AscleAI: A LLM-based Clinical Note Management System for Enhancing Clinician Productivity

Jiyeon Han*
jyjy0808@ewhain.net
Ewha Womans University
Seoul, South Korea

Jimin Park*
parkjimin520@ewhain.net
Ewha Womans University
Seoul, South Korea

Jinyoung Huh
hjy0505@ewhain.net
Ewha Womans University
Seoul, South Korea

Uran Oh†
uran.oh@ewha.ac.kr
Ewha Womans University
Seoul, South Korea

Jaeyoung Do†
jaeyoung.do@snu.ac.kr
Seoul National University
Seoul, South Korea

Daehee Kim
daehee.kim@atto-research.com
Atto Research
Seoul, South Korea

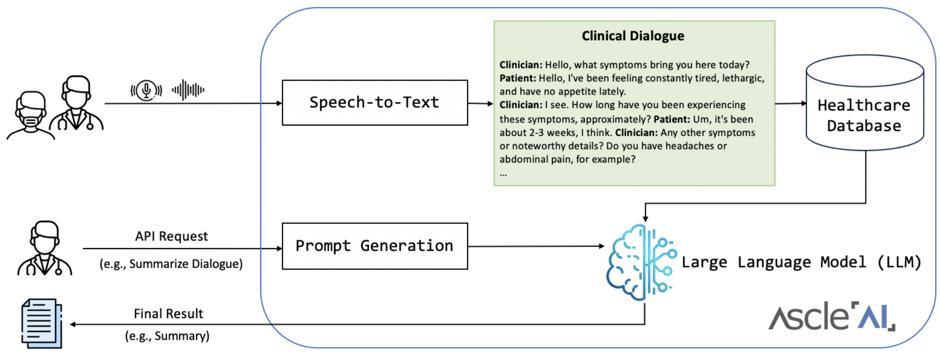


Figure 1: System architecture transcribing a patent-clinician dialogue to text and subsequently summarizing the written dialogues through the utilization of a large language model.

ABSTRACT

While clinical notes are essential to the field of healthcare, they pose several challenges for clinicians since it is difficult to write down medical information, review prior notes, and extract the desired information at the same time while examining a patient. Thus, we designed a system that can automatically generate clinical notes from dialogues between patients and clinicians and provide specific information upon clinicians' query using a Large Language Model (LLM) both in real-time. To explore how this system can be used to support clinicians in practice, we conducted an interview with six clinicians followed by a design probe study with the current version of our system for feedback. Findings suggest that our system has the potential to enable clinicians to write and access clinical notes and examine the patients simultaneously with reduced cognitive loads and increased efficiency and accuracy.

*Both contributed equally to this work as first co-authors.
†The co-corresponding authors.

CHI EA '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0331-7/24/05
https://doi.org/10.1145/3613905.3650784

CCS CONCEPTS

• **Human-centered computing** → **Interactive systems and tools**; • **Applied computing** → **Health informatics**.

KEYWORDS

Large language model, clinical note, design probe, interview

ACM Reference Format:
Jiyeon Han, Jimin Park, Jinyoung Huh, Uran Oh, Jaeyoung Do, and Daehee Kim. 2024. AscleAI: A LLM-based Clinical Note Management System for Enhancing Clinician Productivity. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3613905.3650784

1 INTRODUCTION

Clinical note serves as a fundamental aspect of healthcare practice, aiding clinical reasoning, and facilitating effective communication among healthcare professionals [36]. However, issues such as the omission of crucial patient information and the inadvertent creation of inaccurate records have been identified [35]. Furthermore, the time-intensive nature of creating detailed notes for every patient and every treatment contributes to lower productivity and increased workload for clinicians [3, 7, 29, 30]. The variability in note formats further complicates the review process [12], posing difficulties for clinicians in extracting relevant information [22].

While large language models like ChatGPT have gathered attention in healthcare, recent research highlights their potential various domains such as assistance of decision-making in clinical practice [1, 2, 8, 10, 11, 14, 24, 26, 27]. Yet, privacy concerns arise due to their online connectivity requirements [18, 32, 33].

