AI for Meeting Minutes: Promises and Challenges in Designing Human-AI Collaboration on a Production SaaS Platform

Riku Arakawa*
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
rarakawa@cs.cmu.edu

Hiromu Yakura*
Center for Humans and Machines
Max-Planck Institute for Human Development
Berlin, Germany
hiromu1996@gmail.com

Kei Akuzawa
ACES Inc.
Tokyo, Japan
kei.akuzawa@acesinc.co.jp

Shizuma Kubo
ACES Inc.
Tokyo, Japan
shizuma.kubo@acesinc.co.jp

Abstract

Recent advances in AI technologies, including large language models, have enabled the widespread deployment of automated meeting minute generation at a commercial scale. However, many users continue to take minutes manually. To understand the factors behind this gap, we conducted a case study on a start-up company providing a commercial meeting analysis service. Through detailed observations of the development process over three years and workshops with their designers and developers, we identified key challenges, including discrepancies between user expectations and AI-generated summaries, as well as difficulties in balancing user interaction with automation. Importantly, our study sheds light on factors that have been less emphasized in previous HCI literature, such as the learning curve associated with adopting new technologies for an enterprise product. These insights spotlight the challenges in achieving an effective collaboration between rapidly evolving AI and users, suggesting the increasingly important role of HCI.

CCS Concepts

• **Human-centered computing → Empirical studies in HCI**; *Collaborative interaction*; *Interaction design process and methods.*

Keywords

Meeting minutes, Text summarization, Human-AI collaboration

ACM Reference Format:
Riku Arakawa, Hiromu Yakura, Kei Akuzawa, and Shizuma Kubo. 2025. AI for Meeting Minutes: Promises and Challenges in Designing Human-AI Collaboration on a Production SaaS Platform. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3706599.3706686

*The authors contributed equally to this research.

CHI EA '25, Yokohama, Japan

ACM ISBN 979-8-4007-1395-8/25/04
https://doi.org/10.1145/3706599.3706686

1 Introduction

The ability to take appropriate minutes during meetings has long been regarded as an essential skill for competent business professionals [13]. Beyond merely sharing information, minutes serve as a crucial reference point for stakeholders, helping to accelerate subsequent actions and sometimes clarifying accountability. Recent advances in machine learning have introduced the possibility of automating the process of taking minutes [22]. In addition to significant improvements in speech recognition accuracy [19], the emergence of large language models (LLMs) has enabled computers to transform less-structured transcripts into structured text [17]. Furthermore, the rise of online communication in business settings, spurred by the COVID-19 pandemic, has accelerated the penetration of these technologies [9]. In fact, not only dedicated services like *Otter.ai* but also platforms such as *Zoom* have begun incorporating automatic minute-taking functionalities.

However, we still observe many users manually taking minutes, even in environments where such automatic features are enabled. What drives them to persist in this manual process? We embarked on this case study to shed light on the gap between such user behaviors and the cutting-edge technological paradigm. Specifically, we took a company that provides a meeting recording and analysis service for enterprises in Japan since November 2021 as a field to examine the perspectives of stakeholders involved in the design and development of their products. By comparing their experiences with the HCI research literature, we discussed their challenges and identified open questions in designing AI-driven services to transform conventional workflows at different enterprises. In particular, we base the discussion on the perspective of human-AI collaboration [29], where end-users are not necessarily familiar with the technology. This case study will contribute to introducing more comprehensive perspectives to the stream of HCI research in this realm of rapidly evolving AI technology.

2 Background

We first review HCI literature on the efficient generation of meeting minutes based on AI technologies. Then, we highlight the gap between the research and the current user practice by looking at prior discussions on the role of meeting minutes.

2.1 AI-Driven Meeting Minutes Taking

Recent advancements in speech recognition and LLMs enable automated minutes generation from meeting recordings [24]. Research on document summarization especially plays a pivotal role in transforming unstructured, often lengthy transcripts into structured meeting minutes. In this context, Stiennon *et al.* [26] presented a way to train an AI model for document summarization with the feedback from humans so that their behaviors and human preferences align. These research efforts are leveraged in many industrial tools, like *Zoom*, in which systems present machine-generated minutes once a meeting is completed.

