Augmented Facilitation: Designing a multi-modal Conversational Agent for Group Ideation

Izabel Cvetkovic
University of Hamburg, izabel.cvetkovic@uni-hamburg.de

Maren Gierlich-Joas
Copenhagen Business School, maren.gierlich-joas@uni-hamburg.de

Navid Tavanapour
Universität Hamburg, navid.tavanapour@uni-hamburg.de

Nicole Debowski-Weimann
Volkswagen, nicole.debowski-weimann@volkswagen.de

Eva A. C. Bittner
University of Hamburg, bittner@informatik.uni-hamburg.de

Augmented Facilitation: Designing a multi-modal Conversational Agent for Group Ideation

Completed Research Paper

Izabel Cvetkovic
University of Hamburg
Hamburg, Germany
izabel.cvetkovic@uni-hamburg.de

Maren Gierlich-Joas
Copenhagen Business School
Copenhagen, Denmark
mg.digi@cbs.dk

Navid Tavanapour, Eva Bittner
University of Hamburg
Hamburg, Germany
navid.tavanapour@uni-hamburg.de,
eva.bittner@uni-hamburg.de

Nicole Debowski-Weimann
Volkswagen
Wolfsburg, Germany
nicole.debowski-weimann@volkswagen.de

Abstract

Human facilitators face the challenge to structure and collect relevant insights from collaborative creative work sessions, which can suffer if they face a high workload. Hence, for effective value co-creation in organizational ideation we suggest an facilitation augmentation with a conversational agent (CA). CAs have the ability to support respective collaborative work by documenting and analyzing unstructured data. Following the design science research paradigm, and based on the literature about facilitation and human-AI collaboration, we derive design principles to develop a CA prototype that collects ideas from a group ideation session and displays them back in a structured (multi-modal) manner. We evaluate the CA by conducting four focus groups. Key findings show that the CA successfully distills and enriches information. Our study contributes to understanding the role of CA in augmenting facilitation and it provides guidance for practice on how to integrate these technologies in group meetings.

Keywords: *Augmentation, Conversational Agents, Facilitation, Human-AI Collaboration*

Introduction

Creative work has high relevance for the innovation capability of companies in order to survive in a competitive environment, because innovation processes fuel e.g., emergence of new business models or solution of existing problems (Friesike et al. 2019). To leverage the creative potential in companies, the ideation method is often used. Hereby, teams develop solutions and ideas for a defined problem, thereby co-creating value. Many ideas are generated, then integrated, summarized and finally evaluated (Putman & Paulus 2009). To optimize the results, facilitators monitor interaction between group participants, and control the processes and tasks, among other (Schallmo 2018). While human facilitators deal with all these tasks, team sessions generate a lot of unstructured data which often results in lost insights. Hence, an improved way of documentation needs to be worked on (Chen et al. 2021). Besides human facilitators, digital technologies have been used to support ideation, albeit mostly on individual level (Debowski et al. 2021; Przybilla et al. 2019), where the potential for structuring large amounts of data stayed unexhausted. Furthermore, with the increasing demand for augmentation in the workforce, there is a need for digital innovations that can support collaborative work. For this purpose, it is necessary to investigate how digital innovations could support human facilitators in moderating collaborative creative work.

Automation of partial tasks has become essential, especially at a time when there is a shortage of skilled workers in all areas of the labor market (Seric & Winkler 2020). Recent research has shown that Artificial Intelligence (AI) can successfully support humans by taking over mundane tasks and analyzing and structuring large amounts of data (Dellermann et al. 2019b; Li et al. 2020). Through increased development of machine learning (ML) and Natural Language Processing (NLP) methods, AI-based systems can increasingly perform or augment tasks autonomously. The processing of linguistic input qualifies AI-based technology to support ideation. Applications that process and respond in natural language are known as Conversational Agents (CAs). They support text input, audio input or both (Laumer et al. 2019). The importance of CAs has grown significantly, especially in recent years as the online channel has become inescapable for businesses in many ways. The versatility of the uses of CAs is continuously expanding with the progressive development of technical possibilities (Mockenhaupt 2021). This development leads to a transformation of work, resulting in novel work configurations (Baptista et al. 2020).

While acknowledging that Human-AI Collaboration spans various forms of user interfaces and interaction modes, it's essential to recognize that our research primarily addresses ideation processes centered around generating textual artifacts, which is what a lot of ideation in early stages in practice is focused toward (Daly et al. 2016). This specific focus guides our exploration of AI-based CAs. The rationale for emphasizing CAs lies in their unique ability to seamlessly integrate with existing communication channels and facilitate ideation without the need for specialized hardware or complex interfaces. Furthermore, CAs can potentially be more motivating due to their anthropomorphic features, adaptability to the user, and direct interaction (Diederich et al. 2020), factors that are important in ideation workshops where participants are accustomed to human facilitation. The adaptability of CAs to diverse contexts and communication preferences ensures inclusive participation, a critical element of successful ideation (Poser et al. 2022).

Within this work, we investigate whether a CA could act as a co-facilitator, thereby augmenting the experience while preserving the natural communication. We build upon the concept of Hybrid Intelligence (Dellermann et al. 2019b), and Facilitation Framework (Dickson et al. 1996), to investigate CA's potential in augmenting organizational ideation processes, and hereby, assist in co-creating value. In this process, the CA should support human facilitators and leave cognitive space for the actual value co-creation in teams. We pose the research question (RQ):

How can a CA be designed and implemented to augment group ideation?

Answering this research question, our work contributes to literature in the following manner: The derived design requirements (DRs), design principles (DPs), and design features (DFs) represent prescriptive knowledge for the design of CAs in the context of augmented ideation. On a broader level, the study adds value to literature on human-AI collaboration, focusing on CAs in facilitation, and, thereby, addressing the research streams of changing nature of work. From a practical perspective, the technical artefact can be transferred to different contexts and support human facilitators in ideation tasks.

Theoretical Background

Human-AI Collaboration

Dellermann et al. 2019a propose a concept called Hybrid Intelligence to overcome limitations of humans and AI working independently. They define this concept as the ability to achieve complex goals by combining human and artificial intelligence, resulting in superior results compared to what each of them could achieve on their own. Other terms used to describe this symbiosis are Human-AI teaming, human-in-the-loop or AI-in-the-loop (Siemon 2022). The collaboration between humans and AI is a prerequisite to achieve Hybrid Intelligence, where at least two actors work together towards a shared goal.

Examples of Human-AI collaboration can be seen in various fields. For instance, in the field of education, AI can be used to create personalized learning experiences, while human teachers can provide emotional support and guidance (Maedche et al. 2019). In manufacturing, human workers can collaborate with AI to optimize production processes and reduce errors (Arinez et al. 2020). In finance, AI can help detect fraud, while human experts can interpret the results and make decisions based on the findings (Milana & Ashta 2021). Furthermore, Mahmud et al. 2022 and Anantrasirichai & Bull 2022 have shown that Human-AI collaboration can be especially effective in creative tasks that require both logical and intuitive thinking. By combining the strengths of human and AI intelligence, such as creativity and accuracy, respectively, hybrid teams can generate more innovative and effective solutions.

This collaboration may improve group performance but also raises unresolved questions about the dynamics and characteristics of Human-AI teams. Human-AI collaboration is at an early stage of development, but it is becoming an increasingly important research area in the fields of Human-Computer Interaction and Information Systems. Design issues, such as the design of explainable and transparent AI agents, and the design of a shared workplace with consideration of tasks (Cvetkovic & Bittner 2022; Lubars & Tan 2019), and roles (Siemon 2022) need to be addressed for successful Human-AI collaboration.

Facilitation & Ideation

The term facilitation is used to refer to "a set of different functions and activities that occur before, during, and after a group meeting to assist the group in achieving its goals" (Beranek et al. 1993). Facilitation guides the group through meetings and may use group support systems that provide various facilitation tools. For example, they assist with idea generation or group decision-making (Leimeister 2014). Facilitation is a dynamic process that seeks to positively influence the relationships of the participants along the goal of structuring the group meetings and providing assistance in mastering the tasks in order to support the creation of the group artifacts in the best possible way (Beranek et al. 1993). Thereby, the promotion of communication within the group is essential.

