

## Secondary Publication



Hofeditz, Lennart; Mirbabaie, Milad; Ortmann, Mara

## Ethical Challenges for Human-Agent Interaction in Virtual Collaboration at Work

Date of secondary publication: 12.09.2025

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-110357x

### Primary publication

Hofeditz, Lennart; Mirbabaie, Milad; Ortmann, Mara (2024): Ethical Challenges for Human-Agent Interaction in Virtual Collaboration at Work, in: International journal of human computer interaction, New York: Taylor & Francis, Vol. 40, Nr. 23, pp. 8229–8245, doi: 10.1080/10447318.2023.2279400.

### Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holders.

This document is made available under a Creative Commons license.



The license information is available online:

<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

# Ethical Challenges for Human–Agent Interaction in Virtual Collaboration at Work

Lennart Hofeditz<sup>a</sup> , Milad Mirbabaie<sup>b\*</sup>, and Mara Ortmann<sup>c</sup>

<sup>a</sup>Faculty of Economics and Social Science, University of Potsdam, Potsdam, Germany; <sup>b</sup>Faculty of Business Administration and Economics, Management Information Systems, Paderborn University, Paderborn, Germany; <sup>c</sup>Department of Computer Science and Applied Cognitive Science, University of Duisburg-Essen, Duisburg, Germany

## ABSTRACT

In virtual collaboration at the workplace, a growing number of teams apply supportive conversational agents (CAs). They take on different work-related tasks for teams and single users such as scheduling meetings or stimulating creativity. Previous research merely focused on these positive aspects of introducing CAs at the workplace, omitting ethical challenges faced by teams using these often artificial intelligence (AI)-enabled technologies. Thus, on the one hand, CAs can present themselves as benevolent teammates, but on the other hand, they can collect user data, reduce worker autonomy, or foster social isolation by their service. In this work, we conducted 15 expert interviews with senior researchers from the fields of ethics, collaboration, and computer science in order to derive ethical guidelines for introducing CAs in virtual team collaboration. We derived 14 guidelines and seven research questions to pave the way for future research on the dark sides of human–agent interaction in organizations.

## KEYWORDS

Conversational agents; human–computer interaction; virtual collaboration; ethics; virtual teams; trust



## 1. Introduction

Physically distributed virtual teams are taking on a central and ever-growing role in the increasing globalization of work. It is estimated that over 85% of workers now collaborate in some form of virtual team (Morrison-Smith & Ruiz, 2020). Previous research suggested that applying conversational agents (CAs) to support virtual collaboration can have positive effects on team performance (Benke et al., 2020; Shin et al., 2023). CAs can be defined as a group of automated technologies being able to respond to human user's queries or commands in a natural way (text or speech) in order to solve a goal related task such as scheduling a meeting (Wei et al., 2022), to recommend a product in e-commerce (Rhee & Choi, 2020) or supporting in the recruitment process of an organization (Koivunen et al., 2022). The most popular and artificial intelligence (AI)-enabled CA ChatGPT is also used by many employees in virtual collaboration for brainstorming and the creation of sophisticated texts based on users' questions and prompts (Thorp, 2023).

Previous research merely focused on the positive sides of introducing CAs in virtual collaborative work (Elshan & Ebel, 2020; Stieglitz et al., 2022). Shah et al. (2016) test the conversational abilities of systems such as Cleverbot, Elbot, Eugene Goostman, JFred, and Ultra Hal and concluded that CAs can share personal opinions or mislead people. For CAs such as ChatGPT it is possible that it not only

produces wrong or biased outputs (Zhuo et al., 2023), but also that it produces different answers to the semantically same question. This hallucination increasingly occurs during long-term interactions and might cause confusion among the users (Rodríguez-Cantelar et al., 2023).

Although communication via chat software such as CAs can be used in cases of workplace harassment to hold the harassers legally accountable, it also extends the ethical responsibilities of the employer (Tenório & Bjørn, 2019) including ethical risks such as that information can be viewed and misused by third parties (Lindner, 2020). Since CAs often have AI-based capabilities, it stands to reason that they raise similar ethical challenges than other AI-based systems, e.g., by collecting user data or limiting user autonomy. For example, Richards et al. (2023) explored scenarios around the five ethics principles of beneficence, non-maleficence, justice, explicability, and autonomy and found ethical issues especially regarding people's desire for autonomy in the design of artificial social agents. Due to the social and human nature of a CA, the CA's autonomy of what it can and cannot do is often not transparent (Schuetzler et al., 2021). This also relates to the employee's privacy of their personal data when it is not clear which data will be processed by a CA and for what purpose (Mirbabaie et al., 2022). This could result in uncertainty in virtual collaboration among team members.

**CONTACT** Lennart Hofeditz  [lennart.hofeditz@uni-potsdam.de](mailto:lennart.hofeditz@uni-potsdam.de)  Faculty of Economics and Social Science, University of Potsdam, August-Bebel-Str. 89, 14482 Potsdam, Germany

\*Present address: Faculty of Information Systems and Applied Computer Sciences, University of Bamberg, Bamberg, Germany.

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

In contrast to ethical issues of using AI-based systems in general, applying CAs in virtual collaboration at the workplace also raises unique challenges. In virtual collaboration, CAs can be perceived as both supportive tools and individual team members (Mirbabaie et al., 2021), which could shift dependencies and cause uncertainty about the job responsibility and job security among the employees (Stieglitz et al., 2021). Interacting with a CA in virtual collaboration can also result in alienation of team members (Mirbabaie et al., 2021). Previous research also highlighted cases of discrimination when organizations deploy CAs (Suárez-Gonzalo et al., 2019) or AI-based systems (Dastin, 2022), which could also cause biases within a team. Although some previous works already started to discuss responsibilities, stereotypical CA design and the role of social norms (Rheu et al., 2021), there is little scientific evidence that specifically relates to ethical questions and challenges of the human-agent interaction in digital workplaces and their impact on collaboration between employees. An increasing number of people work in virtual teams, while CAs are being used extensively to support them (Maedche et al., 2019). This raises specific ethical challenges for human-computer interaction in teamwork, such as changes in team dynamics, complex interactions, a shift in responsibilities and trust, and fears about the workplace. Previous research in the area of ethics either focuses on specific application fields or addresses the use of AI in general (cf., AI HLEG, 2019; Jobin et al., 2019). Although these findings might help as a starting point for examining ethical challenges of applying CAs in virtual collaboration, existing guidelines are often too generic or too context specific to address these increasing ethical issues in CA interaction in virtual collaboration at the workplace. Furthermore, ethical norms can differ between different cultural backgrounds (Hafez, 2002). Even the perceived cultural background of a CA can influence whether it is treated like an in-group or out-group member (de Melo & Terada, 2019). It would be therefore evident to start with considering one cultural space. Based on this problem, we derived the following research questions (RQs):

RQ1: Which ethical challenges do researchers see for introducing CAs in virtual collaboration at the workplace in a Western societal context?

RQ2: How could these ethical challenges in virtual collaboration at the workplace be addressed by companies and team leaders in Western countries?

To answer the RQs, we conducted 15 expert interviews with researchers from different Western countries with a European focus to ensure a similar normative ethical and work cultural background (de Melo & Terada, 2019). For this purpose, we first identified central ethical challenges from the literature in the areas of human-computer interaction, ethics, virtual collaboration and CAs and transferred them into an interview guide. We used the five ethical principles of Floridi et al. (2018) as a foundation. Since it is expected that about a quarter of all employees will interact with CAs in the future in virtual collaboration (Maedche et al., 2019), this work is relevant for human-computer interaction research. Our results are relevant for managers

and leaders of collaborative teams, as they may influence the way CAs are used in the context of teamwork and because several of the ethical guidelines are directly addressed to the leaders of virtual teams. Furthermore, the results are relevant for developers of virtual systems used in the context of collaborative work. Finally, the results generally provide a basis for further research in the area of collaboration with CAs. To guide future research, we derived a research agenda for further examining ethical challenges of applying CAs in virtual collaborative work.

### **1.1. Human-agent interaction in virtual collaboration at the workplace**

Virtual collaboration describes the working of team members in companies whose communication is predominantly computer-based instead of face-to-face encounters (Kristof et al., 1995). The key difference of virtual and non-virtual teams is that in physically collaborative teams, members who come from different countries or cities need to come together in one place and work together face-to-face (Orta-Castañón et al., 2018) whereas the members of virtual teams work geographically dispersed at delocalized and decentralized work locations and are often exposed to interdisciplinary, cultural, and linguistic differences (Kauffeld et al., 2016; Kristof et al., 1995; Wainfan & Davis, 2004). Virtual collaboration refers to the use of information and communication technology to support collective interaction between multiple parties involved (Hossain & Wigand, 2003; Kock, 2000). Forming teams based on the professional qualifications of the team members rather than on their local availability represents a competitive advantage for companies (Konradt & Hertel, 2007). The value of collaboration, according to Koch et al. (2015), is that people come together as groups to create value that they could not create as individuals. Besides the world of work is currently consolidating the view that work is more about what is done than the locality in which work is done (Orta-Castañón et al., 2018). According to Seerat et al. (2013), virtual team managers need to ensure powerful communication and collaboration to create an organized and inclusive virtual work environment. However, for the success of virtual teams, it is essential to address challenges that are currently still open and to create framework conditions for virtual collaboration (Kauffeld et al., 2016). Due to the steadily increasing number of virtual teams, it is essential to continue investing in research and the practical implementation of findings in this area (Kauffeld et al., 2016). Virtual teams also face challenges that arise for the teams, the members and the managers on a social and coordinating level (Kauffeld et al., 2016). Due to the geographical, temporal and organizational boundaries, communication and coordination hurdles arise more quickly than in face-to-face collaborations (Kauffeld et al., 2016). In addition, the reduced communication possibilities ensure both a less pronounced occurrence of positive effects, such as increased information processing (Kauffeld et al., 2016), and a neglect of the relationship level between the team members (Straus, 1996). Maintaining group cohesion

is thus more difficult in virtual teams than in physically cooperating teams. This leads to an increased risk of affective conflicts, which can arise due to a lack of familiarity within the team or the formation of sub-groups at the same locations (Kauffeld et al., 2016). In addition, communication via software involves the risk that information can be viewed and misused by third parties (Lindner, 2020).

CAs represent virtual personal assistants that communicate with users via natural language (Gnewuch et al., 2018) and can autonomously perform different tasks for the user (Mirbabaie et al., 2021). In organizations, these tasks can be external such as taking over customer support to relieve employees in their customer interaction (Sands et al., 2021) or internal such as guiding the process of design thinking for employees to generate new ideas (Lembcke et al., 2020). The underlying idea is that users can use CAs with natural language as if they were communicating with another human being. The AI-based systems aim to collaborate with employees and support them in the execution of work-related tasks. Besides, the utility of traditional systems can be increased through the use of AI (Frick et al., 2019). The use of CAs in teamwork brings different benefits, such as an acceleration of work processes (de Vreede & Briggs, 2005), better work results and reduced cognitive burdens (Brachten et al., 2020; Maedche et al., 2016). These potentials lead to an increasing use of AI-based CAs in the workplace (Maedche et al., 2016). CAs are particularly relevant for virtual collaboration as many team members work in physically dispersed locations, which requires teamwork via collaborative tools (Mirbabaie et al., 2021). Virtual teams are also of great benefit in situations such as the Covid-19 pandemic, where even non-physically distributed team members may not work together in one location (Mysirlaki & Paraskeva, 2020). Here, the performance of such virtual teams can be improved by adapting team processes to the opportunities offered by technology (Waizenegger et al., 2020).