Meanwhile, HCI researchers have explored ways to realize mixed-initiative [14], human-AI collaboration by building interactive support systems. For example, Slobodkin *et al.* [25] developed a system that allows users to select parts of meeting transcriptions to be included in minutes and interactively edit the generated minutes. Asthana *et al.* [6] investigated the effectiveness of two different styles of LLM-generated meeting minutes (*i.e.*, important highlights and a structured, hierarchical minutes view) and discussed ways to incorporate user interactions (*e.g.*, editing, sharing) to improve minutes generation. Li *et al.* [15] proposed not only post-processing techniques to avoid hallucination but also design considerations for maintaining users' trust, such as displaying estimated summarization quality and information gain. Lastly, Tsai *et al.* [27] introduced a VR headset for minute taking so that users can adjust LLM-generated minutes in real-time via gaze-based keyword selection. These studies underscore the growing emphasis on user-centered designs based on interactive collaboration with AI.

Such collaboration is also achieved in the *co-pilot* style interaction [21]. As implemented in *Microsoft Teams* and *Zoom*, this approach involves an AI agent participating in a meeting to offer prompt-based interactions to users, such as asking questions and summarizing next actions, both during the meeting and retrospectively. Considering that this allows users who participated to easily refer to the meeting content and users who did not participate to catch up through interactive information seeking, this is expected to be an alternative paradigm for generating meeting minutes.

2.2 Why Do People Take Meeting Minutes Manually?

Despite these advancements, many users are still taking meeting minutes themselves. Researchers have tried to understand the role of meeting minutes, which hints at the user motivation behind their practices of minute-taking. Importantly, as Asthana *et al.* [6] noted, there is debate on what constitutes good meeting minutes due to varying roles of meetings and expectations from participants. In decision-focused meetings, for example, key decisions and action points are prioritized, while collaboration-driven meetings may center more around process-related discussions [11]. Also, too much detail in the minutes can overwhelm participants, creating a cognitive load and potential fixation on ideas that may hinder overall meeting productivity [13]. Furthermore, meetings are deeply influenced by the organizational culture they occur within, and in turn, they help shape that culture [23]. These points suggest that users want to have agency in managing the diverse contextual information to be put into the minutes.

3 Motivation for the Case Study

The above literature suggests the emerging attention to minute generation from both academia and industry, especially regarding how to design a helpful, collaborative system. We thought analyzing the design process in the industry has unique opportunities to understand the stakeholders' experiences and insights, especially during this era of rapidly evolving AI technology. To this aim, we scrutinized the design and development process of a commercialized service. Specifically, this case study happened at *ACES Inc.*[1], which offers *ACES Meet*[2], a SaaS product designed to enhance the efficiency and effectiveness of meeting management on different video-conferencing platforms, including *Microsoft Teams* and *Zoom*. Primarily aimed at the Japanese market, this software utilizes automation tools to streamline repetitive tasks regarding recording the meeting, minute generation, featuring transcription, and video navigation. *ACES Meet* has been operated since November 2021 and is offered for various industrial clients, including automotive, construction, finance, hospitality, and medicine. It has recorded over 1 million meetings as of October 2024.

In this case study, we first reviewed the series of past and concurrent product design meetings conducted by the developers and designers of this service. The recorded meetings happened from November 2021 to August 2024. While they have worked on diverse aspects related to the product (*e.g.*, enhancing the software infrastructure), we focused on tracing the string of decision-making regarding the minute-taking feature. Additionally, between August and September 2024, we held several workshops with employees at *ACES Inc.* These sessions aimed to build a collective understanding of the minute-taking process, drawing on insights from the developers' experiences and past user interviews conducted by the designers.

4 Trace of Decision-Making during the Development of a SaaS Platform

After reviewing recorded meetings over three years and consulting with members involved in key decision-making, we observed multiple iterations in the design process of the minute-taking feature. These modifications were largely influenced by the advancements in AI technologies, particularly LLMs. We identified four major iteration cycles, including the current discussions. In the following subsections, we detail the attempts and reflections at each iteration stage.