In our study, we introduce *AscleAI* , designed to support clinicians in generating clinical notes from dialogues, streamlining note summarization, and seamlessly retrieving information through a conversational interface. Leveraging the impressive power of natural language capabilities in Large Language Models (LLMs), *AscleAI* might assist clinicians in alleviating documentation burdens by promptly providing the most pertinent medical information. Serving as a promising advancement, *AscleAI* suggests an intuitive and effective solution for the simplification of clinical documentation processes by addressing the practical challenges encountered by healthcare professionals in the contemporary medical environment. Additionally, to prioritize privacy, our system utilizes a locally installed LLM.

To assess the potential of LLMs in enhancing the user experience of composing and accessing clinical notes, we conducted a user study involving six participating clinicians. The study comprised an interview aimed at discerning the challenges faced during clinical note documentation, followed by a feedback study where participants used *AscleAI* and provided feedback. The interview responses confirmed that it is inconvenient and inefficient to write a clinical note and find the desired information from past clinical notes while examining patients especially when the note was written by others. From the feedback study, participants found our system useful for reviewing the treatment history of a patient with the summarization feature and for getting the desired information to a satisfactory level via chat. Some participants, however, noted that sentence length in the summaries made reading challenging. Additionally, two participants suggested converting the summary into the SOAP format for enhanced clarity.

The contribution of this paper is described as follows. Firstly, we systematically identify and articulate the challenges associated with writing and reviewing clinical notes. Secondly, we suggest solution in the form of a clinical note-managing system that leverages large language models. This system is designed to streamline and enhance the efficiency of writing and reviewing clinical notes. Lastly, our work integrates valuable insights from clinicians, incorporating their feedback to enhance the design of the proposed system. By examining current challenges, proposing a novel technological solution, and incorporating real-world feedback, our paper contributes to the advancement of clinical note management practices.

2 RELATED WORK

2.1 Clinical Notes and Electronic Medical Records

The shift from paper-based clinical notes to Electronic Medical Records (EMR) has improved healthcare efficiency and documentation completeness[4, 20, 21, 31]. However, EMR systems still face challenges, including errors in information inaccuracies and the inability to reduce the workload of clinicians [3, 5, 7, 29, 30, 35]. Efforts to streamline note creation through customizable templates and natural language processing techniques have been made [6, 9, 15, 17, 19, 28]. Yet, these systems often require additional reviews due to missing or inaccurate information. To address these issues, our proposed system focuses on real-time, efficient, and accurate clinical note generation through dialogues, offering a solution for clinicians to query specific information using a large language model, thus reducing cognitive loads.

2.2 Using Large Language Models for Healthcare Domain

Large language models, such as ChatGPT, have been gaining attention in the healthcare domain [32]. Recent exploration into the application of large language models, such as ChatGPT, in healthcare demonstrates their potential across various domains (e.g., medical education and clinical practice) [1, 2, 8, 10, 11, 14, 24, 26, 27]. For example, ChatGPT has proven valuable in decision-making scenarios, offering insights into diagnoses, prognoses, examinations, and treatment plans, especially in less common clinical cases [1, 24, 27]. However, concerns arise due to ChatGPT's online connectivity requirements, raising privacy issues within the healthcare domain [18, 33]. Our proposed system prioritizes privacy by utilizing a locally installed large language model for clinical note generation, eliminating the need for an online connection.

3 SYSTEM ARCHITECTURE

In this section, we describe *AscleAI* (Fig 1) which is an interactive system designed for efficient communication and information retrieval in healthcare settings. The system integrates an Automatic Speech Recognition (ASR) model and a large language model (LLM) to seamlessly combine the strengths of both technologies for facilitating seamless interactions between clinicians, patients, and the system. Further, our system offers an efficient and intelligent solution for handling large volumes of clinical data while providing users with a user-friendly interface for interactive engagement. The architecture of the proposed system comprises three key stages: Data Capture and Storage, Summarization Process, and Interactive User Interface.

- **Data Capture and Storage** The first stage involves the utilization of OpenAI Whisper [25], a state-of-the-art automatic speech recognition (ASR) model, to transcribe clinician-patient dialogues into textual format. The converted textual conversation is then stored securely in a healthcare database, allowing for easy retrieval and analysis. Clinicians also have an option to store their clinical notes in the database, ensuring a comprehensive repository of pertinent healthcare information.
- **Summarization Process** Upon receiving a user request to summarize a specific dialogue or clinical notes, the system retrieves the relevant textual data from the database. Then a dynamically generated prompt is formulated to instruct the Mistral-7B [13], large language model to generate a concise summary of the provided documents. Additionally, we referred to methodologies based on Prompt Engineering rather than conducting Fine-tuning to enhance performance as much as possible [23]. This process ensures that the summarization is contextually relevant and tailored to the specific information sought by the user. The integration of

Mistral-7B, with its advanced language comprehension capabilities, contributes to the production of coherent and informative summaries.

