed or discussed in-depth independent of the analysis results)?
CA Experience	Did the study collect data on individual user experience with CAs or digital assistants and analyze it? Did researchers explicitly analyze or discuss the data in depth, independent of the analysis results?
Task Experience	Did the study collect data on individual user experience with the task at hand collected and analyzed? Did researchers explicitly analyze or discuss the data in depth, independent of the analysis results?
Personality	Did the study collect data on user personality (e.g., introversion or neuroticism) and analyze it? Did researchers explicitly analyze or discuss the data in depth, independent of the analysis results?
Cognitive Style	Was data on user's cognitive style (e.g., learning style) gathered and analyze (was it explicitly analyzed or discussed in-depth independent of the analysis result?
Context	
Professional task support	CAs for individual task support in a company context (e.g., the CA as a personal assistant, for information search (FAQ), or data analytics).
Team collaboration	CAs in professional team settings (e.g., as a moderator in meetings or for shared task management).
Customer interface	CAs at the customer interface (e.g., for service provision or as additional sales channel).
Private task support	CAs for individual task support (e.g., in-car assistants or for personal time management).
Health	CAs for individual health (e.g., to promote health awareness or to provide initial, individual diagnosis).
Education	CAs for education (e.g., as intelligent tutoring systems).
Multiple	Studies that address multiple contexts.
Generic	Studies that do not define a specific context for the CA.
Other	CAs that do not fit any of the above contexts.
Agent	

Communication Mode	Did the study investigate the CA's communication mode, that is, was it included in the analysis and discussed afterwards?
Embodiment	Did the study investigate the CA's embodiment, that is, was it included in the analysis and discussed afterwards?
(Human) Identity	Did the study investigate the CA's (human) identity, that is, was it included in the analysis and discussed afterwards?
Verbal Communication	Did the study investigate the CA's verbal communication, that is, was it included in the analysis and discussed afterwards?
Non-verbal Communication	Did the study investigate the CA's non-verbal communication, that is, was it included in the analysis and discussed afterwards?
Perception and Outcome	
Perception	Humanness, social presence, competence, authority
Acceptance	Use, intention to use, acceptance, resistance to use
Attitude	Attitude, satisfaction, preference
Performance	Productivity, effectiveness, efficiency
Emotion	Affect, hedonic quality, enjoyment, humor, intrinsic motivation
Trust	Trust, risk, loyalty, security, privacy
Learning	Learning models, learning processes, general training
Ethics	Ethical belief, ethical behavior, ethics
Relationship	Influence, interdependence, interference, agreement/disagreement, persuasiveness
Other	Further topics not included in the list
Complementary	
Article Type	Is the article research in progress/a short paper, or is it completed research/a full paper?
Research Approach	Is the study conceptual or empirical? Studies where, for example, conceptual frameworks are developed and empirically tested are coded as empirical.
Research Method	What method was used in the study? For multi-method studies, the "primary" method is coded.
Unit of Analysis	On what level of analysis was the study conducted?
Theoretical Grounding	Which theories or research backgrounds were used in the study?

Table 13. Coding Guideline

Appendix B. Cluster Analysis

Step	Consolidated Cluster		Coeffi-cient	First Appearance		Next Step	Coeffi-cient Delta	#Cluster
	Cluster1	Cluster 2		Cluster 1	Cluster 2			
252	2	7	677.279	245	225	256	15.007	10
253	8	9	695.607	237	247	254	18.328	9
254	8	15	715.616	253	239	258	20.009	8
255	1	3	735.667	250	243	259	20.051	7
256	2	4	756.878	252	230	259	21.211	6
257	5	19	787.31	251	240	260	30.432	5
258	8	16	820.058	254	248	260	32.748	4
259	1	2	882.41	255	256	261	62.352	3
260	5	8	967.224	257	258	261	84.814	2
261	1	5	1063.046	259	260	0	95.822	1

Table 14. Agglomeration Schedule (last ten steps)