4.1 Iteration 1: Extraction-Based Summary with Heuristics

From November 2021 to July 2022, the team developed a feature that automatically highlights key sentences in the transcription post-meeting. This method, known as extraction-based summarization, contrasts with abstraction-based summarization [30]. The designers and developers collaboratively identified the traits of significant utterances during meetings and implemented a rule-based algorithm to highlight them. They specifically targeted utterances that contained certain keywords (*e.g.*, "next action") or

[1] https://acesinc.co.jp/
[2] https://meet.acesinc.co.jp/

responded to questions. Additionally, they assigned lower scores to utterances at the beginning of meetings, which often relate to initial ice-breaking conversations. The goal was for these automatic highlights to quickly draw users' attention to critical decisions and facilitate the minute-taking process.

The feature was launched as a beta for internal use and a few partner clients in July 2022. Shortly thereafter, the team began receiving feedback on the accuracy of the highlighted utterances. Although the system successfully identified important parts, the occurrence of incorrect predictions—both false positives and false negatives—significantly impacted the user experience. For instance, false negatives forced users to manually review the transcription to ensure no crucial utterances were overlooked. Consequently, the feature did not build user trust and failed to streamline the minute-taking process effectively.

4.2 Iteration 2: LLM-Generated Summary with A Single Prompt

While iteratively refining the accuracy of the highlighting algorithm, the team also explored the potential of abstraction-based summarization to directly aid users, aiming to drastically reduce the effort involved in minute-taking. In November 2022, the release of *ChatGPT* [1] marked a significant recognition of LLM capabilities within the team. Upon the release of its API, they introduced a new feature to their meeting review page in March 2023: AI-generated minutes. This utilized their insights into the characteristics of crucial information for minute generation, gleaned from the development of the highlighting feature. Specifically, they designed a system prompt that directed the LLM to summarize key points, such as the next actions and conclusions from the meeting.

Initially, customer users were excited about this innovative feature. However, challenges arose just a few weeks after its introduction. The primary issue was that the AI-generated minutes did not meet user expectations in terms of quality. Users had hoped for a complete replacement of manual effort with AI, but the minutes often omitted important topics or included incorrect details and hallucinations, such as non-existent participant names. Users also noted that contextual information, such as the purpose of the meeting, was often missing in the AI-generated minutes, which is aligned with the discussion in Section 2.2. They requested the ability to explicitly provide such information through user input to enhance the accuracy and relevance of the summaries. Consequently, users found themselves needing to read or rewatch the meeting to manually correct the generated output.

4.3 Iteration 3: LLM-Generated Summary with Multiple Prompts by Topics

After realizing that the naïve, a single-shot approach using LLMs was not effective, the team contemplated user behavior during minute-taking. They aimed to deconstruct the existing human process and design more tailored AI functions. Recognizing that LLM outputs can be erroneous, they focused on developing AI functions that corresponded to each step of the process, allowing for easy user corrections where necessary. To this end, the designers conducted interviews with multiple users to map out the detailed step-by-step process of minute-taking. They also categorized stakeholders involved in meeting minutes as minute takers, participants, and absentees, and visualized their interactions. Although there were diverse approaches for various types of meetings, the team ultimately identified the abstract structure of the human process involved in minute-taking. The flowchart that illustrates this identified process is shown in Figure 1.

Based on this analysis, they segmented the minute-taking process into topic identification, topic-based summary, and reformatting. To support topic-based summarization, they developed multiple prompt templates tailored for different types of meetings, such as sales and internal meetings. These templates were designed to address specific topics relevant to each type; for example, the sales meeting template included sections for budget, authority, needs, timing, and competitors. Additionally, users were given the advanced option to create custom templates to accommodate the unique cultures and styles of their organizations, as discussed Section 2.2.