- **Interactive User Interface** To enhance user engagement and facilitate further interactions, the system incorporates a friendly user interface that allows individuals to engage in a chat-based interaction. Leveraging the knowledge encapsulated in the generated summaries, users can pose inquiries, seek clarifications, or initiate discussions with the system. This feature empowers clinicians and other users to extract relevant insights from the summarized content and obtain additional information as needed.

3.1 Interface Design

Our system (Fig 2) records clinician-patient conversations, automatically summarizes them, answers questions in LLM, and assists clinicians with writing clinical notes. In the left block (Fig 2.a), you can see their clinical note and the transcript of the voice conversation. If you press the Record button ((Fig 2.b) at the top, your voice will be recorded and transcribed in real-time, and the transcript will be saved as a new mp3 file of the conversation. At the bottom, there are three buttons (Fig 2.c) that allow you to upload a conversation file, upload a note, and generate a summary for the selected file, respectively. The top right block ((Fig 2.d) generates a summary in both English and Korean. Finally, in the bottom right corner, there is a chat box (Fig 2.e) where you can talk to LLM about your medical records and summaries. You can type your question in the textbox and click the 'Enter' keyboard or 'Send' button to start a QnA with LLM.

4 METHOD

To understand how a LLM can be used to improve the accuracy and efficiency of writing and reviewing clinical notes, we conducted a user study with six clinicians. It consists of an interview to identify the challenges of writing and reviewing clinical notes with a current EMR system and a feedback study where participants were asked to use the main features of our system and provide feedback, which are used to assess the potential of *AscleAI* .

4.1 Participants

Six participants (4 male, 2 female) were recruited with snowball sampling. The inclusion criteria for participation was having prior experience in writing and reviewing clinical notes with expertise in the field. Their average age was 37.6 (SD = 4.8), and their average period of working at a medical institution was 11.2 years (SD = 4.5). Two professors (orthopedics and neurosurgery) and four resident doctors (e.g., internal medicine and orthopedics) working at university hospitals in South Korea.

4.2 Apparatus

The study was conducted online using Zoom (https://zoom.us). Each participant joined a video conference using their computers while sharing their screens. The entire session was recorded using the screen recorder with sound. Participants were given a URL to access *AscleAI* using their computers. To demonstrate the functionality of our system during the feedback study, we pre-uploaded a 90-second-long audio recording of a sample dialogue between a patient and a clinician during an actual outpatient treatment provided by the orthopedic of a local hospital. The transcript of the dialogue was prepared using Speech-to-Text (STT) model.[1] It was also pre-uploaded prior to the study sessions.

