Designing Speech-Based Assistance Systems: The Automation of Minute-Taking in Meetings

Research Paper

Anton Koslow[1], Benedikt Berger[2]

[1] LMU Munich School of Management, Munich, Germany
koslow@lmu.de
[2] University of Münster, Department of Information Systems, Münster, Germany
benedikt.berger@uni-muenster.de

Abstract. Advances in speech recognition and processing allow speech-based assistant systems (SBAS) to support or fully automate an increasing number of human tasks. Deploying SBAS in organizations promises economic benefits but may also hold unintended drawbacks of automation, such as negative impacts on employees' competences and professional identity. Using the example of taking meeting minutes, we investigate how SBAS should be designed to balance these benefits and drawbacks. We developed a prototype of an SBAS for minute-taking in online meetings and evaluated its use in an online study. The results show that a higher level of automation improves the capture and processing of information from online meetings but has adverse effects on minute-takers' satisfaction and identification with their work. We derive six design requirements from the results and subsume them under two design principles that describe how automation by SBAS can enhance human work.

Keywords: Automation, speech, digital assistants, design science.

1 Introduction

Advances in artificial intelligence (AI) change and improve how information technology (IT) systems process data, information, and knowledge, enabling the automation of tasks with greater accuracy and efficiency (Raisch & Krakowski, 2021). Modern natural language processing (NLP) applications such as ChatGPT can answer complex queries, summarize documents, and generate texts with writing skills comparable to humans (Rudolph et al, 2023). Based on their capabilities to understand and process human speech (Riedl, 2019), we refer to such IT systems as speech-based assistance systems (SBAS). In an organizational context, SBAS offer the possibility to automate or support time-consuming and labor-intensive tasks, such as writing emails, handling customer inquiries, or transcribing and summarizing meetings (Adam et al, 2021; Jarrahi, 2019).

20th International Conference on Wirtschaftsinformatik,
September 2025, Münster, Germany

While automation using AI systems such as SBAS can potentially increase efficiency and productivity, it can also have unintended consequences and drawbacks (Schwabe & Castellacci, 2020). High levels of automation can limit employees' autonomy, leading to employee de-skilling (Rashid et al, 2017) and reduced identification with work outcomes (Leung et al, 2018). If AI systems threaten employees' skills or professional identity, employees are less inclined to participate in developing and adapting AI systems (Ivanov et al, 2020). This can lead to a lower sense of organizational engagement and accountability (Jarrahi, 2019). To reduce these risks, companies can embed AI systems into workflows while trying to preserve and enhance human work elements (Raisch & Krakowski, 2021). Accordingly, the design of SBAS should seek to reap the benefits while avoiding the pitfalls of automation. Our study contributes to achieving this balance and to answer the following research question (RQ):

RQ: *How can the design of speech-based assistance systems balance the benefits and drawbacks of automation?*

This study follows the design science research methodology (DSRM) by Peffers et al (2007) to investigate the design of SBAS using the example of minute-taking in meetings, a common organizational communication and decision-making task (Mroz et al, 2018). Drawing on four initial user requirements identified in a previous design cycle (Koslow et al, 2021) and informed by existing SBAS, we designed and evaluated an SBAS prototype for minute-taking. Our findings provide six requirements and two overarching design principles for SBAS that emphasize balancing speech recognition and processing capabilities as well as user integration in highly automated processes.

2 Effects of Automation in Organizations

Automation offers companies several benefits, such as lower costs due to the reduced need for manual labor and increased efficiency and productivity (Jarrahi, 2019). It can free employees for more meaningful work, reduce errors, and accelerate processes (Raisch & Krakowski, 2021; Romao et al, 2019). Advances in machine learning, NLP, and computer vision have enabled "intelligent" automation, i.e. AI systems performing cognitive tasks, involving learning, reasoning, and problem-solving, that were previously carried out by humans (Coombs et al, 2020). AI systems using automatic speech recognition and NLP, such as SBAS, can now process human speech (Riedl, 2019), which allows these systems to support or automate tasks involving conversations.

Automation in an organization often requires redesigning tasks and workflows in the day-to-day operations (Coombs et al, 2020). Besides increasing efficiency, reducing costs, and freeing up human labor for more complex tasks, the redesign of work can have negative consequences for employees and organizations (Parker & Grote, 2022; Raisch & Krakowski, 2021). Increasing standardization through automation can reduce organizations' flexibility and opportunities for learning (Parker & Grote, 2022). Reducing employees' work to micro-tasks with little room for creative problem solving can lower employee satisfaction and meaningful work (Schwabe & Castellacci, 2020). High levels of automation without sufficient employee consideration can result in them no longer seeing the contribution of their efforts and, therefore, no longer identifying with their work (Leung et al, 2018). If employees feel increasingly dispossessed in their work or even threatened by automation, they may adapt or reduce their IT system use (Ivanov et al, 2020). In turn, decreasing job satisfaction and performance motivation can reduce the quality of work results and undermine the intended efficiency gains (Spector, 1997). All in all, increasing levels of automation, which can be enabled by advances in AI capabilities for capturing and processing natural language, entail the risk of negative impacts on employees. To balance human performance and automation or automated support (Raisch & Krakowski, 2021), organizations should consider and involve their employees in automation initiatives.

