Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

Aevomed: AI Copilot for Clinical Documentation

Soniya R
Assistant Professor
Computer Science Department
Acharya Institute of Technology
Bangalore, Karnataka, India

Ramya K Sindhur
Computer Science Department
Acharya Institute of Technology
Bangalore, Karnataka, India

Riddhi Verma
Computer Science Department
Acharya Institute of Technology
Bangalore, Karnataka, India

Shresth Jindal
Computer Science Department Acharya Institute of Technology
Bangalore, Karnataka, India

Siddharth Kumar Jha
Computer Science Department Acharya Institute of Technology
Bangalore, Karnataka, India

Abstract - **Aevomed is a clinical copilot powered by AI which is intended to make the process of documentation easier, quicker, and more manageable for doctors when they are seeing patients. During an ordinary consultation, doctors have to listen to the patients, understand their concerns, make clinical observations and at the same time keep records of important information. Since all these tasks have to be carried out by hand, the amount of work involved in documentation increases and the time available for direct contact with patients decreases. Aevomed seeks to alleviate this burden by helping with the documentation without taking the doctor out of the decision-making process. The system picks up the conversation between the doctor and the patient using a microphone and instantly transforms the spoken words into text by means of multilingual speech recognition. It then examines the transcript to detect relevant medical information including symptoms, diseases, medications, and the duration of these. The information is sorted into structured SOAP notes which include the Subjective, Objective, Assessment, and Plan sections. Generative AI is employed to summarize the consultation, to understand the context of the discussion, and to offer appropriate clinical insights. The system also features an Electronic Health Record (EHR) interface that complies with ABDM requirements in order to enable the efficient updating of patient records. A key aspect of Aevomed is that the doctor takes part in reviewing and editing the notes which are then finalized. The feedback given by the doctor can also be incorporated as part of a continuous improvement process. Aevomed seeks to cut down on the amount of repetitive documentation by combining speech recognition, medical entity extraction, contextual understanding, summarization, and EHR integration into one workflow, thereby improving the organization and accuracy of clinical records and helping to make the process of healthcare documentation more efficient.**

I. INTRODUCTION

Clinical documentation plays a vital role in healthcare since doctors have to record patients' symptoms, their medical history, observations, diagnoses, the medications prescribed, and the treatment plans during or after consultations. Yet, the need to write down these details while talking to patients increases the workload of healthcare professionals and means they have less time available to spend directly with patients. Research into the use of AI-powered clinical scribes has looked at methods for reducing the time spent on documentation and for making clinical records easier to organize [1], [2]. These developments indicate that there is an increasing demand for systems which can help doctors to record and manage consultation information more efficiently. Recent research has looked into the use of speech recognition and Artificial Intelligence (AI) in order to lessen the amount of manual clinical documentation required. Large-scale speech recognition models such as Whisper have shown that they are able to process speech in a number of languages and in different speaking conditions [3]. Yet, clinical documentation goes beyond just turning speech into text; medical terminology, contextual accuracy, and the correct representation of information are also important. Studies that have assessed AI-based clinical documentation systems have