Facilitation can be applied by a group member, the group leadership, or an external agent by executing activities. Bostrom et al. 1993 describe such activities in their Facilitation Framework in the context of the process, task and relationship activities. The Relationship (Feel about) activities describe influences during the process (How?) that includes activities towards accomplishing a task (What?). Additionally, Dickson et al. 1996 categorizes facilitation activities in task and interaction interventions. The activities that promote relationships and communication are described as interactional interventions (Dickson et al. 1996). Facilitation activities that pursue the goal of structuring the group meetings and providing assistance towards accomplishing the task are task interventions (Dickson et al. 1996). If a special technology is used in the team meeting, a facilitator's support in using this technology also constitutes a task-related activity (Przybilla et al. 2019).

Ideation sessions require specific manners of facilitation and they consist of different activities (Beranek et al. 1993). While in the opening phase the facilitator familiarizes the participants with each other, introduces the course of the meeting, clarifies the roles of the participants, and lists the goals and rules of the meeting, the facilitator is responsible for encouraging the group to work together and guiding them through the agenda during the execution phase (Beranek et al. 1993). In doing so, all activities should be aligned with the group's goals. The facilitator conducts summaries as needed and identifies decisions (Dickson et al. 1996). All participants are expected to participate during the meeting, so it is important to encourage everyone's participation and strengthen ownership and group responsibility (Leimeister 2014). In the closing phase, facilitation provides a summary of the meeting and identifies future tasks that remain after the meeting (Beranek et al. 1993). In general, it is important to present information to the group in an understandable way throughout the meeting to ensure safe communication and information exchange between participants.

Systems that provide support functions for humans in solving creative tasks are referred to as Creativity Support Systems (CSS) or Group Creativity Support Systems (GCSS). GCSS are used by several people and support the collaboration process between them. Both types of systems can be used in all phases of the ideation, from idea generation, through idea evaluation, to idea decision making (Przybilla et al. 2019). GCSS can also perform facilitation tasks, for example, by executing the pre-scripted the collaboration process. Moreover, by supporting the use of creativity methods and creating an environment that fosters trust, less creatively inclined people are inspired to come up with more creative ideas or their ideas can be developed more deeply (Przybilla et al. 2019). GCSS can promote overall communication between team members and therefore perform essential facilitation tasks (Voigt & Bergener 2013). In this way, they may increase the productivity of collaboration (Willman 2021). In principle, GCSS can provide support in both task and interactional interventions. However, relying on complementary strengths of humans and AI, a hybrid design of facilitation tasks according to the respective strengths could promise better outcomes. Traditional GCSS are not "intelligent" and only novel capabilities allow for hybrid intelligence, giving rise to a new class of GCSS. To this end, CAs have been identified as a potential intelligent addition to GCSS, by taking over specific task interventions in a role of a (co-)facilitator (Debowski et al. 2021).

Conversational Agents as (Co-)Facilitators

In recent years, the increasingly intensive use of instant messengers and rapid advances in AI, particularly NLP and Deep Learning, have led to CAs becoming a widespread technology. They can be assigned to the assistance systems used for AI-driven collaboration between humans and machines (Mockenhaupt 2021). In this context, communication takes place via natural language, via textual or audio channels (Brandtzæg & Følstad 2018). In the former case, they are known as chatbots, in the latter as voice assistants (e.g., Siri, Alexa) (Laumer et al. 2019).

The use of NLP technology is crucial because it influences the output to a high degree (Singh et al. 2019). CAs of this type are also empowered through machine learning and the injection of training data to acquire further knowledge so that response quality can be increased based on past conversations. After the learning phase, the CA is able to recognize intentions and generate responses from a text input (Busse 2021). CAs of this type are, however, susceptible to grammatical errors, as they generate responses dynamically and require large amounts of training and testing data (Ramesh et al. 2017).

Although CAs have been used in variety of domains, their use in collaborative creative work remains relatively unexplored. While CAs have so far mainly communicated with a single person, research has been investigating the extent to which CAs can be used for group collaborations for a few years now (Kim et al. 2021; Winkler et al. 2019). Bittner et al. 2019b; Bittner et al. 2019c showed different possibilities using a taxonomy of how CAs can contribute to the design of a human-machine collaboration. By interviewing professional facilitators, Bittner et al. 2021 were also able to identify key areas of expertise that a CA should cover to support a human facilitator in the context of a facilitation. Toxtli et al. 2018 implemented a CA that can handle task management for individual and group work. They identified particular difficulties and opportunities in CA communication with multiple users. Kim et al. 2020 investigated how CAs can guide and structure discussions and developed corresponding prototypes. Elshan et al. 2022 showed how CAs can relieve teams from innovation blockages. IBM researchers (Tepper et al. 2018) demonstrated a Collabot – a chat assistant that implicitly learns users interests and social ties within a chat group and provides a personalized digest of missed content. By doing so, it assists users in coping with chat information overload by helping them understand the main topics discussed, collaborators, links and resources.

In previous research, CAs have so far replaced humans as facilitators and have not acted as support for a human facilitator to augment creative work. However, the combination of human and CA facilitator has advantages: human can act on the side of interaction interventions, e.g., monitor team dynamics, and control the process dynamically while CA can support tasks and processes by processing and analyzing data quickly and reliably. Hence, it should be investigated to what extent CAs can be used as co-facilitator.

Design Science Research Approach

We followed the DSR methodology by Peffers et al. 2007 to explore a CA's role in supporting a human facilitator in the ideation process. Therefore, we conducted the six iterative DSR stages of understanding the problem, deriving objectives for a solution, designing and developing the artefact, demonstrating and evaluating it, and finally communicating it (Peffers et al. 2007). To ensure both rigor and relevance for our design activities (Hevner et al. 2004), we built our study on literature research, qualitative problem interviews, and confirmatory focus groups. The application of different methods to incorporate insights from practice and scientific knowledge base, and evaluate the generated design knowledge can be found in Table 1.

DSR Step	Method	Purpose
Problem identification	3 expert interviews	Derive design requirements
Problem identification	Literature review	Derive design requirements
Evaluation	4 focus group discussions	Confirm design principles
Evaluation	SWOT analysis	Generate general insights

Table 1. DSR Data Collection Approach

Problem Understanding

In the first stage, the problem space was analysed first via a literature review (see Section Theoretical Background), the results of which were supplemented with semi-structured expert interviews with creative workshops facilitators from a large automotive company (n=3). The semi-structured interviews followed predefined questionnaires, they were recorded, transcribed and analyzed with qualitative content analysis (Mayring & Fenzl 2019). Hereby, a deductive approach was followed with pre-defined categories that were previously identified in the literature. The literature and interview analysis was conducted by four individuals, while accounting for an understanding about (dis)agreements in coding.

Our results reveal that ideation sessions are cumbersome to facilitate as they, among other, generate a lot of unstructured data, often resulting in lost insights, due to improper documentation. Human facilitators, especially novices, often cannot handle this, as they are focused on facilitating tasks, processes, and team interaction (Schallmo 2018). In sum, experts struggle with moderating creative workshops, because they have to handle many aspects previously mentioned (team dynamics, tasks, processes) and, in addition, there is a lot of unstructured data that has to be summarized, prioritized, visualized - in short - enriched and distilled: ”*I must do the whole [thing] then again: make clusters, merge, prioritize [...] and that is mostly many columns, and then someone does not find his contribution there afterwards. Then perhaps he will not cooperate any more, but you need the cooperation of all, even more so in a large enterprise.*” From the expert interviews, we derived 8 user requirements and merged them with existing literature requirements (Fig. 1). In the following, we explain the derivation of DRs, DPs and DFs, which represent our objectives of the solution.

Objectives of a Solution

DRs gathered from existing literature (TRs) and user requirements (URs) gathered from conducting interviews served as a foundation for setting up the respective DPs. Initially, several aspects proposed by papers dealing with automated facilitation, CAs, and human-computer interaction were picked up and put into broader categories to sort, not only alike propositions, but also the large amount of the existing research results. Additionally, the gathered DRs are structured along Dickson et al. 1996 task interventions and interaction interventions.

