Eng. Sci., 2025, **33**, 1377

Engineered Science
DOI: https://dx.doi.org/10.30919/es1377

Framework for Digital Medical Scribe Technology for Healthcare Documentation

Rohi Bindal,[1,#] Prajjwal Dewangan,[1,#] Rahee Walambe,[1, 2,*] Sheela Ramanna[3] and Ketan Kotecha[1, 2,*]

Abstract

Digital medical scribe (DMS) can prove to be a major innovation in the healthcare industry because such an approach can provide a highly useful method for medical documentation and physician-patient interactions. A new era of patient-centered, accurate, and efficient medical scribing has begun with the introduction of cutting-edge technologies like voice recognition, natural language processing, and electronic health records. This is highly relevant as healthcare professionals are facing increasing administrative burdens. In this work, the significance of DMS in lowering provider burnout, streamlining clinical workflows, and improving patient care standards is discussed in detail. A case study of cardiac disorders is considered for demonstration of the proposed DMS method. Through the reduction of time spent on manual notetaking and data entry, the proposed digital solution can help physicians better connect with their patients and result in improved patient satisfaction. The ultimate objective is to improve healthcare for both patients and medical staff by reducing the administrative burden by effectively integrating digital tools into healthcare settings. The extension of this work would focus on the integration of various modalities specific to healthcare data.

Keywords: Digital medical scribes; Automatic speech recognition; Natural language processing; Electronic health report; Healthcare documentation.
Received: 21 March 2024; Revised: 16 August 2024; Accepted: 22 August 2024.
Article type: Research article.

1. Introduction

Owing to the increasing administrative tasks, there is a major concern of health provider burnout in the healthcare sector.[1,2] Work overload, administrative tasks, unavailability of support, and conflicts between work and family life are some of the scenarios quantified by the Maslach Burnout Inventory for burnout. The data entry and oversight requirements make doctors feel stifled by imposing documentation in place of interacting with someone sick. Even though electronic health records (EHRs) have made substantial improvements to health provision, these same systems have increased responsibilities for caregivers, thus taking up most of their working hours.[3-5]

[1] *Symbiosis Institute of Technology, Symbiosis International Deemed University (SIU), Pune, 412115, India*
[2] *Symbiosis Centre for Applied Artificial Intelligence (SCAAI), Symbiosis International Deemed University (SIU), Pune, 412115, India*
[3] *Department of Applied Computer Science, University of Winnipeg, Winnipeg, MB, R3B 2E9, Canada*
[#] *These authors contributed to this work equally.*
**Email:* director@sitpune.edu.in (K. Kotecha), rahee.walambe@sitpune.edu.in (R. Walambe)

The digital medical scribe (DMS) technology is reported to provide favorable outcomes and is supported by numerical information. It is reported that by employing DMS,[6] the documentation time fell by 50 percent while the transcription accuracy went up by 30 %. More than 80 % of healthcare professionals have stated that they were more productive and satisfied with their work as a result of increased precision, quicker procedures, and enhanced communication channels by using DMS.[6] The positive reviews of patients who used these studies indicate that digital scribes help clinic practices significantly. Swiftness in recording and improving precision in information using digital scribes can, therefore, aid in reducing mistakes from manual entry practices as well as make provision for prompt access to vital information, thereby promoting patient welfare and clinical judgment.[7,8]

Assisting patients through digital scribes in entering their medical information can encourage patient engagement, hence improving communication and enhancing diagnosis accuracy.[9-11] Lowering error rates constitutes one of the advantages of a complete medical record, which also enhances continued care. On one hand, it helps in getting effective treatment from physicians. Still, on the other hand, it also assists in ensuring drug adherence as well as satisfaction among patients while controlling their health. Success, however, comes only if issues such as data protection, clear and accurate scribing, clear instructions, *etc.*, are managed. Faster and more accurate communication increases the credibility of employees when delivering healthcare through the assistance of medical staff who are automated digitally.[12,13] Recent studies have shown that medical scribes are becoming increasingly important, especially in the field of orthopaedics. For example, Magyari *et al*.[14] discuss how scribes could be used within orthopaedic clinics while also presenting some ideas about their advantages as well as the problems that come with them. Another study deals with comparing different ways used for recording information, where it was concluded that the presence of scribes enhances the quality of information generated.[15] The role of artificial intelligence (AI) in healthcare documentation is reviewed by Falcetta *et al.*[16], while Tran *et al*.[17] explore the effect of non-lexical sounds on voice recognition software used in clinical recordkeeping. In Ref. [18], the authors investigate the advantages and disadvantages of digital scribe integration, while in Ref. [6], the digital scribes in cancer care settings focusing on physician stress are discussed. Digital scribes who benefit both clinicians and patients may become an integral part of healthcare documentation as research evolves. Table 1 shows some of the studies focusing on the digital technologies implemented in healthcare.

2. Materials and methods

2.1 Data collection

For the purpose of this work, the required dataset was not available in the open domain. Therefore, we created the dataset as per the project requirements. A variety of scripts with diverse dialogues were curated manually by consulting doctors. Following this, audio files were recorded to generate the dataset for testing the transcriber.

2.2 Transcriber and script generation

Fig. 1 illustrates the process flow of the proposed system for integrating technologies derived from ASR and NLP. The main components and it's features of the system are described below:

Real-Time ASR Transcription: The ASR technology translates real-time spoken conversations between patients and healthcare providers into precise text with ease. The whisper Application Programming Interface (API) is used for real-time transcription.

User-Friendly Dashboard: Physicians and patients interact primarily through an easy-to-use dashboard. Real-time access to the transcribed conversations will be possible, enabling quick review and any necessary edits.[29]

Language and Context Recognition: The system is built with the ability to identify and adjust to the language used in medical consultations. It is context-aware, which increases the accuracy by taking into account the technical terms used in the healthcare industry.

Integration with electronic health records (HER): The direct transfer of medical records and transcribed notes, made possible by seamless integration with current EHR systems, will eliminate redundancy and administrative workload.