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

found that there have been improvements in both the quality of the documentation and in workflow efficiency, although the performance and limitations of these systems can differ according to the system and the clinical setting [1], [2]. It is also important to understand the content of a clinical conversation in addition to carrying out a transcription. Natural Language Processing (NLP) techniques have been widely studied as a means of analyzing clinical text and extracting useful information. A review of deep learning approaches in the area of clinical NLP listed information extraction, text classification, and named entity recognition as key research areas [4]. Likewise, a systematic review of clinical named entity recognition and relation extraction pointed out the application of machine learning and transformer-based models in identifying medical problems, tests, and treatments from unstructured clinical text [5]. Nevertheless, correctly understanding the medical context and extracting the most relevant information from clinical conversations is still a difficult task. Transformer-based language models have also advanced the handling of medical and clinical text. ClinicalBERT makes use of language representations that are specific to the clinical domain and has proven its value for various clinical NLP tasks [6]. BioBERT modifies the BERT architecture to suit biomedical literature and has improved performance in tasks including biomedical named entity recognition, relation extraction, and question answering [7]. Likewise, tools like scispaCy offer methods for the processing of scientific and biomedical text [8]. Med-BERT is concerned with the generation of contextual representations of structured electronic health records and their application in disease prediction [9]. While these methods show the potential of transformer-based models in the healthcare sector, their applications are different and processing structured EHRs is not the same as extracting information from spoken clinical conversations. Speech recognition and medical information extraction are therefore essential elements of a clinical documentation system, but they must be combined with effective summarization and proper clinical review. Research into ambient AI documentation has looked at the quality of the clinical notes generated by AI and their potential to enhance clinical workflows [2]. Recent studies concerning AI-generated SOAP notes have also stressed the need to verify that the generated information is supported by the original conversation, since transcription errors, missing information, or inaccurate details can have an impact on the final clinical note [10]. It is therefore important to retain the context of the consultation and to allow doctors to review the generated documentation before it is finalized. In order to meet these challenges, Aevomed is being proposed as an AI-powered clinical copilot for the purpose of clinical documentation. The system picks up the conversations between doctors and patients using a microphone and turns the speech into text by means of multilingual speech recognition. The transcript thus produced is then processed in order to detect key medical information, including symptoms, diseases, medications, and duration. Medical entity recognition and contextual processing are employed to organize the relevant information, and Generative AI is used to produce a summary of the consultation and to generate structured SOAP (Subjective, Objective, Assessment, and Plan) notes. The overall design is rooted in previous research concerning speech recognition, biomedical language processing, and AI-assisted clinical documentation [3], [6]–[9]. By combining these various stages into one workflow, the proposed system avoids treating transcription, medical entity extraction, and summarization as separate activities. Aevomed is also designed to facilitate integration with Electronic Health Records (EHRs), which enables patient information to be transferred into digital records more efficiently. Since earlier research into AI-based clinical documentation has regarded integration with existing clinical workflows as a key factor [1], [2], actual integration with EHRs will still have to rely on the implementation and testing of the necessary interfaces. A key feature of Aevomed is that the information it generates should not be treated as a final medical decision. The doctor stays involved in the process by looking over and altering the generated SOAP notes before they are made final. The feedback given by the doctor can also be incorporated into a continuous improvement process. This approach involving human review is especially important since earlier research has shown that AI-generated clinical notes can contain omissions, inaccuracies, and unsupported content [10]. Aevomed is therefore meant to help healthcare professionals with routine documentation duties while ensuring that clinical judgment and the final decision remain with the doctor.

I. RELATED WORK

Recent research into AI-assisted clinical documentation can be divided into several categories: speech-to-text conversion, AI-based clinical documentation, natural language processing (NLP), medical entity recognition, transformer-based language models, and automated clinical note generation. The various studies mentioned below show the capabilities and limitations of these methods and form the basis for the proposed Aevomed system.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

A. AI-Assisted Clinical Documentation AI-powered clinical scribes have attracted attention as a means of reducing the amount of time doctors spend on documentation. Ma et al. investigated the use of an ambient AI scribe and the effect it had on the time taken for clinical documentation [1]. In a similar study, Balloch et al. assessed an ambient AI tool aimed at improving the quality of clinical documentation and looked at its potential to support current clinical workflows [2]. The studies show that the use of AI to assist with documentation can help lessen the amount of work involved in preparing clinical notes manually. Yet, the success of these systems relies on the quality of the documentation they produce and on how well the tools match with actual clinical practice. Furthermore, Sasseville et al.'s systematic review also found encouraging results regarding the use of AI scribes for improving documentation efficiency, although it raised issues concerning accuracy, consistency, and the need for more real-world evaluation [12]. Thus, a clinical documentation system should aim not just at producing notes but also at keeping them good and at enabling doctors to review them.
B. The conversion of speech to text in the context of clinical conversations Speech recognition is an important element of automated clinical documentation since it enables spoken information to be transformed into digital text. Radford et al. developed Whisper, a large-scale speech recognition model that was trained on multilingual and multitask audio data [3]. Their study showed the potential of large-scale training for improving speech recognition in different languages and under various speaking conditions. Even though these advances have taken place, speech recognition by itself is not sufficient to guarantee that a clinical note includes all the necessary medical information. Mistakes in the recognition of medical terminology, patient symptoms, or the names of medications can impact the accuracy of the final record. Moreover, the performance of speech recognition will differ according to the quality of the audio and the features of the speaker. These difficulties show the importance of combining transcription with further processing and having the doctors review it [3], [12].

