

Article

Abridge: A Mission Driven Approach to Machine Learning for Healthcare Conversation

Sandeep Konam

Abridge AI Inc., Pittsburgh, PA 15219

Shivdev Rao

Abridge AI Inc., Pittsburgh, PA 15219

ABSTRACT

In this brief case study, we describe an approach to structuring and summarizing information from one of the largest untapped sources of data in healthcare delivery — spoken conversations. Abridge's mission is to shift agency to the people and families at the center of those spoken conversations, using bleeding-edge machine learning and human-centered design. The space of conversation understanding is largely untapped and we will discuss our scientific approaches to business challenges that map to the company's mission of helping everyone better understand and follow through on their healthcare conversations.

Journal of Commercial Biotechnology (2021) 26(2), 62–66. doi: 10.5912/jcb994

INTRODUCTION

OUR CORE THESIS is that healthcare is driven by spoken conversations. There are 883.7M [1] spoken doctor-patient conversations every year in the United States. And there are over 2 billion conversations when also accounting for conversations that people have with nurses, pharmacists, care managers, and other healthcare professionals. We believe that the highest leverage care for those with chronic diseases such as diabetes, cardiac disease, and cancer, is delivered via these spoken conversations as opposed to an asynchronous text message or chatbot. In addition, we know that these spoken conversations are actually upstream and less dependent on hegemonic healthcare IT systems such as the Electronic Medical Record.

We founded Abridge with a mission to help people better understand and follow through on those conversations. In addition, and using the same underlying technology, Abridge helps healthcare professionals across Payers, Providers, and Pharma save time in their own professional workflows. That professional value maps to large markets in and of themselves — for example, provider documentation itself is a \$4-10B market [2] in the United States. At a broader and more systemic level, technology that can improve the quality of healthcare conversations can address many of the efficiencies and

waste in the US healthcare system. That waste on aggregate represents costs of \$760–\$930B, representing 25% of total healthcare expenditures [3].

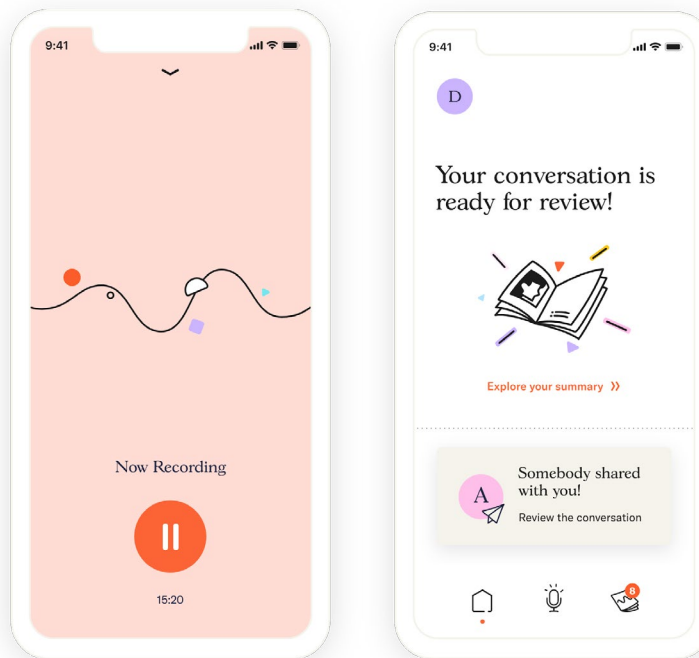
In this case study, we offer a high level description of the end user pain point and solution that we are focused on, and also highlight aspects of the machine learning research that underpin the end user experience we deliver to our users.