Task interventions: The CA shall undergo training on a prepared dataset (UR1) within the specific topic domain of the ideation session to meet the **DR1. Topic Domain Incorporation**. It shall facilitate ideation sessions, including the sub-phases of idea gathering, idea selection, and summary, adhering to the workflow of established ideation methods (UR2, TR1) as per **DR2. Guided Ideation Workflow** (Bittner & Shoury 2019; Strohmann et al. 2018). At the onset of each session, it shall provide a clear explanation of the ideation method and associated rules (TR2) in alignment with **DR3. Session Introduction** (Bittner & Shoury 2019). The CA shall automate routine tasks (TR3) to reduce users' repetitive work and enable them to focus on value-creating activities, fulfilling **DR4. Task Automation** (Oeste-Reiß et al. 2021). To optimize user interaction (TR4) and ensure user efficiency, it shall possess awareness of various user interaction modes and respond to user queries regarding the ideation method in real-time through dedicated functions activated by predefined keywords (UR3, TR5) as dictated by **DR5. Optimized interaction** (Knote et al. 2020) and **DR6. User Query Handling** (Koch et al. 2020; Strohmann et al. 2018). The CA shall also incorporate a memory function (UR5, TR10) to store user contributions for later phases of ideation, especially for result summarization, in accordance with **DR11. Memory Function** (Koch et al. 2020; Strohmann et al. 2018). It shall aggregate the results of the group's contributions and provide a summary in a primarily visual manner (UR6, TR11), e.g., via word clouds, idea one-pagers, or clustering (UR7, TR12) as specified by **DR12. Information Aggregation** (Bittner & Shoury 2019) and **DR13. Information Visualization** (Bittner et al. 2021).

Interaction interventions: The CA shall seamlessly integrate with common collaboration platforms (UR4, TR7) to allow users to interact without the need for application switching, meeting **DR7. Integration with Collaboration Platform** (Bittner & Shoury 2019; Tavanapour & Bittner 2018). It shall communicate with users using natural language via text-based messaging (TR9), ensuring a user-friendly interaction and fulfilling **DR10. Textual Communication** (Cvetkovic et al. 2023; Tavanapour & Bittner 2018). The CA shall encourage user engagement in discussions through the exchange of messages (UR4, TR7) to align with **DR8. User Engagement** (Koch et al. 2020; Strohmann et al. 2018). It shall adapt its language to the user group (TR8) to tailor the communication to users' preferences and understanding, in accordance with **DR9. User-Appropriate Language** (Bittner et al. 2019a; Radziwill & Benton 2017; Tavanapour & Bittner 2018). Lastly, the CA shall incorporate features that emulate human behavior (UR8, TR13) to build user trust, encourage interaction, and reduce the feeling of being constantly observed, as specified by **DR14. Human-like Behavior** (Benke 2020; Gnewuch et al. 2017; Strohmann et al. 2018).

For deriving the DPs, the approach proposed by Gregor et al. 2020 was used. To follow the proposed approch, the DRs were first sorted into categories that could later be transformed into DPs. For that matter, similarities between the DRs were used to sort them accordingly. After the initial sorting of the DRs, the scheme for setting up DPs can be applied. For that, four core concepts are typically included in a DP: the aim, context, mechanism, and rationale, which is extended by a normally constant set of implementers, users, and enactors (Gregor et al. 2020). Using the components 'user', 'aim', 'context', 'mechanism' and 'rationale', we can introduce our main clause for the DPs, here exemplified with DP2:

For teams and facilitators (users) to reduce their cognitive load (aim) in organizational ideation (context), the CA should provide the ability to process natural language of multiple users in real-time and respond in the same way via text via a familiar communication medium (mechanism). This DP draws upon assistance provision and human-like conversing in CA research (rationale).

The created DPs aggregate the DRs into respective categories and reflect the commonly addressed domain.

DP1. Task and process assistance tackles the ideation method and process control. It ensures a predefined process and set of rules based on the ideation method. Respective rules and the underlying methodology of the method should be explained to the users by the CA. In this way, facilitators can be relieved of some mundane tasks, while teams have constant access to this information and can use it without asking for repetitions or additional explanations. This corresponds the assistance provision via speedy assistance and relief from mundane tasks (Moussawi 2018; Waizenegger et al. 2020).

DP2. Natural language communication deals with the processing of user input. As a basis, natural language processing and instant messaging are to be used, to allow the users and the CA to communicate with each other. The input from users should be handled by the CA through the execution of respective functions. These are supposed to ensure the assistance provision and human-like conversing (Mygland et al. 2021).

DP3. Information distilling and enrichment targets the subject of documentation and memory. For ensuring a successful visualization and results at the end of the ideation session, gathered ideas must be saved throughout the process. These functions enable the facilitators to distill and enrich information, as well as facilitation (Mygland et al. 2021).

DP4. Human-likeness is about human-like behavior. The CA should communicate with the users in a manner that is as close to human-human interaction as possible, in order to gain the users' trust and will to interact with the CA and to follow its proposed procedure of the ideation session. This DP is congruent to prior CAs research related to human-like experience and personalization (Mygland et al. 2021).

These DPs are translated into corresponding DFs that serve as a basis for the implementation of the CA. Mostly, they rely on NLP and fine-tuning with domain-specific data as well as incorporating human-like answers. Furthermore, custom scripts enable the clustering and visualization functions. Also, there is an integration into a familiar messenger platform (Slack). To ensure proper intent recognition, the underlying NLP model was enhanced with a BERT featurizer (BERT is a large language model trained on billions of data points) (Devlin et al. 2019).

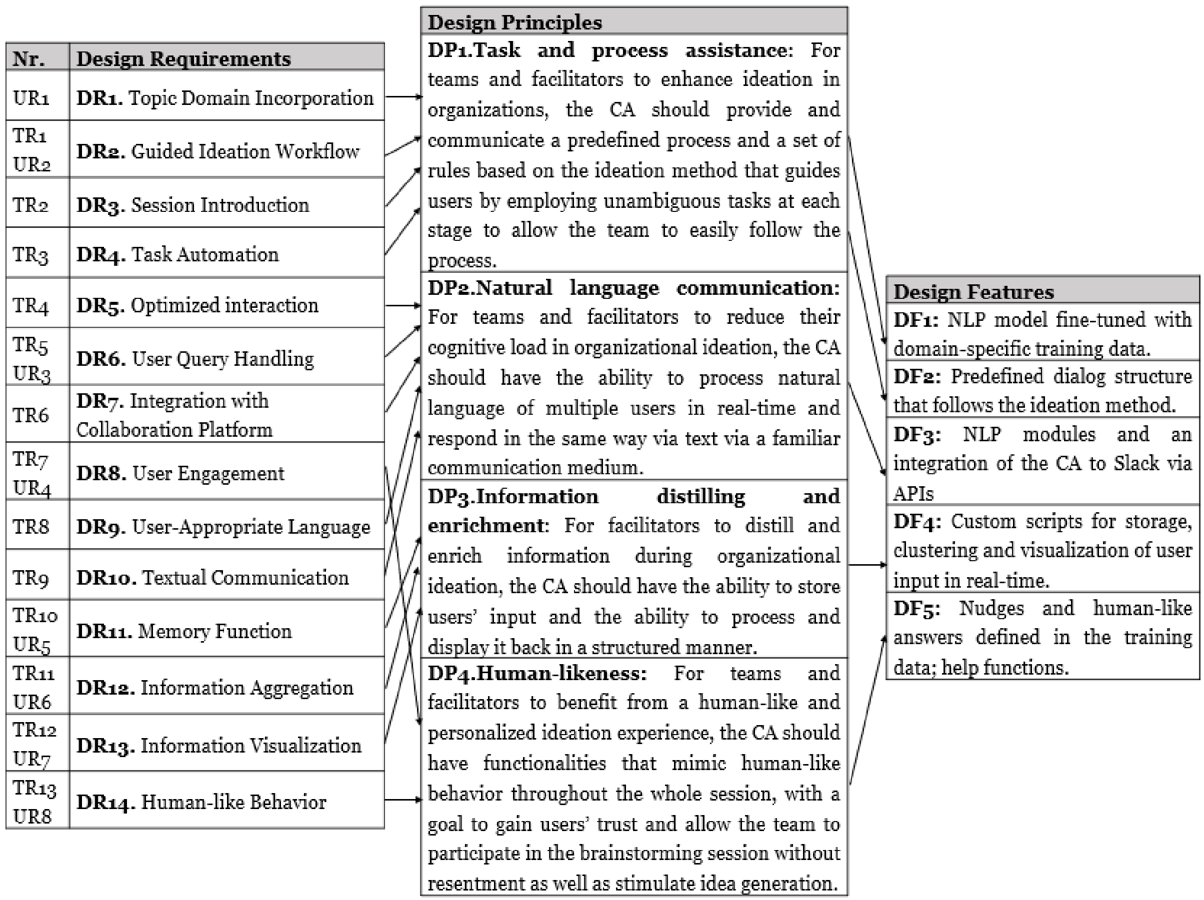

Figure 1. Design Requirements, Principles and Features.

