International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4176 - 4179
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0540

Memora – AI-Powered Smart Meeting Platform

Dr. Prof. Sarita Balshetwar[1], Vaishnavi Deshmukh[2], Snehal Pawar[3], Swaranjali Gaikwad[4]
[1]Guide, Dept. of CSE, Yashoda Technical Campus, Satara, Maharashtra, India
[2,3,4]UG Scholar, Dept. of CSE, Yashoda Technical Campus, Satara, Maharashtra, India
Emails: *csehod_ytc@yes.edu.in[1], deshmukhvaishnavi864@gmail.com[2], snehaldpawar1709@gmail.com[3], swaranjaligaikwad@gmail.com[4]*

Abstract
Memora is an AI-powered smart meeting platform designed to automate and enhance the documentation of meetings. In today's fast-paced professional environment, critical information shared during meetings is frequently lost, action items go untracked, and manual note-taking reduces participant engagement. To address these challenges, Memora captures meeting audio in real-time, converts speech to text using advanced speech-to-text engines, and applies Natural Language Processing (NLP) to generate concise summaries and extract actionable insights. The platform provides a user-friendly dashboard for reviewing transcripts, searching past meetings, and tracking action items. This paper describes the system architecture, technology stack, implementation methodology, and results of testing Memora as an assistive productivity tool for organizations.

Keywords: *Speech-to-Text; Natural Language Processing; Meeting Automation; Action Item Extraction; AI Platform; Meeting Summarization*

1. Introduction
Effective communication and knowledge management are cornerstones of organizational productivity. Meetings constitute a significant portion of work time across industries, yet much of the information exchanged is never properly recorded. Professionals spend hours manually taking notes, and even then, important decisions and action items are often missed or misunderstood. A study by Atlassian found that employees attend an average of 62 meetings per month, with nearly half considered unproductive due to lack of follow-through on action items (Atlassian, 2023). Memora addresses this problem by providing an intelligent, automated meeting assistant. The system captures audio from live microphone input or integrated conferencing platforms such as Zoom and Google Meet, converts the audio into text, and applies NLP algorithms to summarize discussions, highlight key points, and detect action items. All processed data is stored in a searchable database accessible through a centralized dashboard.

1.1.Problem Statement
Current meeting workflows suffer from several recurring issues. First, important points are frequently missed because participants cannot simultaneously engage in discussion and take comprehensive notes. Second, manual summarization after meetings is time-consuming and inconsistent. Third, action items discussed verbally often go untracked, leading to missed deadlines and accountability gaps. Fourth, retrieving information from past meetings requires searching through unstructured notes or replaying lengthy recordings. Memora is designed to eliminate each of these pain points through intelligent automation.

2. Related Work
Several commercial and research solutions have explored automated meeting assistance. Otter.ai provides real-time transcription with basic keyword highlights but offers limited summarization capabilities[1]. Fireflies.ai extends this with action item detection and CRM integrations, though its NLP accuracy in technical domains remains limited (Fireflies, 2023). Microsoft Teams Premium introduced Copilot, an AI assistant that summarizes meetings, but it is restricted to the Microsoft ecosystem and is not available as a standalone tool.
Research by Janin et al. (2003) on the ICSI Meeting Corpus established early benchmarks for automatic speech recognition in meetings. More recently,

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4176 - 4179
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0540

transformer-based models such as BART and T5 have demonstrated strong performance in abstractive summarization tasks (Lewis et al., 2020). Memora builds on these advances by combining a robust STT pipeline with transformer-based NLP to deliver accurate, contextual summaries specifically suited for professional meetings.As shown in Table 1 Comparison of Meeting Assistant Platforms[2]

Table 1 Comparison of Meeting Assistant Platforms

Feature	Otter.ai	Fireflies.ai	Memora
Real-time Transcription	Yes	Yes	Yes
Auto Summary	Limited	Yes	Yes
Action Item Extraction	No	Yes	Yes
Smart Search	Basic	Basic	Advanced
Speaker Identification	Yes	Yes	Planned
Open API Integration	Limited	Yes	Yes

3. System Design and Methodology

The Memora platform is built as a modular pipeline consisting of five layers: Input, Speech-to-Text, NLP Processing, Storage, and Presentation. Each layer is independently scalable and replaceable, allowing for future upgrades without disrupting the overall architecture[3].

3.1.Audio Input Layer

The input layer supports three modes of audio capture: (1) direct microphone input via the browser Web Audio API, (2) integration with Zoom and Google Meet via their respective developer APIs to intercept the audio stream, and (3) file upload for post-meeting processing of recorded audio files. Audio is sampled at 16 kHz mono, the optimal format for speech recognition models, and streamed in real-time chunks to the processing backend[4].

3.2.Speech-to-Text (STT) Layer

The STT layer converts raw audio into text transcripts. Memora employs OpenAI Whisper as the primary transcription engine due to its high accuracy across diverse accents and background noise conditions. Whisper is a transformer-based encoder-decoder model trained on 680,000 hours of multilingual audio (Radford et al., 2022). For real-time use cases, audio chunks are processed incrementally with a 3-second sliding window to minimize latency while maintaining contextual continuity.