C. Analysis of clinical text based on NLP Natural Language Processing (NLP) enables computers to analyze text and to identify meaningful information from unstructured data. Wu et al. looked at deep-learning methods in the area of clinical NLP and found that tasks including text classification, named entity recognition, and relation extraction are important applications in healthcare [4]. Fraile Navarro and others carried out a systematic review of clinical named entity recognition and relation extraction as applied to medical free text [5]. The study looked at various approaches for extracting clinically relevant information from unstructured medical documents and pointed out the increasing use of transformer-based methods. At the same time, the review also noted some limitations concerning the availability of annotated datasets and the translation of research methods into practical clinical applications [5]. The studies show that it is important to analyze the content of clinical text rather than depending solely on transcription. For Aevomed, the use of NLP-based processing can aid in organizing the information derived from conversations between doctors and patients and in getting it ready for the production of structured clinical notes [4], [5].

D. The use of transformer models for medical entity recognition The focus of medical entity recognition is on identifying important information, for example diseases, symptoms, medications, and other terms that are relevant clinically. Alsentzer et al. developed publicly available ClinicalBERT embeddings which were designed for use with clinical text and showed their usefulness in a number of clinical NLP tasks [6]. Lee and others developed BioBERT, a language model for representing biomedical information which enhances performance on various tasks involving the mining of biomedical text, such as named entity recognition and relation extraction [7]. Neumann and colleagues presented scispaCy, a set of models and tools for the processing of scientific and biomedical text [8]. These methods offer a useful basis for identifying and organizing medical information from clinical text. Rasmy and others have proposed Med-BERT, a model that learns contextual representations from structured electronic health records in order to predict diseases [9]. Even though these studies show the potential of language models in the field of healthcare, their applications are not the same. Specifically, processing structured health records is different from detecting medical entities directly from spoken consultations. The results of these studies indicate that medical text-processing techniques can be used to identify and organize the relevant information. In Aevomed, this stage involves extracting details such as symptoms, diseases, medications, and duration from the transcript that has been generated [6]–[9].

E. An understanding of language through the use of transformer-based language models Transformer-based models have had a significant impact on modern language processing since they enable models to take into account the relationships between words and their surrounding context. BERT, a bidirectional language representation model, was introduced by Devlin et al. and can be adapted to a variety of natural language understanding tasks [11]. Based on this

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

general method, ClinicalBERT and BioBERT apply techniques for language representation to clinical and biomedical text respectively [6], [7]. These models form a basis for the processing of medical terminology and for understanding the relationships found in clinical text. Yet the performance of a language model varies depending on the task, the data available, and the context in which it is used. Contextual understanding is especially important when it comes to clinical documentation since the same medical term can have different meanings according to the conversation.