Design and Development

To select a framework for the CA development, we made sure that it satisfies the requirements necessary to implement the above-mentioned DFs. A suitable framework, RASA (Rasa Technologies Inc 2021), was chosen, since it is open-source, models can be fine-tuned, custom functions can be developed, and it can be integrated into many messaging platforms via APIs. The development of the CA prototype was conducted in two iterative cycles, with each cycle significantly influencing the refinement of the DPs and DFs. The initial cycle resulted in an early prototype with limited functionalities, necessitating user feedback gathering and insights. Importantly, we extended the training data through feedback and insights obtained during the first cycle's demonstration and evaluation. This initial version of the prototype lacked a typical ideation phase (idea selection, DR2) and one form of visual summaries (mindmap, DR13). These limitations were addressed in the refined version of the prototype developed during the second cycle. Additionally, we introduced several other features, including the inclusion of an avatar (DR14) and a voting function, based on the feedback and lessons learned from the initial prototype. The dialog flow of the interaction with the CA is showcased in Fig. 2. For visual and textual summaries (clusters) a separate action server was used. This is a part of the RASA framework that offers customization of CA features, i.e., theoretically one could develop any feature supported by their chosen front-end platform. Fig. 2 displays the materialization of DPs. DP2 applies to the whole interaction and is therefore not indicated in the figure.

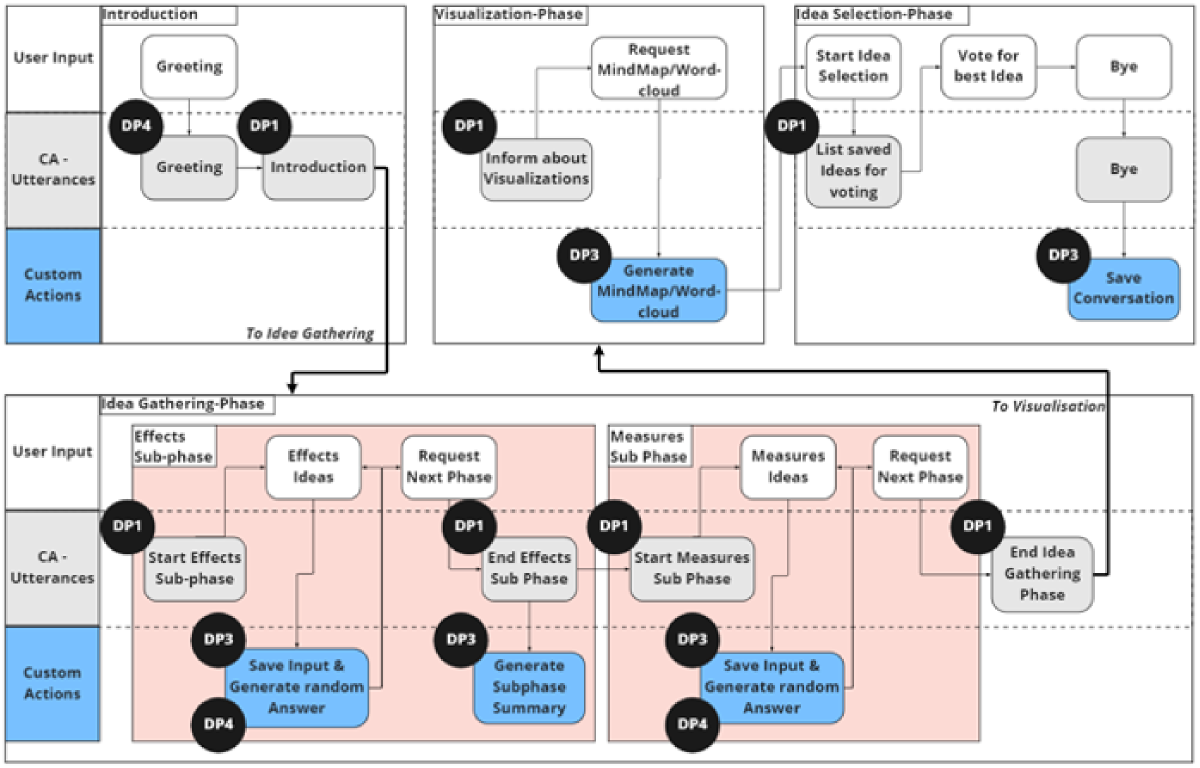

Figure 2. Dialog Flow of the Interaction with the CA.

Demonstration and Evaluation

As part of the demonstration, we performed a user test to assess teams' experiences with the CA during a ideation session. do so, 4 groups of in total 22 participants interacted with the CA: one group of facilitators from a large automotive organization (n=5); and three groups of master and PhD students from the fields of computer science and psychology (n=9;4;4). All participants had previous experience with CAs. Different user types were recruited to ensure complementary and heterogenous groups with both industry and educational perspective, as well as to capture both lay user and facilitator (expert) perspective. We stopped the data collection after reaching saturation, i.e., after repeated answers and no new aspects were mentioned in the evaluation.

The participants were given instructions to greet the CA and await on his instructions. With this, the demonstration began and the CA (Summy) welcomed the participants to the ideation session and explained the task and rules. The task consisted of three phases: idea gathering for two topics (effects of and measures to counter climate change in agriculture), summary of gathered ideas via wordcloud and mindmap, and idea selection via voting. In the first phase, participants had to write answers to the question "Which effects come to your mind regarding the topic 'How will climate change affect water sources, jobs and people in and around the agricultural industry?'". Participants were writing ideas until they could no longer think of anything else, then they proceeded to write their ideas about the measures to counteract these effects. Before advancing to the measures phase, participants got their input clustered (fig. 3, left pane). After typing in the measures, participants could ask for a wordcloud or a mindmap of their contributions (fig. 3, right pane). Finally, Summy reposted all the measures that he previously saved, and participants voted for the best measure. With this, the ideation session was finished.

The demonstration lasted 12 minutes on average, with the shortest one being 10 minutes, and the longest one being 15 minutes long. All participants engaged in the session; help functions were used only in one group. Immediately after the demonstration, groups engaged in focus group discussions.

Figure 3. Demonstration of the Artefact.

The evaluation of the artefact aligns with the proposed framework for evaluation in DSR by Venable et al. 2016. For this purpose, four steps were followed. First, the goals of evaluation were defined. In our case, this is a proof of concept of the artefact by establishing its conformance with proposed DPs. The second step was to define the strategy for evaluation. We opted for a field test of the design artefact with subsequent confirmatory focus groups to establish the utility of the artefact in field use (Tremblay et al. 2010). Third, respective properties to evaluate are defined (Venable et al. 2016) which are our DPs. Lastly, all respective evaluation episodes were designed (Venable et al. 2016).