This update was warmly welcomed by end users, who reported a significant reduction in the effort required to create meeting minutes. Through interviews, the team learned that users had gradually integrated the feature into their minute-taking process, selectively applying it to specific types of meetings. However, users also expressed the need to address occasional inaccuracies in the topic summaries, which still required meticulous checking and manual correction. Additionally, they pointed out that important topics not included in the template were often missed, highlighting the need for an improved topic identification feature. This feedback, in turn, confirmed the team's approach to incremental automation and their strategy of clearly defining the roles of human users and AI in the process.

4.4 Ongoing Discussion: Vision for More Interactivity

As of August 2024, the team continues to implement the other steps in the human workflow identified earlier, such as topic identification and reformatting. Here, they are actively searching for an effective method to compare different LLMs and optimize the development workflow. The rapid evolution of LLMs compels the team to regularly evaluate the best model in terms of quality and cost. This necessitates a well-designed development framework that ensures effective progress while maintaining user satisfaction.

Simultaneously, the team noted the growing interest in Google's *NotebookLM*[3], which offers interactive information extraction via a chat interface. While this interactive approach is appealing, implementing it could significantly alter the user experience and potentially surprise existing users. Indeed, they did a pilot test with a few partner clients and realized that they struggled with asking questions to get desired answers. Consequently, the team is carefully considering the idea of introducing more interactivity.

On the other hand, they are particularly inspired by the feature in *NotebookLM* that highlights the referenced utterances when generating a response to a query. Recognizing that users sometimes struggle to trust the generated summaries fully, the team is proactively looking to adopt this feature to enhance transparency in their LLM-generated summaries.

[3] https://notebooklm.google/

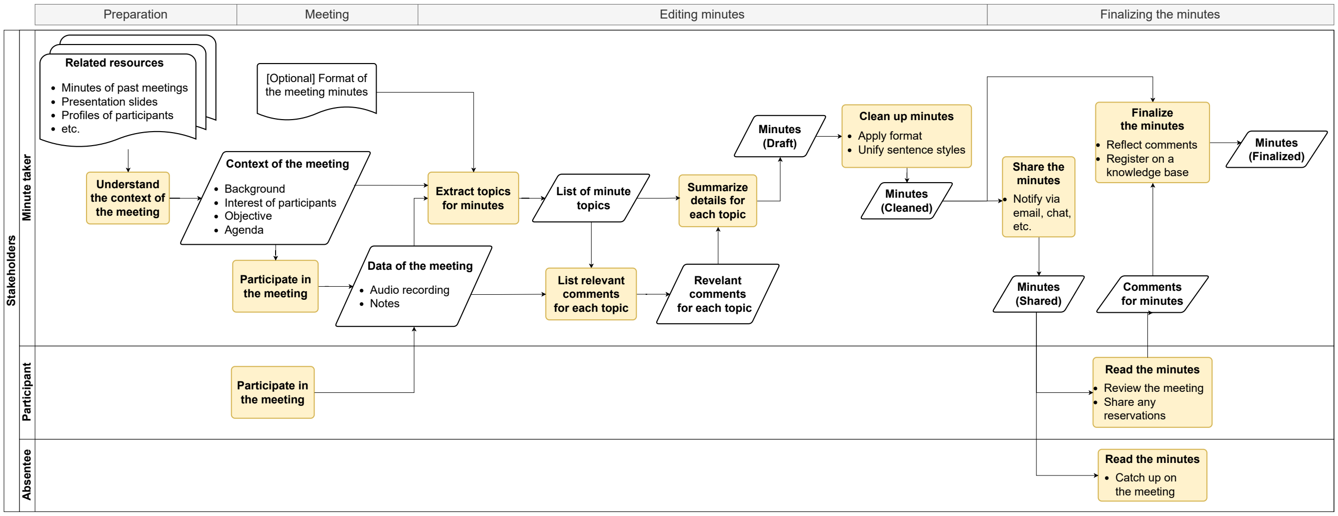

Figure 1: Flowchart of existing human operation in minute-taking. The designers tried to understand the step-by-step process to develop tailored AI functions to facilitate it.

5 Lessons Learned

The three-year product development timeline, enriched by workshops with current designers and developers, reveals multiple iterations influenced by user feedback and shifts in the AI technology landscape. Below, we summarize the key lessons learned from this ongoing process.