This research has been structured into two distinct phases (as shown in Fig. 2), each with its own distinct set of goals and tasks.

2.2.1 Automatic speech recognition (ASR) (Phase 1)

ASR is the first step towards turning spoken conversations into a text format that could be transcribed. This stage was essential in laying the groundwork for further processing and analysis of Whisper API, an open-source resource useful in precisely translating spoken words into written text and serves as the foundation for the textual data used in this work.[30]

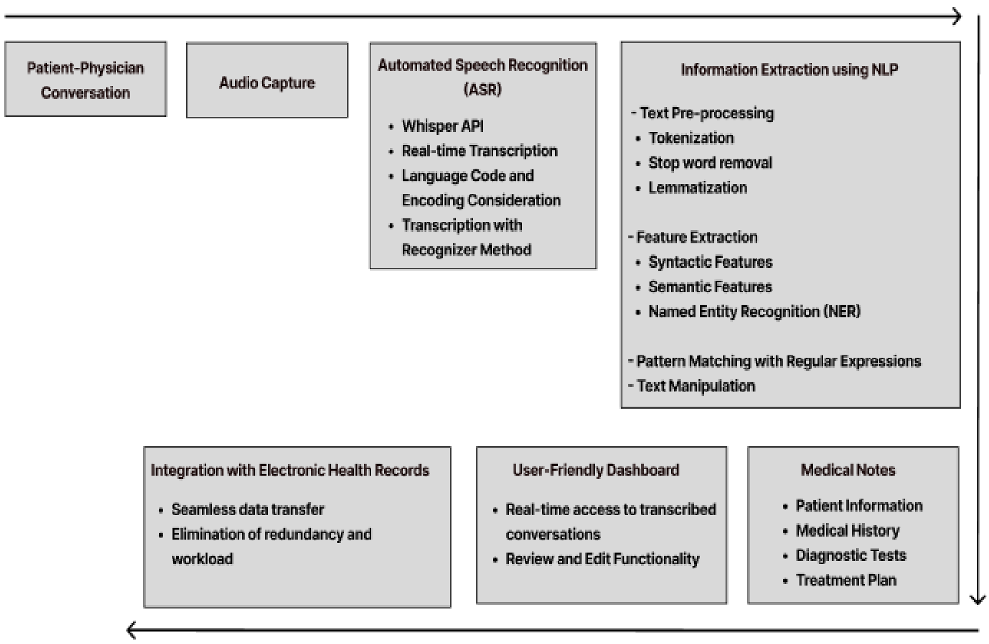

Fig. 1: Proposed system architecture for digital medical scribes.

Table 1: Survey of digital methods used for healthcare.

Paper title	Methodology used	Comparison with the current technologies used
How does medical scribes' work inform the development of speech-based clinical documentation technologies?[17]	Examines how scribes' work can inform speech-based tech development, highlighting gaps in current EHR systems' usability and accuracy.	Supports the integration of natural language processing (NLP) and automatic speech recognition (ASR) to improve documentation, aligning with the proposed architecture.
Medical scribes, provider and patient experience, and patient throughput: a trial in an academic general internal medicine practice.[19]	Shows improved provider and patient experience with scribes, indicating benefits over current EHR systems.	Reinforces the architecture's goal of reducing clinician burnout and improving efficiency.
Challenges of developing a digital scribe to reduce clinical documentation burden.[20]	Identifies challenges in developing digital scribes, such as integration and accuracy issues in existing systems.	Highlights practical barriers the proposed architecture must address, like interoperability and data accuracy.
A sociotechnical multiple perspectives approach to the use of medical scribes: a deeper dive into the scribe-provider interaction.[21]	Explores the complex dynamics between scribes and providers, which current tech often overlooks.	Emphasizes the need for user-friendly design and effective integration, which is crucial for the proposed architecture.
Information Infrastructures in healthcare and emergent data work occupations: The case of medical scribes and clinical documentation improvement specialists (CDIS).[22]	Discusses how scribes adapt to information infrastructures, showing flexibility not present in rigid EHR systems.	Suggests that the proposed architecture should be adaptable and support varied workflows.
The digital scribe.[23]	Proposes a vision for digital scribes, focusing on reducing clinician workload compared to current EHRs.	Aligns with the architecture's focus on using ASR and NLP to streamline documentation.
Scribes, electronic health records, and the expectation of confidentiality.[24]	Discusses confidentiality issues with scribes, a consideration for all digital documentation.	Highlights the importance of data privacy in the proposed architecture.
Potential effects of the electronic health record on the small physician practice: a Delphi study.[25]	Examines EHR impact on small practices, identifying challenges such as cost and complexity.	Suggests the proposed architecture should be cost-effective and simple to implement.
The IT transformation health care needs.[26]	Advocates for comprehensive IT transformation in healthcare, including better EHRs.	Supports the architecture's holistic approach to integrating advanced tech for better outcomes.
Electronic health record logs indicate that physicians split time evenly between seeing patients and desktop medicine.[27]	Highlights the time burden of EHRs on physicians, compared to potential time savings with scribes.	Reinforces the need for the proposed architecture to reduce administrative burden.
Ethical perspectives on recommending digital technology for patients with mental illness.[28]	Emphasizes patient-centered digital technology use, often lacking in current tech implementations.	This aligns with the proposed architecture's goal of enhancing patient engagement and data privacy.

2.2.2 Text transcription (Phase 2)

In the second phase, the focus was on extracting important and useful data from the text transcription. During this stage, a variety of NLP tools and libraries, focusing especially on regular expressions (regex), were used. At this stage, the main tasks were data analysis, sentiment analysis, pattern recognition, and extracting important insights from the textual data.

