WinWin Meets – Investigating the Future of Online Meetings

Martin Žust
marti.zust@gmail.com
Jožef Stefan Institute
Ljubljana, Slovenia

Marko Grobelnik
marko.grobelnik@ijs.si
Jožef Stefan Institute
Ljubljana, Slovenia

Alenka Guček
alenka.gucek@ijs.si
Jožef Stefan Institute
Ljubljana, Slovenia

Adrian Mladenic Grobelnik
adrian.m.grobelnik@ijs.si
Jožef Stefan Institute
Ljubljana, Slovenia

Abstract

Video conferencing is now central to modern collaboration, yet its functionality remains largely limited to passive audio–visual communication. Despite growing investment in artificial intelligence (AI), it is unclear which features truly enhance meetings and how users will adopt them. Here we present WinWin Meets, a Jitsi-based prototype that integrates Whisper transcription and GPT-4o processing to deliver real-time summaries, visual mind maps, and goal-oriented advice. Testing with 16 participants showed strong interest in summaries and mind maps, moderate interest in in-meeting guidance, and a preference for add-on integration. Market research confirmed low organic demand for advanced AI features, with users prioritizing reliable improvements such as automated notes. These results highlight a gap between experimental enthusiasm and everyday adoption, pointing to opportunities for targeted, industry-specific integrations that combine reliability with intelligent support.

Keywords

video conferencing, AI agent, testing, market research, zoom, negotiation, transcription, summarization, advice, meeting notes, AI innovations

1 Introduction

As artificial intelligence advances rapidly, its potential to transform everyday digital tools, particularly video conferencing, has become increasingly apparent. Platforms such as Zoom, Google Meet, and Microsoft Teams have become standard, yet their functionality remains focused on basic communication. A new need is arising for next-generation conferencing, including intelligent assistants, automatic summarization, content analysis, and contextual support. These next-generation systems go beyond passive audio and video transmission to actively support users with intelligent features and real-time analysis [1].

Previous research reveals both promise and challenges. Proactive AI meeting assistants can improve efficiency but need to balance autonomy with what users are willing to accept [1]. Meanwhile, studies of speech-based technology underscore the difficulty of extracting useful outcomes from nuanced group interactions [2]. These perspectives suggest that AI's success in meetings depends on technical feasibility and sensitivity to human collaboration.

With remote meetings now central to how we work, these systems directly impact productivity, collaboration, and organizational culture. This paper explores which functionalities could define the future of video conferencing and how AI may contribute. We combine market trend and user preference analysis, reviews of online discussions, and experimental testing of the WinWin Meets prototype. We explore which features matter to users, examine how AI can support meetings, and assess the potential to improve efficiency, clarity, and structure in digital communication.

2 Background and Related Work

2.1 Overview of Current Video Conferencing Solutions

The video conferencing market is currently dominated by a few major players. Zoom, Microsoft Teams, and Google Meet together account for approximately 94% of global market share, with Zoom alone holding around 56% [3]. While all three platforms are actively investing in artificial intelligence features, their innovation must be carefully balanced with the risk of reputational damage. As established brands, they face more constraints than lesser-known startups, which can afford a higher level of experimental agility. This creates a unique window of opportunity for the emergence of disruptive technologies that have the potential to redefine the video conferencing experience.

Most AI-enabled tools developed recently are not standalone platforms, but integrations designed to work alongside existing services like Zoom, Google Meet, or Microsoft Teams. Notable examples include tl;dv [4], Otter.ai [5], Fathom [6], Fireflies [7], and Sembly AI [8]. These applications primarily offer meeting transcription, and some provide more advanced analytics such as sentiment analysis or participant-level speaking time metrics.

2.2 Limitations of Existing Solutions and Emerging Needs

Despite the growing number of AI integrations, fully independent platforms that natively combine video conferencing with built-in AI features remain rare. These features may include real-time transcription, intelligent meeting summarization, and contextual AI-generated recommendations. This segment remains underdeveloped, presenting a significant opportunity for innovation.