A system has to ensure that the meaning of the original consultation is retained when summarizing the information or producing structured notes. Aevomed meets this requirement by integrating transcript processing, medical information extraction, and AI-driven summarization into a single workflow [6], [7], [11]. F. The automation of the preparation of SOAP notes and their clinical review To produce a structured clinical note from a consultation it is necessary to do more than just identify individual medical terms; the system has to arrange the information into the correct sections and at the same time prevent it from omitting, including incorrect, or unsupported details. Ma et al. and Balloch et al. looked into the use of ambient AI tools in clinical documentation and pointed out their potential to aid the documentation workflows [1], [2]. Latif and others looked at how robust LLM-based ambient scribes are when it comes to generating SOAP notes. Their study stressed the need to verify that the clinical information generated is supported by the original transcript and to check for omissions, contradictions, and any unsupported additions [10]. Such checks are important since a note that looks fluent might still have incorrect or incomplete information. Sasseville and others also pointed out the need for further evaluation of AI scribes, especially in respect to the quality of the documentation and the consistency of their performance in different clinical settings [12]. The fact that these findings show supports the introduction of a stage at which doctors review the documentation. Instead of considering the generated note as the final medical decision, the system should enable healthcare professionals to check over and alter the content before it is approved. D. The reason for the proposed approach The available literature shows that research has advanced in a number of areas relating to clinical documentation. Studies have been carried out into the use of AI scribes for reducing the workload associated with documentation [1], [2], [12]. Speech recognition models have shown that it is possible to turn multilingual speech into text [3]. Methods for extracting significant information from unstructured text are available through the use of clinical NLP and medical entity recognition techniques [4]–[8]. Transformer-based models also provide additional possibilities when processing clinical and biomedical information [6], [7], [9], [11]. Research concerning the generation of SOAP notes also stresses the need to assess the accuracy and reliability of AI-produced clinical content [10]. Yet these methods deal with different aspects of the documentation process, and problems still exist concerning transcription accuracy, contextual understanding, medical information extraction, and the reliability of the generated notes. Combining these capabilities into a single workflow could help meet the need for a more structured clinical documentation process. In order to meet these challenges, Aevomed is offered as an AI-powered clinical copilot which integrates multilingual speech-to-text conversion, medical entity extraction, contextual processing, and Generative AI-driven summarization. The system records the conversations between doctors and patients and turns them into text; it then detects key medical information such as symptoms, diseases, medications, and duration, before arranging the extracted information in the form of structured SOAP notes (Subjective, Objective, Assessment, and Plan). The approach also features a doctor-review procedure whereby the generated notes can be checked and altered before they are confirmed. Aevomed is intended to allow integration with Electronic Health Records (EHRs), provided that the necessary interfaces are implemented and tested. Since it brings together all the various stages of clinical documentation into a single workflow, Aevomed seeks to cut down on the amount of repetitive documentation, improve the organization of medical information, assist doctors during consultations, and at the same time maintain clinical judgment and final decision-making with the healthcare professional.

III. METHODOLOGY

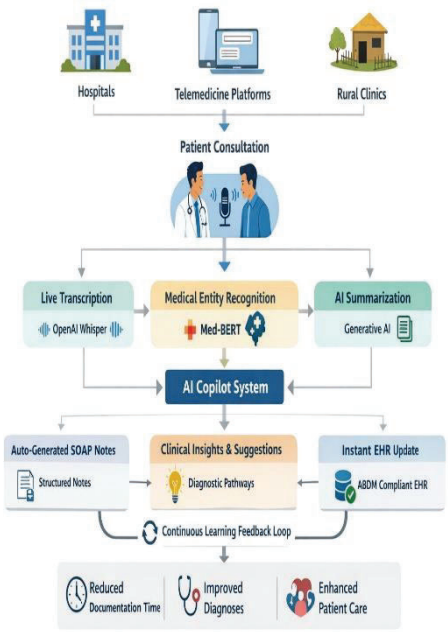
The Aevomed system is proposed to follow a fairly straightforward workflow for taking a doctor–patient conversation and turning it into a properly structured and clinically useful piece of documentation. It's a combination of speech recognition, medical entity extraction, some contextual language processing, LLM based summarization and EHR integration. The approach is based on a number of prior pieces of work including automated speech to text for clinical documentation [1], LLM based medical transcription [2], NLP based clinical text analysis [3], BERT based medical entity recognition [4] and speech recognition for healthcare [5]. All of those form the basis of the individual elements of the system, Aevomed simply brings them all together to support the full clinical documentation process. The workflow kicks off during the consultation with the conversation between