2.3 System design

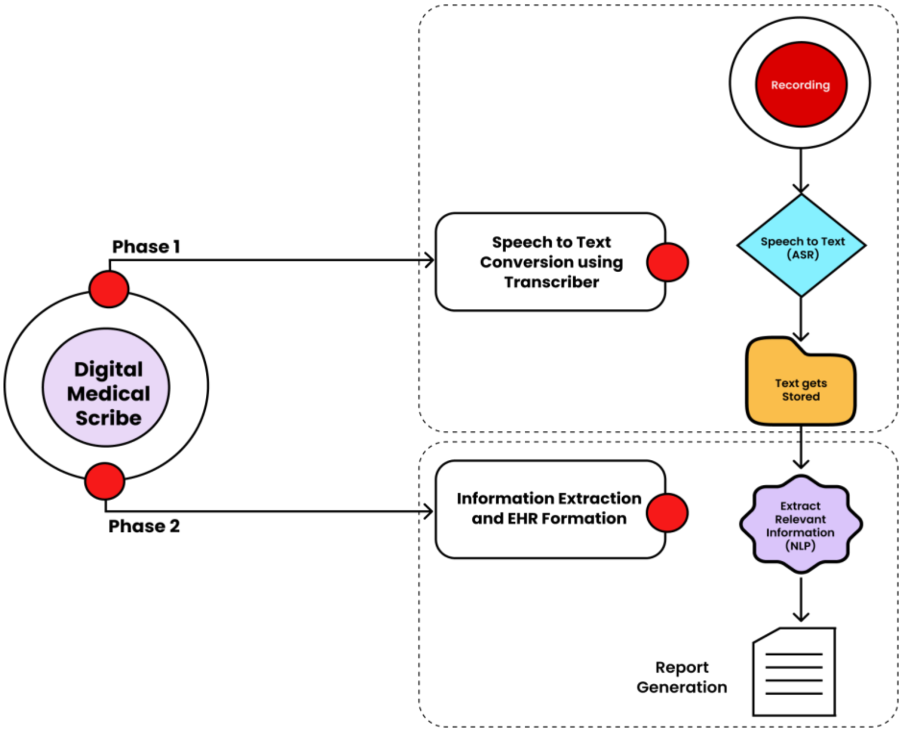

Fig. 2: Phases in crafting the digital medical scribe system.

The phases of the design of the system are described below:

2.3.1 Automated speech recognition (Phase I)

An essential part of turning live conversations between doctors and patients into written text is the automatic speech recognition (ASR) component.[31] Fig. 3 illustrates the workflow.

API Call and Transcriber App Launch: An API call triggers the start of the ASR process. The Transcriber App is launched upon making the call. With its ability to quickly and accurately translate spoken words into written text, this app is a vital resource for distilling the main points of doctor-patient conversations.

Audio File Reading: The audio file containing the ongoing conversation can be read by the Transcriber App. This is accomplished by utilising the 'io' library, which makes sure that the audio stream is correctly processed and set up for transcription.

Language Code and Encoding Consideration: The ASR process considers the language used during the conversation to guarantee accurate and contextually relevant transcriptions. Information encoding is also taken into account to guarantee data integrity and compatibility. These vital parameters travel with the audio file as it is being transcribed.

Transcription with the Recognizer Method: The 'Recognizer' method is the core of the ASR process. This method uses the power of the Whisper API to accomplish the difficult task of transcribing the audio into text after receiving the audio file along with the language code and encoding information. The Whisper API is a dependable option for this stage of our project because of its sophisticated algorithms and models that are made to interpret speech correctly.

Dashboard Display: The transcribed text is returned by the Whisper API's Recognizer method after it has finished its task. The dashboard for our project then seamlessly incorporates this useful textual output. The real-time display of the

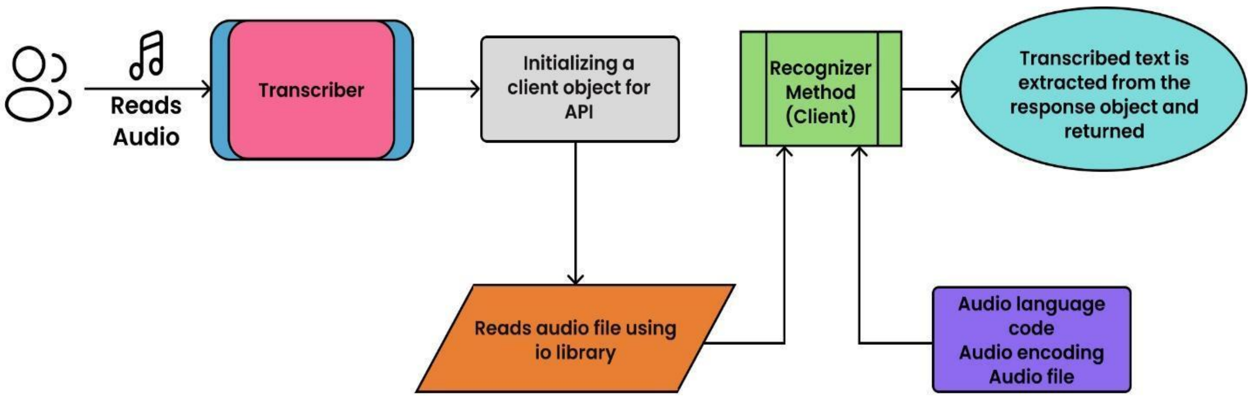

Fig. 3: The workflow of the transcriber in digital medical scribe systems.

transcribed conversation takes place at the dashboard, which functions as a central hub. In the end, this improves the quality of healthcare services and makes sure that no important information is missed during the consultation by giving doctors and patients an easy-to-use interface to follow and review the conversation.[32]

2.3.2 Whisper API architecture

The OpenAI Whisper API is based on a transformer encoder-decoder architecture (given in Fig. 4) with a number of modifications that make it well-suited for speech recognition. Encoder: The encoder receives audio and converts it into a number of hidden states. The audio is first converted into a log-Mel spectrogram, which is a representation of the audio that is compressed in terms of both time and frequency. The log-Mel spectrogram is then passed through a layer stack of a convolutional neural network (CNN). The CNN layers are used to extract the key features from the log-mel spectrogram that are needed for voice recognition.

The output of the CNN layers is then subjected to a stack of self-attention layers.[28] Self-attention allows the model to detect long-range dependencies in the input sequence. This is important for speech recognition because a word's meaning can be affected by words that have been spoken earlier in a phrase.