Semi-structured discussions were directed via an interview guideline, which slightly differed across the two user groups: for facilitators, we asked if some aspects would support them as facilitators; for students, we asked if it would support them as ideation teams. The guideline consisted of a few questions for each DP, followed by spontaneous follow-up questions depending on the answers. The discussions lasted approximately 32 minutes on average. Following the focus group discussions, all participants filled out SWOT analysis templates (Strengths, Weaknesses, Opportunities, Threats) on the chatbot's ability to assist facilitators/teams in (moderating) organizational ideation processes. The resulting qualitative data was analysed with MAXQDA, using Mayring's (Mayring & Fenzl 2019) qualitative content analysis method. Categories (codes) were created based on the DPs and corresponding sub-categories based on the relating DRs. Then, categories were mapped to corresponding text.

Results

In the following, the answers from the focus group discussions are summarized for each DP. We use illustrative quotes to provide insights into the evaluation. To assess the fitness of DP1, we asked how the CA made the assignment clear, and how it guided the group through the process. Furthermore, we asked in what sense the CA relieved the groups - as facilitators vs. as ideation teams. Participants agree that the assignment explanation was direct, clear, and sufficient. "*I think it was pretty clear, because it said [...]*". Two participants highlighted the need for more elaborate explanations with examples. "*I was missing the BECAUSE [...] you are working on this topic now BECAUSE...*". While on the one hand, some participants agree that the explanations were good for guiding them through the process "*We had in general a good introduction and guidance*", on the other hand, some participants state that one part of the process was confusing "*[...] but the way it was phrased, I didn't really realize we were supposed to have these two phases separate.*" Another point was stressed - the CA should have made it clearer that it is enough when one person addresses him. Suggestions for improvement were mentioned, such as a list with commands, or colored parts of the text for a better overview. The topic of relief through the CA was thoroughly discussed, which resulted in 25 codes. Participants mentioned both aspects that relieved them as well as what overwhelmed them. For the former, the saving of arguments and their clustering was mentioned as particularly relieving: "*that was really very impressive, so wow*" - a facilitator said about clusters. The effects that they produce along with mindmap and wordcloud were related to time-saving, the processing and summarizing are very elaborate, as well as impulses to take in regard different perspectives, and showing off the relevance of certain topics: "*we do that to [determine the relevance of specific topics] quite often with voting... Then [with the mindmap] you also see the importance of a topic. If there are 6 effects on one topic, then it seems to be particularly significant [...] and I actually found that quite good.*" The rather overwhelming aspects concerned the CA's answers to each single idea: "*So I thought as a user that it's just an artefact. It's not meant as a feedback or appreciation.*" Furthermore, the nature of the exercise in the group chat settings produced a lot of text, so that some participants couldn't keep up with what other people were posting while also thinking of ideas themselves. They suggested a blend-in function, where one can choose whether to see the contributions of others.

DP2 focuses on the correctness of input processing and the influence of CA in focusing on the actual value-adding task. When it comes to the input processing and corresponding reaction from the CA, participants mostly agreed that the CA correctly recognized their input, except in few cases, where the CA replied with a fallback answer. "*And it's not like this instant [response], like everyone gets the answer right after their message [...], maybe he actually waited to make it more natural. So there's a lot of thought process behind that.*" One facilitator suggested that few funny answers from the CA might be appropriate. Several participants further wish that there were more categories in the clusters. Regarding the support to focus on the actual value-adding task, both positive and negative aspects were identified. For the former, participants discussed that the way in which CA helped focus on value-adding task is by leading the task, moderating, saving input, classifying, "*taking notes in the background.*"; "*So we didn't have to think about all the other stuff coming, but just focus on the [idea gathering] phase.*" Paraphrased from a facilitator: "*[...] the benefit is that I can focus on the content, since I do not have to explain things in-between. Especially when there are many people in the workshop, and some repeatedly ask for the same instruction.*" Few participant argued that the benefit was hard to see in this short user test, but that they definitely see the potential of CA being helpful with the mindmaps, wordclouds and clusters to orient the process. On the negative side, participants mentioned the fast pace of the assignment, which acted distracting: "*It was a bit difficult because there was so much input from different users and it was hard to see who said what.*"

DP3 deals with distilling and enriching information by saving, processing and providing teams' input in a structured manner. Most participants valued that by saving their input, the CA supported by taking notes in the background, saved time, and structured the input. A facilitator further praised the CA for saving the arguments from previous phase regardless of the fact that other teammates already started with a new phase. On the other hand, few participants argued that the saving function did not support much in this context, since the exercise was short. Had it been longer, the potential of this function would be more visible. One of the means for distilling information the CA used were clusters (see Fig. 3). This was a highly praised feature by both user groups, but facilitators especially emphasized its helpfulness: "*really strong very, very surprised that it was so fast and so good, kind of summarized.*" One participant said the categories inspired him by providing a summarized overview of the covered topics, hinting that the team should take in regard other topics as well. A discussion was led about whether there should have been more categories in the cluster. However, a suggestion to think about uncovered topics by CA would have been helpful. The CA used wordcloud and mindmap as another means for distilling and enriching information. Both visualizations were differently perceived by the participants: while some found the wordcloud helpful and inspiring "*I think word cloud, for example, could be a good source of inspiration for further ideas to brainstorm.*", others preferred the mindmap, and vice versa. While some participants didn't understand the idea behind the mindmap, mostly because they did not use its implications to subsequently choose the best idea, due to very fast proceeding in the next phase (chosen by participants themselves); others could make several conclusions with the help of it. For example, they could see which topics might be more relevant, due to being mentioned more frequently than others, or they could also see the relations between different levels: ”*It was interesting to see that for the effects, we had 3 clusters (agriculture, population, and politics) and for the measures, we had only agriculture and politics – meaning the measures have to be taken on another level to account for different effects*.” Some participants further mentioned that they would like to have another level on the mindmap where all individual contributions are also shown, with a possibility to blend this in/out. In a discussion whether the usage of the mindmap should be explained, a facilitator stated ”*There should be a task in this phase so to say, hey, look at the content here, does it make you think of something regarding the importance [of shown topics] [...] it should be an explicit task*.” On the other hand, other participant raised concerns that this might introduce bias.

DP4 was evaluated by asking about CA’s human-likeness and stimulation to generate ideas. Participants agreed that the CA was not human-like, since its answers were fast, very structured and repetitive. However, this mostly did not negatively affect the participants: ”*it was bot-like, but that didn’t affect me*.” They also agree that CA’s behavior mostly did not stimulate idea generation: ”*so the summaries, if at all, stimulated me a bit, otherwise mindmap and wordcloud have the potential, but here it wasn’t the case*”. Further, participants say ”*it was more of taking notes in the background*”, ”*being a moderator, process assistance*.”

Besides evaluating DPs, we also performed a SWOT analysis of the prototype to carve out its strengths, weaknesses, opportunities, and threats. The most frequently mentioned points are displayed in the Fig. 4.

Strengths	**Weaknesses**
• Clusters the content quickly and understandable (n= 10) • Different displays of summarized information (n=8) • Stimulates new ideas (n=7) • Clear structure and organization (n=6) • Guided support (n=4) • Time-efficient (n=4)	• Repetitive answers (n=9) • Not human-like (n=6) • Overload of information (n=6) • Advancing to next phase decided by one team member (n=4) • Time pressure (n=3) • Cannot recognize keywords not related to the subject (n=2)
Opportunities	**Threats**
• Source of inspiration (n=7) • Moderation assistance (n=5)	• Clustering might introduce bias due to pre-conceived topics (n=4) • People don't interact with each other (n=4) • Can make users more passive in the process(n=2)

Figure 4. SWOT Analysis Results.

Discussion

Our study responds to a pressing need in the realm of collaborative creative work. As the volume of unstructured data generated during ideation processes continues to rise, capturing valuable insights becomes increasingly challenging (Chen et al. 2021). Our research offers a solution by proposing design principles for a CA that serves as a (co-)facilitator, documenter, and summarizer of information. To address the RQ on how to design and implement a CA to augment group ideation, we designed, developed and evaluated a CA along four proposed DPs. In the discussion, we reflect on the artifact itself before discussing it in the context of prior works and connecting our findings to scholarly debates.