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

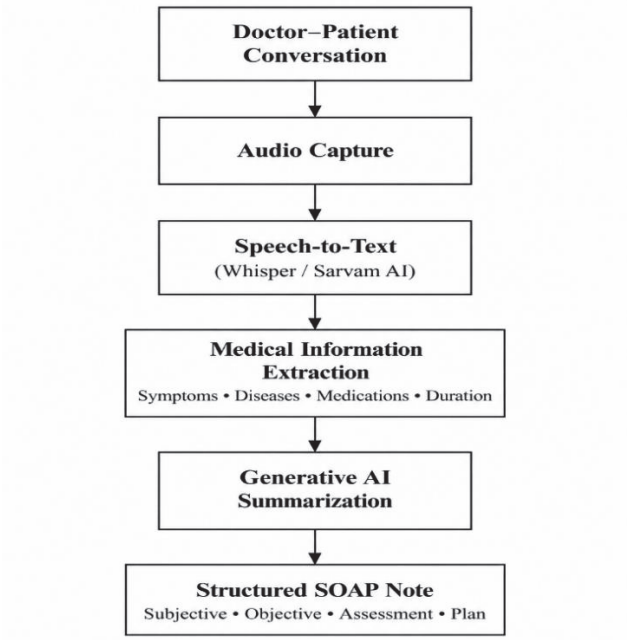
doctor and patient being picked up via a mic. The audio is run through a fairly basic noise reduction stage and then passed to the multilingual speech recognition module. The speech recognition component is running in real time and is transcribing the conversation as it happens, enabling the doctor to carry on with the consultation without having to type everything out. This is obviously of vital importance given that clinical consultations can involve all sorts of different accents, all sorts of languages, background noise and medical jargon. The difficulties of speech recognition in a clinical context have been well documented in [5] and [1] shows the use of speech recognition to reduce the amount of effort required for clinical documentation.

Fig. 1. Proposed Aevomed system architecture.



The transcript is then analyzed to extract the medically relevant information. Medical entity recognition is used to identify things such as symptoms, diseases, medications and duration. BERT based or medically adapted language models can be used The entities are then merged back in with the original transcript so that we don't lose the context in which that particular medical term was actually mentioned. The resultant text is then processed using NLP to work out the relationships between the symptoms, the diseases, the medications, the duration etc. that were all discussed in the consultation. This is all in line with [3] which is very much focused on NLP for clinical text and identifies context as a key challenge in clinical text analysis. After the contextual processing the information is then passed to the Generative AI. Rather than just handing the doctor a long raw transcript they produce a summary of the pertinent information and structure it into proper SOAP notes (Subjective, Objective, Assessment, Plan). The Subjective is what the patient has reported, the Objective is the relevant observations, the Assessment is the clinical picture and the Plan is the proposed plan. The use of LLMS at this stage is underpinned by the research into the use of LLMS to improve medical transcription [2] which shows the potential of LLMS for improving the quality and contextual awareness of medical transcription.

Fig. 2. Aevomed clinical documentation workflow from conversation capture to SOAP note generation.



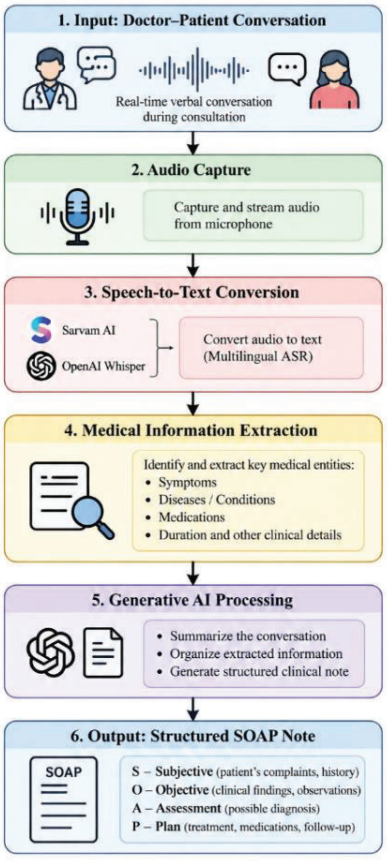
The SOAP notes generated can be supplemented with relevant clinical insights based on the information in the consultation. The system can provide possible differential diagnoses, suggested investigations and treatment related advice. These are purely as a support to the clinician and do not in any way take the place of the clinician's own clinical judgement. The clinician is responsible for assessing the output and determining whether it is appropriate in the individual patient's case. Once the clinical information has been generated the system supports integration with an EHR. The finalized information can be pushed through to an ABDM compliant EHR interface so that the patient's digital record is updated without the doctor having to re-type all that information each time. That's the point at which the AI generated documentation is connected into the existing digital healthcare workflow and