4.3 Procedure

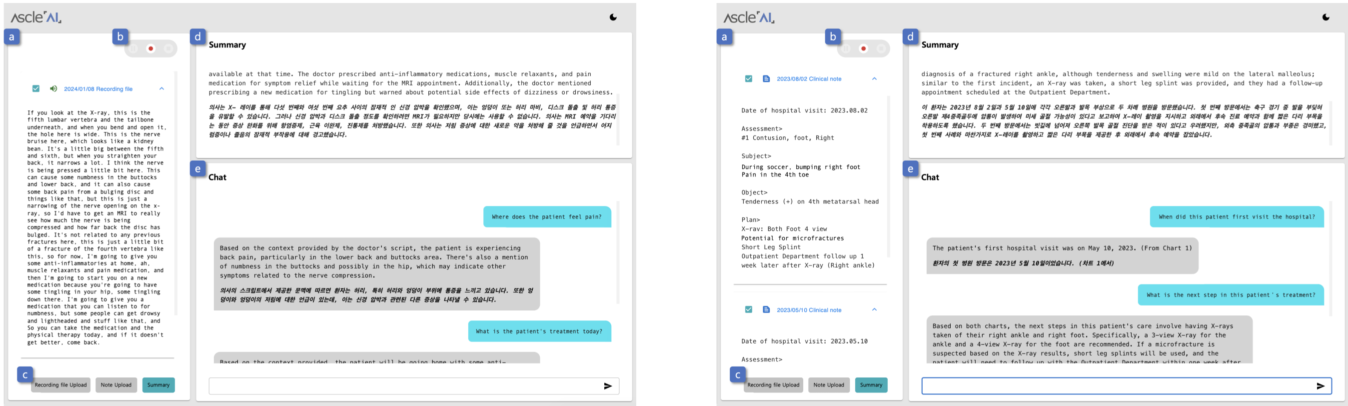
The study session began by signing the consent form and collecting participants' demographic information. During the interview, we asked participants about the content and format of note entries, difficulties encountered while writing, accessing, and reviewing clinical notes. During the feedback study,we briefly explained the components of *AscleAI* . Next, we asked participants to use our system first with a clinician-patient dialogue and then with multiple clinical notes as input. As for the dialogue, to be specific, we played an audio recording of a sample clinician-patient dialogue for participants to get familiar with the data and then we instructed participants to click 'summary' button at the bottom of the left panel in the user interface and read the summarized results displayed at the top in the right panel as shown in Figure 2a. After that, we instructed participants to use the chat feature to ask two specific topics (i.e., the location where the patient feels pain, and the treatment plan) using their own words, and two free topics of their choice (refer to the bottom part of the right panel in Figure 2a. Lastly, we collected subjective feedback such as the perceived accuracy of the transcribed dialogue, the summarized result of the dialogue, and the answers generated upon participants' queries via chat. The procedure was the same for the multiple clinical notes except that participants were asked to skim through multiple clinical notes of the same patient instead of listening to the clinician-patient dialogue and reading the transcribed version of it as shown in Figure 2b. At the end of the user study, we conducted an additional interview asking about their demand for the automation of creating clinical notes from a clinician-patient dialogue and the recommendation of treatment plan through treatment algorithm.

5 FINDINGS

5.1 Creating Clinical Notes

5.1.1 What is Written in Clinical Notes and When. To support the creation of clinical notes, we asked participants what they write in clinical notes and found that all participants wrote four types of information following Subjective-Objective-Assessment-Plan (SOAP) notes [34]: *Subjective data* (e.g., patients medical history, symptoms and behaviors), *Objective data* (e.g., examination results), the *Assessments*, and *treatment Plans*. We further asked which section(s) they consider more important than the others, there were various answers: *Assessment* (N=3), *Objective* (N=2), and *Plan* (N=2), and one stated that it depends on the patients; as for new patients, they focus on symptoms and medical history, while returning patients, they emphasize implementing plans and outcomes. Some participants also mentioned that they add previous examination results and surgical history (N=3), or copy and paste information from previous notes (N=1) when creating a new clinical note to related past notes.

[1] insanely-fast-whisper - https://github.com/Vaibhavs10/insanely-fast-whisper



(a) A clinician-patient dialogue as input

(b) Clinical notes as input

Figure 2: User interface screenshots for different types of input: (a) an audio recording of a clinician-patient dialogue and (b) multiple clinical notes of the same patient over time.

Then we asked the time the notes are created to know when the automatic notes should be provided to clinicians. As a result, three participants stated that they created the notes simultaneously while examining patients, while two stated that they roughly prepared the notes in advance and added necessary information during the examination. One participant mentioned that it depends on whether the patient is new or not. Note that all clinical notes are created immediately after the examination of each patient.

5.1.2 Difficulties in Writing Clinical Notes. We also asked participants about difficulties when creating clinical notes. Four participants mentioned that it is inconvenient to create the note and have a conversation with the patient simultaneously. As a result, they were concerned about not being able to concentrate on the patient, creating insufficient content or even omitting important information when creating notes especially when the conversation with patient becomes extensive. Their strategies include making corrections right after the end of the examination by recalling the conversation (N=2) or using templates with specified sections (i.e., S, O, A, P) to reduce the time spent writing repetitive wordings (N=1).

5.2 Reviewing Clinical Notes

5.2.1 What is Reviewed in Clinical Notes and When. All participants emphasized the importance of reviewing patients' past clinical notes as an essential reference for the examination. Indeed, they all reported that they review the patient's previous clinical notes or summarize them in a note before every examination. Except for one who reviewed all previous clinical notes, the rest stated that they only reviewed the subset. To understand what to focus on when summarizing the prior notes, we asked what they prioritize, and the results are as follows: primary symptoms (N=3), examination results (N=3), patient's medical history such as currently prescribed medications and underlying diseases (N=3); note that multiple answers were allowed. When asked about when they review prior notes, all but one participant, who quickly checked the notes before the patient arrived, and reviewed the notes while checking up on the patients.