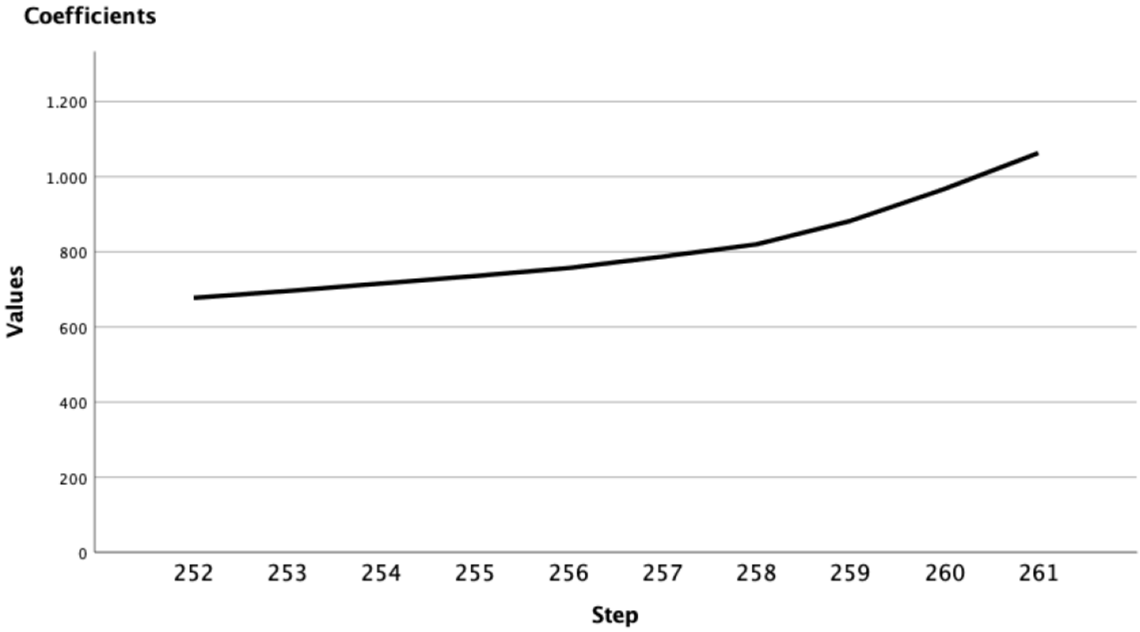


Figure 3. Scree Diagram (last ten steps)