With the advances that are taking place in the continuous development of AI-based CAs, CAs are becoming a part of everyday life. It is foreseeable that they will also become a key element for the future of work (Maedche et al., 2019), as AI has the potential to make machines the smart collaboration partners of its users in the future (Seeber et al., 2020; Strohmman et al., 2018) and mean numerous benefits for different areas in organizations that focus on collaboration between users in professional practice (Frick et al., 2019).

### **1.2. The role of ethics in virtual collaboration at the workplace**

A problematic issue when considering virtual collaboration in companies is that while ethics is often recognized as an important issue in human collaboration, ethical principles are not always explicitly included in the design of organizational virtual collaboration. Human issues and ethical issues in the context of technology are often considered independently from each other (Hofeditz, Clausen, et al., 2022). Here, there is a need to incentivize clear baseline assessments so

that companies have a more concrete sense of the need to seriously and concretely incorporate ethical principles into collaboration (Chatterjee et al., 2009; Hofeditz, Mirbabaie, et al., 2022). It is also an open question what ethical rules should be followed when virtual teams operate globally and there are differentiating ethical principles depending on the project or company location. These are issues that need to be addressed to ensure effective and efficient human-computer collaboration at work (Saat & Salleh, 2010). Stephanidis et al. (2019) raised six challenges for human-computer interaction and discussed them with a group of 32 experts. They identified human-technology symbiosis, human-environment interactions, ethics, privacy and security, well-being, health and eudaimonia; accessibility and universal access, learning and creativity, and social organization and democracy as the most important challenges for human-computer interaction. Although these provide a good starting point for highlighting societal issues in human-computer interaction, they are neither focusing on virtual collaboration at work and the potential impact on employees and teamwork, nor are they taking into account effects of human-like characteristics of CAs.

The ubiquitous use of CAs in everyday life, and especially in professional contexts, makes a critical evaluation of the impact of CAs essential (Ruane et al., 2019). Ruane et al. (2019) point out that the ethical concerns that arise in the use of CAs or conversational AI must always be considered contextually. Considering and addressing existing ethical challenges and identifying emerging ethical problems from the fields of AI and collaboration is particularly relevant as CAs are often AI-enabled and increasingly used in collaboration. The interdisciplinary nature of the topic requires looking at and linking ethical challenges from different fields.

According to Meyer von Wolff et al. (2019), principles for the design of CAs in the digital workplace need to be established to address the lack of exploration of CAs in virtual collaboration. One way to do this is to identify the concrete opportunities that arise from the use of smart technologies (Seeber et al., 2020). Different data protection issues may also arise, such as surveillance and lack of accountability and transparency (Mateescu & Nguyen, 2019). The risk of misuse of data by third parties is also addressed, with an associated risk to workers' privacy and data protection (Lindner, 2020). The literature therefore raises the question of what types of transparency are required on the part of employers (Floridi et al. 2018; Mateescu & Nguyen, 2019). In this context, it is also necessary considering social risks that arise from working in virtual teams for the team and its members and especially when working with technologies with a certain degree of autonomy (Floridi et al. 2018). Ozmen Garibay et al. (2023) raised six main ethical challenges for the design and use of AI-based technologies such as CAs: (1) the technology is centered in human well-being, (2) it is designed in a responsible way, (3) users' privacy is respected, (4) it matches human-centered design principles, (5) there is some degree of human governance and oversight, and (6) it interacts with humans and respects their

cognitive capacities. In the context of virtual collaborative work, there are more specific challenges that need to be taken into account. One example is that the virtual nature of collaboration can limit team communication and cause a neglect of the relational level between team members, as well as cohesion and trust (Hofeditz et al., 2021; Straus, 1996). Besides, user autonomy is a relevant aspect in the interaction with CAs that finds much application in the literature, which also raises the question of the role of accountability in the use of CAs in virtual team collaboration (Floridi et al. 2018). For AI-based systems, there are several frameworks that aim to guide future research and practice (Almeida et al., 2020; European's Commission: High Level Expert Group (HLEG), 2019; Kaplan & Haenlein, 2020; Paraman & Anamalah, 2023; Richards et al., 2023; Shneiderman, 2020). Floridi et al. (2018) provided an ethical framework for AI consisting of the five principles beneficence, non-maleficence, autonomy, justice, and explicability. Many previous works took these very general principles as a starting point for further investigation. As an illustration, the study conducted by Richards et al. (2023) investigated the importance of each principle in different context scenarios. The research revealed notable ethical concerns, particularly related to individuals' autonomy desires when it comes to designing artificial social agents.

As these principles are the most common general ethics principles for AI-based technologies, we took them as foundational framework for the identification of challenges and risks in the introduction of CA and human-agent interaction in virtual collaboration.

## 2. Materials and methods

We transferred the problems and challenges identified in the literature analysis into an interview guide in order to make them comprehensible and answerable for the experts (Kaiser, 2014). The interview guide consisted of a total of 18 main questions, which are grounded on the challenges and problems identified in the literature. We further grouped them thematically according to the five ethical principles of Floridi et al. (2018), some virtual collaboration specific questions (e.g., Gutwin & Greenberg, 2002) and general questions regarding the introduction and use of CAs resulting in the following categories:

Introduction and use of CAs (five main questions, e.g., What tasks should not be handed over to a CA?), beneficence and affective conflicts (three main questions: e.g.: What new pressures might the introduction of CAs place on team members?) (Floridi et al., 2018), justice, solidarity, and workspace awareness (four questions, e.g., What threats to team solidarity can arise from the use of CAs?) (Floridi et al., 2018; Gutwin & Greenberg, 2002), Non-maleficence and privacy (two questions: e.g., What is the value of the employee's right to privacy and data protection when using a CA that continuously collects and analyses data?) (Floridi et al., 2018), explicability (one question: What types of transparency are required on the part of the employer?) (Floridi et al., 2018), accountability/autonomy (two

questions: e.g., What role does accountability play in the use of CAs?) (Floridi et al., 2018), and concluding remarks (e.g., do you would like to add something that we missed so far at the end?). The complete interview guideline can be found in the [Appendix](#).

The interviews were semi-structured (B. Saunders et al., 2018; J. F. Saunders et al., 2020), as this approach gave the interviewees the flexibility to respond to the questions as they think appropriate (Fylan, 2005). There were neither predetermined answers nor time constraints. Where it was thematically appropriate, the order of the interview guide was deviated from. This allowed the experts to answer the questions openly, to introduce their own aspects and to ask follow-up questions (Kaiser, 2014). The quantitative research method of expert interviews was chosen because through expert interviews, causes for problematic issues can be developed, backgrounds can be tapped into and ambiguities can be removed (Röbken & Wetzel, 2017). Furthermore, the open nature of expert interviews offers a deep information content and makes it possible to open up new and previously unknown facts (Röbken & Wetzel, 2017). In the context of the present research, researchers from the fields of ethics, collaboration, and computer science were selected as experts. Researchers from these fields were explicitly approached, as ethical challenges in the use of CAs in virtual collaboration has aspects of each of these three research areas and requires an interdisciplinary approach. As the topic is a fairly new and interdisciplinary field of research, the expertise of experts here can both bring new knowledge and lead to a deeper understanding of the topic.

### 2.1. Selection of the interview partner and interview conduction

The interview partners were selected on the basis of their expertise in the fields of ethics, computer science, or collaboration. When selecting the persons, attention was paid to whether the respective expert is or has been involved in research, publications, or teachings that fit the topic of the present work and therefore have an in-depth understanding of the respective research areas. The interview partners are professors and PhDs in different countries with a Western cultural background with a focus on Europe to be able to draw conclusions, which are more generalizable for one cultural space with similar ethical norms. [Table 1](#) provides an overview of the participants and their respective research areas.

In total, 15 researchers from five different countries and three continents were interviewed in order to cover a broad sample and obtain multi-layered data. With nine interviews, the majority were conducted with people who are resident or working in Germany. The remaining six interviews were conducted with researchers currently practicing in the United States, Australia, Norway, and Austria. Care was taken to ensure that each research area was covered by several experts in order to obtain a comprehensive and multi-layered view of the topic. Ten of the interviews were conducted in German and five in English. The interviews were subsequently

**Table 1.** Description of the interview partners.

ID	Professional position or expertise	Country	Expert in		
			Computer science	Collaboration	Ethics
E1	Associate Professor in information systems and researcher for communication technologies for team collaboration	USA	X	X	
E2	Researcher in the field of digitalization in companies and virtual teams	Germany		X	
E3	Professor in the field of innovation in society	Australia	X		
E4	Professor for business information systems	Germany	X		
E5	Researcher in human-computer interaction and computer science	Norway	X		
E6	Professor in the fields of management and technology	Austria	X	X	
E7	Professor for business information systems	Germany	X		
E8	Professor for information science and information ethics	Germany	X		X
E9	Professor in the field of philosophy, business ethics and digitalization	Germany	X		X
E10	Researcher at the intersection of artificial intelligence and human behavior	Germany	X	X	
E11	Professor for digital transformation and business information systems	Germany		X	
E12	Associate Professor in the fields of media and information	USA	X		
E13	Professor in the fields of leadership and organization	Germany		X	
E14	Professor for the field of software development and collaboration engineering	USA	X	X	
E15	Researcher in the field of AI ethics and Information and media ethics	Germany	X		X

conducted virtually via Zoom in most cases, and one interview took place on the telephone. 14 of the interviews were each conducted in one piece whereas interview 15 was conducted on two dates a few weeks apart due to the interview partner's schedule. Expert 6 was the only one to receive the interview guide questions in advance of the interview upon request. Audio recordings of the interviews were made and transcribed later. The duration of the interviews varied from 24 to 97 minutes whereby the average duration of an interview was 51 minutes. The high variance in the length of the interviews is on the one hand based on the comprehensiveness with which the respective experts answered the questions, and on the other hand on the inclusion of their own experiences and additional aspects.

## 2.2. Qualitative content analysis

After the interviews, the audio recordings were transcribed on the basis of transcription rules by Rädiker and Kuckartz (2019). Subsequently, we conducted a qualitative content analysis based on Mayring (2015) with the transcripts. In order to be able to systematically analyze the expert interviews and extract statements relevant to answering the RQs, we coded and categorized the content of the transcriptions. The aim of this procedure is to establish a category system by means of which the experts' statements can be systematically analyzed. Only statements that are relevant for answering the RQs are included in the categories. We used MAXQDA<sup>1</sup> as a coding for the analysis. The transcriptions of the 15 expert interviews served as communication material.