Decoder: The encoder outputs a sequence of text tokens, which the decoder receives along with its concealed states. The decoder is also a stack of layers designed to facilitate self-attention. However, the decoder also includes a few additional components that are specific to voice recognition.\

As an example, the decoder has a feature that helps train the model to align sounds and text. This is important because the length of an audio utterance can be influenced by the accent and speaking rate of the speaker.

The decoder also has an element that helps the model generate text that follows the language model. This is significant because the language model can assist in fixing output errors and enhancing text fluency.

Training: The Whisper API is trained on a large set of diverse audio files. The dataset contains audio from various sources, including meetings, lectures, and podcasts. In addition, the audio is labeled with the relevant written transcript.

The Whisper API was trained using a technique called "weak supervision." When supervision is weak, perfect labeling of the training data is not required. This is important because it can be expensive and time-consuming to obtain precisely labeled data. Rather, the Whisper API is trained through a process called "self-training." A model that is self-training is first trained on a small set of labeled data. After that, the model is used to generate transcripts for a significant amount of unlabelled data. The model is then further trained using the transcripts that are generated. The model is iterated through this process until it converges. This indicates that the model is no longer performing better on the labeled data.[4]

Inference: Once trained, the Whisper API can be used to transcribe new audio. To be transcribed, new audio must first be converted into a log-Mel spectrogram. The log-Mel spectrogram is then processed by the encoder and decoder. The output of the decoder is a sequence of text tokens. After that, a language model is applied to the text tokens to generate the final output.

The Whisper API is a powerful voice recognition tool that has many applications. Early results are encouraging, despite the fact that it is still in development.

2.3.3 Natural language processing (Phase II)

The primary points and supplementary test result data are extracted by the DMS using an NLP technique from the patient-physician conversation. The steps in the NLP approach

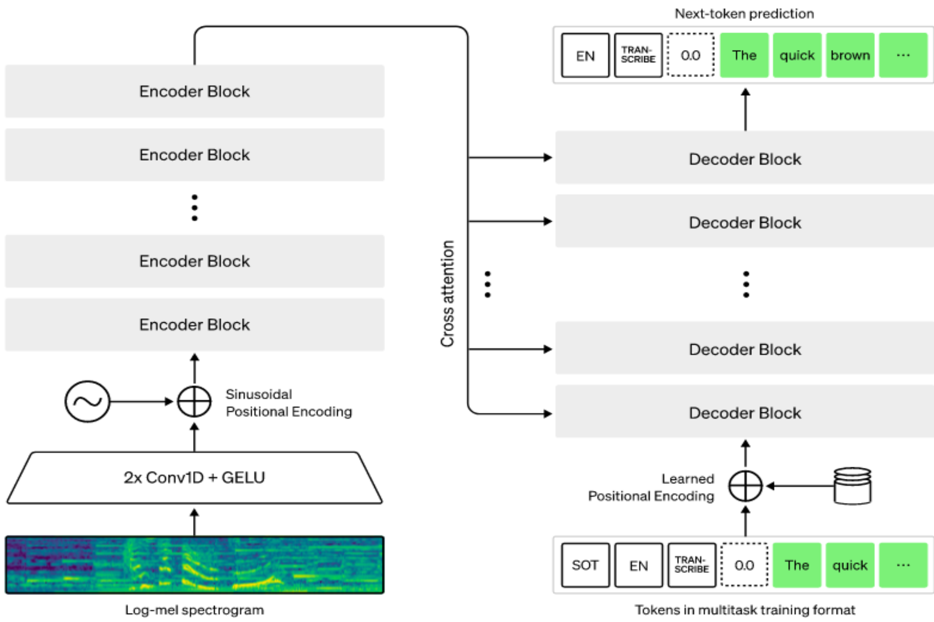

Fig. 4: The architecture of the OpenAI Whisper API.

are as follows:

Text pre-processing: To cut down on noise and ambiguity, the conversation's text is pre-processed. Tasks like tokenization, stop word removal, and lemmatization are involved in this.

Tokenization: To enable additional analysis, the text is divided into discrete tokens, like words or sentences. Algorithms such as the Penn Treebank tokenizer or libraries like NLTK (Natural Language Toolkit) or SpaCy can be used for tokenization.

Stop word removal: Words that are frequently used but don't add anything to the text, like "the," "and," or "is," are eliminated to cut down on background noise. Stop word removal can be applied with pre-made lists of words to avoid or by utilizing statistical techniques to find words that are frequently used but don't provide much information.

Lemmatization: To normalize variations and increase analysis accuracy, words are reduced to their dictionary or base forms. Words are reformed into their root forms using lemmatization algorithms like the WordNet Lemmatizer or stemming strategies like Porter Stemming

Extraction of features: The pre-processed text is used to extract features. These characteristics are the text's syntactic organization, the frequency of particular words or phrases, or the existence of particular entities.

Syntactic Features: To comprehend the linguistic patterns used in the conversation, the syntactic structure of the text-which includes parts of speech, grammatical relationships, and sentence structures-is examined. Syntactic characteristics are extracted using dependency parsing techniques, like the Stanford Dependency Parser or spaCy's dependency parser.

Semantic Features: To determine the context and meaning of words and phrases used in the discussion, semantic analysis techniques are applied. Word meanings and semantic associations can be compared by using word embeddings, such as Word2Vec or GloVe, which represent words as dense vectors in a continuous semantic space.

Named Entity Recognition (NER): To recognize and categorise named entities in the text, such as medical phrases, symptoms, drugs, or procedures, NER algorithms are used. NER models are used to extract significant things from the conversation. These models are trained on annotated datasets or pre-trained language models such as Bidirectional Encoder Representations from Transformers (BERT).[33]

Regular Expressions-Based Approach: In addition to NLP techniques, regular expressions are used to extract specific patterns and information from the conversation text.

Pattern Matching: Regular expressions are character patterns that are used to find and extract text that meets predetermined criteria. This process is known as pattern matching. To find trends in test findings, medical conditions, or treatment plans discussed throughout the conversation, custom regular expressions are created.