3 Speech-Based Assistance System Design

3.1 Design Approach and Context

As a specific instance of intelligent automation, this study addresses the design of SBAS for minute-taking during meetings. Employees spend much time in meetings to share information and experiences, discuss work steps, processes, and outcomes, and make decisions (Mroz et al, 2018). Hence, meeting summarization is a common and vital component of organizational communication, collaboration, and decision-making (Standaert et al, 2021). Meeting documentation aims at capturing a meeting's essential information (Niforatos et al, 2018) and making it accessible to other employees for later review but constitutes a cumbersome task.

To investigate how SBAS should be designed to reduce efforts while avoiding unintended negative consequences for minute-takers, we follow the DSRM process by Peffers et al (2007). This structured approach has guided us through two design cycles, as illustrated in Figure 1. The first design cycle preceded this study and explored how SBAS capabilities can reduce cognitive and time efforts for recording and documenting business meetings (Koslow et al, 2021). For this purpose, we developed a prototype for transcribing and processing meeting conversations. We demonstrated and discussed this prototype in exploratory focus groups with experienced professionals who take meeting minutes regularly as part of their daily work (Koslow et al, 2021). Analyzing the focus group discussions on recurring issues during this task and on shortcomings of existing solutions for its support yielded the identification of 24 user needs and 23 user goals. We summarized these in needs and goals in ten technical capabilities of SBAS required to effectively assist minute-takers. The first four requirements, namely reliable speech recognition (R1), reliable speaker recognition (R2), summarization of meeting content (R3), and structuration of meeting minutes (R4), focus on system capabilities with high automation potential for assisting minute-takers (Koslow et al, 2021).

Providing automated, reliable speech recognition (R1), i.e., live transcription, should reduce the amount of manual effort in capturing meeting content. Applications like Zoom or Otter.ai offer live or post-hoc transcription of virtual meetings. Speech support has also been shown to reduce cognitive effort and increase work efficiency, such as in e-teaching, where live speech-to-text transcription in online lectures can assist students to follow more effectively (Shadiev et al, 2014). In meetings with multiple participants, identifying speakers is crucial for minute-takers, making reliable speaker recognition (R2) an important requirement, as it enables accurate attribution of statements in the minutes. Most video conferencing solutions detect speakers via user logins and audio activity, whereas applications such as Otter.ai use audio analysis for this purpose (Hansen & Hasan, 2015). With the automated transcription of oral communication in meetings comes the necessity of support in processing the resulting long transcripts. Therefore, summarizing meeting transcripts (R3) is another important requirement for SBAS. This capability should ensure that relevant details are provided to all meeting stakeholders, especially if they could not attend (Mroz et al, 2018). Applications such as Otter.ai already use generative AI capabilities to create such summaries. Furthermore, SBAS should be able to structure the meeting content (R4) based on inductive analyses or specific templates, such as meeting goals or agendas (Mroz et al, 2018).

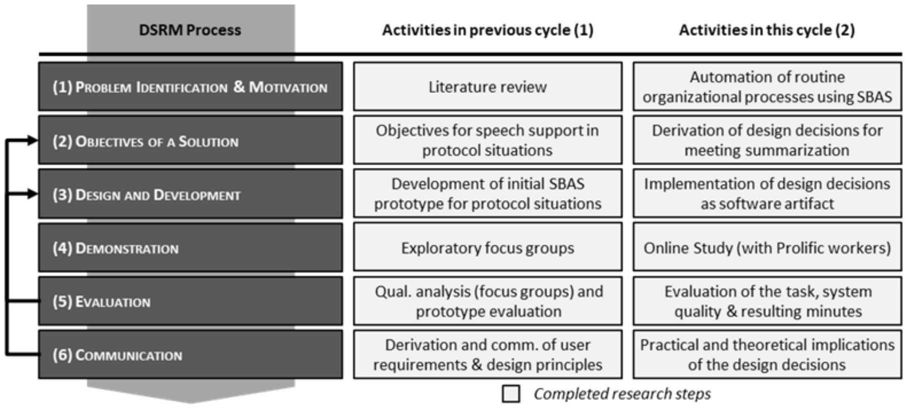

Figure 1. DSRM Process by Peffers et al (2007) and activities in this design cycle.