5.1 Nuanced Feelings for AI-Generated Minutes

We observed that AI-generated minutes could be perceived as ineffective without carefully designing the user interaction process. As can be seen in Figure 1, minute-taking is not merely taking notes during meetings but also prioritizing topics based on various contextual information and making a consensus among participants. From this, we can imagine that naïve introduction of AI-generated minutes is perceived as an addition of an independent stakeholder, inherently making consensus more complex. Also, minute takers need to consider various contextual information, which is often not incorporated into the single-shot generation approach. Therefore, preparing topics based on meeting types and generating a topic-by-topic summary is a practical approach identified so far. This solution maintains simplicity in user experience while allowing for user input, which has been well-received. This result highlights the effectiveness of clarifying the role of the end-user and AI, which has been emphasized in the human-AI collaboration research literature in diverse domains, including human assessment [4], coaching [3, 5] and pathology diagnosis [12].

5.2 Challenges for Interactive Minute-Taking Support

Meanwhile, the product team is conservative in adopting interactive approaches, like those reviewed in Section 2. One of the largest concerns is the learning cost of users. Such interactive designs inherently involve the action from the user side, which goes beyond the conventional workflow of taking minutes (see Figure 1). This demands the service to onboard and provide support for a user, resulting in a longer time until the user is impressed by the power of AI for the first time. It is crucial from the perspective of industrial stakeholders, as this leads to a non-optimal user journey with a higher bounce rate. Considering that the target users are not necessarily familiar with AI technologies, we can expect that nudging changes in their behavioral patterns would not be straightforward. This point also applies to the *co-pilot*-like approach. For instance, *NotebookLM* is a notable benchmark for interactively organizing information from meeting recordings. However, it is often said that asking the right question is much harder than finding the right answer [10], which is aligned with their pilot user feedback. Given that, the prompt-based interaction with co-pilot agents would not provide beneficial results for a wide spectrum of users at this point.

Moreover, the cost incurred by interaction with AI models is another big challenge. This is because, regardless of whether a self-hosted model or an external API is employed, utilizing LLMs poses significantly higher financial costs than conventional web services. Also, while open-ended, fine-grained exploration opportunities are important for these human-AI collaborative settings, this introduces much uncertainty in cost estimation. If the service can employ pay-as-you-go billing specifically for the usage of AI models, it would not be a large obstacle. However, changing the price structure is difficult, especially for the B2B domain. These points hindered the company from implementing real-time, interactive collaboration opportunities with AI for minute-taking.

5.3 Limitations

While this analysis uncovers ongoing discussions on how to design a human-AI collaborative system in the industry, the lessons presented in this section are based on observations from a small and medium-sized Japanese enterprise. Unlike large technology companies, the limited resources available to this company influenced their decision-making process. Additionally, the cultural context plays a significant role in shaping technological adoption and user literacy [20], which is important to consider when interpreting the results of this study.

6 Open questions

Finally, based on what we learned throughout the case study, we advocate the following open questions.

6.1 How Can Developers Reflect Rapid Advancements in LLMs into Their Service?

The current advancements in AI technology are unparalleled, presenting unique challenges for industrial practitioners within the ecosystem. A primary concern is avoiding drastic changes to the user experience, given the presence of established customers. This consideration influences various decision-making processes. For example, updating LLMs behind the service necessitates extensive testing to preserve expected behavior [7, 18]. Additionally, while we did not elaborate much in this case study for its business-specific aspect, the changes in API costs can impact business revenue directly. Consequently, this presents a unique need and opportunity for HCI researchers to bridge the gap between end-user needs and expectations and the rapid evolution of technology, all while maintaining trust in services, particularly for enterprises developing third-party applications.

6.2 How Can the Minute-Taking Service Incorporate Broader Meeting Context?

Meetings often rely on implicit contexts such as the roles of participants and their previous interactions, yet there has been no effective method for LLMs to integrate this information into their outputs. Currently, LLMs primarily process textual information, such as meeting transcriptions, but the multimodal behaviors of participants are equally crucial. Researchers in social signal processing and HCI have long studied human behavior during dialogues to improve the understanding of meeting dynamics [16, 28]. However, the knowledge derived from human behavior analysis is often implicit and sensitive [2, 4]; relying solely on machine learning-based approaches for analysis could have unintended consequences. Recent HCI studies have begun developing user-centered systems that utilize these behavioral signals, anticipating that such approaches will influence industrial product design.