In the first step of the qualitative content analysis, we determined the units of analysis. These included individual words and phrases, as well as individual and consecutive sentences if they belong together in terms of content. In the second step, the main categories of the category system were determined. The main categories were deductively derived based on the 18 theory-driven questions of the interview guideline. The first two main questions were combined into one category, so that the category system contains 17 main categories. This resulted in the first version of the category

system, which is the third step. The main categories were defined in the fourth step and assigned anchor examples. In addition, coding rules were defined, for example, that units of analysis that address themes from several categories can be assigned multiple times. We calculated an interrater agreement by using a built-in function of MAXQDA. For this, three of the transcribed interview materials were used as a sample that was coded by a second coder to calculate an intercoder reliability according to Cohen's kappa (Cohen 1960). With a code overlap at segments of at least 80%, the intercoder reliability was 60.54%, which corresponds to a kappa value of 0.605. According to Landis and Koch (1977), this can be considered "good". This suggests that the coding system is perceived and interpreted in the same way by different people, which illustrates its validity. In the fifth step, the units of analysis were assigned to categories. When relevant units of analysis addressed topics that were not covered by existing categories, new categories were formed in the sixth step, which led to a partial repetition of some previous steps. Thus, the category system created in the third step was supplemented by the respective new category. The corresponding categories were also defined and assigned anchor examples (step 4). A total of 1013 analysis units were assigned to the categories, with units per interview ranging from 33 to 103. On average, 68 analysis units were assigned per transcription. The variance can thereby be explained in the length and the relevance of the statements related to the topic. In the seventh step, the main statements of each category were summarized and serve as a basis for further analysis, interpretation and answering the RQs in the eighth step.

## 3. Results

The chapter summarizes the results from the interviews, which have been deductively assigned to the five ethical principles of Floridi et al. (2018). In each result category, the respective categories are first defined, then the experts' statements are thematically summarized, sorted and underpinned with anchor examples.

### 3.1. Category 1: beneficence

Beneficence deals with the conditions that need to be in place for well-being, common good, and dignity to exist within team collaboration. In the course of this, the stresses and social risks that can arise from the introduction of a CA within a virtual team are discussed and the possibilities for counteracting them.

The subcategory burdens deals with affective conflicts of team members that can arise when introducing CAs in teams. Reasons for the emergence of burdens can be dysbalances that arise from the disruption of existing team structures (E9). These disruptive changes cause teams to split into those members who adapt to the change and those who show skepticism and an unwillingness to embrace the introduction of technology (E6–E7, E9). In addition, adaptation difficulties can occur as the introduction of CAs does not make it clear to users who is responsible (E10) and what exactly the technology is to be used for (E11). Another inhibiting factor that was mentioned is the different speed with which users learn to use CAs (E12). Expert 9 describes this as follows: "Most of the time there is a kind of architecture that has been set up, so that there could be dysbalances or different opinions about how good or bad such an assignment is, or willingness to get involved in something like that." In addition, data protection burdens become relevant when using CAs. For example, uncertainty about the CA's level of knowledge and information can be a burden for team members (E1). Other fears are that working hours (E4) and data (E4, E7, E15) will be monitored through the use of CAs. Expert 14 summarizes this as follows: "It's nice if people see my outcomes, but they don't have to see exactly how I get to my outcomes. (...) So I think transparency and performance evaluation are the biggest fears or areas of discomfort that people will experience." One way to prevent the burdens is to clearly delineate what the CA is used for and what it is not used for, thus counteracting adaptation difficulties (E10–E11). However, concrete preventive measures could not be determined at this stage (E4) and would depend on the context of use (E6). Team members do not have "to change their way of collaboration because of chatbots, but that could of course change as CAs get more advanced or get different capabilities" (E5).

One social risk that may arise when team members adapt to the CA is that the transition to the new technology may be more difficult for older individuals than for younger ones, and they may be more at risk of being isolated from the team as a result (E12). There is also a risk of sub-group formation (E13), for example into technology-enthusiastic and -rejecting users (E12). Expert 12 summarizes this as follows: "Certainly for the old people who are resistant to change. (...) You're 60 years old and you didn't grow up with the Internet and all of a sudden you're expected to speak to this agent (...) So that's going to be hard." Impairments of social interaction can occur as the use of CAs, depending on whether the CA replaces human team members (E14), promotes the isolation of team members (E1, E2), less interaction (E1), and small talk (E2) takes place between team members and team communication is

negatively affected by adaptation difficulties (E10). Other experts are of the opinion that the CA cannot replace tasks with a lot of social interaction and therefore there would be no risk through isolation (E4, E6, E9, E15). An anchor examples come from expert 2: "I don't think we should build more technical barriers between people." (E2)

### 3.2. Category 2: justice, solidarity, and workspace awareness

The category justice addresses aspects that must be in place to promote prosperity and justice and to maintain solidarity. For this reason, the following identifies which threats to solidarity can occur within a team using a CA. Some experts believe that the introduction of a CA into a team is not a threat to solidarity (E2, E7) if the CA only acts as a machine teammate and control remains with the human team members (E6). A sense of belonging to the team would not be easily destroyed by a CA and whether the team culture works or not is not dependent on the CA (E15). A CA could also have a positive effect on the team (E8). However, threats could occur if team members get along better with the CA than with their human team members (E9) and if a team member loses his or her value within a team if the CA is asked for advice instead of him/her (E3). Moreover, since the CA has an extroverted personality due to his ability to answer questions directly, introverts might be increasingly excluded from the team and a split of the team into introverted and extroverted team members might take place (E6). An anchor example comes from Expert 3: "You start thinking things like: 'I'm not gonna bother [my team colleague] for the fifth time in a day, I shouldn't.' 'I just ask the bot.' The person loses their place in the team. They lose their value." In addition, the focus of the CA is more on efficiency than on social aspects and thus allows less spontaneity, brainstorming and collaboration (E1). In addition, creativity is lost when working with the CA and no social connections characterized by politeness can develop, as the CA is merely given orders (E1). Threats are also possible if an information vacuum is created for team members due to a lack of clarification of responsibility, scope of action and functionality (E10). To counter threats to solidarity from the use of CAs, autonomy must remain with team members and democratization must take place, giving team members a say and the opportunity to co-design the Ca (E6). In addition, the CA should only be used for routine decision-making and not yet for management and leadership problems (E1). In addition, it is important to create transparency about support areas, data processing and data protection (E11) as well as to clarify responsibility and scope for action (E10–E11) and functionality (E10). The anchor example for this sub-category comes from expert 6: "Employees simply want to have more autonomy, to have flexibility. And accordingly, I can imagine that everything that comes to machine teammates, that the team (...) actually wants to have a lot to say about it, and maybe even to the extent that each team can design their team member as they want."

### 3.3. Category 3: non-maleficence and privacy

The category of non-maleficence addresses data protection and security in the use of CAs in virtual collaboration as well as the avoidance of violations of personal privacy. The claim of non-maleficence is to do nothing harmful and to avoid misuse. According to the experts, data protection and privacy have a high priority (E2, E13), but the protection in organizations (E6) and the existing GDPR regulations (E3) are not sufficient. Besides, regulations cannot be set across the board for all areas but must be decided on the basis of the respective use cases (E9, E12).

Abuse of users occurs when their data or systems are misused (E2, E4), hacked (E2, E8), manipulated (E4), tracked (E10), or spied on (E15). Further possibilities of misuse are the violation of privacy (12, 15), identity theft (E12) and profiling, e.g., regarding the working hours of the users (E2, E11). Collecting data that has nothing to do with teamwork and office work but is related to work is also problematic (E14). The CA can be a source of misuse of user data if the data are used to manipulate team members (E7) or to outmanoeuvre other team members using the data (E4), as well as if it provides verification of user absence from work during working hours (E11). The anchor example for this area comes from expert 14: "Well, it becomes abused the moment it falls outside of what has been approved or what is acceptable or what is unethical."

Concrete expert's recommendations for preventing misuse through the use of a CA depend on the respective situation (E1) and the type of organization (E5). However, existing data protection rules must be taken into account (E2, E4), respectively, existing rules must be applied and adapted to the CA (E1). In addition, there must be laws on the types of data that employers can collect or use (E12). It must be ensured that data are not misused inside or outside a company (E11). Data must be encrypted (E2) and protected (E12) and personal data such as names and addresses must be automatically masked (E5). There must also be clarification about who has access to data and who should have it (E8, E11), as well as transparency and communication about where data go, how they are collected, processed, and stored, and what is done with it (E2, E5, E6, E10, E11, E14). In addition, consideration must be given to whether the use of the CA poses risks of violating the privacy of users (E5) and that a company may not transfer certain knowledge to the CA (E8). An anchor example for this area comes from expert 6: "But at the same time, we know that organizations often don't have the capacity or the knowledge to ensure real security for their data. And maybe it's safer to store all that with Microsoft or with Google. But is that what the average employee wants?" (E6).

### 3.4. Category 4: explicability

The category explicability enables the fulfillment of the other principles through comprehensibility and traceability as well as transparency and interpretability. One challenge in the context of transparency is that users are numb to information boxes on data processing and often accept them without

understanding or questioning them (E4). The information regarding data processing may also be less transparent and more difficult to access with a CA than with a textual agent, where the information is visually available (E5). Besides it has to be transparency about exactly what data are collected (E14), how it is encrypted (E4), what purpose it serves (E2, E4, E9, E14), what it does exactly (E7, E11), when it is deleted (E4), how secure the tool and data used are (E6) and whether biases can occur (E6). A company must clearly communicate which data are used and which are not, as the user of the CA is not obliged to provide information (E10). Depending on the purpose of the data, the user may or may not consent to the use of the same data (E14). It is also necessary to make it transparent to the user that s/he is not talking to a real person but to a CA (E1, E4–E5, E7, E9). The anchor example for these aspects come from expert 14: "It has to be transparent to me what data is collected. And it has to be transparent to me where the data is going to be used for what reason. Because based on that second part I may agree or disagree with it for the same data" (E14).

Regarding the development and use of CAs, the experts recommend complying with the existing GDPR guidelines (E4) and using systems that are already certified (E7). In addition, ethical training for the developers of CAs makes sense (E6). Ethical aspects must already be questioned in the design of CAs and not be acted upon according to what is best for the company that wants to use the CA (E6). The principle of ethics by design must be fulfilled here (E6). It must be transparent why the CA is asking the user certain questions and making certain recommendations (E5). The explanations must be designed in such a way that everyone can understand them immediately (E11). More detailed explanations about the use of the CA should then be offered, adapted to the respective target groups (E11). The explanations should be such that all users are able to make well-founded decisions with the information provided (E11). Furthermore, it is important for the users of CAs to establish a corporate culture and a basis of trust with the company, as they have to rely on the information provided by the company regarding transparency aspects (E15). The anchor example for this sub-category comes from expert 15: "The transparency that a company can provide is of course always one that is based on trust, because it is quite clear that I cannot check whether it is not being stored and where it is being stored and by whom it is being processed. And I think it is important to establish a certain corporate culture on both sides."