The second design cycle, presented in this paper, builds on and extends these insights. The (1) *problem identification & motivation* of this design cycle addresses the potential benefits and drawbacks of automation in minute-taking. Whether SBAS can reduce workload effectively depends on their ability to meet these four requirements, while still allowing for human review and correction. Thus, we investigate the effects of using an SBAS for minute-taking on minute-takers' work and their work perception.

3.2 Deriving Design Decisions

Our (2) *objective* is to investigate how SBAS can enhance minute-takers' performance or reduce their efforts while avoiding unintended consequences. Minute-takers perform four primary activities: a) capturing important meeting content; b) assigning the content to agenda topics; c) assigning speakers to the captured content; d) summarizing the meeting content in final minutes (Koslow et al, 2021; Mroz et al, 2018). To evaluate the required automation level of an SBAS to support these activities, we define three incremental support stages in minute-taking and implement each of them in a separate version of our SBAS prototype (see Table 1).

No automation (NA) support: This version represents the reference case of taking minutes without automation support, as is still common in analog on-site meetings. In this case, the minute-takers only have a pen and paper or a laptop/tablet with a word processing program to note key points during the meeting. The minute-taker usually finalizes the minutes after the meeting based on their notes and recollections (Niforatos et al, 2018).

Partial automation (PA) support: The second version represents partial automation support for minute-takers. This prototype provides automated live transcription (R1), enabling real-time monitoring and correction of speech input through speech-to-text processing (Kushalnagar et al, 2015). Additionally, this prototype partially automates the assignment of topics to the meeting agenda. While minute-takers must still decide on the subject of the current meeting discussion, ongoing statements are automatically assigned to selected topics on the agenda to support the structuring of the meeting minutes according to R4. However, speaker attribution and summarizing the meeting content based on the transcript to create the final meeting minutes remain manual tasks.

High automation (HA) support: To further increase the level of automation, we added features to the partially automated version. The high automation design of our SBAS prototype offers minute-takers additional support in identifying and assigning speakers. Once minute-takers have assigned a speaker to a captured statement, the prototype recognizes and assigns the speaker automatically in the further process, in line with R2. After the meeting, minute-takers receive further support in finalizing the minutes using a text summary feature according to R3. The system creates a summary proposal of the key discussion points, which minute-takers can review and revise.

Table 1. Levels of automation in the SBAS prototype versions

Task	No Automation (NA)	Partial Automation (PA)	High Automation (HA)
Content capture	Manual	Automated	Automated
Topic assignment	Manual	Assisted	Assisted
Speaker recognition	Manual	Manual	Assisted
Content summary	Manual	Manual	Assisted

3.3 Implementation of Design Decisions

We developed a prototype for the (3) *design and development* of an SBAS for minute-taking of online meetings. Building on a word processing web application, we created the three prototype versions with varying levels of automation designed to support the minute-taking process as outlined by Mroz et al (2018). Similar to videoconference invitations, a 'Prepare' interface provides an overview of the meeting details, such as schedule, participants, and agenda, in all prototype versions. From the 'Prepare' interface, users can move on to the start of the meeting.

Once the minute-taker joins the meeting, they access a ‘Take Minutes’ interface in the version without automation support. This interface features a text editor on the left for minute-taking and a video area on the right displaying the meeting to be recorded. After the meeting, users can finish and save the minutes in the text editor before exiting the application. Instead of the ‘Take Minutes’ interface, the prototype versions with partial and high automation support proceed with a ‘Recording’ interface. The left side of the ‘Recording’ interface shows the meeting video and a text field for ‘Preliminary Records’. During the meeting, spoken statements are continuously displayed in this field to help the minute-taker follow the video and read the live text. The output is limited to six lines, as recommended by Kushalnagar et al (2015). Completed statements move as chat-style entries to the ‘Transcript’ area on the right of the video.

Next to the ‘Preliminary Records’ field, the two prototype designs with automation assistance contain a list of current agenda topics. Selecting a topic automatically assigns all incoming statements to this agenda topic in the transcript. In the partially automated prototype version, users can correct statements in the transcript and manually assign speaker names to them. In contrast, the users of the high automation version can use simulated speaker recognition. However, users must initially assign speakers once by clicking on a placeholder (e.g., P1, P2, etc.) on the right side of a completed statement in the ‘Transcript’ area and selecting the appropriate name from a pop-up list with meeting participants. After the meeting, the users of both automated prototype versions can save the transcript and proceed to the following ‘Notes & Editing’ interface, where they can continue preparing the actual meeting minutes. In this interface, the ‘Transcript’ area moves to the left side of the application. The left side provides a ‘Notes’ text field with the same text editor as in the version without automation support. Users of the partial automation prototype can edit the ‘Notes’ field directly or copy and paste content from the transcript. While users of the partial automation prototype must summarize the meeting transcript manually, users of the high automation version get assistance by a ‘Create Notes’ feature. This feature generates a summary proposal in the ‘Notes’ field, which users can revise before saving the minutes and leaving the meeting.