Appendix C. Studies over Time

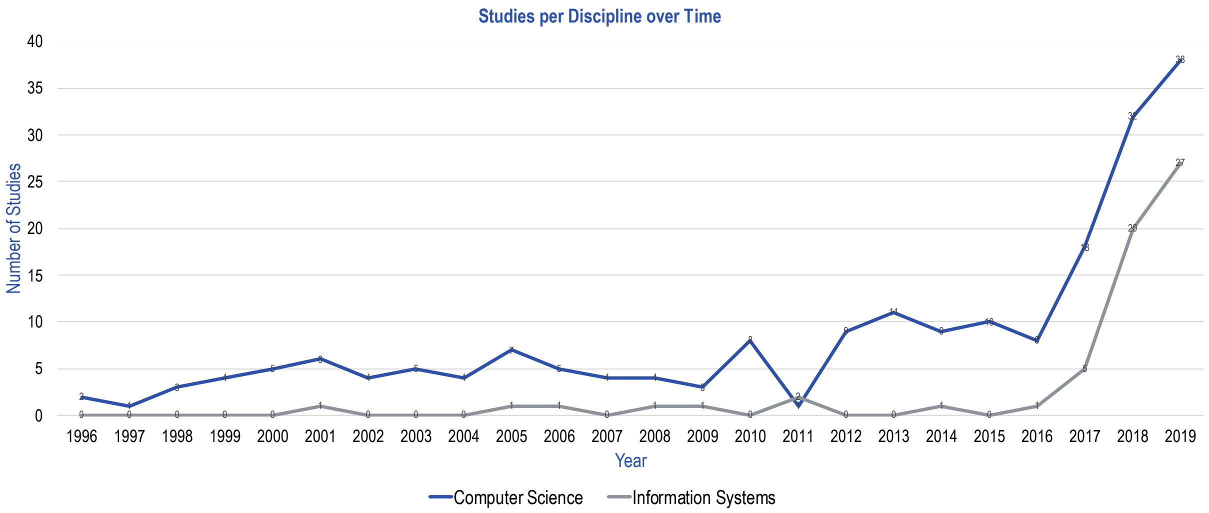


Figure 4: Studies over Time by Discipline

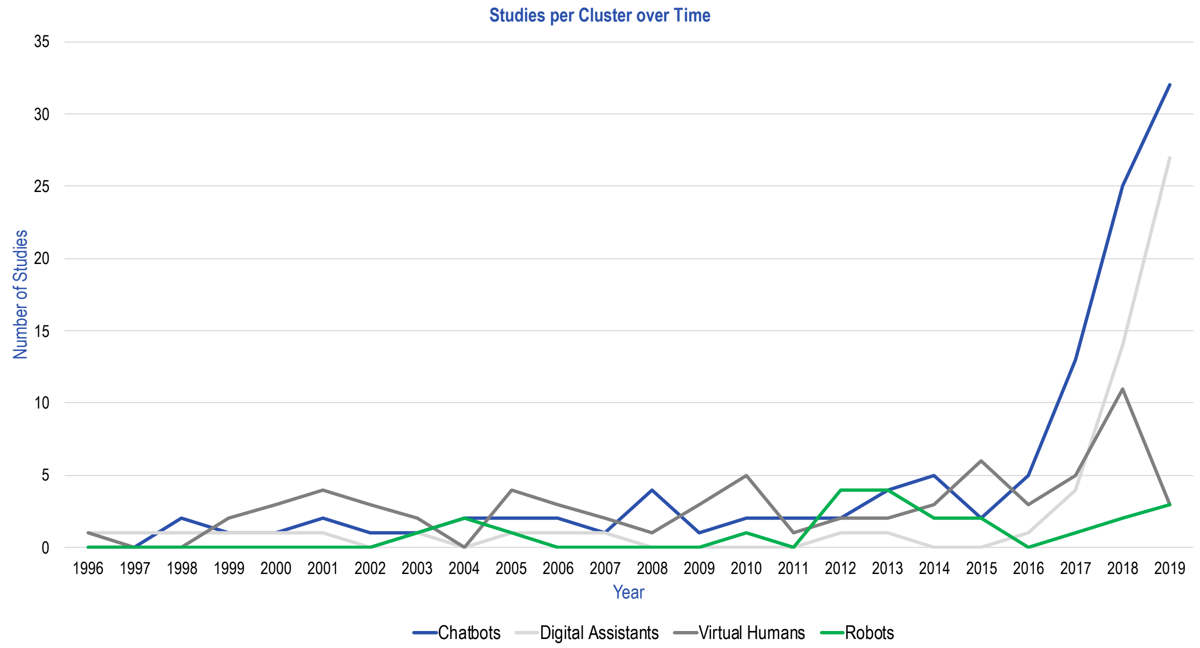


Figure 5: Studies over Time by Research Stream (Cluster)

Stephan Diederich is a former research associate at the Chair of Information Management of the University of Göttingen and a senior consultant at a large management consultancy concentrating on the digital transformation of organizations. His research focuses on conversational agents, in particular chatbots, within an organizational context, and has been published in journals such as Business and Information Systems Engineering and AIS Transactions on Human-Computer Interaction as well as IS conferences like the International Conference on Information Systems or the European Conference on Information Systems.

Alfred Benedikt Brendel is an associate professor of Business Informatics, esp. Intelligent Systems and Services, at the Technische Universität Dresden, Germany. His research is concerned with the design of information systems in the domains of healthcare, transportation, and digital workplace. Alfred's research has been published or is forthcoming in leading IS journals, such as Journal of Information Technology, Business & Information System Engineering, Information Systems Frontiers, and AIS Transactions on Human-Computer Interaction.

Stefan Morana is a junior professor of Digital Transformation and Information Systems at the Saarland University. His research focuses on the human-centered design of interactive systems for the digital transformation from the perspective of the individual, organizations, and society. More specifically, he investigates the design of assistant systems and conversational interfaces supporting the individual usage of information systems. His research has been published in journals such as the Journal of the Association for Information Systems, Decision Support Systems, International Journal of Human-Computer Studies, Business & Information Systems Engineering, Internet Research, AIS Transactions on Human-Computer Interaction, and Communications of the Association for Information System as well as major information systems conferences.

Lutz Kolbe is a professor of Information Systems and leads the Chair of Information Management at the University of Göttingen. His research focuses on the management of information and information technology as a crucial factor for sustainable business success. Lutz manages the research and project activities of three groups covering the areas of digital transformation, digital health, and smart mobility. His publications have appeared in several IS journals and conferences, such as the European Journal of Information Systems, the Journal of the Association for Information Systems, and the Information Systems Journal.