PROBLEM AND SOLUTION

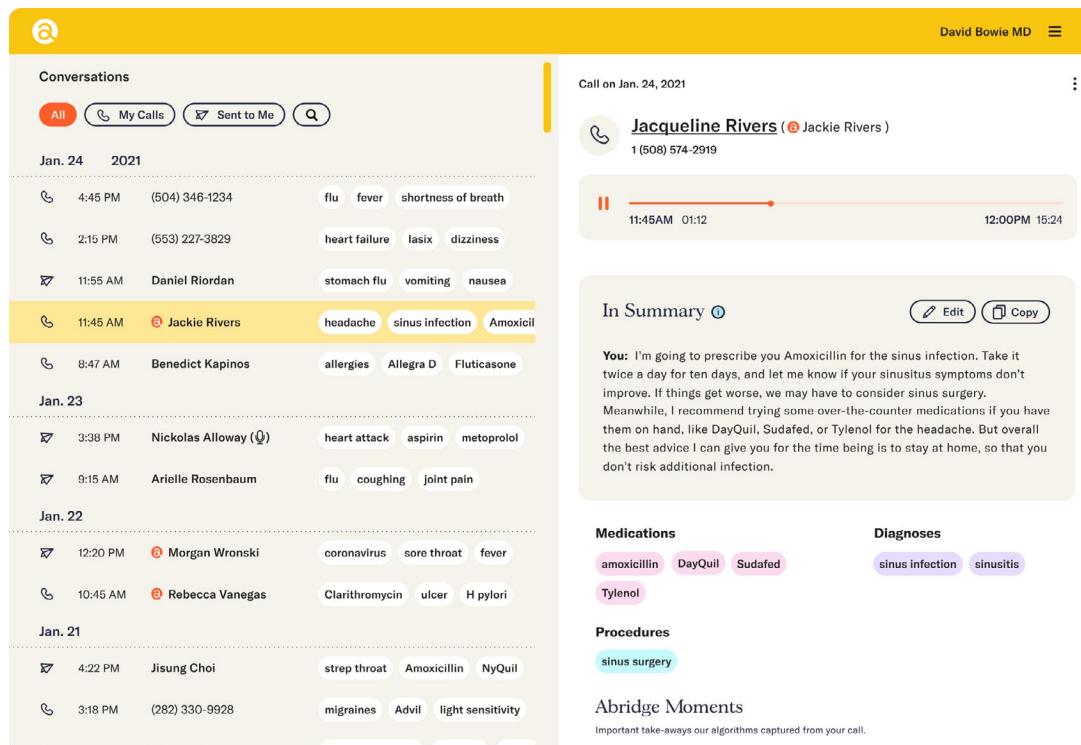
People forget up to 80% of what they hear in medical conversations [4]. In fact, studies show that half of all patients walk away from medical conversations unclear on what they were just told, unless they took notes or had someone accompany them. Poor recall, understanding, and follow-through leads to poor outcomes, especially given that healthcare is powered in large part by these spoken conversations. For example, adherence to care plans can be as low as 50% for chronic disease patients, and poor adherence in diabetes alone costs \$25B annually.

Our solution includes a mobile phone application that any person can download to immediately begin recording their health related conversations.

On the healthcare professional side, Abridge can integrate with any modern telemedicine solution and



Abridge App



Abridge Professional Dashboard

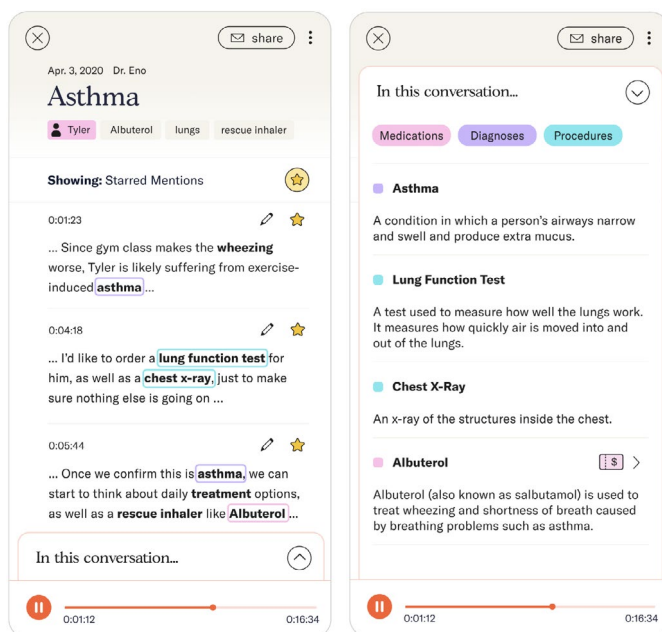
also offers a full stack solution for telephone and video calls. Patients can access those professional – initiated Abridge conversations via the consumer mobile application as above. In this way, when Abridge is a part of the conversation, patients as well as healthcare professionals and their associated enterprise systems can benefit.