3.4 Design Demonstration

For the (4) *demonstration* of our SBAS prototype with different levels of automation, we conducted an online vignette study, in which participants had to take minutes of a scripted fictional online meeting using our SBAS prototype for minute-taking. We mimicked the meeting using a 3.5-minute video recording of a fictive online business meeting. We cropped the video sourced from an online course on minute-taking and adapted it to the format of a Zoom meeting. All participants had to perform the same four main tasks, as summarized in Table 1. The study followed a 3 x 1 between-subject experimental design. The participants were assigned randomly to one of three automation groups (NA, PA, and HA). Before the interaction with one of the three prototype versions, a scenario description and video tutorials introduced the assigned prototype design and the functionality it offered in detail. These videos showed and explained the range of features provided to the group to perform the assigned tasks. Comprehension questions followed the tutorial videos.

We recruited 339 participants for the study on the crowdsourcing platform Prolific. As the actors in our recorded meeting spoke professional, easy-to-understand British English, we limited the distribution of our study to UK participants. Eight participants were excluded from the study due to reported major technical issues. The remaining 331 participants had a mean age of 37.6 years (SD = 12.734; range: 18–78), with 166 female (50.2%), 160 male (48.3%), and 5 non-binary (1.5%) respondents. 43.8% of the participants had no previous experience in taking meeting minutes, while 56.2% had at least three months or more of experience in total. Random assignment to the groups resulted in 111 participants in the NA group, 113 participants in the PA group, and 107 participants in the HA group. Each participant received a fixed £4.00 payment (£9.90/h) and we awarded a £4.00 bonus for the best minutes created in each group according to an external evaluation of the minutes.

3.5 Design Evaluation

Our (5) *evaluation* focuses on the participants' perception of the task, the system they used, and their satisfaction and identification with their contribution to the result. We also contrast the participants' responses with an external evaluation of the performed task and the quality of the minutes. The participants responded to questionnaires before and after the interaction with one of the prototype versions for minute-taking. The pre-study questionnaire started with an introduction to the study procedure, followed by a technical pre-check regarding the compatibility of the device used for the study. After the study description, questions on demographics (age, gender, degree, occupation) and participants' experience with minute-taking (none to more than three years) followed. The post-study questionnaire included measurement scales assessing the result, the task, the participant's performance, the minute application, the scenario, potential technical difficulties, and the overall study. The primary scales are perceived satisfaction with the result (Nicolaou & McKnight, 2011), perceived identification with the result (Leung et al, 2018), perceived task performance (Tsai et al, 2007), perceived cognitive effort for the task (Li et al, 2011), and perceived efficiency of the system (Vandenbosch & Ginzberg, 1996). We measured all constructs with three to five items each on a 7-point Likert scale from 1: "strongly disagree" or "very poor" to 7: "strongly agree" or "very good". Additionally, three student assistants conducted an external evaluation of the provided minutes. As suggested by Iskender et al (2021), we aligned the evaluation criteria through a joint meeting before the evaluators assessed the completeness (Lentz & De Jong, 1997) and relevance (Iskender et al, 2021) of the minutes. Additionally, the evaluators checked how well the participants assigned the statements to the topics and the speakers. For the external evaluation of the minutes, we used a 5-point Likert scale from 1: "very good" to 5: "very poor" for each criterion.

4 Results

Before the main analysis, we used chi-square tests to ensure that the groups did not statistically differ (all $p > 0.1$) in our nominal control variables (gender, education, occupation, and experience). To assess group differences in our metric control variables, we conducted one-way Analyses of Variances (ANOVAs). The results do not show significant differences between groups in age, tutorial quality (Gallivan et al, 2005), perceived system complexity (Geissler et al, 2001), perceived system control (Collier & Sherrell, 2010), perceived cognitive effort for the system use (Bulgurcu et al, 2010), and perceived cognitive effort through IT overload (Ahuja & Thatcher, 2005), with $p = 0.079$ for the latter and $p > 0.1$ for all other constructs. These results indicate that our variation in automation did not have unintended effects on the cognitive effort of the participants.