5.2.2 Difficulties in Reviewing Clinical Notes. All participants mentioned experiencing difficulties when finding the desired information efficiently while checking up on the patients. The reasons include notes not being well-organized (N=2), unfamiliar abbreviations (N=1), When facing such challenges, participants mentioned resolving them by directly searching for examination results (N=4) or by directly inquiring with the patient (N=1).

All participants also mentioned communication issues when reviewing notes written by other clinicians due to having different format (N=2), different abbreviations (N=2), or having to understanding subjective opinions with objective interpretation (N=2). In addition, all participants said that they found it inconvenient and time-consuming to review multiple notes, which is necessary when the patients had multiple treatments from them (N=2), or received treatments in other departments (N=2).

5.3 The Assessment of *AscleAI*

To assess the perceived performance and usefulness of *AscleAI* in summarizing a conversation between a patient and a clinician and multiple documents, and acquire desired information via chat, we collected subjective ratings as shown in Figure 3. The metrics include the representativeness of the summarized content (*Representative*), how well the summary is reflecting important information (*Reflecting Information*), how much they think our system can relieve the inconvenience compared to the conventional way (*Improvement*), willing to use our system (*Willingness to Use*), the usefulness (*Usefulness*) and the overall satisfaction (*Satisfaction*).

5.3.1 Summarization of a Dialogue and Multiple Clinical Notes. Two participants stated that *AscleAI* effectively filtered out unimportant parts from the conversation and provided a concise summary. However, two other participants found it inconvenient that the system did not convert the summary into SOAP format, and

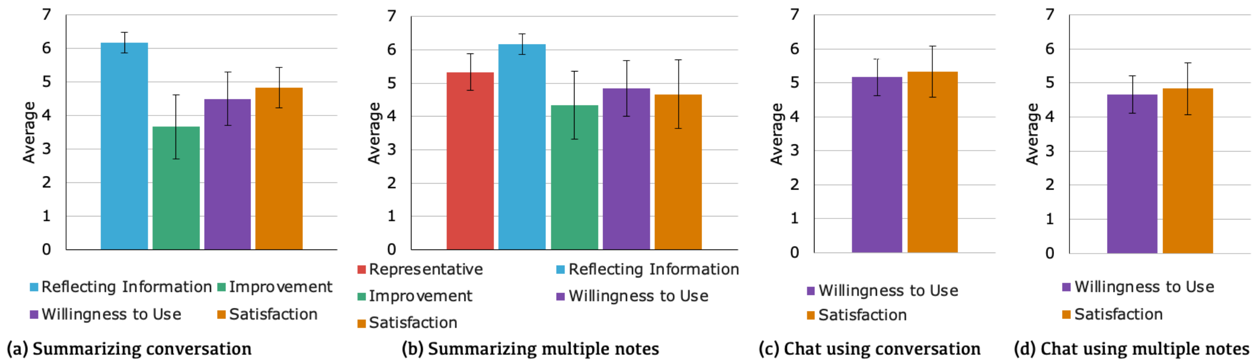


Figure 3: Average scores with summarizing (a, b) and chat-based questioning and answering (c, d) (1-7; 7 is the best). Error bars indicate standard errors (N=6).

the other two participants expressed discomfort with the sentence-based summarization as opposed to the bullet-pointed version. Additionally, four participants mentioned that the summaries could be useful to confirm if and what necessary information was explained to patients. One participant also mentioned that it could be useful to check treatments received from other departments. Participants also wished for additional information such as patient's medical history mentioned by patient (N=2), and side effects of prescribed medications (N=1). One also asked for more structured information (N=1).

Participants appreciated that the summary was concise and contained necessary information. They found it helpful for understanding the treatment history quickly convenient for viewing multiple notes in one place. Suggestions include adding indications of related notes, reducing the length for speed reading, and using medical terms (N=1 for each).

5.3.2 Chat-Based Questioning and Answering. The most common feedback was that the answers provided by the system were satisfactory (N=3). One participant mentioned that the model could be helpful in summarizing the conversation during the clinic (N=1). Two participants mentioned that the LLM seemed to understand the notes well, Participants also appreciated the use of language similar to actual medical consultations, the description being easy to understand even though it included notes written by a different department and showing predictions (N=1 each).