POWERED BY ARTIFICIAL INTELLIGENCE

Machine learning powers many of the key features that help patients and healthcare professionals alike derive more value from their conversations. Since inception, the company has invested heavily into Artificial Intelligence challenges that map to the mission of helping people better understand and follow through on their medical conversations. At this time, the company has published 10 papers [5] on spoken medical conversation AI. These papers are centered on challenges around transcribing, summarizing, classifying, and extracting relevant information from medical conversations. In the following sections, we present an overview of some key mission-driven machine learning challenges.

1. BETTER UNDERSTANDING:

To help people better understand their medical conversation, our machine learning algorithms transcribe, extract and highlight the key clinical concepts, and define the medical jargon at a consumer reading level. Complex medical terminology, accents, interruptions, overlapping speech, false starts, and filler words like “umm” and “okay” all make it harder for an algorithm to track a conversation correctly [6]. Abridge algorithms need to accurately capture the words in each conversation before they can determine which parts of the conversation are important to people’s health. That’s why we tackle challenges and contribute to research in Automatic Speech Recognition (ASR), the field of machine learning dedicated to the transcription of speech. We also use machine learning to adapt, or correct, off-the-shelf ASR systems to improve the transcription accuracy of medical terminology. We’ve trained our algorithms to focus more on medical concepts, and to understand relevant bits of context that might be spread across each conversation [7] [8]. The output of our ASR system — the transcript — is passed through our clinical concept extraction pipeline, which highlights medications, diagnoses, and procedures. Some of these key medical terms are then linked with concise explanations from our trusted content partners, including the National Library of Medicine and the Mayo Clinic.



[On the left] Transcribed medical parts of the conversation with clinical concepts highlighted, [On the right] Definitions curated with help from our trusted partners such as the Mayo Clinic

2. BETTER FOLLOW THROUGH:

To help people better follow through on their care plan, we built a machine learning model [9][10] that can classify utterances from medical conversations according to (i) whether they were more likely spoken by a doctor or patient, and (ii) where they might be classified into specific sections of a doctor's note that a patient would benefit from understanding. We adopted a widely-accepted SOAP Note template that contains:

- **Subjective:** The “story” from the patient about why they are visiting.
- **Objective:** The objective record of the doctor's physical examination and review of diagnostic results.
- **Assessment:** A summary of the doctor's decision-making process and diagnoses.
- **Plan:** The doctor's next steps for the patient based upon their Assessment.

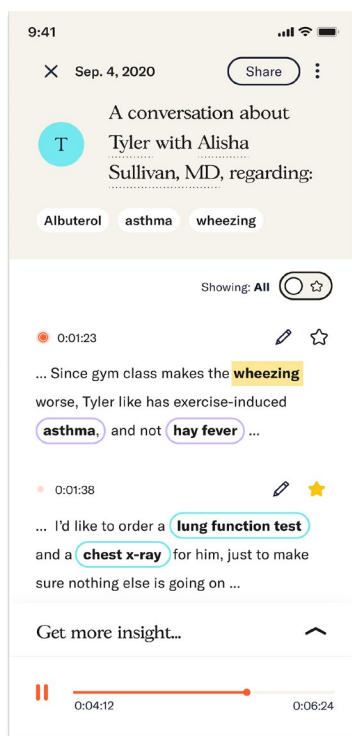
Using the above four classes, we formulated a multi-label classification problem and built a classifier that can identify clinically relevant parts of the conversation. For this classifier to perform well on ASR transcripts as input, we also developed a method for mapping human annotations from a clear, high-quality signal (the human transcript) to a noisier signal (the ASR transcript). Training

our models on the ASR dataset made our systems more robust to the types of noise injected by ASR systems.

In addition to the above classification work, we also tackled research challenges around extracting structured information from the conversations. We focused primarily on two information extraction challenges so far: 1) Medication Regimen extraction [11] [12] and 2) Appointment extraction [13]. These systems can help our users in medication adherence and in keeping their appointments.

In the medication regimen extraction work, we specifically focus on frequency, route of the medication and any change in the medication's dosage or frequency. For example, given the conversation excerpt and the medication “Fosamax” as shown in the figure below, the model needs to extract the spans “one pill on Monday and one on Thursday”, “pill” and “you take” for attributes frequency, route and change, respectively.

In the appointment extraction work, we focus on extracting the appointment reason and time spans from medical conversations as shown in the figure below. The reason span refers to a phrase that corresponds to diagnostics, procedures, follow-ups, and referrals. The time span refers to a phrase that corresponds to the time of the appointment.



Care plan “starred” by our machine learning classifier

DR: Limiting your alcohol consumption is important, so, and, um, so, you know, I would recommend vitamin D¹ to be taken¹. Have you had Fosamax² before?