The external evaluation of the meeting minutes aims to assess how well the study participants fulfilled their assigned tasks, including assigning important statements to the corresponding speakers and agenda topics and summarizing all essential and relevant content in the final minutes. The results in Table 2 show significant main effects of the level of automation on all four rating variables (all $p \leq 0.001$). Post-hoc pairwise comparisons with Bonferroni corrections show that, for the speaker assignment task, meeting minutes produced with high automation resulted in better results than minutes produced with partial or no automation (both $p < 0.001$). In contrast, partial automation did not significantly improve speaker assignments compared to no automation ($p = 0.265$), but it did significantly enhance the assignment of statements to agenda topics ($p < 0.001$). The other two pairwise comparisons of topic assignment scores show no significant differences ($p > 0.05$). The results further show that minute contents can be captured more completely with increasing automation (all $p < 0.001$). We also observe significant differences in evaluating the relevance of the minutes' contents for all prototype versions (all $p < 0.001$). However, the assessment of the relevance of the minutes content decreases significantly for the version with partial automation support and only improves significantly with high automation support.

As summarized in Table 3, the analysis of the results of the user evaluations based on ANOVAs shows significant effects of automation on users' satisfaction ($p < 0.001$) and identification ($p = 0.02$) with the resulting minutes as well as perceived efficiency of the system ($p < 0.001$). The influence of automation levels on perceived task performance and cognitive effort for the task is also significant (both $p < 0.001$). Satisfaction with the resulting minutes is significantly higher in the prototype's partial and high automation versions compared to the no automation version. However, identification with the resulting minutes is significantly higher for partial automation than for high automation ($p = 0.016$). The perceived task performance is significantly higher with automation support than without automation support (both $p < 0.001$). At the same time, assessments of perceived cognitive effort decrease significantly with increasing levels of automation (all $p < 0.05$). Lastly, users perceived the efficiency of the system to be significantly higher in both automated versions than in the non-automated one, with no difference between partial and high automation.

Table 2. External evaluation of the meeting minutes.

Main effects of the level of automation (between-subjects effects)		**Pairwise comparisons between levels of automation**			
		NA (avg.)	**PA** (avg.)	**HA** (avg.)	Significance of Differences (p)
Speaker Assignment Task	F = 24.165 p < 0.001	2.679	2.983	1.771	NA to PA (0.265) NA to HA (< 0.001) PA to HA (< 0.001)
Topic Assignment Task	F = 6.764 p = 0.001	1.806	1.379	1.641	NA to PA (< 0.001) NA to HA (0.514) PA to HA (0.084)
Completeness of the Minutes Content	F = 159.178 p < 0.001	3.239	2.138	1.206	NA to PA (< 0.001) NA to HA (< 0.001) PA to HA (< 0.001)
Relevance of the Minutes Content	F = 51.309 p < 0.001	1.930	2.437	1.356	NA to PA (< 0.001) NA to HA (< 0.001) PA to HA (< 0.001)
Note: NA (No Automation); PA (Partial Automation); HA (High Automation). Mean values on a 5-point Likert scale (from 1: "very good" to 5: "very poor").					

5 Discussion and Design Principles

The evaluation of our SBAS prototype shows that increasing levels of automation can reduce cognitive effort and partially improve result quality in minute-taking. However, more automation does not linearly increase performance and satisfaction with work results. Our findings call for thoughtful user integration when designing highly automated systems. Based on this, we reflect on the requirements derived before and subsume them under overarching design principles for SBAS for minute-taking.

Automated support for minute-taking can increase minute-takers' satisfaction with the created results and the systems' perceived efficiency. The design evaluation also shows that the perceived cognitive effort of minute-taking decreases with higher levels of automation support. Similar to the results of Kushalnagar et al (2015), automated speech-to-text capabilities (R1) with live transcription can help minute-takers to track, capture, and process speech. Partly automated support to structure the captured content (R4) by assigning it to predefined topics improves the completeness of the captured conversation content and enhances perceived task and system performance.

Table 3. User self-evaluation of the results, system, and task.