Meanwhile, one participant suggested shorter responses similar to the summary. Participants also pointed out that it contained inaccurate information.

6 DISCUSSION

6.1 Needs of Clinicians When Examining Patients

The results of the interview showed that clinicians face challenges of reviewing notes promptly while engaging in conversation with patients simultaneously, which could lead to the omission of important details or adding inaccurate information as pointed out by Weiner et al. [35]. Moreover, we confirmed that it is difficult for clinicians to concentrate on the patient while writing clinical notes as found in previous studies [3, 7, 29, 30]. With the summary of the conversation between patient-clinician based on the transcribed dialogues generated by our system can be used to enable clinicians to check for any errors and make corrections as they write clinical notes. In addition, the chatting feature of our system can provide easy access to the desired information from a number of prior clinical notes efficiently without having to read every note. By using our system, we expect to improve the quality of patient examination.

6.2 The Potential of Our System

While we did not have empirical evidence to show the performance or the usability of our system, our study results collected by the target users (i.e., clinicians from various departments) suggest the potential usefulness of using a LLM model for supporting clinicians for reviewing dialogues and prior clinical notes and acquiring desired information with two main features: summarization and question-and-answering-like chat.

For our system to be used in practice, we plan to pay extra attention to deal with privacy concerns as clinical notes contain very sensitive information about patients. Our future work includes developing a fast but lightweight model trained with more domain-specific data where personal information is encrypted to be used in real-time on a local server. Moreover, as LLM is also capable of prediction, our next step is to provide treatment or prescription recommendations.

6.3 Guidelines for Utilizing LLMs for Clinicians

The result of the user study indicates that clinicians need to access multiple clinical notes, which are often complex and diverse in format. Additionally, we identified communication issues that often occur among clinicians due to different abbreviations by the department. To support efficient access and accurate understanding of the content of clinical notes, we recommend mapping medical terms and abbreviations or reflecting relationships between medical data into a Knowledge Graph, and then further training this into the LLMs as suggested by Li et al. [16]. In addition, as clinical notes often contain image data such as MRI, CT scans, or prescription images, it is challenging to comprehensively understand and process medical information solely through text. Therefore, a multimodal approach capable of processing and analyzing both text and image data should be considered for better performance.

Moreover, to ensure clinicians check the generated content and correct errors if any, we strongly suggest generating clinical notes no later than the end of the examination. Lastly, adding a summary of prior notes or indicating related notes at the end of each note for cross-references can be helpful for clinicians to understand the treatment history of the patient at ease.

6.4 Limitations

We conducted a user study involving 6 participants, and we are aware of the limitation imposed by the small sample size. This may not be representative, and further research with a larger number of participants may be necessary to capture more accurate and comprehensive behaviors. Additionally, conducting the experiment remotely might not adequately reflect various factors in the actual usage environment. We need to explore methods to generalize the results through testing in real-world environments and consider a more diverse range of conditions in our experiments. Furthermore, our system may not be highly reliable for consideration by clinicians. Therefore, additional future improvements and validations are necessary to enhance its reliability.

7 CONCLUSION

Clinical notes play a crucial role in healthcare, aiding in clinical reasoning and communication among clinicians. Despite its importance, it poses challenges such as the omission of patient information. Moreover, the clinician's workload increased because of the time-intensive nature of note creation. Our study introduces *AscleAI* , a tool designed to manage clinical notes by employing Large Language Model (LLM). In the future, we will enhance the model's performance and reliability in summarizing clinical notes while emphasizing support clinicians in composing objective assessments. Furthermore, we will focus on recommending future treatment plans, including suggestions on which tests or medications to prescribe. Through these advancements, we aspire to achieve more effective and comprehensive clinical note generation.

ACKNOWLEDGMENTS

This work was supported by the Basic Research Lab Program through the National Research Foundation of Korea (NRF-2021R1A4 A1032582) and by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. RS-2022-00155966, Artificial Intelligence Convergence Innovation Human Resources Development (Ewha Womans University)).

REFERENCES

[1] Chiwon Ahn. 2023. Exploring ChatGPT for information of cardiopulmonary resuscitation. *Resuscitation* 185 (2023).

[2] Stephen R Ali, Thomas D Dobbs, Hayley A Hutchings, and Iain S Whitaker. 2023. Using ChatGPT to write patient clinic letters. *The Lancet Digital Health* 5, 4 (2023), e179–e181.