### 3.5. Category 5: accountability/autonomy

In the category, it is considered how the aspect of accountability is to be handled in the team and who is considered responsible for the actions of a CA. Furthermore, autonomy in the cooperation with the CA is addressed as well as a possible dependence on technology. There is a diffusion of responsibility as to whether a newly introduced CA equivalent to a new human team member is as responsible for his/her own actions as the team member (E10) and who is liable for the CA's statements (E1–E2). The allocation of responsibility is an

important question (E5) and difficult decision (E6–E7), which is still in process and not clarified (E1) and is an ethical issue (E8). Another expert is of the opinion that the current legal situation is sufficient for the use of CAs (E10). The question of who is responsible for the decisions and possible wrong decisions of an agent is not clear (E4), unclear (E2) and not finally clarified (E11). However, a corresponding clarification and legal situation is urgently needed (E11). However, a CA must always remain an assistant and should not be allowed to make decisions (E3). In the area of human resources, however, a CA can make more objective decisions than a human being, since no evaluation takes place according to sympathy (E11). An anchor example comes from expert 3: "We've got to allow those conversational agents to be assistive, but not to be the decision makers. That's where I draw the line, so the human makes the decision with the assistance of the conversational robots."

The experts saw different instances being responsible for the actions of a CA. The system itself is not seen as responsible for itself (E10). If no programming errors were committed (E15), the developers and programmers of a CA are also generally not seen as responsible (E3, E6). However, if something is intentionally unethical by design, the blame and responsibility lie with the programmer (E4). Usually, the producers of a system are considered responsible according to our interviewees (E8, E10, E12), respectively, those who earn the most money from using the system (E12). In addition, especially those who have decided to use the system in the company (E2, E7, E10, E15) and implement the output of the system (E10), i.e., who are at the top of the company hierarchy (E6), should be held responsible. As a rule, these are the leaders of companies (E2, E8), respectively managers (E7), and top managers (E7). The user himself would not bear any responsibility *per se* (E6). This would change, however, if a user does not take into account crucial information that was available to him/her through the CA, resulting in severe consequences (E14). In addition, it would be important to ensure that a user should not pass on responsibility to the company in the event of deliberate misconduct or careless handling when interacting with a CA (E10, E14). Furthermore, if a CA only has a support function in decision-making, the user bears the responsibility (E7). The anchor example comes from expert 7: "You always have to consider who ultimately makes the decision and (...) to what extent it is supported by the system. As long as people ultimately make the decision, I think they should also be responsible. If systems make decisions, then those who made the decision to use these systems are responsible. So I think (...) it's not the one who programs, because he is not responsible that somewhere a system is used the way it is used" (E7).

Further examples for the ethical principles, key themes, and illustrative passages can be found in [Table 2](#).

## 4. Discussion

### 4.1. Ethical guidelines and an agenda for future research

Our experts raised concerns and made suggestions for each of the five ethics principles of beneficence, non-maleficence,

autonomy, justice, and explicability when CAs are introduced to virtual collaboration at the workplace. In contrast to Floridi et al. (2018), with virtual collaboration at the workplace, we focused on one specific domain with specific challenges that even increase in importance. We found differences and similarities to Richards et al. (2023), which focused on more general moral issues in the context of human-agent interaction.

We found one special aspect for CAs being introduced in virtual collaboration at the workplace for the principle of beneficence. There was disagreement among experts as to whether the introduction and use of a CA in a virtual team context leads to increased isolation of team members. While some experts suggest that fewer interactions among team members and less small talk are the triggers for increased isolation (E1–E2), other experts believe that team members are not at risk of isolation through the use of CAs (E4, E6, E9, E15). Research is also not unanimous on this. Some scholars suggest that especially the virtual nature of collaboration and lack of social partnerships (Niederfranke & Drewes, 2017) can lead to psychosomatic complaints (Charalampous et al., 2019) and professional and social isolation (Cascio, 2000; Charalampous et al., 2019). Other research demonstrated how CA can positively impact employees' psychological wellbeing through the interaction (Aymerich-Franch & Ferrer, 2022). One study found an effect of CA interaction on a decreased social identity within a team (Mirbabaie et al., 2021). Experts suggest that especially the different adaptation of users to a CA can lead to subgrouping and the exclusion of older or more technology-shy users (E12–E13). The literature argues that measures need to be taken to prevent work-related impairments such as social isolation (Lengen et al., 2021). One approach of applying this could be that designers are trained in human-centered design principles (Ozmen Garibay et al., 2023). According to both the literature and the experts, these measures would have to be taken by the managing authorities of the virtual teams (Lengen et al., 2021). There also was disagreement among experts about the relevance of social interaction among teammates. While some experts believe that polite, social interaction, and face-to-face interactions are important and should not be replaced by the use of a CA (E1–E2, E7), another expert believes that a restriction on social interaction should be made if it limits the efficiency of teamwork (E9). However, there is a consensus in the literature that maintaining group cohesion and communication within a team is hugely important for the interpersonal character and success of a team (Kauffeld et al., 2016; Konradt & Hertel, 2007). The literature shows that team members should always have the opportunity to interact face-to-face and form informal contacts and networks (Pyöriä; 2009) and the future of AI in knowledge work should focus on collaborative approaches (Sowa et al., 2021). According to the literature, if such cohesion is not present, familiarity in the team may be lost (Kauffeld et al., 2016).

According to the experts, trust within a team is very relevant for the psychological security of the team members (E13). This aspect is supported by findings from the

**Table 2.** Key themes and illustrative passages from the interviews.

Ethical principle	Themes	Illustrative passages from the interviews
Beneficence	Social risks	"I can see that that would happen if people work with AI instead of with other humans. So, if I have another counterpart that I work with all the time, but that counterpart is replaced by an AI (...)." (E14)
	Prevent social risks	"No, I don't, because where there is technology to be abused, it will be abused." (E3)
Justice, solidarity, and workspace awareness	Threats to solidarity	"So, if performance goals remain team goals, then it's much more likely that they're also will develop a level of team coherence and identification with the team versus if your objectives are purely individual, then you stop caring more about the team. So that is a design choice that organizational designers have to make." (E14)
	Information deficits	"But we know that the functions, i.e., the functional component of cooperation, is always strengthened where the personal, the social component is also strong, and there is of course a clear connection. The more distributed, the more difficult." (E13)
Non-maleficence and privacy	Workspace awareness	"So, I don't think we should put up more technical barriers between people." (E2)
	Types of abuse	"So I think there's, as with all software and as with all data that we collect, there's an infinite range of possibilities for mischief." (E4)
Explicability	Burdens	"So, I think transparency and performance evaluation are the biggest fears or areas of discomfort that people will experience." (E14)
Accountability/autonomy	Responsibility	"When systems make decisions, then those who have made the decision to use these systems are responsible. So, I don't think that it is, except in the case of blatant programming errors or deliberate programming errors, but I think it is not the one who programs, because he is not responsible for a system being used somewhere in the way it is used. It's the person who makes the decision to use it." (E7)
	Avoid dependencies	"So explainability in AI will be important to avoid dependency." (E5) "I think it's a bit of give and take. So, you take away a bit of autonomy when it comes to performing a certain task, but then you get additional autonomy again through another task because you can now make a better decision or because you now have more information, you have better security, things like that." (E6) "And I don't know if it should be avoided, quite honestly, this dependence of technology. Or if it's just the next step in what we're doing anyway (...) how we're just evolving." (E6) "I think technology, and especially technology that works, always makes us dependent in a certain way." (E9) "This complaining about the loss of true and correct and good and canonical knowledge is as old as humanity itself. So, I think that's pointless. I think we have to look at what exactly is being lost and then consider: Is that something that affects our autonomy in a very special way? And then we have to see whether we can perhaps substitute it in some other way, or possibly find another solution somehow." (E15)

literature. A trusting environment for dealing with the CA must also be created (Hofeditz et al., 2021; Strohmman et al., 2018). However, experts argue that building trust is significantly complicated by the virtual nature of collaboration and that, in addition, increasing trust in a CA's abilities could bring about a reduction in trust in team members' abilities as less reliance is placed on the expertise of team members (E6). The literature also shows that it is not possible to predict with certainty what impact a CA may have on a team, as this differs depending on the use of the CA and the existing team climate (Seeber et al., 2020). From the above aspects, we derived the following RQs and ethical guidelines (EGs):

RQ1: How can the maintenance of social interaction, trust and well-being within a virtual team using a CA be ensured?

EG1: Virtual team managers must ensure that all team members have a similar level of knowledge and starting point when adapting to a CA.

EG2: Ethical problems must be identified at the concrete use case and ethical guidelines have to be developed application-specific considering the needs of the users of the CA at the workplace.

There was a consensus among experts that data protection and privacy have a high priority (E2, E13). They argue that it need to be ensured that there are no conflicting interests between the organizations and individual employees' interests. When it comes to methods to prevent and mitigate misuse, experts' opinions differ on the extent to which regulations already in place are sufficient to prevent misuse of user data. This results in two different views: some experts are of the opinion that regulations to protect against data misuse within companies, respectively, GDPR regulations, are sufficiently extensive and that no further regulations are necessary (2, E4–E6, E9–E10). In the literature, compliance with the principles of data protection (Henderson et al., 2018; Kaul, 2021) and privacy (Kaul, 2021; Rothenberger et al., 2019) is considered as an essential challenge that is ethically relevant (Ozmen Garibay et al., 2023; Stephanidis et al., 2019). Other experts believe that this is not sufficient and that existing policies and guidelines need to be expanded, adapted and detailed to the use of a CA (E1, E3). In the literature, Floridi et al. (2018) recommend that an assessment of existing regulations needs to take place to see if their ethical principles are sufficient to create a legal

framework that makes sense for the technology being used. Expert 6 raises the point that while principles exist, it will take years to apply them meaningfully in practice. According to the experts, an adaptation of these guidelines is necessary, as the recording of the voice of the user of a CA collects biometric, sensitive data (E15), which cannot be easily anonymized and these aspects are not sufficiently taken into account by current recommendations and regulations. According to the experts, it should be taken into account above all that users have a right to have their data forgotten (E7, E15). It has also been shown in the literature that protecting privacy when dealing with AI systems is a challenge that needs to be countered with application-specific analyses (Rothenberger et al., 2019). Since no clear direction on how to deal with abuse by a CA in a team context can be drawn from the findings in the literature and the interviews with the experts, we formulated the following RQ and ethical guidelines for using CAs at the workplace:

RQ2: What adjustments to existing privacy policies need to be made to counteract data misuse through the use of a CA in virtual teamwork?

EG3: Organizational CAs must not store user data and personal information that is not directly related to virtual team and office work or essential for the performance of work tasks.

EG4: It must be clear to users at all times which of their data is collected by the CA, when, by whom, how and for what purpose.