3.3.NLP Processing Layer

Once the transcript is available, it passes through the NLP pipeline, which performs three tasks. First, abstractive summarization is performed using a fine-tuned BART model to compress long transcripts into concise meeting summaries. Second, action item extraction uses a Named Entity Recognition (NER) model combined with dependency parsing to identify task assignments, deadlines, and responsible parties. Third, keyword extraction and topic modeling are performed using TF-IDF scoring and LDA to generate searchable tags for each meeting[5].

3.4.Storage Layer

Processed meeting data, including the full transcript, summary, action items, and keywords, is stored in MongoDB. MongoDB's flexible document model is well suited for storing variable-length transcripts and nested action item structures. Audio files are stored in an AWS S3 bucket with presigned URLs for secure access. A secondary Elasticsearch index is maintained over the MongoDB records to support fast full-text and keyword-based search queries[6].

3.5.Presentation Layer

The frontend is built with React.js and provides three primary views: the Dashboard, the Live Transcript view, and the Meeting History view. The Dashboard displays a summary of all recent meetings, upcoming action items, and search functionality. The Live

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4176 - 4179
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0540

Transcript view shows the real-time transcription stream during an active meeting. The Meeting History view allows users to browse past meetings, view their summaries, and search for specific topics or keywords across all recorded sessions.

4. Technology Stack

Table 2 summarizes the technologies used across each layer of the Memora platform.

Table 2 Technology Stack of Memora

Layer / Component	Technology
Frontend	React.js, Tailwind CSS
Backend API	Node.js with Express.js
Speech-to-Text	OpenAI Whisper API
NLP / Summarization	BART (HuggingFace Transformers)
Action Item Extraction	spaCy NER + Dependency Parsing
Primary Database	MongoDB (Atlas)
File Storage	AWS S3
Search Index	Elasticsearch
Authentication	JWT + OAuth 2.0
Meeting Integration	Zoom SDK, Google Meet API

5. Results And Discussion

5.1.Results

The Memora prototype was tested on a dataset of 50 recorded meetings ranging from 15 to 90 minutes in length, spanning technical, managerial, and academic discussion types. The STT layer achieved a Word Error Rate (WER) of 8.3% on clear audio and 14.7% on audio with moderate background noise, comparable to the performance reported for Whisper on similar benchmarks (Radford et al., 2022). Summarization quality was evaluated using ROUGE-1 and ROUGE-L scores against human-written reference summaries, yielding scores of 0.42 and 0.38 respectively, which are consistent with state-of-the-art abstractive summarization models. Action item extraction achieved a precision of 81% and a recall of 76% on the test set, with the primary source of errors being implicit action items stated without a named assignee. The search functionality returned relevant results with a mean average precision (MAP) of 0.87 across keyword queries. Average end-to-end latency from speech input to transcript display was 3.8 seconds, which is within acceptable bounds for near-real-time use.

5.2.Discussion

The results confirm that Memora can reliably automate the documentation of meetings in most real-world conditions. The WER increase under noisy conditions highlights a known limitation of current STT systems; integrating noise suppression preprocessing could further reduce this gap. The action item recall of 76% indicates that approximately one in four action items expressed implicitly—without a direct verb-object-assignee structure—is missed, suggesting that a more context-aware extraction model could improve coverage. User feedback collected from 20 test participants rated the platform's ease of use at 4.3/5 and the usefulness of generated summaries at 4.1/5. Participants noted that the automatic action item list reduced post-meeting follow-up time by an estimated 40%. These findings support the hypothesis that AI-assisted meeting management can significantly improve professional productivity.

Conclusion

This paper presented Memora, an AI-powered smart meeting platform that automates transcription, summarization, and action item extraction from meetings. The system achieved strong performance across all key metrics, including an 8.3% WER for clear speech, ROUGE-1 of 0.42 for summarization, and 81% precision in action item detection. User evaluation confirmed meaningful improvements in

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4176 - 4179
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0540

meeting productivity. Future work will focus on implementing speaker diarization for identifying individual speakers, multi-language support, sentiment analysis to assess meeting tone, and integration with calendar and task management systems such as Google Calendar and Jira. The platform demonstrates that intelligent automation of meeting workflows is both technically feasible and practically valuable for modern organizations.

Acknowledgements

The authors would like to thank the faculty of the Department of Computer Science and Engineering at Yashoda Technical Campus, Satara for their guidance and support throughout this project. Special thanks to the participants who took part in user testing and provided valuable feedback. This work was carried out as part of the undergraduate final year project program and received no external funding.

References

[1]. Atlassian. (2023). The state of meetings. Atlassian Teamwork Lab Report.
[2]. Janin, A., Baron, D., Edwards, J., Ellis, D., Gelbart, D., Morgan, N., & Shriberg, E. (2003). The ICSI Meeting Corpus. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 1, 364–367.
[3]. Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., & Zettlemoyer, L. (2020). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. Proceedings of ACL 2020, 7871–7880.
[4]. Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust speech recognition via large-scale weak supervision. arXiv preprint arXiv:2212.04356.
[5]. Fireflies.ai. (2023). AI meeting assistant overview. Retrieved from https://fireflies.ai/
[6]. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of NAACL-HLT 2019, 4171–4186.