While major platforms like Zoom have started introducing their own AI assistants (e.g., Zoom AI Companion [9]), they must innovate cautiously to protect their reputation and user base. This creates space for new companies to develop more ambitious

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
Information Society 2025, Ljubljana, Slovenia
© 2025 Copyright held by the owner/author(s).
https://doi.org/https://doi.org/10.70314/is.2025.sikdd.14

AI-first conferencing tools, unrestricted by established brand expectations or legacy user commitments. However, innovating in markets where most users are already committed to existing platforms has notable downsides. Only about 2.5% of people are actively seeking new alternatives, with the majority being reluctant to change [10].

3 Development of WinWin Meets

3.1 Overview

As part of our research, we developed WinWin Meets, an AI-based alternative to Zoom. The application maintains familiar functionality, allowing users to start or join meetings just as they would expect. The key difference comes before entering the meeting room, where users can define their meeting goals. Once inside, they find a familiar interface with standard video conferencing features.

These core functionalities are provided through an integration with Jitsi [11], an open-source video conferencing platform. It supports screen sharing, microphone and camera toggling, chat-based communication, polls, and many other standard features.

Beyond the familiar main meeting window found in applications like Zoom, WinWin Meets adds a dedicated panel on the right side of the screen for the WinWin Agent. This panel features two main buttons: Summarize and Give Advice.

The Summarize button generates meeting summaries up to the current moment, particularly useful for late arrivals. Hovering reveals three options: Short Text, Long Text, and Mind Map. While the text options provide traditional summaries of varying length, the Mind Map offers a quicker and more accessible visual overview. The idea behind the mind map is based on the observation that modern workplace attention is highly fragmented, with a median focus duration of just 40 seconds on any screen [12].

The Give Advice button offers guidance on how to achieve the goals specified before the meeting. These goals can also be adjusted during the meeting by clicking the Manage Goals button in the top right corner. Hovering over the Give Advice button reveals three options: Short Text, Medium Text, and Long Text, which provide advice in different levels of detail.

Once the meeting concludes, a meeting report is quickly generated. The report includes all key points, action items, a meeting timeline, and the list of participants. Users can also generate a mind map from the final meeting content.

3.2 System Architecture and Implementation

The frontend of the application was developed in Cursor [13], with assistance from Claude 3.7 Sonnet [14] and GPT-4o [15]. It is built using the React 19 framework [16]. We aim for a clean and minimalistic design that intuitively guides the user through each step of the interface.

In the meeting room interface, we integrated Jitsi via its iframe API. Jitsi integration is straightforward, and the platform allows the use of its hosted servers for up to 25 active monthly users free of charge, which was sufficient for our prototype testing.

The backend is built in Python, using the FastAPI framework [17]. For transcription, we integrated Whisper [18], and for natural language processing tasks (such as summarization and advice generation), we used GPT-4o. The backend exposes several endpoints, including:

- Transcription
- Advice generation
- Meeting summarization
- Health monitoring
- Meeting notes
- File uploads

The WinWin Agent dynamically adapts to the language selected by the user. In this prototype, we supported English, German, and Slovene, allowing users to interact with the summarization and advice features in their preferred language.

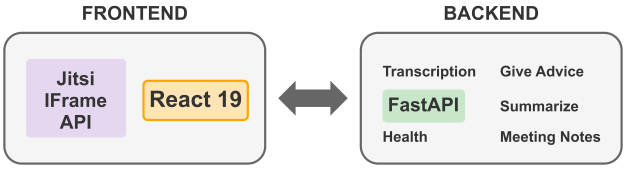


Figure 1: System architecture of the WinWin Meets application

In this prototype version, we did not use any persistent database; all data is stored locally. Additionally, user authentication is not yet implemented, as the focus was on demonstrating core functionalities.

4 Testing and User Insights

To evaluate the usefulness and usability of WinWin Meets, we conducted a structured user testing process involving 16 participants. Testing sessions were held in small groups of 2 to 4 participants, each lasting approximately 15 minutes. Participants simulated realistic discussions—including casual exchanges and role-play scenarios such as negotiations or political debates—to test all implemented functionalities. The following sections present our testing results, with key findings shown in Figure 2.

4.1 Test Coverage

Participants explored all key features, including the three variants of the Summarize function (Short Text, Long Text, and Mind Map), the three formats of the Give Advice function (Short, Medium, Long), and the Meeting Notes feature. After each session, they completed an anonymous survey with both multiple-choice and open-ended questions to assess usefulness and provide feedback.