Text Manipulation: In order to extract numerical data, format dates, or find keywords related to medical concepts, regular expressions are also used to modify and transform text. Tools and Libraries: Various tools and libraries are utilized to implement NLP techniques and regular expressions-based approaches.

The Natural Language Toolkit, or NLTK, is a collection of libraries and tools for NLP tasks such as syntactic analysis, stop word removal, tokenization, and lemmatization.

SpaCy: With pre-trained models for tokenization, part-of-speech tagging, dependency parsing, and named entity recognition, SpaCy is a potent NLP package.

Stanford NLP Toolkit: The Stanford NLP Toolkit is a collection of powerful NLP tools created by the Stanford NLP Group. These tools include named entity recognizers, taggers, and parsers.

Regular Expressions: Regular expressions are supported by programming languages such as Python. They may also be used to manipulate text and match patterns thanks to libraries such as the re-module.

The system can efficiently extract key points and additional test result data from patient-physician conversations by utilizing regular expressions-based approaches, NLP techniques, and tools and libraries designed for text processing and analysis. This helps to produce thorough and precise EHR reports.

The project extracts additional test result data and the main points from the conversation using a regular expressions-based approach. Regular expressions are an effective tool for text manipulation and search. Regular expressions are used by the code to match particular patterns in the text and extract the needed data.[34]

Once the main points and additional test result data have been extracted from the conversation, the code uses them to generate an EHR report. The EHR report includes patient information, medical history, diagnostic tests, and treatment plans.

3. Results and discussion

Test results on the primary dataset show that the digital medical scribe performs well. Over a wide range of medical scenarios, the scribe has proven to be accurate and efficient in handling complex medical terminology.

The scribe's strong capabilities are demonstrated by its ability to operate efficiently on the generated dataset. The dataset has offered a difficult but realistic testing ground because it was thoughtfully created to incorporate a variety of medical terminologies and complex dialogues. The positive outcomes show that the scribe is ready to assist medical professionals by accurately deciphering and processing medical jargon.

A noteworthy accomplishment is the scribe's effective performance on this dataset, which highlights its versatility and dependability in actual medical situations. The digital scribe is well-positioned to improve healthcare professionals' workflow and efficiency by reliably identifying and deciphering complex medical terminologies. This will ultimately lead to better patient care and higher standards for medical practice.

This successful evaluation of the self-created dataset further establishes the scribe's potential to become an indispensable tool within the healthcare ecosystem by confirming its ability to thrive in the complex and varied world of medical terminology and dialogues.

The Transcriber Dashboard, designed with a minimalist interface, primarily serves as the platform for displaying transcribed text generated from spoken healthcare conversations. Alongside this key function, the dashboard offers essential features that enhance its utility and adaptability in clinical settings.

Language Selection: Our Transcriber Dashboard excels with an advanced language selection tool. By specifying the language spoken in patient-physician interactions, you ensure highly accurate transcriptions. The system detects the spoken language and applies tailored models and algorithms, boosting accuracy in diverse language scenarios. This capability shines in multilingual healthcare settings, where multiple languages may be used. Our research demonstrates that this feature significantly enhances transcription accuracy, making our digital scribe system highly adaptable and effective in capturing accurate medical information. It’s worth noticing that 65% of Whisper training data is assigned to specialized ASR for English only, while multilingual accelerates speech recognition applications for 17%, with the rest being English translation functionalities. For the moment, Whisper formally backs up 99 languages.

Font Size Customization: The Transcriber Dashboard has an adjustable font size feature that makes the text easier to read. Users can change the font size to their liking, which makes it helpful for people with different vision preferences or accessibility needs. This feature has been well-received by users, especially those who have specific vision preferences or accessibility requirements. By allowing users to adjust the font size, the system prioritizes inclusivity and user satisfaction, making it easier and more efficient to use for medical documentation workflows.

Speech Rate Adjustment: Our dashboard now includes a customizable slider that lets you control the speech rate in real time. This means you can easily adjust the transcription speed to match the pace of the conversation, reducing transcription errors. Whether people are speaking quickly or slowly, you can adjust the speech rate to align with their natural cadence, making the transcription more accurate and easier to understand. This helps streamline medical documentation workflows and improve user experience. Speech-rate control takes place in real-time. The speed at which speech is produced is monitored using the Whisper API via the analysis of the length of time between words uttered and syllables; methodologies such as the Speech Tempo Analysis Algorithm are employed in order to determine the speed at which something is spoken. To make sure that transcriptions are reliable under conditions where people talk quickly or slowly, the program adjusts the rate of speech accordingly, ensuring an accurate representation. An actual-time correction is crucial. That is because it maintains accuracy and clarity when recording messages for medical purposes. It also enables the system to cope with variations in speed during live consultations that involve medics.

The ultimate goal of the research project was to generate EHR reports that could seamlessly integrate with existing healthcare systems. Through the utilization of Digital Medical Scribes and the Whisper API, spoken conversations were successfully transcribed. These were then channeled into the creation of EHR reports.

Table 2: Sample conversation between a physician and a patient.

Character	Conversation
Patient:	Hello, Doctor. My name is Rohil, I've been experiencing some discomfort in my chest and shortness of breath. It's been going on for a few weeks now.
Physician:	Physician: I'm sorry to hear that. Can you describe the chest discomfort? Is it a sharp pain or more like pressure?
Patient:	It's more like a pressure in the center of my chest. Sometimes, it feels like a squeezing sensation, and it often radiates to my left arm.
Physician:	I see. That's concerning. Have you noticed if these symptoms occur during physical activity or at rest?
Patient:	It seems to happen more often when I'm active, like when I'm climbing stairs or walking briskly. But it's also happened when I'm just sitting at my desk.
Physician:	Okay, that's important information. How about other symptoms like dizziness, nausea, or excessive sweating during these episodes?
Patient:	Yes, I've felt dizzy and sweaty a couple of times when it happened. No nausea, though.
Physician:	Thank you for sharing that. Given your symptoms, I'd like to order some tests. We should do an electrocardiogram (ECG) to check your heart's electrical activity and a stress test to see how your heart responds to physical exertion. We'll also schedule a blood test to check your cholesterol levels, cardiac enzymes, and hemoglobin content.