[3] Brian G Arndt, John W Beasley, Michelle D Watkinson, Jonathan L Temte, Wen-Jan Tuan, Christine A Sinsky, and Valerie J Gilchrist. 2017. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. *The Annals of Family Medicine* 15, 5 (2017), 419–426.

[4] Omar Ayaad, Aladeen Alloubani, Eyad Abu ALhajaa, Mohammad Farhan, Sami Abuseif, Ahmad Al Hroub, and Laila Akhu-Zaheya. 2019. The role of electronic medical records in improving the quality of health care services: Comparative study. *International journal of medical informatics* 127 (2019), 63–67.

[5] Kevin J Bennett and Christian Steen. 2010. Electronic medical record customization and the impact upon chart completion rates. *Fam Med* 42, 5 (2010), 338–42.

[6] B Briggs and H Carter-Templeton. 2014. Electronic health record customization: a quality improvement project. *On-Line Journal of Nursing Informatics* 18, 3 (2014).

[7] L Byrne, J Bottomley, and A Turk. 2016. British Medical association survey of GPs in England. *London: British Medical Association* (2016).

[8] Gunther Eysenbach et al. 2023. The role of ChatGPT, generative language models, and artificial intelligence in medical education: a conversation with ChatGPT and a call for papers. *JMIR Medical Education* 9, 1 (2023), e46885.

[9] Carrie Lee Gardner and Patricia F Pearce. 2013. Customization of electronic medical record templates to improve end-user satisfaction. *CIN: Computers, Informatics, Nursing* 31, 3 (2013), 115–121.

[10] John Giorgi, Augustin Toma, Ronald Xie, Sondra Chen, Kevin An, Grace Zheng, and Bo Wang. 2023. Wanglab at mediqa-chat 2023: Clinical note generation from doctor-patient conversations using large language models. In *Proceedings of the 5th Clinical Natural Language Processing Workshop*. 323–334.

[11] Takanobu Hirosawa, Yukinori Harada, Masashi Yokose, Tetsu Sakamoto, Ren Kawamura, and Taro Shimizu. 2023. Diagnostic accuracy of differential-diagnosis lists generated by generative pretrained transformer 3 chatbot for clinical vignettes with common chief complaints: A pilot study. *International journal of environmental research and public health* 20, 4 (2023), 3378.

[12] Gretchen M Hultman, Jenna L Marquard, Elizabeth Lindemann, Elliot Arsoniadis, Serguei Pakhomov, and Genevieve B Melton. 2019. Challenges and opportunities to improve the clinician experience reviewing electronic progress notes. *Applied clinical informatics* 10, 03 (2019), 446–453.

[13] Albert Q Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, et al. 2023. Mistral 7B. *arXiv preprint arXiv:2310.06825* (2023).

[14] Eunkyung Jo, Daniel A Epstein, Hyunhoon Jung, and Young-Ho Kim. 2023. Understanding the benefits and challenges of deploying conversational AI leveraging large language models for public health intervention. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.

[15] Jeffrey G Klann and Peter Szolovits. 2009. An intelligent listening framework for capturing encounter notes from a doctor-patient dialog. *BMC medical informatics and decision making* 9 (2009), 1–10.

[16] Linfeng Li, Peng Wang, Jun Yan, Yao Wang, Simin Li, Jinpeng Jiang, Zhe Sun, Buzhou Tang, Tsung-Hui Chang, Shenghui Wang, et al. 2020. Real-world data medical knowledge graph: construction and applications. *Artificial intelligence in medicine* 103 (2020), 101817.

[17] Steven Y Lin, Tait D Shanafelt, and Steven M Asch. 2018. Reimagining clinical documentation with artificial intelligence. In *Mayo Clinic Proceedings*, Vol. 93. Elsevier, 563–565.

[18] Jialin Liu, Changyu Wang, and Siru Liu. 2023. Utility of ChatGPT in clinical practice. *Journal of Medical Internet Research* 25 (2023), e48568.

[19] Lientje Maas, Mathan Geurtsen, Florian Nouwt, Stefan Schouten, Robin Van De Water, Sandra Van Dulmen, Fabiano Dalpiaz, Kees Van Deemter, and Sjaak Brinkkemper. 2020. The Care2Report System: Automated Medical Reporting as an Integrated Solution to Reduce Administrative Burden in Healthcare.. In *HICSS*. 1–10.