PT: I think my mum did.

DR: Okay, Fosamax², you take² one pill² on Monday and one on Thursday².

DR: Do you use much caffeine?

PT: No, none.

DR: Okay, this is³ Actonel³ and it's one tablet³ once a month³.

DR: Do you get a one month or a three months supply in your prescriptions?

An utterance window from a medical conversation annotated with medications and associated attributes: change, route and frequency

DR: And the thing with angiogram , uh, um, um, we already scheduled, uh, the days I do it is Mondays and on Fridays. That's the best time for me to do the angiogram.

DR: Um, yeah, I, I guess I could make it Monday next week.

DR: Or next, or Friday?

DR: Which one is better for you?

PT: Um, well, it's already - Today is, uh, uh, Tuesday.

DR: Yeah, let's do it as soon as possible.

PT: Okay, let's do it on Monday.

An utterance window from a medical conversation annotated with appointment reasons and time span

CONCLUSION

In this case study, we cover the founding thesis around which we started Abridge and briefly discuss the market, user pain points, and the patient centered solution currently being used across the United States. We specifically focus on the machine learning challenges we've been tackling in our effort to transcribe, classify, extract, and understand the medical conversations exchanged between patients and healthcare professionals. In future editions, we hope to cover regulatory challenges, go-to-market strategy and product adoption.

REFERENCES

1. CDC/National Center for Health Statistics <https://www.cdc.gov/nchs/fastats/physician-visits.htm>. Page last reviewed: April 14, 2021
2. U.S. Transcription Market Size, Share & Trends Analysis Report, <https://www.grandviewresearch.com/industry-analysis/us-transcription-market>. Report ID: GVR-4-68038-404-8. Published Date: Mar, 2020
3. Shrank, W. H., Rogstad, T. L. & Parekh, N. Waste in the US Health Care System: Estimated Costs and Potential for Savings. *JAMA*. 2019 Oct 15;322(15):1501-1509. <https://doi.org/10.1001/jama.2019.13978>. PMID: 31589283.
4. Kessels, R. P. (2003). Patients' memory for medical information. *Journal of the Royal Society of Medicine* 96(5):219-222.
5. Abridge, <https://www.abridge.com/machine-learning/publications>. Retrieved June 8, 2021.
6. Konam, S. "When will AI be ready to really understand a conversation?" <https://www.fastcompany.com/90641568/when-will-ai-be-ready-to-really-understand-a-conversation>. Published Date: May 28, 2021.
7. Mani, A., Palaskar, S., Meripo, N. V., Konam, S. & Metze, F. "ASR Error Correction and Domain Adaptation Using Machine Translation," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 6344-6348, <https://doi.org/10.1109/ICASSP40776.2020.9053126>.
8. Mani, A., Palaskar, S. & Konam, S. "Towards understanding ASR error correction for medical conversations," In *ACL Workshop on Natural Language Processing for Medical Conversations*, 2020, pp. 7-11, <https://doi.org/10.18653/v1/2020.nlpmc-1.2>
9. Schloss, B. & Konam, S. "Towards an Automated SOAP Note: Classifying Utterances from Medical Conversations. In *Proceedings of the 5th Machine Learning for Healthcare Conference*, 2020, pp. 610-631, PMLR.
10. Ferracane, E. & Konam, S. (2020). "Towards Fairness in Classifying Medical Conversations into SOAP Sections. *AAAI Workshop: Trustworthy AI for Healthcare*, 2021
11. Selvaraj S.P. & Konam S. (2021). Medication Regimen Extraction from Medical Conversations. In: Shaban-Nejad A., Michalowski M., Buckeridge D.L. (eds) *Explainable AI in Healthcare and Medicine. Studies in Computational Intelligence*, vol 914. Springer, Cham. https://doi.org/10.1007/978-3-030-53352-6_18
12. Patel, S., Selvaraj S.P. & Konam, S. (2020). Weakly Supervised Medication Regimen Extraction from Medical Conversations. In *Proceedings of the 3rd Clinical Natural Language Processing Workshop* (pp. 178-193). Association for Computational Linguistics.
13. Meripo, N. & Konam, S. (2021). Extracting Appointment Spans from Medical Conversations. In *Proceedings of the Second Workshop on Natural Language Processing for Medical Conversations* (pp. 41-46). Association for Computational Linguistics.