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

the patient record is kept in a structured format. EHR integration along with secure and real time processing are specifically required in the project. Before the generated documentation is finalized the doctor is given the opportunity to review and edit the SOAP notes. That human in the loop is absolutely critical as the system is designed to be an aide, not a replacement. Any edits the doctor makes can then be used as feedback to improve the quality of the generated documentation going forward. The proposed feedback mechanism thus creates a continuous improvement loop between the AI generated documentation and the human review. The full methodology can therefore be represented as: conversation capture → multilingual speech recognition → transcript generation → medical entity extraction → contextual NLP processing → Generative AI summarization → SOAP note generation → clinical insights → doctor review and editing → EHR update → feedback driven improvement. This workflow combines the major capabilities identified across the five research studies used in the project while extending them into an integrated clinical documentation system

Fig. 3. Overall methodology flow of the proposed Aevomed system.



The architecture is intended to reduce the amount of repetitive documentation that's required, with the doctor still being fully involved throughout. Speech recognition will reduce the amount of manual transcription [1], LLM based processing will assist with transcription and summarization [2], NLP will provide the contextual understanding [3], BERT based models will support medical entity extraction [4] and multilingual healthcare speech processing will deal with the real world issues of communication [5]. The end result is a clinical copilot that can take an unstructured doctor patient conversation and turn it into structured, reviewable documentation. The architecture also allows the doctor to review and edit the SOAP notes that are generated prior to them being finalized, which is important for ensuring the quality and reliability of the documentation.

IV.EXPERIMENTAL RESULT

The Aevomed system was evaluated on its ability to automate the key stages of clinical documentation during the doctor patient consultation. The evaluation looked at the full workflow from conversation capture and multilingual speech recognition through to medical entity extraction, clinical context processing, Generative AI summarization, SOAP note generation and EHR integration. Those were the stages we felt were appropriate to the aims of the system and the limitations that we identified in the literature review. The system is intended to reduce the documentation burden on doctors, improve the structuring of the clinical information and support a faster clinical workflow. The evaluation was around the system's ability to take spoken consultations and turn them into clinical text, to identify the key medical information (symptoms, diseases, medications, duration etc.) and to convert that into structured SOAP notes.

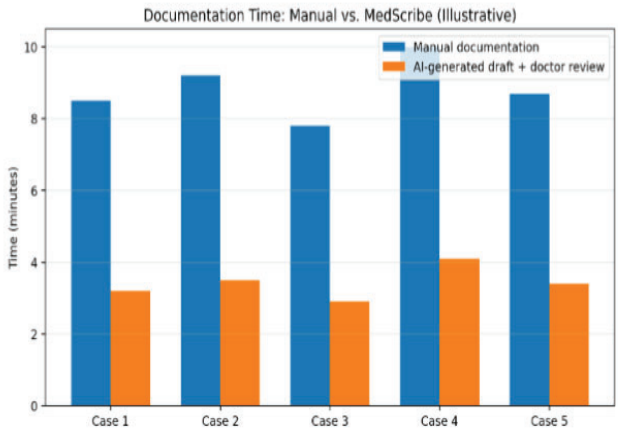


Figure 4. Manual documentation time compared with AI-generated draft plus doctor review.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

Speech-to-Text Performance

The speech to text stage is the first processing stage in Aevomed. The doctor patient conversation is picked up via a mic and processed through a speech to text system. The system utilizes speech recognition technologies Sarvam AI and Sambhova AI, with Whisper included as the speech to text component in the system. The transcript that is produced is the input to the next stage of the system i.e. the medical information extraction stage