4.2 Key Findings

General Usefulness
Most participants recognized the potential of AI-enhanced meetings. In fact, 87.5% responded *Yes* when asked whether AI could help them achieve meeting goals, while the remaining 12.5% answered *Maybe*.

Summarize Feature
The Summarize function was considered useful by 81.3% of participants. Preferences were split almost evenly: nearly half favored the Short Text, another 43.8% opted for the Mind Map, while only 12.5% selected the Long Text variant.

Give Advice Feature
When choosing advice length, participants showed a clear preference for medium-length suggestions:

- 50% selected Medium
- 25% chose Short
- 25% chose Long

Meeting Notes Feature
Participants emphasized three expectations for meeting notes:

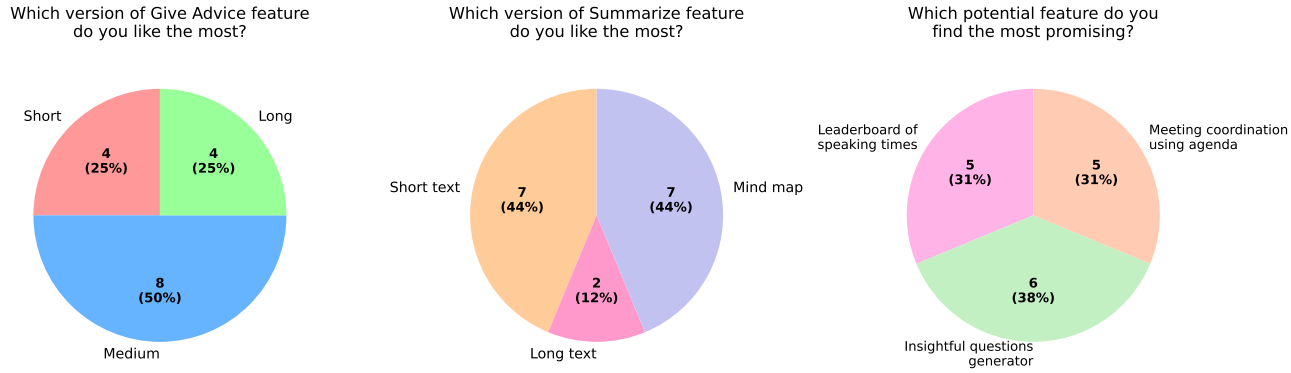


Figure 2: User survey results (n=16) comparing preferences for existing features (Give Advice and Summarize) and ranking of proposed new features for application WinWin Meets

- High reliability (timestamps, content accuracy)
- Fast post-meeting availability
- Stable performance across sessions

4.3 Ideas for Additional Features

Among the proposed additions, the insightful question generator attracted the most interest (37.5%), while the speaking time leaderboard and agenda-based coordination were equally valued (31.3% each). Participants also suggested several custom features, including personal notes, live transcription export, cloud synchronization, calendar integration, live translation with tone analysis, and domain-specific modes for law, sales, or education.

4.4 Integration Preferences

A clear majority (68.8%) preferred to use WinWin Meets as an add-on to existing platforms, while only 31.2% supported a standalone application.

4.5 Use Cases by Industry

Participants identified several promising domains for WinWin Meets, such as negotiation and sales, legal and consulting services, corporate meetings, academic events, client feedback sessions, NGO coordination, and specialized contexts like logistics, mergers and acquisitions, or trade deals.

5 Market Research and Trend Analysis

Beyond developing and testing WinWin Meets, we conducted market research to understand user needs and expectations in the video conferencing space. Our approach combined online surveys, social media engagement, search trend analysis, and reviews of blog posts and user forums. This investigation aimed to reach a wider audience than application testing alone could provide. The resulting quantitative and qualitative insights complement rather than replace our user testing results.

5.1 Survey and Social Media Feedback

Informal polls and surveys were conducted on platforms such as Facebook and Reddit. In a Facebook group focused on digital tools (GrowthHacking Slovenia), a poll asking users which feature they would most like to add to Zoom revealed that over 60% of respondents preferred having meeting notes generated at the end of a call as we can see in Figure 3. In contrast, only two respondents selected a real-time AI assistant. This suggests a clear user preference for simple and familiar enhancements over more complex and unfamiliar innovations.