Additionally, the technology faces challenges with accurately transcribing unique words (*e.g.*, proper noun), as they may not be present in the training data. This adversely affects the minute-taking process. Designers have noted that specific terms and company names are frequently transcribed incorrectly, severely impacting the user experience. To address this, the team is attempting to train speech-to-text models using synthetic data for these words. This issue is significant even in Japanese, a language with extensive speech data, suggesting it is a critical challenge for industry professionals in different languages.

6.3 How Does Optimal Human-AI Collaboration Shift Over Time?

An interesting development is the company's decision to limit user agency by providing predefined scenarios for minute-taking. This approach has been effective, likely because end novice users appreciate the streamlined experience of engaging with set interactions that offer limited flexibility. However, as users become more accustomed to AI-driven services, their expectations will evolve, as emphasized by Cockburn *et al.* [8]. This suggests a potential shift in the ideal balance of mixed-initiative interactions [14]. For instance, the company initially introduced AI-generated minutes using a single template, but now they offer multiple templates with specific topics for different types of meetings, such as sales, and the users can also customize the topics. This progression underscores the need for a longitudinal study to examine how collaborative structures adapt over time within the current AI paradigm, particularly focusing on industrial services that engage real end users across various sectors.

7 Conclusion

This case study sheds light on the promises and challenges involved in designing human-AI collaborative systems by taking the example of meeting minute generation. Informed from the development process of a commercial service, we identified key challenges, including user expectations for high-quality AI-generated content and the inherent difficulties in balancing user interaction with automation. Our findings emphasize the importance of understanding user behavior and backgrounds in AI-driven services, particularly in scenarios where technological familiarity varies widely among users. Also, we highlighted the perspectives of professional designers and developers that were less emphasized in the past literature, such as the tension between the evolving LLM's capability and the user learning curve for enterprise products. Looking forward, this case study raises several questions about the future of human-AI collaboration, suggesting that the evolution of AI capabilities will likely shift the balance between automation and user input over time, creating new demands for alternative design strategies.

As AI technologies, including LLMs, will continue to advance at an increasing pace, the opportunities and challenges in designing effective human-AI interactions can occur universally. Here, we argue that capturing the perspectives and decision-making processes of practitioners during this transformative period would comprise an important work of the HCI community. In particular, inviting not only large technology firms but also various organizations in different cultural and business contexts enriches the depth and diversity of the knowledge in this community. These traces of real-world decision-making are more than a retrospective report, as they can offer a foundational asset for both researchers and practitioners in facing future waves of technological innovation. Given that the role of the HCI community in bridging the gap between theoretical research and practical applications will become increasingly important, we hope that this case study will contribute to a step forward in this direction.

Acknowledgments

This research was supported in part by JST PRESTO Grant Number JPMJPR246B.