The issue of autonomy is defined as an important ethical principle in the context of AI and CAs in the literature (AI HLEG, 2019), but it can also pose ethical challenges (Rothenberger et al., 2019). In this context, our experts also address how autonomous a team can remain when a CA is introduced. The experts' responses can be interpreted to mean that the autonomy of the team depends on which task area the CA takes on. If the CA controls conversations, interactions and discussion, the team is more likely to lose autonomy to the CA (cf., E6). Assisting tasks, in contrast, are perceived as less detrimental to autonomy if the final decision remains with the user (E15). There is thus agreement among the experts and in the literature that the final control over the user-CA interaction must remain with the user in order to avoid a restriction of autonomy. Four of the experts agree that CA work can involve a dependency on technology (E2, E9, E11, E12). However, there is a disagreement on whether such dependency has negative effects and should be avoided. While two experts argue, on the one hand, a dependency on CA can lead to a loss of users' skills and a loss of learning to solve problems (E3, E10), on the other hand, it is discussed that users can gain capacities for other tasks (E6, E9) and that this can lead to other forms of independence and new possibilities for action (E9). This is supported by a recent study that found that virtual collaboration with CAs like ChatGPT generally has a positive impact on productivity and complements rather than replaces workers' skills (Noy & Zhang, 2023). The study also found that virtual collaboration with the CA could decrease inequality between employees by benefiting the productivity of employees with lower abilities. Therefore, it seems that

deploying CAs in virtual collaboration at the workplace can mitigate ethical issues such as inequality while it, at the same time, can positively impact productivity of employees.

In our study, there was disagreement on possible avoidances of dependence on CA. According to Hendrickx et al. (2021), users must always have control and transparency over their own data. One expert also argues that control and autonomy must be given back to users through the nature of CA design (E6) and that decisions must be made at most with the support of a CA, but not by the CA. Sankaran et al. (2020) further emphasize the need to consider human autonomy as an important component of the user experience. They also suggested that increasing independence of CAs may lead to a perceived threat to users' human autonomy and that there is a need to explore how much control and independence may be handed over to CAs. It is also important to distinguish between autonomy that is given to the CA by design and by the user. Previous research raised concerns that users might voluntarily give up their autonomy to a CA and allow it to respond on the user's behalf (Richards et al., 2023). They suggested to ensure that a CA should never transform from a consultant to a decision maker. In our study, the experts also identify this as an area where further consideration is needed. They argue that it is necessary to look at which user capabilities are limited by the increasing independence of the CA and then weigh up whether these aspects limit the autonomy of the team in a particular way (cf., E15). We propose the following RQ and guidelines to better address these aspects:

RQ3: In which areas of virtual team collaboration is it necessary to limit the independence of the CA from the team in order not to endanger the autonomy of the team and the individual team members?

EG5: The highest level of control and autonomy must not be held by the CA, but by the user or the team.

EG6: It always need to be ensured that the final decision is made by a human team member and not by a CA.

There is disagreement among the experts as to whether a CA should be given decision-making power. In addition, there is disagreement about which authority bears responsibility for decisions, actions, and consequences made under the influence or with the help of a CA. The literature also highlights responsibility, accountability, and responsibility as relevant ethical principles (AI HLEG, 2019; Jobin et al., 2019; Kaul, 2021; Mateescu & Nguyen, 2019; Shahriari & Shahriari, 2017). However, uncertainties and inconsistencies currently exist regarding the interpretation and prioritization of these principles (Jobin et al., 2019). Rothenberger et al. (2019) identify the question of who is responsible for the actions and consequences of an AI system as not yet finally clarified. They name the manufacturers of the respective technologies, the developers or programmers, or the users as possible responsible entities. The experts take a different view: The experts agree that users and programmers are not responsible for the actions and consequences of the CA as long as they have not deliberately misbehaved and disregarded information or committed programming errors respectively (E3, E6, E10, E14, E15). Moreover, they agree

that the CA is not responsible for its own actions. However, some experts also see the producers of CAs as responsible (E8, E10, E12). The majority of experts, however, see the companies or managers who have decided to use the CA and benefit from its output as responsible (E2, E7–E8, E10, E12). However, a clear answer cannot be determined. However, Rothenberger et al. (2019) emphasize that the existence of a person responsible for the actions and consequences of an AI system is necessary. Floridi et al. (2018) also emphasize the risk of not having a human being declared responsible for the outcomes of an AI. Zou and Schiebinger (2018) further emphasize that there must be a responsible person who cannot escape accountability. Based on these aspects, the following RQ and ethical guidelines arise:

RQ4: Which entity can be held accountable for the actions, decisions and consequences that result from a CA or with the help of a CA in virtual team collaboration?

EG7: The organization which deploys a CA in virtual collaboration should always ensure that there is a visible administrator or product owner that is accountable for the CAs outputs and functions.

EG8: It need to be ensured that there is at least one programmer and designer of a CA which cannot escape accountability to secure that errors causing ethical challenges can be fixed quickly.

In the following, we will discuss which aspects of virtual teamwork with a CA can endanger solidarity in the team and what influence trust within a team has on the team members. While some experts see no threat to solidarity through the use of a CA in the context of virtual teamwork (E2, E7), Luengo-Oroz (2019) argues that the ethical principle of solidarity receives little attention in existing ethical AI guidelines and often goes unmentioned. Solidarity and justice or fairness are mentioned as ethical principles in the literature (AI HLEG, 2019; Jobin et al., 2019; Kaul, 2021), but only in a small percentage of articles (cf., Luengo-Oroz, 2019). While experts consider that CAs at most foster a toxic team climate but do not trigger it (4), the literature argues that aspects harmful to solidarity, such as prejudice and bias, also occur in conversational interfaces. This necessitates an assessment of possible threats. According to Luengo-Oroz (2019), potential risks and harms of AI need to be assessed before new technologies are deployed. There is also disagreement among experts as to whether trust can be built in virtual teams to the same extent as in analogue teams. While some experts see no challenges in this (E1, E12), others argue that implicit information and physical presence of team members are necessary for building trust in the team (E2, E4, E8–E11). There is a need to identify how best to address the problem of trust building in virtual teams (cf., Hossain & Wigand, 2006) as trust is one highly important component in successful human-agent interaction (Hofeditz, Nissen, et al., 2022). However, no proposals are mentioned to combat specific threats to solidarity. For this reason, we propose the following RQ and guidelines:

RQ5: How can the building of trust within virtual teams using a CA be strengthened?

EG9: Virtual team leaders need to regularly check if there are threats to solidarity within their team due to the use of a CA.

EG10: The CA should regularly suggest in-person meetings of the team members to strengthen trust among the team members.

In the following, the results from the expert interviews and findings from the literature are discussed in order to identify which specific types of understandability, interpretability, and comprehensibility must be given in order to fulfill the principle of explicability. Transparency is considered particularly ethically relevant in the literature and represents one of the main requirements for trusting systems (AI HLEG, 2019; Kaul, 2021; Rothenberger et al., 2019; Ruane et al., 2019; Sağlam & Nurse, 2020). Accordingly, creating a trusting and transparent environment for users of CAs is a high priority (Strohmann et al., 2018). This relevance of a basis of trust between companies and users of CAs is supported by the experts, as users need to rely on the company's statements regarding transparency aspects (E15). On the one hand, previous literature argues that those individuals or entities in companies responsible for data processing should be open and honest about how they process users' personal data (Sağlam & Nurse, 2020). Frick et al. (2019), on the other hand, argue that the technology itself must be designed to be transparent about what personal data are processed, when how and where. The experts expand on these aspects by adding that it must also be made clear who has access to user data (E8, E11), how specific what purpose it serves (E2, E4, E9, E14), how it is encrypted (E4) and when it is deleted (E4), and what data are explicitly not collected (E10). The experts see the companies as the responsible authority for providing information about these aspects. According to Rheu et al. (2021), this process should already start with the designers of the CAs that need to consider the ethical consequences of each design element and take responsibility. The users of the CAs or the members of the virtual teams respectively are not themselves obliged to provide information (E10). The aspects on necessary characteristics of transparency were also elaborated by Sağlam and Nurse (2020) and Frick et al. (2019) and can be summarized and extended by the experts' statements to the following RQ and guidelines:

RQ6: How can the processing of personal data of team members be explained to the users in order to increase transparency?

EG11: Companies must always make transparent, in a way that is appropriate for the team members, which data is collected when, by whom and for what purpose, and how CA decisions are reached.

EG12: Individuals in organizations which are responsible for data processing and data protection need to be trained in understanding which data is collected via the CA.

In their study, Rothenberger et al. (2019) elaborated on the need for informedness in certain situations about users interacting with an AI instead of a real human. This aspect is also elaborated by several experts. They emphasize the need for users of CAs to always be aware and made transparent that they are interacting with a technology and not a

human (E1, E4–E5, E7, E9) and that identification and clarification are necessary in this area (E1). However, unlike Rothenberger et al. (2019), the experts emphasize that informedness in this regard is not only required in certain situations, but should always be present.

According to the literature on virtual collaboration, explainability within the work of virtual teams is more difficult to implement than in face-to-face teams due to spatial, temporal, and organizational boundaries (Kauffeld et al., 2016). However, there is disagreement among experts regarding this view. While some experts are of the opinion that the virtual nature of collaboration makes it more difficult to perceive subtexts and implicit information of team members (E2, E4, E13), others argue that given information is permanently available (E7). In questions of explainability, this information could then be consulted again. The experts continue to raise the issue that users have become numb to existing ways of being informed about data processing, such as information boxes, and accept them without engaging with them (E4). This challenge is also addressed by Hendrickx et al. (2021), who find that current consent forms are not sufficient to address users' privacy concerns, as users have no possibility to influence them afterwards. The experts are also of the opinion that a one-time consent to data processing is not sufficient and that there should be regular updates, information and reminders about what exactly is collected by the CA (E3, E6). In this regard, they raise the question of how often relevant aspects should be explained and made transparent, and in what way they should be brought to the attention of users (E4). These aspects lead to the following RQ and ethical guidelines:

RQ7: How regularly and by what means should users of CAs in virtual collaboration be informed about the use of their data in order to ensure comprehensive informedness?

EG13: Users must be always given transparency that they are interacting with a CA instead of a human interaction partner.

EG14: There need to be an informed consent at multiple times clarifying the use of employees' data through the interaction with the CA in virtual collaboration.

## 4.2. Implications

The first RQ addresses the ethical challenges workers face when using CAs for digital collaboration in the workplace. In the context of this article, a variety of challenges were identified from the literature and contextualized with results from the expert interviews. There was a consensus in the literature and among the experts regarding key ethical challenges. Therefore, the first RQ can be answered in the sense that employees are confronted with challenges of adaptation to the technology when using CAs in the work context, as well as comprehensive and transparent information about data used and data protection, and questions of the distribution of control and decision-making between user, team, and CA. Other challenges may include limitations on solidarity as well as user wellbeing and trust building. To answer the second RQ, which addresses how ethical

challenges can be addressed by companies and virtual team leaders, 14 ethical guidelines for team leaders and companies were identified based on the expert interviews and literature. The guidelines, based on the five ethical principles of Floridi et al. (2018) for an ethical use of AI, form practice-related implications for the ethical use of CAs in team collaboration.