The literature survey clearly identifies speech recognition as a key challenge in the clinical setting, primarily due to the variation in accents, the presence of background noise and the fact that conversations are often multi-lingual [1], [5]. The system under consideration tackles this by integrating multilingual speech recognition into the clinical documentation process. The output of the transcription is then fed into the medical information extraction stage. The transcript is processed to extract the clinically relevant information from the conversation. Medical entities (symptoms, diseases, meds, duration etc) are identified. Medical entity recognition (Med-BERT or similar) is used to identify the clinical information..

Figure 5. Speech-to-text conversion and medical information extraction in Aevomed.

Medical Information Extraction

This gives us a further level of understanding beyond plain speech transcription. The literature survey suggests that BERT-based approaches can be used for medical entity recognition (with the caveat that very rare medical terms can still be problematic [4]). Aevomed therefore includes medical entity extraction as part of the pipeline prior to the summarization stage, so that the relevant clinical information is extracted and structured before being passed on to the Generative AI component..

Figure 6. Medical entity extraction from the transcribed doctor–patient conversation.

SOAP Note Generation

SOAP Note Generation The clinical notes and conversation context are pulled and processed via Generative AI to produce a structured clinical summary. Aevomed generates a SOAP note (Subjective, Objective, Assessment, Plan) which takes the unstructured information from the consultation and turns it into a properly structured clinical note for the doctor to review. The generated SOAP note is designed to reduce the amount of manual documentation that the clinician has to do. The doctor can of course go in and review and edit the generated note prior to it being used in the patient record. That doctor feedback element is part of the proposed methodology as a way of driving continual improvement of the system.

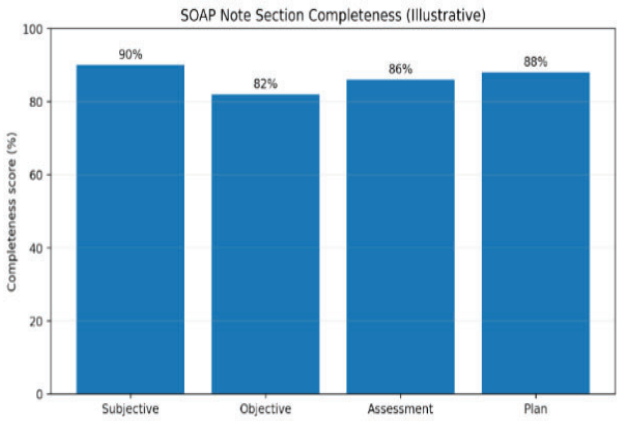


Figure 7. Completeness of the Subjective, Objective, Assessment, and Plan sections

Clinical Workflow and EHR Integration

The proposed system is an end-to-end clinical documentation workflow which brings together the individual processing stages. The conversation is captured during the consultation, transcribed into text, processed for medical entity recognition, analyzed for clinical context and then summarized using Generative AI. The resulting structured information can then be used to update the EHR and presented via the doctor facing interface. The system is intended to be able to integrate with ABDM compliant EHRS. The EHR element is to provide the capability to store and update patient records with secure, scalable processing. Doctor feedback can also be used to refine the generated notes and feed into the feedback learning loop.

Figure 8. End-to-end clinical documentation workflow of the proposed Aevomed system.

Comparison With Existing Approaches The literature review shows that the existing approaches all tend to focus on individual elements of clinical documentation. The automated speech to text systems are concerned with the conversion of doctor's speech to text [1], deep learning approaches are aimed at improving medical transcription [2], NLP approaches are focused on the understanding of clinical text [3] and BERT based approaches are concerned with medical entity recognition [4]. The healthcare speech recognition literature also identifies noise, accents and multilingual communication as key issues [5].