Main effects of the level of automation (between-subjects effects)		**Pairwise comparisons between levels of automation**			
		NA (avg.)	**PA** (avg.)	**HA** (avg.)	Significance of Differences (p)
Overall Satisfaction with the Resulting Minutes	$F = 32.293$ $p < 0.001$	4.292	5.427	5.136	NA to PA (< 0.001) NA to HA (< 0.001) PA to HA (0.151)
Identification with the Resulting Minutes	$F = 3.942$ $p = 0.02$	4.730	4.979	4.483	NA to PA (0.466) NA to HA (0.497) PA to HA (0.016)
Perceived Task Performance (Task Evaluation)	$F = 15.928$ $p < 0.001$	4.411	5.257	5.184	NA to PA (< 0.001) NA to HA (< 0.001) PA to HA (1.000)
Perceived Cognitive Effort for the Task	$F = 17.536$ $p < 0.001$	5.223	4.544	4.075	NA to PA (0.001) NA to HA (< 0.001) PA to HA (0.049)
Perceived Efficiency of the System	$F = 16.28$ $p < 0.001$	5.173	5.876	6.098	NA to PA (< 0.001) NA to HA (< 0.001) PA to HA (0.571)

Note: NA (No Automation); PA (Partial Automation); HA (High Automation). Mean values on a 7-point Likert scale (from 1: “strongly disagree/very poor” to 7: “strongly agree/very good”).

Considering the relevance of the captured conversation content (Iskender et al, 2021), our results show that users will include more non-relevant content if speech assistance is partially automated. This result indicates that providing live transcripts can make it harder for minute-takers to judge important content if they are not assisted in summarizing it. Raising the automation level further by providing automatic transcript summarization (R3) can enhance the completeness and relevance of the captured content. While automation of speech capture alone does not impact the performance in speaker assignment, the addition of speaker recognition (R2) does. Accordingly, live transcription does not sufficiently enhance minute-takers’ focus for attribution of verbal statements. This result indicates that partial automation of a primary task doesn't necessarily enhance performance in a secondary task. For systems with partially automated speech support, this observation implies that designers should account for interdependencies between simultaneous sub-tasks (Skaugset et al, 2016).

Several contextual factors may have affected our findings and should be considered when interpreting the results. The short, scripted 3.5-minute online meeting scenario likely reduced task complexity and limited natural behavior, as participants could not interact or ask questions. Since the four agenda topics were covered quickly, artificial time pressure potentially encouraged quick task execution over thoughtful engagement. Furthermore, participants only used the SBAS prototype once after watching a tutorial, which may have limited confident use of the system and impaired secondary tasks such as the speaker assignment. Additionally, combining multiple automation features per prototype limited our ability to isolate the effects of specific system capabilities. Given the recruitment of participants via a crowdsourcing platform, some participants may have been distracted, multitasking, or rushing through the task, potentially lowering their attention or sense of responsibility. Lastly, novelty effects, especially under high automation, may have inflated perceived usefulness or simplicity without reflecting long-term utility.

Applied to the design of SBAS, our results provide evidence that implementing our initially identified design requirements (R1 to R4) help to reduce users' workload and increase performance if fulfilled jointly. The speech-to-text feature (R1) and automatic speaker recognition (R2) relieve minute-takers of the burden of capturing orally shared information but result in long transcripts that require processing. Consequently, the speech recognition capabilities of SBAS reduce existing efforts at the cost of creating new ones. To reduce the overall effort for minute-taking, SBAS must also support the structuring (R4) and summarizing (R3) of captured content to ease the processing of transcripts into meeting minutes. We therefore synthesize these four requirements in an overarching design principle (DP1) for SBAS based on the approach proposed by Chandra et al (2015):

DP1 – Balance of Speech Recognition and Processing Capabilities: Speech-based assistance systems must provide capabilities to process, summarize, and structure speech that match its capabilities to recognize and record speech to reduce their users' overall effort.

Our findings indicate that designers of an SBAS for minute-taking must not only consider technological requirements, but also the integration of the minute-takers in highly automated workflows. The high automation prototype offered editable meeting-minute proposals to ease the workload and improve quality (which we incentivized with an additional reward for the best meeting minutes). However, our results indicate many participants often accepted the proposals with little to no changes. While this behavior saves effort, it risks cognitive complacency, which means that minute-takers no longer question the system's output (Jarrahi, 2019). Accepting the generated minute proposals without review increases the chance of capturing incomplete or irrelevant information, risking the loss of important details despite automated transcription. This observation leads to our fifth requirement (R5) for the design of a highly automated SBAS for minute-taking:

R5: To reliably identify and fix errors in captured content, minute-takers require support from speech-based assistance systems, preventing them from accepting automatically created content unchecked.

Another result of the design evaluation is that participants were more satisfied with their minutes under partial automation than under high automation, suggesting they valued their contributions more when automation was limited. Leung et al (2018) showed that high automation can reduce identification with work results if personal contribution feels undervalued. Similarly, high automation support in minute-taking leads to lower identification with the results than partial support. One reason could be that minute-takers feel bored while performing the task due to the significantly lower perceived cognitive effort under high automation support and become more detached from the task and the work result (Cummings et al, 2016). Low identification with a task or work result can negatively impact motivation, performance, and productivity, highlighting that their integration and effort play an important role in defining the level of support by automation. To mitigate this effect, users should receive active support in recognizing their contribution and the value of their effort in achieving a work result to make their work meaningful. We formalize this support in our sixth requirement (R6) for designing a highly automated SBAS for minute-taking:

R6: To identify with their work results while working with highly automated support, minute-takers need speech-based assistance systems to support them in recognizing the value of their contributions.