The guidelines G1, G2, G5, and G6 are directly addressed to the leaders of virtual teams, since a deeper knowledge of the respective team members is necessary for the implementation of these guidelines, respectively, the team leaders must have a certain closeness and possibility to control the team. G3, G4, G7, and G8 are practical guidelines for the use of CAs in a business context. They are addressed to companies, as the implementation of the respective guidelines is within the authority of the respective company regulations. As the guidelines are ultimately aimed at preventing ethical problems within teamwork, this concerns the users of CAs to a large extent. In this context, the results can be relevant for workers in the gig economy and remote workers, too, if CAs are used to support the mostly virtual work here. The guidelines can also be relevant for developers of CAs if the guidelines presented are already taken into account during development. The guidelines should be implemented by companies and team leaders, in order to ensure an ethical use of CAs in the team and to prevent ethical problems, for example, to improve the performance and productivity of the team (cf., Frick et al., 2019) and to reduce the cognitive burden on team members (cf., Brachten et al., 2020). Where no consensus could be found in the discussion of literature and the experts' statements, questions were raised or open questions from the research could not be answered, RQs were set up. Seven RQs emerge that address relevant open questions regarding the use of CAs in a team context, oriented toward the five ethical principles of Floridi et al. (2018). These are aimed at researchers who wish to deepen research into the use of CAs in the virtual workplace and answer open RQs. Further research is a necessity in that not all ethical challenges that may arise in the use of CAs in the virtual team have been comprehensively addressed. The increasing relevance of the topic calls for further research to promote the ethical use of CAs in the team. The RQs developed add value to further research in that they highlight unresolved ethical issues that are considered particularly important and worthy of clarification by both the literature and renowned experts.

## 4.3. Limitations and future research

This research comes with some limitations that provide opportunities for future research. The perspective of additional experts would be necessary to validate the developed guidelines. We mainly focused on a European perspective to ensure a similar normative ethical and work cultural background (de Melo & Terada, 2019). Although we suggest that future research should address our research agenda to derive more knowledge on this topic, another cultural background may require different guidelines and future RQ. In this

**Table 3.** Summary of key future research.

Areas/ethical principles	Possible research questions	Research directions
<i>Beneficence</i>	RQ1: How can the maintenance of social interaction, trust, and well-being within a virtual team using a CA be ensured?	Investigating the impact of different communication modalities (text-based, voice-based, avatar-based) facilitated by CAs on social interaction within virtual teams. Designing interventions or guidelines to promote trust-building activities and foster team cohesion in CA-mediated virtual teamwork. Exploring the influence of team diversity (cultural, linguistic, etc.) on social interaction and well-being in virtual teams utilizing CAs
<i>Non-maleficence and privacy</i>	RQ2: What adjustments to existing privacy policies need to be made to counteract data misuse through the use of a CA in virtual teamwork?	Examining the legal and ethical implications of data privacy in the context of CAs used in virtual teamwork. Designing robust security measures and encryption techniques to safeguard personal data transmitted and stored during virtual team collaboration. Investigating the impact of privacy policy transparency and user control mechanisms on users' trust and willingness to share data with CAs.
<i>Accountability/autonomy</i>	RQ3: In which areas of virtual team collaboration is it necessary to limit the independence of the CA from the team to avoid endangering the autonomy of the team and individual team members?  RQ4: Which entity can be held accountable for the actions, decisions, and consequences resulting from a CA or with the help of a CA in virtual team collaboration?	Investigating the optimal balance between autonomous decision-making by CAs and team decision-making in different stages of virtual team collaboration. Designing adaptive CA interfaces that allow team members to customize and influence the decision-making processes of CAs. Exploring the impact of CA transparency and explainability on team members' acceptance and trust in CA-generated decisions. Examining the legal frameworks and liability regulations surrounding CA usage in virtual team collaboration. Investigating the roles and responsibilities of stakeholders (organizations, developers, team members) in assuming accountability for CA actions and decisions. Designing governance models and contractual agreements to allocate responsibility and mitigate potential risks in CA-mediated virtual teamwork.
<i>Explicability</i>	RQ5: How can the building of trust within virtual teams using a CA be strengthened?  RQ6: How can the processing of personal data of team members be explained to users to increase transparency?	Investigating the impact of transparency, explainability, and accountability mechanisms in CAs on trust formation within virtual teams. Designing interventions and training programs to enhance team members' trust in CAs and foster positive perceptions of their reliability and competence. Exploring the influence of team communication strategies and team-building activities on trust development in CA-mediated virtual teamwork. Developing clear and user-friendly interfaces that provide detailed information about the data processing practices of CAs in virtual team collaboration. Investigating the effectiveness of different communication methods (visualizations, notifications, privacy dashboards) in conveying data processing information to users. Designing educational materials and privacy literacy programs to enhance users' understanding and awareness of data privacy in CA-mediated virtual teamwork.
<i>Justice, solidarity, and workspace awareness</i>	RQ7: How regularly and by what means should users of CAs in virtual collaboration be informed about the use of their data to ensure comprehensive awareness?	Investigating the optimal frequency and timing of data usage notifications and updates to keep virtual team members informed and aware. Designing user-centered communication channels (email, in-app notifications, interactive interfaces) that effectively convey data usage information to users. Exploring the impact of personalized and customizable data usage notifications on users' comprehension and trust in CA-mediated virtual teamwork.

context, the experience of practitioners could also bring added value. As researchers tend to be more familiar with the implications of guidelines, only the knowledge of researchers was included in the present work. However, validation of the guidelines by developers or employees working in virtual collaboration with and without CAs could be an added benefit. Another limitation is that, overall, there were fewer ethicists among the experts than researchers from the fields of collaboration and computer science. This

was due to the fact that ethicists were less willing to participate in an interview than researchers from the other fields. One of the reasons given for this was that they felt not familiar with the technical factettes of this topic. This demonstrates the problem that some researchers do not work enough in an interdisciplinary and cross-disciplinary way. However, an interdisciplinary perspective is highly valuable for addressing ethical issues that affect such intertwined research areas as virtual collaboration with CAs at the

workplace. Future research could seek further views from experts in the field to validate and, if necessary, extend the present findings.

We summarized our key suggestions for future research in Table 3.

## 5. Conclusions

In summary, this article contributes to research by providing guidelines for the use of CAs in virtual collaboration and an agenda for future research. The ever-growing number of virtual teams in the work context (Morrison-Smith & Ruiz, 2020) and the increasing capabilities of AI-based CAs (Schwartz et al., 2016) highlight the relevance of this. The research agenda identified contributes to future research by providing a concrete framework for in-depth research of existing and relevant ethical challenges in the use of CAs in teams. Guided by the five ethical principles of beneficence, non-maleficence, autonomy, justice, and explicability by Floridi et al. (2018), which were specified for the context of virtual collaboration at the workplace, 14 ethical guidelines for entrepreneurs and team leaders and a research agenda of seven open questions for further research were elaborated. Through the guidelines, the five principles are specified for the use of CAs in virtual team collaboration. The RQs also allow for more in-depth research that is guided by the five principles but specifically addresses unresolved challenges in the use of CAs in teams.

## Note

1. <https://www.maxqda.com/>

## Disclosure statement

There are no competing financial or other interests to declare.

## Funding

We did not receive any funding for this work.

## ORCID

Lennart Hofeditz  <http://orcid.org/0000-0002-5552-9599>

## Data availability statement

A pseudonymized form of the transcripts can be made available to interested parties upon request. Since even after pseudonymization it cannot be guaranteed that the interview partners cannot be identified on the basis of their statements and their job title, we refrain from making the transcripts openly accessible. All other materials, such as the interview guide, are freely accessible in the appendix.

## References

AI HLEG. (2019). *Ethics guidelines for trustworthy AI*. Futurium. <https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html>

- Almeida, P., Santos, C., & Farias, J. S. (2020). Artificial intelligence regulation: A meta-framework for formulation and governance. In *Proceedings of the 53rd Hawaii International Conference on System Sciences* (pp. 5257–5266). Scholarspace. <https://doi.org/10.24251/HICSS.2020.647>
- Aymerich-Franch, L., & Ferrer, I. (2022). Investigating the use of speech-based conversational agents for life coaching. *International Journal of Human-Computer Studies*, 159(March), 102745. <https://doi.org/10.1016/j.ijhcs.2021.102745>
- Benke, I., Knierim, M. T., & Maedche, A. (2020). Chatbot-based emotion management for distributed teams: A participatory design study. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2), 1–30. <https://doi.org/10.1145/3415189>
- Brachten, F., Brünker, F., Frick, N. R. J., Ross, B., & Stieglitz, S. (2020). On the ability of virtual agents to decrease cognitive load: An experimental study. *Information Systems and e-Business Management*, 18(2), 187–207. <https://doi.org/10.1007/s10257-020-00471-7>
- Cascio, W. F. (2000). Managing a virtual workplace. *Academy of Management Perspectives*, 14(3), 81–90. <https://doi.org/10.5465/ame.2000.4468068>
- Charalampous, M., Grant, C. A., Tramontano, C., & Michailidis, E. (2019). Systematically reviewing remote e-workers' well-being at work: A multidimensional approach. *European Journal of Work and Organizational Psychology*, 28(1), 51–73. <https://doi.org/10.1080/1359432X.2018.1541886>
- Chatterjee, S., Sarker, S., & Fuller, M. (2009). A deontological approach to designing ethical collaboration. *Journal of the Association for Information Systems*, 10(3), 138–169. <https://doi.org/10.17705/1jais.00190>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Dastin, J. (2022). *Amazon scraps secret AI recruiting tool that showed bias against women*. *Ethics of data and analytics*. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>
- de Melo, C. M., & Terada, K. (2019). Cooperation with autonomous machines through culture and emotion. *PLOS One*, 14(11), e0224758. <https://doi.org/10.1371/journal.pone.0224758>
- de Vreede, G.-J., & Briggs, R. O. (2005). Collaboration engineering: Designing repeatable processes for high-value collaborative tasks. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences* (p. 17c). Scholarspace. <https://doi.org/10.1109/HICSS.2005.144>
- Elshan, E., & Ebel, P. (2020). Let's team up: Designing conversational agents as teammates. In *Proceedings of the International Conference on Information Systems, ICIS 2020*. AISel.
- European's Commission: High Level Expert Group (HLEG). (2019). *Ethics guidelines for trustworthy AI Shaping Europe's digital future*. <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Frick, N. R. J., Brünker, F., Ross, B., & Stieglitz, S. (2019). Towards successful collaboration: Design guidelines for AI-based services enriching information systems in organisations. In *Proceedings of the 30th Australasian Conference on Information Systems*, arXiv: 1912.01077. arXiv.
- Fylan, F. (2005). Semi-structured interviewing. In J. Miles & P. Gilbert (Eds.), *A handbook of research methods for clinical and health psychology* (pp. 65–78). Oxford University Press.
- Gnewuch, U., Morana, S., Maedche, A. (2018, December). Towards designing cooperative and social conversational agents for customer service. In *Proceedings of the International Conference on Information Systems 2017*. AISel.
- Gutwin, C., & Greenberg, S. (2002). A descriptive framework of workspace awareness for real-time groupware. *Computer Supported*