References

[1] Open AI. 2022. ChatGPT. https://openai.com/index/chatgpt/
[2] Riku Arakawa, Kiyosu Maeda, and Hiromu Yakura. 2024. ConverSearch: Supporting Experts in Human Behavior Analysis of Conversational Videos with a Multimodal Scene Search Tool. *ACM Transactions on Interactive Intelligent Systems* (2024), In press. https://doi.org/10.1145/3709012
[3] Riku Arakawa and Hiromu Yakura. 2019. REsCUE: A framework for REal-time feedback on behavioral CUEs using multimodal anomaly detection. In *Proceedings of the 2019 ACM CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 572. https://doi.org/10.1145/3290605.3300802
[4] Riku Arakawa and Hiromu Yakura. 2023. AI for human assessment: What do professional assessors need?. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 378:1–378:7. https://doi.org/10.1145/3544549.3573849
[5] Riku Arakawa and Hiromu Yakura. 2024. Coaching Copilot: Blended Form of an LLM-Powered Chatbot and a Human Coach to Effectively Support Self-Reflection for Leadership Growth. In *Proceedings of the 2024 ACM Conversational User Interfaces*. ACM, New York, NY, 2:1–2:14. https://doi.org/10.1145/3640794.3665549
[6] Sumit Asthana, Sagih Hilleli, Pengcheng He, and Aaron Halfaker. 2023. Summaries, Highlights, and Action items: Design, implementation and evaluation of an LLM-powered meeting recap system. *arXiv* abs/2307.15793 (2023), 1–29. https://doi.org/10.48550/ARXIV.2307.15793
[7] Gagan Bansal, Besmira Nushi, Ece Kamar, Daniel S. Weld, Walter S. Lasecki, and Eric Horvitz. 2019. Updates in Human-AI Teams: Understanding and Addressing the Performance/Compatibility Tradeoff. In *The 33rd AAAI Conference on Artificial Intelligence*. AAAI Press, Washington, DC, 2429–2437. https://doi.org/10.1609/AAAI.V33I01.33012429
[8] Andy Cockburn, Carl Gutwin, Joey Scarr, and Sylvain Malacria. 2014. Supporting Novice to Expert Transitions in User Interfaces. *Comput. Surveys* 47, 2 (2014), 31:1–31:36. https://doi.org/10.1145/2659796
[9] Evan DeFilippis, Stephen Michael Impink, Madison Singell, Jeffrey T. Polzer, and Raffaella Sadun. 2022. The Impact of COVID-19 on Digital Communication Patterns. *Humanities and Social Sciences Communications* 9, 1 (2022). https://doi.org/10.1057/s41599-022-01190-9
[10] Terry Fadem. 2008. *Art of Asking, The: Ask Better Questions, Get Better Answers*. FT Press, Upper Saddle River, NJ.
[11] Jennifer L. Geimer, Desmond J. Leach, Justin A. DeSimone, Steven G. Rogelberg, and Peter B. Warr. 2015. Meetings at work: Perceived effectiveness and recommended improvements. *Journal of Business Research* 68, 9 (2015), 2015–2026. https://doi.org/10.1016/j.jbusres.2015.02.015
[12] Hongyan Gu, Jingbin Huang, Lauren Hung, and Xiang 'Anthony' Chen. 2021. Lessons Learned from Designing an AI-Enabled Diagnosis Tool for Pathologists. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 10:1–10:25. https://doi.org/10.1145/3449084
[13] Joanna Gutmann. 2001. *Taking Minutes of Meetings*. Kogan Page, London, UK.
[14] Eric Horvitz. 1999. Principles of Mixed-Initiative User Interfaces. In *Proceeding of the ACM CHI '99 Conference on Human Factors in Computing Systems*. ACM, 159–166. https://doi.org/10.1145/302979.303030
[15] Daniel Li, Thomas Chen, Alec Zadikian, Albert Tung, and Lydia B. Chilton. 2023. Improving Automatic Summarization for Browsing Longform Spoken Dialog. In *Proceedings of the 2023 ACM CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 106:1–106:20. https://doi.org/10.1145/3544548.3581339
[16] Werner Liebregts, Pourya Darnihamedani, Eric Postma, and Martin Atzmueller. 2019. The Promise of Social Signal Processing for Research on Decision-Making in Entrepreneurial Contexts. *Small Business Economics* 55, 3 (2019), 589–605. https://doi.org/10.1007/s11187-019-00205-1