Besides the SBAS capabilities, the engagement of the assisted user in the task determines the quality of the work results. As minute-takers are responsible for accurately recording and summarizing meetings, designers of highly automated SBAS should consider users' decision-making competence and personal commitment to the task. Our results indicate that cognitive complacency (R5) and low satisfaction and identification with result (R6), can undermine minute-takers' commitment when they collaborate with highly automated SBAS. While R5 and R6 address distinct risks (unchecked automation and declining task identification), they share a focus on user engagement and accountability. With an emphasis on active review (R5) and the motivational value of recognizing personal contributions (R6), these two requirements form the basis of our second design principle (DP2):

DP2 – User Integration in Highly Automated Processes: Speech-based assistance systems must actively and meaningfully integrate users to motivate their engagement in the assisted task.

6 Conclusion

Our findings have key implications for designing human-AI collaboration in knowledge-intensive tasks. Simply increasing automation may not yield proportional gains in perceived performance or satisfaction. Instead, SBAS should augment human expertise while ensuring users stay actively involved in shaping the output. Building on R5 and R6, further design efforts should support active review and user identification with co-created outcomes, especially under high automation and for outcomes with high organizational impact. Overall, this study contributes to a deeper understanding of how AI systems can support knowledge work without undermining user agency, which is essential for shaping human-AI collaboration in organizational practices.

References

Adam, M., Wessel, M. & Benlian, A. (2021) AI-Based Chatbots in Customer Service and Their Effects on User Compliance. *Electronic Markets*, *31*(2), 427-445.

Ahuja, M. & Thatcher, J. (2005) Moving Beyond Intentions and Toward the Theory of Trying: Effects of Work Environment and Gender on Post-Adoption Information Technology Use. *MIS Quarterly*, *29*(3), 427-459.

Bulgurcu, B., Cavusoglu, H. & Benbasat, I. (2010) Information Security Policy Compliance: An Empirical Study of Rationality-Based Beliefs and Information Security Awareness. *MIS Quarterly*, *34*(3), 523-548.

Chandra, L., Seidel, S. & Gregor, S. (2015) Prescriptive Knowledge in IS Research: Conceptualizing Design Principles in Terms of Materiality, Action, and Boundary Conditions, *Proceedings of the 48th Hawaii International Conference on System Sciences*. Kauai, Hawaii, United States.

Collier, J. & Sherrell, D. (2010) Examining the Influence of Control and Convenience in a Self-Service Setting. *Journal of the Academy of Marketing Science*, *38*(4), 490-509.

Coombs, C., Hislop, D., Taneva, S. K. & Barnard, S. (2020) The Strategic Impacts of Intelligent Automation for Knowledge and Service Work: An Interdisciplinary Review. *The Journal of Strategic Information Systems*, *29*(4), 101600.

Cummings, M. L., Gao, F. & Thornburg, K. M. (2016) Boredom in the Workplace: A New Look at an Old Problem. *Human Factors*, *58*(2), 279-300.

Gallivan, M. J., Spitler, V. K. & Koufaris, M. (2005) Does Information Technology Training Really Matter? A Social Information Processing Analysis of Coworkers' Influence on IT Usage in the Workplace. *Journal of Management Information Systems*, *22*(1), 153-192.

Geissler, G., Zinkhan, G. & Watson, R. (2001) Web Home Page Complexity and Communication Effectiveness. *Journal of the Association for Information Systems*, *2*(1), Article 2.

Hansen, J. H. & Hasan, T. (2015) Speaker Recognition by Machines and Humans: A Tutorial Review. *IEEE Signal Processing Magazine*, *32*(6), 74-99.

Iskender, N., Polzehl, T. & Möller, S. (2021) Reliability of Human Evaluation for Text Summarization: Lessons Learned and Challenges Ahead, *Proceedings of the Workshop on Human Evaluation of NLP Systems (HumEval)*. Online.

Ivanov, S., Kuyumdzhiev, M. & Webster, C. (2020) Automation Fears: Drivers and Solutions. *Technology in Society*, *63*, 101431.

Jarrahi, M. H. (2019) In the Age of the Smart Artificial Intelligence: AI's Dual Capacities for Automating and Informating Work. *Business Information Review*, *34*(4), 178-187.