Recalling the DPs, DP1 addresses task and process assistance. The evaluation shows that this DP displays pretty high conformance with the artefact. It could however be extended to account for individual differences in understanding task descriptions, by e.g., adding a read more button, so as to not overwhelm everyone with long texts. Further, team members should have the ability to blend in/out the contributions of others, in case these are overwhelming them. Facilitators further agreed that the goal of the exercise should be explained too, along with some examples.

DP2 focuses on natural language communication and partly conforms with the artefact. While users' input was mostly processed correctly, CA reactions to single ideas were not always satisfactory. Student participants suggested either no feedback or content-dependent feedback to ideas. Facilitators also argued for content-dependent feedback and questioning of contributions. While this is necessary, we argue that questioning and providing context-dependent feedback would need excessive training and could still turn out to be unsatisfactory, which is why these kinds of tasks should be left for the human facilitator. This aligns with the principles of Hybrid Intelligence (Dellermann et al. 2019b), which state that the tasks in a human-AI ensemble should be split according to respective strengths and skills. Our idea for the CA is not to automate creative sessions, but rather augment with its (e.g., data processing) abilities. To improve the natural communication, DP2 should further be extended with a response delay throughout the session. Previous research has shown that dynamic delay calculated based on the complexity of the response and complexity of the previous message increases the perception of social presence and overall satisfaction with the interaction (Gnewuch et al. 2018).

DP3 represents the core functionalities of the CA which aims to distill and enrich information. This DP mostly conforms with the artefact, but the short nature of the user test and conflicting team dynamics did not allow for its proper usage. However, the potential of the respective DFs was recognized. Of the three types of summaries provided by the CA, participants preferred the clusters, especially facilitators. More categories were wished for, however this would again create a clutter of information and go against the principles related to distilling information (Mygland et al. 2021). Some participants needed a "user manual" for the mindmaps, hereby a facilitator argues that its logic should be explicitly explained. Others argue that this would introduce bias. This opens an interesting avenue for further research.

DP4 does not seem to conform with the prototype. It was neither perceived as human-like, nor did it stimulate idea generation. During focus groups, participants mentioned that the lack of human-likeness did not affect the interaction negatively. The CA was perceived as a note taker, moderator and process assistance. This implicates that perceived competence might be more relevant than perceived anthropomorphism, and goes in line with the findings of Ciechanowski et al. 2019, who found that too much human-likeness produces uncanny valley effects (Mori et al. 2012), and less negative affect in cooperation with a simpler text chatbot. However, the results of the SWOT analysis imply that CA's lack of human-likeness was perceived as a weakness (n=6). Given its role as a co-moderator, it is highly relevant to uncover further insights on this debating topic via further studies and use as input for further DSR cycles.

Results of the SWOT analysis reflect the assumptions of Hybrid Intelligence (Dellermann et al. 2019b). While the CA performed well in analyzing and structuring data, it displayed repetitive patterns and did not account for team dynamics, which might be an indication for the strength of the CA to support task interventions and its limited potential to facilitate interaction interventions (Debowski et al. 2021). Therefore, based on the strengths (Figure 4) we can deduce to augment task interventions (Dickson et al. 1996) for the facilitation activities process and task of the facilitation framework (Bostrom et al. 1993) with CAs and handover the interaction interventions (Dickson et al. 1996) for the relationship activity (Bostrom et al. 1993) to human facilitators to enable Hybrid Intelligence according to (Dellermann et al. 2019b). This is another indication that the CA should serve as an augmentation to human facilitator instead of an automation, if applicable.

Besides reflecting on the artefact's value, we want to embed our research in the context of prior literature. Our research extends the range of contexts in which AI is applied. It introduces the concept of AI-powered facilitation within creative team environments, thereby advancing the scope frontier (Berente et al. 2021). By compensating for the limits of human memory and providing cognitive assistance, our CA contributes to the evolving landscape of AI applications. In terms of endowments from the delegation framework to agentic IS artifacts (Baird & Maruping 2021), our artifact compensates for the limits of human memory by aggregating and distilling data using the mechanisms of cognitive appraisal by e.g., taking over a task when humans have no capacity to complete it at current time in such fast manner. Furthermore, the insights from this work could serve as a starting point for identifying the willingness-to-delegate in the specific context of creative teamwork (Baird & Maruping 2021).

Moreover, there is a tradition of human-AI collaboration with CAs (e.g., Poser et al. 2022; Sowa et al. 2021).

With this study we show that CAs have a potential to successfully enrich facilitation of collaborative creative work and thereby co-create value by by applying task interventions such as providing assistance with facilitation, distilling and enriching information. In this way, CAs can take over tedious tasks of documentation and summarization of information as well as task explanations and process assistance. In line with Stoeckli et al. 2018, our CA also increases the expressive power of textual communication.

From the specific context of facilitating ideation (Poser et al. 2022; Strohmann et al. 2017), we also learned that it is crucial to take team dynamics into account and elaborate on the logic behind specific CA functions, as the ability of technology to co-create value is contingent on the levels of social connectivity between the human users (Breidbach et al. 2013). Future studies could explore whether additional human moderation of team dynamics would lead to even better actualization of CA functions on the side of process assistance and information enrichment.

Finally, we exemplify how the use of autonomous tools opens unprecedented opportunities for creative problem solving (Seidel et al. 2018). Our research lays the foundation for future work designs that incorporate AI-powered CAs as collaborative tools. It exemplifies the relevance of design theories in anticipating the implications of autonomous agents in the context of future work (Kane et al. 2021). Our study is not just about introducing a technical artifact but about reshaping how work is accomplished through human-AI collaboration.

Contribution and Conclusion

Collaborative work produces a lot of unstructured data and often many insights get lost. To address this issue, we proposed principles for a CA to support human facilitators by taking over the part of documenting and summarizing information. Following an established DSR methodology (Peffers et al. 2007), this study builds upon prior knowlege (Möller et al. 2022) and user requirements to provide initial evidence toward proof of concept, by creating an instantiation of the CA and then evaluating its potential. The results reflect broad potential for CAs in supporting creative collaborative work.

Our work contributes to both theory and practice: for the former, we provide prescriptive knowledge in form of DRs, DPs, and DFs for a CA that facilitates value co-creation in group ideation. With and beyond that, we add to the knowledge base on the topic of human-AI collaboration in creative work and CAs in group settings. Integrating CAs for augmentation in organizational teams paves the path towards novel work scenarios. Thereby, we link our study to the research stream on future of work and we illustrate how a CA can co-create value alongside human teams. For the latter, we provide a technical artefact that is reusable as well as adaptable. It can be integrated with other channels, e.g., Microsoft Teams, custom summarization features can be reused or integrated in another software, such as intelligent dashboards or whiteboards (more frequently used tool for digital collaborative creative teamwork). For users, this artefact can be transferred to different contexts and support human facilitators. For developers working on similar technical solutions, it provides a guidance on how to support ideation processes.

Our research comes with limitations. Even though we tested the prototype with experts, we could not observe a role-out in an organization with respective boundary conditions and contextual factors. With performing a short user test, we did not account for team dynamics that affected the interaction in a way that not all features were fully exploited. Moreover, our artifact is designed to facilitate text-based ideation sessions and, thereby, it represents one specific instance of a broader phenomenon. However, at the current state, it does not take voice-based ideation sessions into account. To account for these limitations, it would be interesting to perform a case study in an organization with a real business problem to better understand team dynamics and the influence of the organizational context. Furthermore, we see great value in quantifying the results achieved with the help of the CA. Hereby, one could move from a proof of concept to a proof of value by experimentally comparing results of ideation sessions with and without a CA. Adding another layer of human facilitation and forming a Hybrid Intelligence (Dellermann et al. 2019b) is another relevant research venue that could be born out of this research. Finally, research beyond this proof of concept is needed to explore CA's potential to augment facilitation. The path from conceiving new systems to creating standalone value involves many steps and can require many studies (Nunamaker et al. 2017). In the process of designing a CA for augmented facilitation and testing a limited prototype, we did not demonstrate that the CA concept has identified and overcome all significant obstacles and is prepared for wider deployment. Instead, this study has offered a proof of concept, which includes outlining the issue and a potential remedy as well as finding the evidence that the concept has potential, albeit only at a limited scope. This foundation should be improved upon by subsequent studies.