[20] Gregory Makoul, Raymond H Curry, and Paul C Tang. 2001. The use of electronic medical records: communication patterns in outpatient encounters. *Journal of the American Medical Informatics Association* 8, 6 (2001), 610–615.

[21] Nir Menachemi and Taleah H Collum. 2011. Benefits and drawbacks of electronic health record systems. *Risk management and healthcare policy* (2011), 47–55.

[22] Stéphane M Meystre, Guergana K Savova, Karin C Kipper-Schuler, and John F Hurdle. 2008. Extracting information from textual documents in the electronic health record: a review of recent research. *Yearbook of medical informatics* 17, 01 (2008), 128–144.

[23] Harsha Nori, Yin Tat Lee, Sheng Zhang, Dean Carignan, Richard Edgar, Nicolo Fusi, Nicholas King, Jonathan Larson, Yuanzhi Li, Weishung Liu, et al. 2023. Can generalist foundation models outcompete special-purpose tuning? case study in medicine. *arXiv preprint arXiv:2311.16452* (2023).

[24] Aditya kumar Purohit, Aditya Upadhyaya, and Adrian Holzer. 2023. Chatgpt in healthcare: Exploring ai chatbot for spontaneous word retrieval in aphasia. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing*. 1–5.

[25] Alec Radford, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. 2023. Robust speech recognition via large-scale weak supervision. In *International Conference on Machine Learning*. PMLR, 28492–28518.

[26] Arya S Rao, Michael Pang, John Kim, Meghana Kamineni, Winston Lie, Anoop K Prasad, Adam Landman, Keith Dryer, and Marc D Succi. 2023. Assessing the utility of ChatGPT throughout the entire clinical workflow. *medRxiv* (2023), 2023–02.

[27] Mary Sabry Abdel-Messih and Maged N Kamel Boulos. 2023. ChatGPT in clinical toxicology. *JMIR Medical Education* 9 (2023), e46876.

[28] Benjamin Schloss and Sandeep Konam. 2020. Towards an automated SOAP note: classifying utterances from medical conversations. In *Machine Learning for Healthcare Conference*. PMLR, 610–631.

[29] Christine Sinsky, Lacey Colligan, Ling Li, Mirela Prgomet, Sam Reynolds, Lindsey Goeders, Johanna Westbrook, Michael Tutty, and George Blike. 2016. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. *Annals of internal medicine* 165, 11 (2016), 753–760.

[30] Ming Tai-Seale, Cliff W Olson, Jinnan Li, Albert S Chan, Criss Morikawa, Meg Durbin, Wei Wang, and Harold S Luft. 2017. Electronic health record logs indicate that physicians split time evenly between seeing patients and desktop medicine. *Health affairs* 36, 4 (2017), 655–662.

[31] Paul C Tang, Michael P LaRosa, and Susan M Gorden. 1999. Use of computer-based records, completeness of documentation, and appropriateness of documented clinical decisions. *Journal of the American Medical Informatics Association* 6, 3 (1999), 245–251.

[32] Augustin Toma, Senthujan Senkaiahliyan, Patrick R Lawler, Barry Rubin, and Bo Wang. 2023. Generative AI could revolutionize health care—but not if control is ceded to big tech. *Nature* 624, 7990 (2023), 36–38.

[33] Changyu Wang, Siru Liu, Hao Yang, Jiulin Guo, Yuxuan Wu, and Jialin Liu. 2023. Ethical considerations of using ChatGPT in health care. *Journal of Medical Internet Research* 25 (2023), e48009.

[34] Lawrence L. Weed. 1964. Medical records, patient care, and medical education. *Irish Journal of Medical Science* 39, 6 (June 1964), 271–282. https://doi.org/10.1007/bf02945791

[35] Saul J Weiner, Shiyuan Wang, Brendan Kelly, Gunjan Sharma, and Alan Schwartz. 2020. How accurate is the medical record? A comparison of the physician's note with a concealed audio recording in unannounced standardized patient encounters. *Journal of the American Medical Informatics Association* 27, 5 (2020), 770–775.

[36] Sharon M Weyer, Nancy Konrad, Diane Esola, Meredith A Goodwin, Kurt C Stange, and Susan A Flocke. 2005. Features of medical records in community practices and their association with preventive service delivery. *Medical care* 43, 1 (2005), 28–33.