Similar sentiment was observed on Reddit (r/Zoom and r/remotework), where posted polls received limited engagement. Among the few responses, a general disinterest in AI-based meeting assistance was evident, with some users explicitly selecting “None of those”.

5.2 Search Behavior and Online Interest Trends

Public search trends were analyzed using tools such as Answer the Public [19], Answer Socrates [20], AlsoAsked [21], and Ubersuggest [22]. These platforms provided insight into the types of questions users search for on Google, YouTube, and Reddit. The analysis showed minimal interest in AI-enhanced conferencing features. Instead, users were more focused on improving the efficiency and effectiveness of their meetings.

Popular search queries we found included:

- What are the 3 C’s of effective meetings?
- What is the 10-10-10 rule for meetings?
- How can I take better meeting notes?
- What are the 5 P’s of meeting productivity?
- How to extend the 40-minute limit on Zoom?
- Is Google Meet better than Zoom?
- Is Zoom free to install and use?

These patterns confirm that users are primarily concerned with meeting outcomes and platform reliability, rather than with novel AI-driven functionalities.

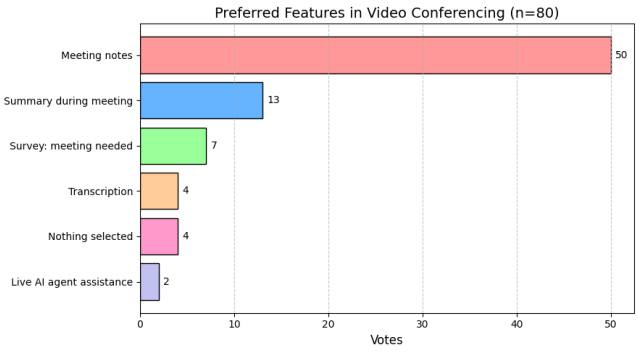


Figure 3: Distribution of 80 votes for preferred video conferencing features from our informal polling.

5.3 Forum Discussions and Deep-Search Insights

Using tools like Grok [23] and Floth [24], we conducted a deeper exploration of online discussions and feedback. The most frequently mentioned user pain points include:

- Low video quality and unstable connections
- Privacy concerns (e.g., Zoom bombing, data storage policies)
- Psychological fatigue from constant camera presence
- Lack of end-to-end encryption and transparency
- Poor UX from interface changes (e.g., Google Meet "floating bubbles", Webex chat restrictions)
- Discomfort with platform claims over recorded content

User feedback highlights a desire for reliable, simple, and secure platforms with minimal friction in setup and usage.

5.4 Conclusions from Market Research

Our market analysis reveals several key trends:

(1) Users strongly prefer practical features like note-taking and agenda management over complex AI-based tools.
(2) Popular search queries suggest a need for structured meeting frameworks and productivity strategies.
(3) Persistent dissatisfaction exists around technical reliability, interface design, and data privacy.
(4) Open-source alternatives offer control and security but are hindered by usability and cost barriers.

Overall, the market exhibits demand for video conferencing improvements that enhance meeting effectiveness and reduce user burden, rather than introducing new technical complexity.

6 Discussion

There are two primary approaches to understanding user preferences: direct inquiry and behavioral observation. Direct questioning suffers from significant limitations, including social desirability bias where respondents provide socially acceptable rather than genuine answers, and the fact that approximately 95% of human decisions occur subconsciously as discussed in [25]. Observational methods capture the unconscious preferences that drive actual user behavior, providing more accurate insights into real-world usage patterns.

These methodological considerations explain our contradictory findings. While 87.5% of WinWin Meets participants believed AI could help achieve meeting goals, market research revealed minimal organic interest in AI-enhanced conferencing. This divergence reflects the difference between conscious evaluation in controlled environments versus unconscious behavioral preferences that emerge during natural usage. Additionally, our testing participants were primarily young AI researchers, likely more receptive to AI features than typical users.

Our research uncovered widespread "Zoom fatigue", indicating that users have reached cognitive saturation with current video conferencing complexity. The strong preference for meeting notes over real-time AI assistance (60% versus minimal interest) demonstrates users' desire for post-meeting value without additional in-meeting cognitive burden. This psychological context explains why solutions that prioritize seamless integration over feature prominence tend to gain market traction [26].