[17] Bonan Min, Hayley Ross, Elior Sulem, Amir Pouran Ben Veyseh, Thien Huu Nguyen, Oscar Sainz, Eneko Agirre, Ilana Heintz, and Dan Roth. 2024. Recent Advances in Natural Language Processing via Large Pre-trained Language Models: A Survey. *Comput. Surveys* 56, 2 (2024), 30:1–30:40. https://doi.org/10.1145/3605943
[18] Moses Openja, Foutse Khomh, Armstrong Foundjem, Zhen Ming Jiang, Mouna Abidi, and Ahmed E. Hassan. 2024. Studying the Practices of Testing Machine Learning Software in the Wild. *ACM Transactions on Software Engineering and Methodology* in press (2024), 1–63. https://doi.org/10.1145/3680463
[19] Rohit Prabhavalkar, Takaaki Hori, Tara N. Sainath, Ralf Schlüter, and Shinji Watanabe. 2024. End-to-End Speech Recognition: A Survey. *IEEE ACM Transactions on Audio, Speech, and Language Processing* 32 (2024), 325–351. https://doi.org/10.1109/TASLP.2023.3328283
[20] Martin Prause. 2019. Challenges of Industry 4.0 Technology Adoption for SMEs: The Case of Japan. *Sustainability* 11, 20 (2019), 5807. https://doi.org/10.3390/su11205807
[21] Tori Reddy Dodla. 2024. Integrate AI with Microsoft Co-Pilot. In *Mastering Knowledge Management Using Microsoft Technologies*. Apress, Berkeley, CA, 117–138. https://doi.org/10.1007/979-8-8688-0372-7_5
[22] Virgile Rennard, Guokan Shang, Julie Hunter, and Michalis Vazirgiannis. 2023. Abstractive Meeting Summarization: A Survey. *Transactions of the Association for Computational Linguistics* 11 (2023), 861–884. https://doi.org/10.1162/tacl_a_00578
[23] Cliff Scott and Joe Allen. 2022. Toward an organizational theory of meetings: Structuration of organizational meeting culture. *Organizational Psychology Review* 13, 4 (2022), 506–531. https://doi.org/10.1177/20413866221127249
[24] Kartik Shinde, Tirthankar Ghosal, Muskaan Singh, and Ondrej Bojar. 2022. Automatic Minuting: A Pipeline Method for Generating Minutes from Multi-Party Meeting Proceedings. In *Proceedings of the 36th Pacific Asia Conference on Language, Information and Computation*. Association for Computational Linguistics, Stroudsburg, PA, 691–702.
[25] Aviv Slobodkin, Niv Nachum, Shmuel Amar, Ori Shapira, and Ido Dagan. 2023. SummHelper: Collaborative Human-Computer Summarization. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Stroudsburg, PA, 554–565. https://doi.org/10.18653/V1/2023.EMNLP-DEMO.50
[26] Nisan Stiennon, Long Ouyang, Jeff Wu, Daniel M. Ziegler, Ryan Lowe, Chelsea Voss, Alec Radford, Dario Amodei, and Paul F. Christiano. 2020. Learning to Summarize from Human Feedback. In *Advances in Neural Information Processing Systems*, Vol. 33. Curran Associates, Red Hook, NY, 3008–3021.
[27] Hsin-Ruey Tsai, Shih-Kang Chiu, and Bryan Wang. 2024. GazeNoter: Co-Piloted AR Note-Taking via Gaze Selection of LLM Suggestions to Match Users' Intentions. *arXiv* abs/2407.01161 (2024), 1–18. https://doi.org/10.48550/ARXIV.2407.01161
[28] Alessandro Vinciarelli, Maja Pantic, Dirk Heylen, Catherine Pelachaud, Isabella Poggi, Francesca D'Errico, and Marc Schröder. 2012. Bridging the Gap between Social Animal and Unsocial Machine: A Survey of Social Signal Processing. *IEEE Transactions on Affective Computing* 3, 1 (2012), 69–87. https://doi.org/10.1109/T-AFFC.2011.27
[29] Dakuo Wang, Elizabeth F. Churchill, Pattie Maes, Xiangmin Fan, Ben Shneiderman, Yuanchun Shi, and Qianying Wang. 2020. From Human-Human Collaboration to Human-AI Collaboration: Designing AI Systems That Can Work Together with People. In *Extended Abstracts of the 2020 ACM CHI Conference on Human Factors in Computing Systems*. ACM, 1–6. https://doi.org/10.1145/3334480.3381069
[30] Jin-ge Yao, Xiaojun Wan, and Jianguo Xiao. 2017. Recent advances in document summarization. *Knowledge and Information Systems* 53, 2 (2017), 297–336. https://doi.org/10.1007/S10115-017-1042-4