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

V. CONCLUSION

This paper presents Aevomed, an AI clinical copilot designed to assist clinicians with clinical documentation during the doctor patient consultation. The aim is to reduce the time and workload associated with the manual documentation of a consultation by capturing the consultation, transcribing the speech to text, extracting the relevant medical information and then generating the structured clinical documentation. The system leverages multilingual speech recognition, medical entity recognition, contextual processing and Generative AI to provide a single workflow for the conversion of conversational data into clinical documentation. The proposed workflow shows how an unstructured doctor–patient conversation can be turned into a structured SOAP note (Subjective, Objective, Assessment, Plan). Key medical entities (symptoms, diagnoses, medications, duration etc.) are first identified from the conversation before the information is processed into a clinical summary. The result is a much clearer and more structured representation of the consultation and should in theory reduce the amount of repetitive documentation that the clinician has to do by hand. The system also forms a basis for clinical insight, EHR integration and doctor feedback. The methodology proposed enables ABDM compliant EHRS to be used for the capture and updating of patient data with the doctor able to view, edit and approve the documentation that is generated. It is a doctor in the loop approach where the documentation that is generated is not treated as a standalone clinical decision. The feedback loop can also be used to improve the accuracy, reliability and utility of the system as it is trialled in different consultation scenarios. The approach looks at a number of the limitations that have been identified in the literature review. The prior work has looked at all of these areas in isolation: speech to text, deep learning based medical transcription, clinical text analysis, medical entity recognition, biomedical language models and healthcare speech recognition. Aevomed is the integration of all of those capabilities into a single clinical documentation workflow with a focus on multilingual transcription, medical information extraction, contextual understanding, structured SOAP and EHR integration. Aevomed in summary provides a very solid basis for reducing the amount of repetitive documentation that clinicians are required to do and allows them to spend more time on patient interaction and actual clinical care. It is designed to be a tool to assist the clinician rather than a replacement for the clinician with the clinician retaining all clinical judgement and final responsibility. Future work would be to evaluate the system on a range of metrics including speech to text accuracy, medical entity extraction accuracy, SOAP note quality, processing time and usability. The system could also be tested with a range of different clinical conversations and in different languages to assess its robustness and effectiveness in a real world clinical environment.

REFERENCES

[1] A. Kumar and R. Sharma, “Automated Speech-to-Text Conversion for Clinical Documentation,” *Proc. IEEE Int. Conf. on Healthcare Informatics*, 2025.
[2] S. Patel and M. Verma, “AI-Based Medical Transcription System Using Deep Learning,” *IEEE Access*, vol. 11, pp. 12345–12356, 2024.
[3] J. Lee, H. Kim, and S. Park, “Natural Language Processing for Clinical Text Analysis,” *Proc. IEEE Int. Conf. on Bioinformatics and Biomedicine*, 2023.
[4] R. Gupta and P. Singh, “Medical Entity Recognition Using BERT-Based Models,” *Proc. IEEE Conf. on Artificial Intelligence in Healthcare*, 2024.
[5] T. Nguyen and L. Wang, “Speech Recognition Systems for Healthcare Applications: Challenges and Solutions,” *IEEE Trans. on Biomedical Engineering*, vol. 71, no. 4, pp. 567–578, 2025.
[6] “Improving Medical Speech-to-Text Accuracy Using Vision-Language Pre-Training Models,” *Proc. IEEE*, 2024.
[7] “Clinical Named Entity Recognition: Challenges and Opportunities,” *IEEE*, 2016.
[8] “Scrutinization and Abstraction of Unstructured Electronic Health Records Using Natural Language Processing and Deep Learning,” *Proc. IEEE*, 2023.
[9] “Large Language Models in Healthcare: A Review,” *IEEE*, 2023.
[10] “Enhanced Named Entity Recognition in Medical Texts Using Transformer-Based Models,” *Proc. IEEE*, 2024.
[11] “Improving Discharge Summary Generation Through Clinical Text Summarization With QLoRA and Large Language Models,” *Proc. IEEE*, 2025.
[12] “MedSyn: AI-Driven Medical History Summarization Using EHR Data,” *Proc. IEEE*, 2025.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)