Koslow, A., Berger, B., Morana, S. & Hess, T. (2021) Enhancing Knowledge Externalization with Speech-Based Protocol Systems: A Design Science Approach, *Proceedings of the 25th Pacific Asia Conference on Information Systems*. Dubai (Virtual Conference), United Arab Emirates.

Kushalnagar, R. S., Behm, G. W., Kelstone, A. W. & Ali, S. (2015) Tracked Speech-to-Text Display: Enhancing Accessibility and Readability of Real-Time Speech-to-Text, *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*. Lisbon, Portugal.

Lentz, L. & De Jong, M. (1997) The Evaluation of Text Quality: Expert-Focused and Reader-Focused Methods Compared. *IEEE Transactions on Professional Communication*, *40*(3), 224-234.

Leung, E., Paolacci, G. & Puntoni, S. (2018) Man Versus Machine: Resisting Automation in Identity-Based Consumer Behavior. *Journal of Marketing Research*, *55*(6), 818-831.

Li, N., Harris, T. B., Boswell, W. R. & Xie, Z. (2011) The Role of Organizational Insiders' Developmental Feedback and Proactive Personality on Newcomers' Performance: An Interactionist Perspective. *Journal of Applied Psychology*, *96*(6), 1317-1327.

Mroz, J. E., Allen, J. A., Verhoeven, D. C. & Shuffler, M. L. (2018) Do We Really Need Another Meeting? The Science of Workplace Meetings. *Current Directions in Psychological Science*, *27*(6), 484-491.

Nicolaou, A. I. & McKnight, D. H. (2011) System Design Features and Repeated Use of Electronic Data Exchanges. *Journal of Management Information Systems*, *28*(2), 269-304.

Niforatos, E., Laporte, M., Bexheti, A. & Langheinrich, M. (2018) Augmenting Memory Recall in Work Meetings: Establishing a Quantifiable Baseline, *Proceedings of the 9th Augmented Human International Conference*. New York, United States.

Parker, S. K. & Grote, G. (2022) Automation, Algorithms, and Beyond: Why Work Design Matters More Than Ever in a Digital World. *Applied Psychology*, *71*(4), 1171-1204.

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.

Raisch, S. & Krakowski, S. (2021) Artificial Intelligence and Management: The Automation-Augmentation Paradox. *Academy of Management Review*, *46*(1), 192-210.

Rashid, M., Clarke, P. M. & O'Connor, R. V. (2017) Exploring Knowledge Loss in Open Source Software (OSS) Projects, *Proceedings of the 17th International Conference on Software Process Improvement and Capability Determination*. Palma de Mallorca, Spain.

Riedl, M. O. (2019) Human-Centered Artificial Intelligence and Machine Learning. *Human Behavior and Emerging Technologies*, *1*(1), 33-36.

Romao, M., Costa, J. & Costa, C. J. (2019) Robotic Process Automation: A Case Study in the Banking Industry, *Proceedings of the 14th Iberian Conference on Information Systems and Technologies*. Coimbra, Portugal.

Rudolph, J., Tan, S. & Tan, S. (2023) ChatGPT: Bullshit Spewer or the End of Traditional Assessments in Higher Education? *Journal of Applied Learning and Teaching*, *6*(1), 342-363.

Schwabe, H. & Castellacci, F. (2020) Automation, Workers' Skills and Job Satisfaction. *PLoS ONE*, *15*(11), e0242929.

Shadiev, R., Hwang, W.-Y., Chen, N.-S. & Huang, Y.-M. (2014) Review of Speech-to-Text Recognition Technology for Enhancing Learning. *Journal of Educational Technology & Society*, *17*(4), 65-84.

Skaugset, L. M., Farrell, S., Carney, M., Wolff, M., Santen, S. A., Perry, M. & Cico, S. J. (2016) Can You Multitask? Evidence and Limitations of Task Switching and Multitasking in Emergency Medicine. *Annals of Emergency Medicine*, *68*(2), 189-195.

Spector, P. E. (1997) *Job Satisfaction: Application, Assessment, Causes, and Consequences*. Thousand Oaks, California, USA: Sage Publications, Inc.

Standaert, W., Muylle, S. & Basu, A. (2021) How Shall We Meet? Understanding the Importance of Meeting Mode Capabilities for Different Meeting Objectives. *Information & Management*, *58*(1), 103393.

Tsai, W.-C., Chen, C.-C. & Liu, H.-L. (2007) Test of a Model Linking Employee Positive Moods and Task Performance. *Journal of Applied Psychology*, *92*(6), 1570.

Vandenbosch, B. & Ginzberg, M. J. (1996) Lotus Notes® and Collaboration: Plus ça change. *Journal of Management Information Systems*, *13*(3), 65-81.