The results demonstrate the capability of our system to generate comprehensive and accurate EHR reports that serve as a vital resource for healthcare providers. This outcome is a significant step forward in streamlining the documentation process, improving healthcare quality, and reducing the administrative burdens often faced by clinicians. Table 2 shows a sample conversation between a patient and a physician. Table 3 shows the report generated by the software. The overall quality and usefulness of the project’s output are improved by enabling the processing and interpretation of the transcribed text efficiently by integrating NLP and regex libraries.

Table 3: Report generated by the software.

Electronic Health Record Report
Patient Information: ● Name: None ● Symptom: None ● Chest Discomfort: Pressure in the center of my chest. Sometimes, it feels like a squeezing sensation. ● Occurs During Physical Activity: Yes ● Dizziness: Yes Medical History: ● Chief Complaint: None ● Past Medical History: None reported. ● Medications: List of current medications, if any. ● Allergies: None reported. Diagnostic Tests: ● Electrocardiogram (ECG): Available ● Stress Test: Available ● Blood Test: Available ● LDL Cholesterol: 150 mg/dL ● Cardiac Enzymes: Elevated ● Hemoglobin Content: 13.5 g/dL Treatment Plan: ● Further evaluation needed: Coronary angiography is recommended. ● Medication Management: Prescription of medications as necessary. ● Lifestyle modifications: Discuss diet, exercise, and stress management to support recovery.

Rigorous validation measures were considered to assess the accuracy and precision of the transcription and documentation strategies in the DMS system. Independent trying become performed, related to healthcare professionals and experts in clinical transcription. They meticulously compared the transcribed text generated via the DMS with the authentic audio recordings of patient-medical doctor interactions. This thorough evaluation enabled the identification of any discrepancies or mistakes in transcription, allowing for an evaluation of their impact on medical documentation accuracy.

Furthermore, we benchmarked the performance of the DMS against enterprise standards, such as hooked-up metrics, which include word errors fee (WER), and calculated the proportion of incorrectly transcribed words and adherence to unique scientific terminology standards and guidelines. By comparing the gadget's performance in opposition to those benchmarks, precious insights were gained into its accuracy and precision in shooting and transcribing medical dialogue.

A number of issues were faced while implementing DMS, specifically the difficulties in incorporating cutting-edge technology into clinical practice. Multiple challenges have been faced that call for careful consideration and calculated solutions, such as negotiating the complexities of compatibility with current EHR systems and guaranteeing the precision and dependability of transcription and documentation. Maintaining patient confidentiality and data security in the face of legal mandates such as the Health Insurance Portability and Accountability Act (HIPAA) has proven to be a complex task that calls for constant attention to detail and adherence to industry best practices. Targeted training and engagement programs are also required to overcome the significant challenge of fostering user acceptability and adoption amid worries about job displacement and skepticism about technology's ability to appropriately capture clinical nuances. Strategic planning and careful resource management are needed to balance the installation and maintenance costs with your budget and competing goals. Transparency, accountability, and cooperation with legal professionals and regulatory agencies are further stressed by ethical issues pertaining to patient-provider communication and the legal ramifications of documentation errors. In spite of these obstacles, the proposed research confirms the revolutionary potential of digital scribe technology to boost provider satisfaction, improve patient care, and create operational efficiencies in healthcare environments. We are unwavering in our commitment to achieving the full benefits of digital scribe technology in clinical practice by proactively addressing these issues and iteratively improving our system based on user feedback and real-world insights.

The main aim of research on the digital medical scribe project should be to reduce healthcare costs while also improving outcomes for permanent infusions. In addition, it is important to promote the development of more accurate methods of natural language processing that will allow for better integration across languages and encourage multimodal processing of different types of health-related information. Long-term clinical impacts can only be ascertained through longitudinal studies, while individual healthcare provider-based adaptations will not be made without personalized adaptation research. Unless there are established interoperability standards, there is no possibility for easy data exchange between platforms, so as we develop new systems, we must also consider moral aspects like privacy, consent and effects on the labor force.[35] Maximized benefits of the digital medical scribes, enhanced patient outcomes, optimized healthcare delivery, and reduced provider burden in the evolving healthcare technology landscape by advancing these areas.

4. Conclusion

Significant outcomes have been realized in the digital transformation of the healthcare sector thus far. These improvements can change medical record-keeping and treatment stages, bringing about high-level, precise, and patient-centered care. Moreover, lessening the workload for persons working in the health sector is a step towards job satisfaction and better outcomes for patients.[36] In this work, we provide a complete account of dataset curation and audio report recording procedures that foster transparency and credibility in the proposed method. In offering our findings, it's vital to underscore the discussion of the challenges and obstacles encountered. The advantages of this work are significant. As healthcare carriers and institutions increasingly undertake this generation, we anticipate a future in which standardized medical documentation, advanced data analytics, and more holistic personal care become the norm. However, it is vital for all stakeholders, from policymakers to generation developers, to work collaboratively to deal with demanding situations related to facts protection, interoperability, and personnel education. Overcoming these boundaries will be instrumental in understanding the overall capacity of virtual clinical scribe technology. For this purpose, different healthcare systems must allow for the interchange of information by making them interoperable; this way, doctors will have access to complete information concerning their patients, helping them make correct decisions.[37] In order to take full advantage of digital scribe technology, thereby increasing its usefulness, healthcare providers should always continue educating themselves on a regular basis. Incorporating better user interfaces that help patients access and make contributions to their records may also be a consideration for future versions in order to facilitate a more interactive, patient-centered approach.[38]

Fig. 5 shows the future roadmap for the DMS technology. Incorporation of image recognition and sensor data analysis into the DMS could go a long way in simplifying processes as well as enhancing diagnostic precision. This will be a gateway to monitoring patients from far distances and individualized treatment, thus enabling the provision of better health care services more conveniently. It is very important to improve data integrity using modern access controls, instant audit trails, and continual feedback mechanisms. Future systems can integrate AI, meaning they will make use of sophisticated data validation through automated rules applied together with machine learning models for error detection and rectification on the fly. Furthermore, cross-referencing with external databases as well as post-entry validation tools can significantly help maintain its accuracy hence making medical records less prone to errors and of high quality.