- Cooperative Work (CSCW), 11(3-4), 411-446. <https://doi.org/10.1023/A:1021271517844>
- Hafez, K. (2002). Journalism ethics revisited: A comparison of ethics codes in Europe, North Africa, the Middle East, and Muslim Asia. *Political Communication*, 19(2), 225-250. <https://doi.org/10.1080/10584600252907461>
- Henderson, P., Sinha, K., Angelard-Gontier, N., Ke, N. R., Fried, G., Lowe, R., & Pineau, J. (2018). Ethical challenges in data-driven dialogue systems. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 123-129). Association for Computing Machinery. <https://doi.org/10.1145/3278721.3278777>
- Hendrickx, I., van Waterschoot, J., Khan, A., ten Bosch, L., Cucchiari, C., & Strik, H. (2021, April 13-17). Take back control: User privacy and transparency concerns in personalized conversational agents. *Joint Proceedings of the ACM IUI 2021 Workshops*, College Station, Texas, USA.
- Hofeditz, L., Clausen, S., Rieß, A., Mirbabaie, M., & Stieglitz, S. (2022). Applying XAI to an AI-based system for candidate management to mitigate bias and discrimination in hiring. *Electronic Markets*, 32(4), 2207-2233. <https://doi.org/10.1007/s12525-022-00600-9>
- Hofeditz, L., Mirbabaie, M., Luther, A., Mauth, R., & Rentemeister, I. (2022). Ethics guidelines for using AI-based algorithms in recruiting: Learnings from a systematic literature review. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS)*. ScholarSpace. <https://doi.org/10.24251/HICSS.2022.018>
- Hofeditz, L., Mirbabaie, M., Stieglitz, S., & Holstein, J. (2021). Do you trust an AI-journalist? A credibility analysis of news content with AI-authorship. In *Proceedings of the European Conference on Information Systems (ECIS)*. AISel.
- Hofeditz, L., Nissen, A., Schütte, R., & Mirbabaie, M. (2022). Trust me, I'm an influencer! A comparison of perceived trust in human and virtual. In *European Conference on Information Systems (ECIS)* (pp. 1-11). AISel.
- Hossain, L., & Wigand, R. T. (2003). Understanding virtual collaboration through structuration. In *Proceedings of the 4th European Conference on Knowledge Management* (pp. 475-484).
- Hossain, L., & Wigand, R. T. (2006). ICT enabled virtual collaboration through trust. *Journal of Computer-Mediated Communication*, 10(1), JCMC1014. <https://doi.org/10.1111/j.1083-6101.2004.tb00233.x>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399. <https://doi.org/10.1038/s42256-019-0088-2>
- Kaiser, R. (2014). *Qualitative experteninterviews*. Springer Fachmedien Wiesbaden.
- Kaplan, A., & Haenlein, M. (2020). Rulers of the world, unite! The challenges and opportunities of artificial intelligence. *Business Horizons*, 63(1), 37-50. <https://doi.org/10.1016/j.bushor.2019.09.003>
- Kauffeld, S., Handke, L., & Straube, J. (2016). Verteilt und doch verbunden: Virtuelle Teamarbeit. *Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie (GIO)*, 47(1), 43-51. <https://doi.org/10.1007/s11612-016-0308-8>
- Kaul, A. (2021). *Virtual assistants and ethical implications*. IntechOpen.
- Koch, M., Schwabe, G., & Briggs, R. O. (2015). CSCW and social computing. *Business & Information Systems Engineering*, 57(3), 149-153. <https://doi.org/10.1007/s12599-015-0376-2>
- Kock, N. (2000). Benefits for virtual organizations from distributed groups. *Communications of the ACM*, 43(11), 107-112. <https://doi.org/10.1145/353360.353372>
- Koivunen, S., Ala-Luopa, S., Olsson, T., & Haapakorpi, A. (2022). The march of chatbots into recruitment: Recruiters' experiences, expectations, and design opportunities. *Computer Supported Cooperative Work (CSCW)*, 31(3), 487-516. <https://doi.org/10.1007/s10606-022-09429-4>
- Konradt, U., & Hertel, G. (2007). *Management virtueller Teams. Von der Telearbeit zum virtuellen Unternehmen*. Beltz.
- Kristof, A. L., Brown, K. G., Sims, H. P., & Smith, K. A. (1995). The virtual team: A case study and inductive model. In M. M. Beyerlein, D. A. Johnson, & S. T. Beyerlein (Eds.), *Advances in interdisciplinary studies of work teams: Knowledge work in teams* (pp. 229-253).
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159. <https://doi.org/10.2307/2529310>
- Lembcke, T.-B., Diederich, S., & Brendel, A. B. (2020). Supporting design thinking through creative and inclusive education facilitation: The case of anthropomorphic conversational agents for persona building. In *Proceedings of the Twenty-Eighth European Conference on Information Systems*. AISel.
- Lengen, J. C., Kordsmeyer, A.-C., Rohwer, E., Harth, V., & Mache, S. (2021). Soziale Isolation im Homeoffice im Kontext der COVID-19-Pandemie. *Zentralblatt für Arbeitsmedizin, Arbeitsschutz und Ergonomie*, 71(2), 63-68. <https://doi.org/10.1007/s40664-020-00410-w>
- Lindner, D. (2020). Chancen und Risiken durch virtuelle Teams. In *Virtuelle teams und homeoffice* (pp. 9-12). Springer Gabler.
- Luengo-Oroz, M. (2019). Solidarity should be a core ethical principle of AI. *Nature Machine Intelligence*, 1(11), 494. <https://doi.org/10.1038/s42256-019-0115-3>
- Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., & Söllner, M. (2019). AI-based digital assistants. *Business & Information Systems Engineering*, 61(4), 535-544. <https://doi.org/10.1007/s12599-019-00600-8>
- Maedche, A., Morana, S., Schacht, S., Werth, D., & Krumeich, J. (2016). Advanced user assistance systems. *Business & Information Systems Engineering*, 58(5), 367-370. <https://doi.org/10.1007/s12599-016-0444-2>
- Mateescu, A., & Nguyen, A. (2019). *Algorithmic management in the workplace* (pp. 1-15). Data & Society Research Institute.
- Mayring, P. (2015). Qualitative Inhaltsanalyse. In *Qualitative Inhaltsanalyse: Grundlagen und Techniken (12. Auflage)*. Beltz.
- Meyer von Wolff, R., Hobert, S., & Schumann, M. (2019). How may i help you? - State of the art and open research questions for chatbots at the digital workplace. In *Proceedings of the 52nd Hawaii International Conference on System Sciences* (pp. 95-104). ScholarSpace. <https://doi.org/10.24251/HICSS.2019.013>
- Mirbabaie, M., Brendel, A. B., & Hofeditz, L. (2022). Ethics and AI in information systems research. *Communication of the Association for Information Systems*, 50(1), 38. <https://doi.org/10.17705/1CAIS.05034>
- Mirbabaie, M., Stieglitz, S., Brünker, F., Hofeditz, L., Ross, B., & Frick, N. R. J. (2021). Understanding collaboration with virtual assistants - The role of social identity and the extended self. *Business & Information Systems Engineering*, 63(1), 21-37. <https://doi.org/10.1007/s12599-020-00672-x>
- Morrison-Smith, S., & Ruiz, J. (2020). Challenges and barriers in virtual teams: A literature review. *SN Applied Sciences*, 2(6), 1096. <https://doi.org/10.1007/s42452-020-2801-5>
- Mysirlaki, S., & Paraskeva, F. (2020). Emotional intelligence and transformational leadership in virtual teams: Lessons from MMOGs. *Leadership & Organization Development Journal*, 41(4), 551-566. <https://doi.org/10.1108/LODJ-01-2019-0035>
- Niederfranke, A., & Drewes, M. (2017). Neue Formen der Erwerbstätigkeit in einer globalisierten Welt: Risiko der Aushöhlung von Mindeststandards für Arbeit und soziale Sicherung? *Sozialer Fortschritt*, 66(12), 919-934. <https://doi.org/10.3790/sfo.66.12.919>
- Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *SSRN Electronic Journal*, (1745302). Available at SSRN 4375283. <https://doi.org/10.2139/ssrn.4375283>
- Orta-Castañón, P., Urbina-Coronado, P., Ahuett-Garza, H., Hernández-de-Menéndez, M., & Morales-Menendez, R. (2018). Social collaboration software for virtual teams: Case studies. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 12(1), 15-24. <https://doi.org/10.1007/s12008-017-0372-5>
- Ozmen Garibay, O., Winslow, B., Andolina, S., Antona, M., Bodenschatz, A., Coursaris, C., Falco, G., Fiore, S. M., Garibay, I., Grieman, K., Havens, J. C., Jirotko, M., Kacorri, H., Karwowski, W., Kider, J., Konstan, J., Koon, S., Lopez-Gonzalez, M., Maifeld-Carucci, I., ... Xu, W. (2023). Six human-centered artificial intelligence grand challenges. *International Journal of Human-Computer Interaction*, 39(3), 391-437. <https://doi.org/10.1080/10447318.2022.2153320>