Acknowledgements

This research was funded by the German Federal Ministry of Education and Research (BMBF), in the context of the project HyMeKI (01IS20057). We thank the students Florentin Hitzke, Daniel Rupprecht, Fabian Burkhardt, and Niclas Trelle for their support with the literature review, conducting and analyzing the expert interviews, and designing and evaluating the prototype in the first iteration.

References

Anantrasirichai, N. & Bull, D. (2022). Artificial intelligence in the creative industries: a review. *Artificial intelligence review* (*55*), 589–656.

Arinez, J. F., Chang, Q., Gao, R. X., Xu, C. & Zhang, J. (2020). Artificial intelligence in advanced manufacturing: Current status and future outlook. *Journal of Manufacturing Science and Engineering* (*142*:11).

Baird, A. & Maruping, L. M. (2021). The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts. *MIS quarterly* (*45*:1).

Baptista, J., Stein, M.-K., Klein, S., Watson-Manheim, M. B. & Lee, J. (2020). Digital work and organisational transformation: Emergent Digital/Human work configurations in modern organisations. *The Journal of Strategic Information Systems* (*29*:2), 101–618.

Benke, I. (2020). Towards Design Principles for Trustworthy Affective Chatbots for Virtual Teams. In: *ECIS 2020 Proceedings*.

Beranek, P. M., Beise, C. M. & Niederman, F. (1993). Facilitation and group support systems. In: *Proceedings of the Twenty-sixth Hawaii International Conference on System Sciences*. Vol. 4, pp. 199–207.

Berente, N., Gu, B., Recker, J. & Santhanam, R. (2021). Managing artificial intelligence. *MIS quarterly* (*45*:3).

Bittner, E., Küstermann, G. & Tratzky, C. (2019)a. The Facilitator is a Bot: towards a Conversational Agent for Facilitating Idea Elaboration on Idea Platforms. In: *Proceedings of the 27th European Conference on Information Systems (ECIS)*.

Bittner, E., Mirbabaie, M. & Morana, S. (2021). Digital Facilitation Assistance for Collaborative, Creative Design Processes. In: *Proceedings of the 54th Hawaii International Conference on System Sciences*.

Bittner, E., Oeste-Reiss, S., Ebel, P. & Söllner, M. (2019)b. Mensch-Maschine-Kollaboration: Grundlagen, Gestaltungsherausforderungen und Potenziale für verschiedene Anwendungsdomänen. *HMD Praxis der Wirtschaftsinformatik* (*56*:1), 34–49.

Bittner, E., Oeste-Reiss, S. & Leimeister, J. M. (2019)c. Where is the Bot in our Team? Toward a Taxonomy of Design Option Combinations for Conversational Agents in Collaborative Work. In: *Proceedings of the 52nd Hawaii International Conference on System Sciences*.

Bittner, E. & Shoury, O. (2019). Designing Automated Facilitation for Design Thinking: A Chatbot for Supporting Teams in the Empathy Map Method. In: *Proceedings of the 52nd Hawaii International Conference on System Sciences*.

Bostrom, R. P., Anson, R. & Clawson, V. K. (1993). Group facilitation and group support systems. *Group support systems: New perspectives* (*8*), 146–168.

Brandtzæg, P. B. & Følstad, A. (2018). Chatbots: changing user needs and motivations. *Interactions* (*25*), 38–43.

Breidbach, C. F., Kolb, D. G. & Srinivasan, A. (2013). Connectivity in Service Systems: Does Technology-Enablement Impact the Ability of a Service System to Co-Create Value? *Journal of Service Research* (*16*:3), 428–441.

Busse, M. (2021). Maschinelles Lernen vs. Regelbasierte Erkennung - was ist Besser Geeignet für chatbots? https://www.kiko.bot/blog/allgemein/maschinelles-lernen-vs-regelbasierter-chatbot/, Accessed on 20.05.2022.

Chen, R., Demko, M., Byrne, D. & Louw, M. (2021). Probing Documentation Practices: Reflecting on Students' Conceptions, Values, and Experiences with Documentation in Creative Inquiry. In: *Creativity and Cognition*. Association for Computing Machinery.

Ciechanowski, L., Przegalinska, A., Magnuski, M. & Gloor, P. (2019). In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems* (*92*), 539–548.

Cvetkovic, I. & Bittner, E. (2022). Task Delegability to AI: Evaluation of a Framework in a Knowledge Work Context. In: *Proceedings of the 55th Hawaii International Conference on System Sciences*.

Cvetkovic, I., Rosenberg, V. & Bittner, E. (2023). Conversational Agent as a Black Hat: Can Criticising Improve Idea Generation? In: *Proceedings of the 56th Hawaii International Conference on System Sciences*.

Daly, S. R., Seifert, C. M., Yilmaz, S. & Gonzalez, R. (2016). Comparing Ideation Techniques for Beginning Designers. *ASME Journal of Mechanical Design* (*138*:10), 101–108.

Debowski, N., Tavanapour, N. & Bittner, E. (2021). Einsatz eines virtuellen Kollaborators in analogen & digitalen Workshops im organisationalen Kontext. *Informatik Spektrum* (*44*), 170–177.

Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S. & Ebel, P. (2019)a. The future of human-AI collaboration: a taxonomy of design knowledge for hybrid intelligence systems. In: *Proceedings of the 52nd Hawaii International Conference on System Sciences*.

Dellermann, D., Ebel, P., Söllner, M. & Leimeister, J. M. (2019)b. Hybrid Intelligence. *Business & Information Systems Engineering* (*61*:5), 637–643.

Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *ArXiv* (*abs/1810.04805*).

Dickson, G. W., Lee-Partridge, J. E., Limayem, M. & Desanctis, G. L. (1996). Facilitating computer-supported meetings: A cumulative analysis in a multiple-criteria task environment. *Group Decision and Negotiation* (*5*), 51–72.

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

Elshan, E., Zierau, N., Janson, A. & Leimeister, J. M. (2022). How Conversational Agents Relieve Teams from Innovation Blockages. *Academy of Management Proceedings* (*2022*:1).

Friesike, S., Flath, C. M., Wirth, M. & Thiesse, F. (2019). Creativity and productivity in product design for additive manufacturing: Mechanisms and platform outcomes of remixing. *Journal of Operations Management* (*65*:8), 735–752.

Gnewuch, U., Morana, S., Adam, M. T. P. & Maedche, A. (2018). Faster is Not Always Better: Understanding the Effect of Dynamic Response Delays in Human-Chatbot Interaction. In: *Proceeding of the 26th European Conference on Information Systems (ECIS)*.

Gnewuch, U., Morana, S. & Maedche, A. (2017). Towards Designing Cooperative and Social Conversational Agents for Customer Service. In: *Proceedings of the 38th International Conference on Information Systems (ICIS)*.

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

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

Kane, G. C., Young, A. G., Majchrzak, A. & Ransbotham, S. (2021). Avoiding an oppressive future of machine learning: A design theory for emancipatory assistants. *Mis Quarterly* (*45*:1), 371–396.

Kim, S., Eun, J., Oh, C., Suh, B. & Lee, J. (2020). Bot in the bunch: Facilitating group chat discussion by improving efficiency and participation with a chatbot. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–13.

Kim, S., Eun, J., Seering, J. & Lee, J. (2021). Moderator chatbot for deliberative discussion: Effects of discussion structure and discussant facilitation. *Proceedings of the ACM on Human-Computer Interaction* (*5*:CSCW1), 1–26.

Knote, R., Janson, A., Söllner, M. & Leimeister, J. M. (2020). Value co-creation in smart services: a functional affordances perspective on smart personal assistants. *Journal of the Association for Information Systems* (*22*:2), 418–458.

Koch, J., Taffin, N., Beaudouin-Lafon, M., Laine, M., Lucero, A. & Mackay, W. E. (2020). ImageSense: An Intelligent Collaborative Ideation Tool to Support Diverse Human-Computer Partnerships. *Proceedings of the ACM on Human-Computer Interaction* (*4*:CSCW1).