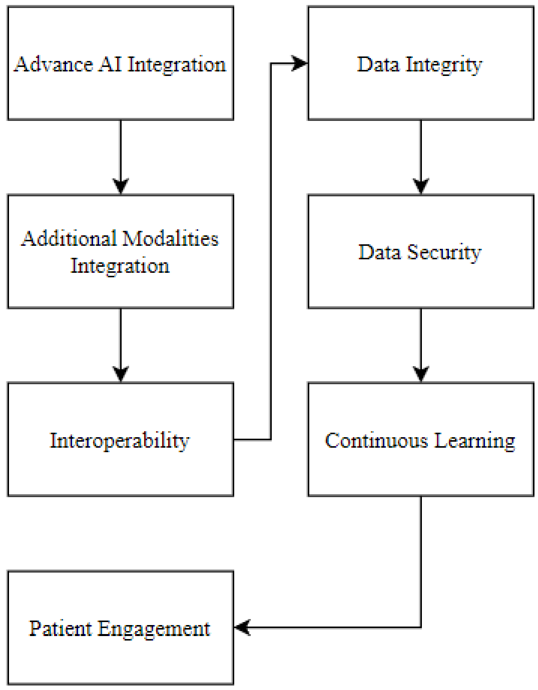

Fig. 5: Future roadmap.

Acknowledgements

This work is supported by the Shastri Scholar Travel Subsidy Grant (SSTSG) 2023-24 and the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant No: 194376.

Conflict of Interest

There is no conflict of interest.

Supporting Information

Not applicable.

References

[1] D. A. Fisher, The conundrum of digital medicine, *Orthopedics*, 2017, **40**, 200-202, doi: 10.3928/01477447-20170627-02.

[2] J. Wang, M. Lavender, E. Hoque, P. Brophy, H. Kautz, A patient-centered digital scribe for automatic medical documentation, *JAMIA Open*, 2021, **4**, ooab003, doi:

10.1093/jamiaopen/ooab003.
[3] M. M. van Buchem, H. Boosman, M. P. Bauer, I. M. J. Kant, S. A. Cammel, E. W. Steyerberg, The digital scribe in clinical practice: a scoping review and research agenda, *NPJ Digital Medicine*, 2021, **4**, 57, doi: 10.1038/s41746-021-00432-5.
[4] J. Wang, M. Lavender, E. Hoque, P. Brophy, H. Kautz, A patient-centered digital scribe for automatic medical documentation, *JAMIA Open*, 2021, **4**, ooab003, doi: 10.1093/jamiaopen/ooab003.
[5] S. Ghatnekar, A. Faletsky, V. E. Nambudiri, Digital scribe utility and barriers to implementation in clinical practice: a scoping review, *Health and Technology*, 2021, **11**, 803-809, doi: 10.1007/s12553-021-00568-0.
[6] O. T. Nguyen, K. Turner, D. Charles, O. Sprow, R. Perkins, Y. R. Hong, J. Y. Islam, N. Khanna, A. Alishahi Tabriz, J. Hallanger-Johnson, J. Bickel Young, C. E. Moore, Implementing digital scribes to reduce electronic health record documentation burden among cancer care clinicians: a mixed-methods pilot study, *JCO Clinical Cancer Informatics*, 2023, **7**, e2200166, doi: 10.1200/CCI.22.00166.
[7] J. P. Avendano, D. O. Gallagher, J. D. Hawes, J. Boyle, L. Glasser, J. Aryee, B. M. Katt, Interfacing with the electronic health record (EHR): a comparative review of modes of documentation, *Cureus*, 2022, **14**, e26330, doi: 10.7759/cureus.26330.
[8] M. A. Bos, The future of the electronic health record: testing a speech commanded interface in combination with a smartwatch (Master's thesis, University of Twente), 2023.
[9] National Academies of Sciences, National Academy of Medicine and Committee on Systems Approaches to Improve Patient Care by Supporting Clinician Well-Being, Taking action against clinician burnout: a systems approach to professional well-being, 2019.
[10] B. D. Tran, K. Latif, T. L. Reynolds, J. Park, J. Elston Lafata, M. Tai-Seale, K. Zheng, "Mm-hm," "uh-uh": are non-lexical conversational sounds deal breakers for the ambient clinical documentation technology? *Journal of the American Medical Informatics Association*, 2023, **30**, 703-711, doi: 10.1093/jamia/ocad001.
[11] D. Fraile Navarro, A. B. Kocaballi, M. Dras, S. Berkovsky, Collaboration, not confrontation: understanding general practitioners' attitudes towards natural language and text automation in clinical practice, *ACM Transactions on Computer-Human Interaction*, 2023, **30**, 1-34, doi: 10.1145/3569893.
[12] D. V. Woodcock, R. Pranaat, K. McGrath, J. S. Ash, The evolving role of medical scribe: variation and implications for organizational effectiveness and safety, *Studies in Health Technology and Informatics*, 2017, **234**, 382-388, doi: 10.3233/978-1-61499-742-9-382
[13] A. Croatti, S. Montagna, A. Ricci, A personal medical digital assistant agent for supporting human operators in emergency scenarios, *Agents and Multi-Agent Systems for Health Care*, 2017, 59-75, doi: 10.1007/978-3-319-70887-4_4.
[14] R. Magyari, F. Secomandi, Service blueprinting for better collaboration in human-centric AI: the design of a digital scribe for orthopedic consultations, *International Journal of Design*, 2023, **17**, 63.
[15] R. Pranaat, V. Mohan, M. O'Reilly, M. Hirsh, K. McGrath, G. Scholl, D. Woodcock, J. A. Gold, Use of simulation based on an electronic health records environment to evaluate the structure and accuracy of notes generated by medical scribes: proof-of-concept study, *JMIR Medical Informatics*, 2017, **5**, e30, doi: 10.2196/medinform.7883.
[16] F. S. Falcetta, F. K. de Almeida, J. C. S. Lemos, J. R. Goldim, C. A. da Costa, Automatic documentation of professional health interactions: A systematic review, *Artificial Intelligence in Medicine*, 2023, **137**, 102487, doi: 10.1016/j.artmed.2023.102487.
[17] B. D. Tran, Y. Chen, S. Liu, K. Zheng, How does medical scribes' work inform development of speech-based clinical documentation technologies? A systematic review, *Journal of the American Medical Informatics Association*, 2020, **27**, 808-817, doi: 10.1093/jamia/ocaa020.
[18] N. A. Pennell, A. P. Dicker, C. Tran, H. S. L. Jim, D. L. Schwartz, E. J. Stepanski, mHealth: mobile technologies to virtually bring the patient into an oncology practice, *American Society of Clinical Oncology Educational Book*, 2017, **37**, 144-154, doi: 10.1200/EDBK_176093.
[19] J. Heckman, K. J. Mukamal, A. Christensen, E. E. Reynolds, Medical scribes, provider and patient experience, and patient throughput: a trial in an academic general internal medicine practice, *Journal of General Internal Medicine*, 2020, **35**, 770-774, doi: 10.1007/s11606-019-05352-5.
[20] J. C. Quiroz, L. Laranjo, A. B. Kocaballi, S. Berkovsky, D. Rezazadegan, E. Coiera, Challenges of developing a digital scribe to reduce clinical documentation burden, *NPJ Digital Medicine*, 2019, **2**, 114, doi: 10.1038/s41746-019-0190-1.
[21] S. Corby, J. A. Gold, V. Mohan, N. Solberg, J. Becton, R. Bergstrom, B. Orwoll, C. Hoekstra, J. S. Ash, A sociotechnical multiple perspectives approach to the use of medical scribes: a deeper dive into the scribe-provider interaction, *AMIA Annual Symposium Proceedings AMIA Symp*osium, 2019, **2019**, 333-342.
[22] C. Bossen, Y. Chen, K. H. Pine, Information Infrastructures in healthcare and emergent data work occupations: The case of medical scribes and CDIS, Infra health 2019-Proceedings of the 7th International Workshop on Infrastructure in Healthcare 2019.
[23] E. Coiera, A. Kocaballi, J. Halamka, L. Laranjo, The digital scribe, *NPJ Digital Medicine*, 2018, **1**, 58, doi: 10.1038/s41746-018-0066-9.
[24] P. M. Wangenheim, Scribes, electronic health records, and the expectation of confidentiality, *The Journal of Clinical Ethics*, 2018, **29**, 240-243, doi: 10.1086/jce2018293240.
[25] C. C. Sines, G. R. Griffin, Potential effects of the electronic health record on the small physician practice: a Delphi study, *Perspectives in Health Information Management*, 2017, **14**, 1-6.
[26] N. R. Sahni, R. S. Huckman, A. Chigurupati, D. M. Cutler, The IT transformation health care needs, *Harvard Business Review*, 2017, **95**, 128-138.
[27] M. Tai-Seale, C. W. Olson, J. Li, A. S. Chan, C. Morikawa, M. Durbin, W. Wang, H. S. Luft, Electronic health record logs