- Paraman, P., & Anamalah, S. (2023). Ethical artificial intelligence framework for a good AI society: Principles, opportunities and perils. *AI & Society*, 38(2), 595–611. <https://doi.org/10.1007/s00146-022-01458-3>
- Pyyriä, P. (2009). Virtual collaboration in knowledge work: From vision to reality. *Team Performance Management*, 15(7/8), 366–381. <https://doi.org/10.1108/13527590911002140>
- Rädiker, S., & Kuckartz, U. (2019). *Analyse qualitativer Daten mit MAXQDA*. Springer Fachmedien Wiesbaden.
- Rhee, C. E., & Choi, J. (2020). Effects of personalization and social role in voice shopping: An experimental study on product recommendation by a conversational voice agent. *Computers in Human Behavior*, 109(2019), 106359. <https://doi.org/10.1016/j.chb.2020.106359>
- Rheu, M., Shin, J. Y., Peng, W., & Huh-Yoo, J. (2021). Systematic review: Trust-building factors and implications for conversational agent design. *International Journal of Human-Computer Interaction*, 37(1), 81–96. <https://doi.org/10.1080/10447318.2020.1807710>
- Richards, D., Vythilingam, R., & Formosa, P. (2023). A principlist-based study of the ethical design and acceptability of artificial social agents. *International Journal of Human-Computer Studies*, 172(2023), 102980. <https://doi.org/10.1016/j.ijhcs.2022.102980>
- Röbken, H., & Wetzel, K. (2017). *Qualitative und quantitative Forschungsmethoden*. Carl von Ossietzky Universität.
- Rodríguez-Cantelar, M., Estecha-Garitaogitia, M., D'Haro, L., Matía, F., & Córdoba, R. (2023). Automatic detection of inconsistencies and hierarchical topic classification for open-domain chatbots. *Applied Sciences*, 13(16), 9055. <https://doi.org/10.20944/preprints202306.1588.v1>
- Rothenberger, L., Fabian, B., & Arunov, E. (2019). Relevance of ethical guidelines for artificial intelligence – A survey and evaluation. In *Proceedings of the European Conference on Information Systems (ECIS) 2019* (pp. 1–11). AISel.
- Ruane, E., Birhane, A., & Ventresque, A. (2019). *Conversational AI: Social and ethical considerations* (pp. 104–115). AICS.
- Saat, R. M., & Salleh, N. M. (2010). *Issues related to research ethics in e-research collaboration* (pp. 249–261). Springer. [https://doi.org/10.1007/978-3-642-12257-6\\_15](https://doi.org/10.1007/978-3-642-12257-6_15)
- Sağlam, R. B., & Nurse, J. R. C. (2020). Is your chatbot GDPR compliant? In *Proceedings of the 2nd Conference on Conversational User Interfaces* (pp. 1–3). Association for Computing Machinery. <https://doi.org/10.1145/3405755.3406131>
- Sands, S., Ferraro, C., Campbell, C., & Tsao, H. Y. (2021). Managing the human-chatbot divide: How service scripts influence service experience. *Journal of Service Management*, 32(2), 246–264. <https://doi.org/10.1108/JOSM-06-2019-0203>
- Sankaran, S., Zhang, C., Funk, M., Aarts, H., & Markopoulos, P. (2020). Do I have a say? In *Proceedings of the 2nd Conference on Conversational User Interfaces* (pp. 1–3). <https://doi.org/10.1145/3405755.3406135>
- Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity*, 52(4), 1893–1907. <https://doi.org/10.1007/s11135-017-0574-8>
- Saunders, J. F., Eaton, A. A., & Aguilar, S. (2020). From self(ie)-objectification to self-empowerment: The meaning of selfies on social media in eating disorder recovery. *Computers in Human Behavior*, 111(May), 106420. <https://doi.org/10.1016/j.chb.2020.106420>
- Schuetzler, R. M., Giboney, J. S., Grimes, G. M., & Rosser, H. K. (2021). Deciding whether and how to deploy chatbots. *MIS Quarterly Executive*, 20(1), 1–15. <https://doi.org/10.17705/2msqe.00039>
- Schwartz, T., Zinnikus, I., Krieger, H., Christian, B., Pirk, G., Folz, J., Kiefer, B., Hevesi, P., & Christoph, L. (2016). Hybrid teams: Flexible collaboration between humans, robots and virtual agents. In M. Klusch, R. Unland, O. Shehory, A. Pokahr, & S. Ahrndt (Eds.), *German Conference on Multiagent System Technologies* (pp. 131–146). Springer International Publishing.
- Seeber, I., Bittner, E., Briggs, R. O., de Vreede, T., de Vreede, G. J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & Söllner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. *Information & Management*, 57(2), 103174. <https://doi.org/10.1016/j.im.2019.103174>
- Seerat, B., Samad, M., & Abbas, M. (2013). Software project management in virtual teams. In *Proceedings of the Science and Information Conference* (pp. 139–143). IEEE.
- Shah, H., Warwick, K., Vallverdú, J., & Wu, D. (2016). Can machines talk? Comparison of Eliza with modern dialogue systems. *Computers in Human Behavior*, 58(May), 278–295. <https://doi.org/10.1016/j.chb.2016.01.004>
- Shahriari, K., & Shahriari, M. (2017). IEEE standard review – Ethically aligned design: A vision for prioritizing human wellbeing with artificial intelligence and autonomous systems. In *IHTC 2017 – IEEE Canada International Humanitarian Technology Conference 2017* (pp. 197–201). IEEE. <https://doi.org/10.1109/IHTC.2017.8058187>
- Shin, D., Kim, S., Shang, R., Lee, J., & Hsieh, G. (2023). IntroBot: Exploring the use of chatbot-assisted familiarization in online collaborative groups. In A. Schmidt, K. Väänänen, T. Goyal, P. O. Kristensson, A. Peters, S. Mueller, J. R. Williamson, & M. L. Wilson (Eds.), *CHI '23: ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–13). Association for Computing Machinery. <https://doi.org/10.1145/3544548.3580930>
- Shneiderman, B. (2020). Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy human-centered AI systems. *ACM Transactions on Interactive Intelligent Systems*, 10(4), 1–31. <https://doi.org/10.1145/3419764>
- Sowa, K., Przegalinska, A., & Ciechanowski, L. (2021). Cobots in knowledge work. *Journal of Business Research*, 125(November), 135–142. <https://doi.org/10.1016/j.jbusres.2020.11.038>
- Stephanidis, C., Salvendy, G., Antona, M., Chen, J. Y. C., Dong, J., Duffy, V. G., Fang, X., Fidopiastis, C., Fragomeni, G., Fu, L. P., Guo, Y., Harris, D., Ioannou, A., Jeong, K.-a., Konomi, S., Krömker, H., Kurosu, M., Lewis, J. R., Marcus, A., ... Zhou, J. (2019). Seven HCI grand challenges. *International Journal of Human-Computer Interaction*, 35(14), 1229–1269. <https://doi.org/10.1080/10447318.2019.1619259>
- Stieglitz, S., Frick, N. R. J., Mirbabaie, M., Hofeditz, L., & Ross, B. (2021). Recommendations for managing AI-driven change processes: When expectations meet reality. *International Journal of Management Practice*, 16(4), 407–433. <https://doi.org/10.1504/IJMP.2023.132074>
- Stieglitz, S., Hofeditz, L., Brünker, F., Ehnis, C., Mirbabaie, M., & Ross, B. (2022). Design principles for conversational agents to support emergency management agencies. *International Journal of Information Management*, 63(April), 102469. <https://doi.org/10.1016/j.ijinfomgt.2021.102469>
- Straus, S. G. (1996). Getting a clue the effects of communication media and information distribution on participation and performance in computer-mediated and face-to-face groups. *Small Group Research*, 27(1), 115–142. <https://doi.org/10.1177/1046496496271006>
- Strohmann, T., Fischer, S., Siemon, D., Brachten, F., Lattemann, C., Robra-Bissantz, S., & Stieglitz, S. (2018, December). Virtual moderation assistance: Creating design guidelines for virtual assistants supporting creative workshops. In *Proceedings of the 22nd Pacific Asia Conference on Information Systems – Opportunities and Challenges for the Digitized Society: Are We Ready?*, PACIS 2018. AISel.
- Suárez-Gonzalo, S., Mas-Manchón, L., & Guerrero-Solé, F. (2019). Tay is you. The attribution of responsibility in the algorithmic culture. *Observatorio (OBS\*)*, 13(2), 1–14. <https://doi.org/10.15847/obsOBS13220191432>
- Tenório, N., & Bjørn, P. (2019). Online harassment in the workplace: The role of technology in labour law disputes. *Computer Supported Cooperative Work (CSCW)*, 28(3–4), 293–315. <https://doi.org/10.1007/s10606-019-09351-2>
- Thorpe, H. H. (2023). ChatGPT is fun, but not an author. *Science (New York, N.Y.)*, 379(6630), 313. <https://doi.org/10.1126/science.adg7879>
- Wainfan, L., & Davis, P. K. (2004). *Challenges in virtual collaboration: Videoconferencing, audioconferencing, and computer-mediated communications*. Rand Corporation.
- Waizenegger, L., McKenna, B., Cai, W., & Bendz, T. (2020). An affordance perspective of team collaboration and enforced working from home during COVID-19. *European Journal of Information Systems*, 29(4), 429–442. <https://doi.org/10.1080/0960085X.2020.1800417>

- Wei, Y., Lu, W., Cheng, Q., Jiang, T., & Liu, S. (2022). How humans obtain information from AI: Categorizing user messages in human-AI collaborative conversations. *Information Processing & Management*, 59(2), 102838. <https://doi.org/10.1016/j.ipm.2021.102838>
- Zhuo, T. Y., Huang, Y., Chen, C., & Xing, Z. (2023). Exploring AI ethics of ChatGPT: A diagnostic analysis. <http://arxiv.org/abs/2301.12867>
- Zou, J., & Schiebinger, L. (2018). AI can be sexist and racist—It's time to make it fair. *Nature*, 559(7714), 324–326. <https://doi.org/10.1038/d41586-018-05707-8>

## About the authors

**Lennart Hofeditz** is a research associate at the SAP-endowed research group of Professor Stefan Stieglitz at the University of Potsdam, Germany. He is a PhD candidate in the field of Information Systems with the focus on ethical responsibility related to the application of virtual agents in organizations.

**Milad Mirbabaie** is a Junior Professor for Information Systems & Digital Society at Paderborn University, Germany. He has published in reputable journals such as JIT, BISE, ELMA, ISF, JDS, Internet Research, IJIM, and IJHCI. His work focuses on Sociotechnical Systems, AI-based Systems, Social Media, Digital Work, and Crisis Management.

**Mara Ortmann** is a professional working as an account manager at a large corporation in Germany. She studied Applied Cognition and Media Sciences at the University of Duisburg-Essen where she completed her Master of Science and started working with the research group of Professor Stefan Stieglitz.

## Appendix

### (A) Interview guideline

#### Introductory questions

1. Introduction of the expert
  - Please briefly introduce yourself and your current job.
2. Where have you already encountered conversational agents of team collaboration in your work/everyday life?
  - For which task areas has it been used?

#### Pre-conditions and use

3. What are the (ethical) prerequisites for the use of conversational agents in the digital workplace?
  - What factors support/inhibit the use of virtual assistants?
  - What are the advantages/disadvantages of their use?
4. What are the most reasonable areas to use conversational agents as teammates?
  - Are there areas where they should not be used?
  - If so, why should they not be used in these areas?

5. What tasks should not be handed over to a conversational agent? *Beneficence and affective conflicts*
6. What new pressures might the introduction of conversational agents place on team members?
7. What social risks (for example, isolation, lack of familiarity) could arise within the team from the use of a conversational agent?

- How could they be counteracted? (In the case of spatial distance: e.g., increased isolation, lack of familiarity, building of a sense of “we”, formation of subgroups)
8. How can these pressures and risks be prevented?

#### *Justice, solidarity, and workspace awareness*

9. What threats to team solidarity can arise from the use of conversational agents?
  - How can these be avoided?
10. Which information deficits arise from the fact that team members do not physically work together?
  - How can the deficits be compensated?
11. How do team members without direct physical/virtual/verbal contact gather information about virtual teammates and their virtual workspace?
12. In the context of which virtual collaboration activities is knowledge about team members and the workspace particularly relevant?

#### *Non-maleficence and privacy*

13. What is the value of the employee's right to privacy and data protection when using a conversational agent that continuously collects and analyses data?
14. What types of abuse can occur when using the technology?
  - How can these be avoided or how can they be counteracted?

#### *Explicability*

15. What types of transparency are required on the part of the employer?

#### *Accountability/autonomy*

16. What role does accountability play in the use of conversational agents?
  - Who holds the responsibility in case of wrong decisions?
17. How can worker autonomy be maintained?
  - How can dependence on technology be avoided? (unlearning to solve problems themselves, forgetting the solution path)

#### *Closing remarks*

18. What other aspects can you think of that are or could become important for the use of conversational agents in digital collaboration now or in the future?
  - Are there any questions or aspects that you expected but that I did not ask? Would you like to address them?