Our findings suggest distinct pathways for AI-enhanced video conferencing innovation. Industry-specific applications such as negotiations, sales, and legal consultations represent focused market segments where specialized AI features deliver measurable value propositions. The 68.8% preference for add-on integration over standalone applications indicates a market opportunity in enhancing existing platforms rather than replacing them, as demonstrated by successful tools like Fathom and Otter.ai. Although there is room for breakthrough products, any new solution must be at once reliable, easy to use, and meaningfully smarter than current tools—a difficult balance as existing platforms already invest heavily in their core features.

The emphasis on reliability and customizable AI assistance reveals that AI features must meet higher performance standards than traditional features. Users consistently prioritize dependable functionality over advanced capabilities, suggesting that product development should focus on perfecting core AI functions before expanding feature sets. Future research should examine longitudinal adoption patterns and explore how user acceptance evolves as AI capabilities mature and become more familiar in workplace contexts.

7 Acknowledgements

The research described in this paper was supported by the TWON project, funded by the European Union under Horizon Europe, grant agreement No 101095095.

References

[1] Rutger Rienks, Anton Nijholt, and Paulo Barthelmess. 2009. Pro-active meeting assistants: attention please! *Ai & Society*, 23, 2, 213–231.
[2] Moira McGregor and John C Tang. 2017. More to meetings: challenges in using speech-based technology to support meetings. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, 2208–2220.
[3] T3 Technology Hub. 2024. Market share of videoconferencing software worldwide in 2024, by program. Statista. Graph. (Apr. 2024). Retrieved Jan. 13, 2025 from https://www.statista.com/statistics/1331323/videoconferencing-market-share/.
[4] tldx Solutions GmbH. 2025. Tl;dv. https://tldv.io/. Accessed: August. (2025).
[5] Otter.ai, Inc. 2025. Otter.ai. https://otter.ai/. Accessed: August. (2025).
[6] 2025. Fathom. https://fathom.video/. Accessed: August. (2025).
[7] 2025. Fireflies. https://fireflies.ai/. Accessed: August. (2025).
[8] 2025. Sembly ai. https://www.sembly.ai/. Accessed: August. (2025).
[9] Zoom Video Communications. 2025. Zoom ai companion. https://www.zoom.com/en/ai-assistant/. Accessed: August. (2025).
[10] Everett M Rogers, Arvind Singhal, and Margaret M Quinlan. 2014. Diffusion of innovations. In *An integrated approach to communication theory and research*. Routledge, 432–448.
[11] 8x8, Inc. 2025. Jitsi. https://jitsi.org/. Accessed: August. (2025).
[12] Gloria Mark, Shamsi T. Iqbal, Mary Czerwinski, Paul Johns, and Akane Sano. 2016. Neurotics can't focus: an in situ study of online multitasking in the workplace. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 1739–1744.
[13] Anysphere Inc. 2025. Cursor. https://cursor.sh/. Accessed: August. (2025).
[14] Anthropic. 2025. Claude 3.7 sonnet. https://www.anthropic.com/news/claude-3-7-sonnet. Accessed: August. (2025).
[15] OpenAI. 2025. Gpt-4o. https://openai.com/index/hello-gpt-4o/. Accessed: August. (2025).
[16] Meta Open Source. 2025. React. https://react.dev/. Version 19. Accessed: August. (2025).
[17] Sebastián Ramírez. 2025. Fastapi. https://fastapi.tiangolo.com/. Accessed: August. (2025).
[18] OpenAI. 2025. Whisper. https://openai.com/research/whisper. Accessed: August. (2025).
[19] NP Digital. 2025. Answer the public. https://answerthepublic.com/. Accessed: August. (2025).
[20] 2025. Answer socrates. https://answersocrates.com/. Accessed: August. (2025).
[21] Candour. 2025. Alsoasked. https://alsoasked.com/. Accessed: August. (2025).
[22] Neil Patel Digital. 2025. Ubersuggest. https://neilpatel.com/ubersuggest/. Accessed: August. (2025).
[23] xAI. 2025. Grok. https://grok.x.ai/. Accessed: August. (2025).
[24] 2025. Floth. https://floth.ai/. Accessed: August. (2025).
[25] Gerald Zaltman. 2003. *How Customers Think: Essential insights into the mind of the market*. Harvard Business Press.
[26] Fred D Davis. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319–340.