indicate that physicians split time evenly between seeing patients and desktop medicine, *Health Affairs*, 2017, **36**, 655-662, doi: 10.1377/hlthaff.2016.0811.

[28] M. Bauer, T. Glenn, S. Monteith, R. Bauer, P. C. Whybrow, J. Geddes, Ethical perspectives on recommending digital technology for patients with mental illness, *International Journal of Bipolar Disorders*, 2017, **5**, 6, doi: 10.1186/s40345-017-0073-9.

[29] K. Ostherr, The medical futures lab: An applied media studies experiment in digital medical humanities, 2017.

[30] A. Edwards, L. Kanner, S. Tewar, L. Pesce, M. Leyser, The value of adding scribe services to 2 distinct pediatric subspecialties in the era of the electronic medical record, *Clinical Pediatrics*, 2024, **63**, 341-349, doi: 10.1177/00099228231174849.

[31] M. Lam, S. Sabharwal, The role of scribes in orthopaedics, *JBJS Reviews*, 2023, **11**, e22.00247, doi: 10.2106/JBJS.RVW.22.00247.

[32] N. G. Menon, A. Shrivastava, N. D. Bhavana, J. Simon, Deep Learning based Transcribing and Summarizing Clinical Conversations, 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC). November 11-13, 2021. Palladam, India. IEEE, 2021.

[33] N. H. Crampton, Ambient virtual scribes: Mutuo Health's AutoScribe as a case study of artificial intelligence-based technology, *Healthcare Management Forum*, 2020, **33**, 34-38, doi: 10.1177/0840470419872775.

[34] G. Gellert, S. Webster, J. Gillean, E. Melnick, H. Kanzaria, Should US doctors embrace electronic health records? *BMJ*, 2017, **356**, j242, doi: 10.1136/bmj.j242.

[35] P. Karimi, P. Bora, A. Martin-Hammond, Scribe: improving older adults' access to medical instructions from patient-physician conversations, Proceedings of the 18th International Web for All Conference. Ljubljana Slovenia. ACM, 2021.

[36] P. G. Brindley, Distracted doctoring: returning to patient-centered care in the digital age. do we have time for this? *Anesthesia and Analgesia*, 2018, **126**, 719, doi: 10.1213/ANE.0000000000002696.

[37] P. A. Stokes, Scribal attribution across multiple scripts: a digitally aided approach, *Speculum*, 2017, **92**, S65-S85, doi: 10.1086/693968.

[38] S. Michie, L. Yardley, R. West, K. Patrick, F. Greaves, Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop, *Journal of Medical Internet Research*, 2017, **19**, e232, doi: 10.2196/jmir.7126.

Publisher's Note: Engineered Science Publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.