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

Leimeister, J. M. (2014). Collaboration Engineering: IT-gestützte Zusammenarbeitsprozesse systematisch entwickeln und durchführen. Springer Berlin Heidelberg.

Li, M. M., Bronner, E., Peters, C. & Leimeister, J. M. (2020). Künstliche Intelligenz und menschliche Kompetenz zur Automatisierung und Personalisierung von Dienstleistungen am Beispiel des Support. In: *Automatisierung und Personalisierung von Dienstleistungen: Methoden – Potenziale – Einsatzfelder*. M. Bruhn & K. Hadwich (eds.). Springer Fachmedien Wiesbaden, pp. 235–251.

Lubars, B. & Tan, C. (2019). Ask not what AI can do, but what AI should do: Towards a framework of task delegability. In: *Advances in Neural Information Processing Systems*. Vol. 32.

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

Mahmud, B. U., Hong, G. Y. & Fong, B. (2022). A Study of Human-AI Symbiosis for Creative Work: Recent Developments and Future Directions in Deep Learning. *ACM Trans. Multimedia Comput. Commun. Appl.* (*Just Accepted*).

Mayring, P. & Fenzl, T. (2019). Qualitative inhaltsanalyse. In: *Handbuch Methoden der empirischen Sozialforschung*. Springer, pp. 633–648.

Milana, C. & Ashta, A. (2021). Artificial intelligence techniques in finance and financial markets: a survey of the literature. *Strategic Change* (*30*:3), 189–209.

Mockenhaupt, A. (2021). Digitalisierung und künstliche Intelligenz in der Produktion. Springer Wiesbaden.

Möller, F., Strobel, G., Schoormann, T. & Hansen, M. (2022). Unveiling the Cloak: Kernel Theory Use in Design Science Research. In: *Proceedings of the 43rd International Conference on Information Systems (ICIS)*.

Mori, M., MacDorman, K. F. & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & automation magazine* (*19*:2), 98–100.

Moussawi, S. (2018). User experiences with personal intelligent agents: a sensory, physical, functional and cognitive affordances view. In: *Proceedings of the 2018 ACM SIGMIS Conference on Computers and People Research*, pp. 86–92.

Mygland, M. J., Schibbye, M., Pappas, I. O. & Vassilakopoulou, P. (2021). Affordances in human-chatbot interaction: a review of the literature. In: *Conference on e-Business, e-Services and e-Society*. Springer, pp. 3–17.

Nunamaker, J. F., Twyman, N. W., Giboney, J. S. & Briggs, R. O. (2017). Creating High-Value Real-World Impact through Systematic Programs of Research. *MIS Quarterly* (*41*:2), 335–351.

Oeste-Reiß, S., Bittner, E., Cvetkovic, I., Günther, A., Leimeister, J. M., Memmert, L., Ott, A., Sick, B. & Wolter, K. (2021). Hybride Wissensarbeit. *Informatik-Spektrum* (*44*), 148–152.

Peffers, K., Tuunanen, T., Rothenberger, M. A. & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems* (*24*:3), 45–77.

Poser, M., Küstermann, G. C., Tavanapour, N. & Bittner, E. A. C. (2022). Design and Evaluation of a Conversational Agent for Facilitating Idea Generation in Organizational Innovation Processes. *Information Systems Frontiers* (*24*), 771–796.

Przybilla, L., Baar, L., Wiesche, M. & Krcmar, H. (2019). Machines as Teammates in Creative Teams: Digital Facilitation of the Dual Pathways to Creativity. In: *Proceedings of the 2019 on Computers and People Research Conference*. Association for Computing Machinery, pp. 94–102.

Putman, V. L. & Paulus, P. B. (2009). Brainstorming, brainstorming rules and decision making. *The Journal of creative behavior* (*43*:1), 29–40.

Radziwill, N. M. & Benton, M. C. (2017). Evaluating Quality of Chatbots and Intelligent Conversational Agents. arXiv: 1704.04579 [cs.CY].

Ramesh, K., Ravishankaran, S., Joshi, A. & Chandrasekaran, K. (2017). A survey of design techniques for conversational agents. In: *International conference on information, communication and computing technology*. Springer, pp. 336–350.

Rasa Technologies Inc (2021). RASA: Open source conversational AI. https://rasa.com/.
Schallmo, D. (2018). Jetzt Design Thinking anwenden: In 7 Schritten zu kundenorientierten Produkten und Dienstleistungen. In: Wiesbaden: Springer Fachmedien Wiesbaden.
Seidel, S., Berente, N., Lindberg, A., Lyytinen, K. & Nickerson, J. V. (2018). Autonomous tools and design: a triple-loop approach to human-machine learning. *Communications of the ACM* (*62*:1), 50–57.
Seric, A. & Winkler, D. (2020). COVID-19 could spur automation and reverse globalisation–to some extent. https://shorturl.at/rNQ04, Accessed on 18.06.2022.
Siemon, D. (2022). Elaborating team roles for artificial intelligence-based teammates in human-AI collaboration. *Group Decision and Negotiation* (*31*:5), 871–912.
Singh, A., Ramasubramanian, K. & Shivam, S. (2019). Building an enterprise chatbot: Work with protected enterprise data using open source frameworks. O'Reilly.
Sowa, K., Przegalinska, A. & Ciechanowski, L. (2021). Cobots in knowledge work: Human–AI collaboration in managerial professions. *Journal of Business Research* (*125*), 135–142.
Stoeckli, E., Uebernickel, F. & Brenner, W. (2018). Exploring affordances of slack integrations and their actualization within enterprises-towards an understanding of how chatbots create value. In: *Proceedings of the 51st Hawaii International Conference on System Sciences*.
Strohmann, T., Fischer, S., Siemon, D., Brachten, F., Lattemann, C., Robra-Bissantz, S. & Stieglitz, S. (2018). Virtual Moderation Assistance: Creating Design Guidelines for Virtual Assistants Supporting Creative Workshops. In: *PACIS 2018 Proceedings. 80*.
Strohmann, T., Siemon, D. & Robra-Bissantz, S. (2017). brAInstorm: Intelligent Assistance in Group Idea Generation. In: *Designing the Digital Transformation*. A. Maedche, J. vom Brocke & A. Hevner (eds.). Cham: Springer International Publishing, pp. 457–461.
Tavanapour, N. & Bittner, E. (2018). Automated Facilitation for Idea Platforms: Design and Evaluation of a Chatbot Prototype. In: *Proceedings of the International Conference on Information Systems (ICIS)*.
Tepper, N., Hashavit, A., Barnea, M., Ronen, I. & Leiba, L. (2018). Collabot: Personalized Group Chat Summarization. In: *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. Association for Computing Machinery, pp. 771–774.
Toxtli, C., Monroy-Hernández, A. & Cranshaw, J. (2018). Understanding chatbot-mediated task management. In: *Proceedings of the 2018 CHI conference on human factors in computing systems*, pp. 1–6.
Tremblay, M. C., Hevner, A. R. & Berndt, D. J. (2010). The Use of Focus Groups in Design Science Research. In: *Design Research in Information Systems: Theory and Practice*. Springer US, pp. 121–143.
Venable, J., Pries-Heje, J. & Baskerville, R. (2016). FEDS: a framework for evaluation in design science research. *European journal of information systems* (*25*:1), 77–89.
Voigt, M. & Bergener, K. (2013). Enhancing Creativity in Groups - Proposition of an Integrated Framework for Designing Group Creativity Support Systems. In: *Proceedings of the 46th Hawaii International Conference on System Sciences*, pp. 225–234.
Waizenegger, L., Seeber, I., Dawson, G. & Desouza, K. (2020). Conversational agents-exploring generative mechanisms and second-hand effects of actualized technology affordances. In: *Proceedings of the 53rd Hawaii international conference on system sciences*.
Willman, J. (2021). Creating a Chatbot. In: *Modern PyQt*. Springer, pp. 261–282.
Winkler, R., Neuweiler, M. L., Bittner, E. & Söllner, M. (2019). Hey Alexa, Please Help Us Solve This Problem! How Interactions with Smart Personal Assistants Improve Group Performance. In: *Proceedings of the International Conference of Information Systems (ICIS)*.