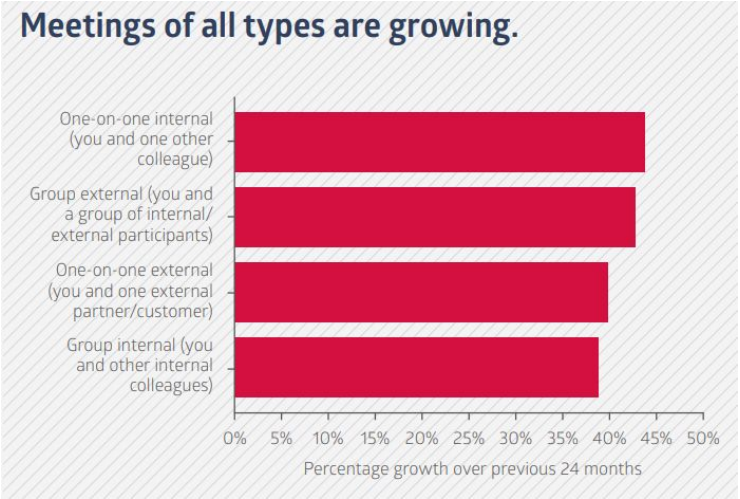
Transcriber

ML-Based Live audio transcription and webinar service

Motivation

According to the results of the *“Streamlined Video Conferencing Strategy Drives Business Results”* study, the video conferencing is growing in usage, in the number of solutions, and in overall complexity for end users and IT alike already for a number of years.

Additionally, online meetings have their fair share of “Zoom fatigue”. Our solution provides automatic meeting minutes generated on the fly as the participants speak.



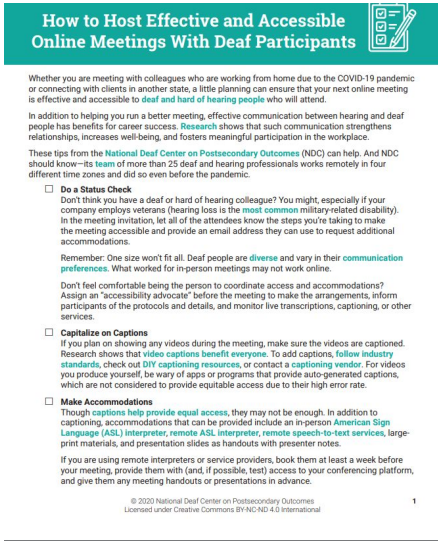
Ovum, Conference 2.0

Motivation

Nowadays, due to the the Covid-19 many people are forced to work from home, but even before the pandemic people needed to e.g. connect with their clients from another country or state.

Therefore it is important to plan the video online meetings to ensure that they are hosted in an effective and accessible to deaf and hard of hearing people in case if they will attend.

According to National Deaf Center on Postsecondary Outcomes, one of the key points of their tips on *HOW TO HOST EFFECTIVE AND ACCESSIBLE ONLINE MEETINGS WITH DEAF PARTICIPANTS* is providing an audio transcript.



How to Host Effective and Accessible Online Meetings With Deaf Participants

Whether you are meeting with colleagues who are working from home due to the COVID-19 pandemic or connecting with clients in another state, a little planning can ensure that your next online meeting is effective and accessible to deaf and hard of hearing people who will attend.

In addition to helping you run a better meeting, effective communication between hearing and deaf people has benefits for career success. Research shows that such communication strengthens relationships, increases well-being, and fosters meaningful participation in the workplace.

These tips from the National Deaf Center on Postsecondary Outcomes (NDC) can help. And NDC should know—its team of more than 25 deaf and hearing professionals works remotely in four different time zones and did so even before the pandemic.

- ☐ **Do a Status Check**
  Don't think you have a deaf or hard of hearing colleague? You might, especially if your company employs veterans (hearing loss is the most common military-related disability). In the meeting invitation, let all of the attendees know the steps you're taking to make the meeting accessible and provide an email address they can use to request additional accommodations.

  Remember: One size won't fit all. Deaf people are diverse and vary in their communication preferences. What worked for in-person meetings may not work online.

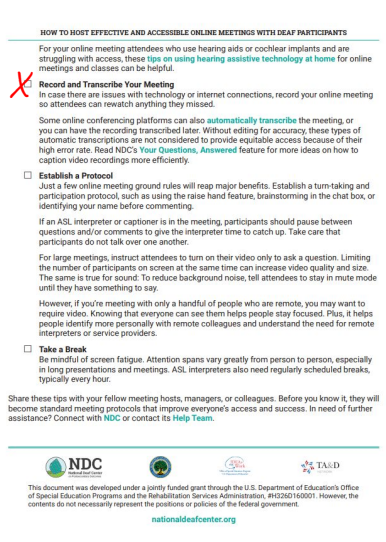
  Don't feel comfortable being the person to coordinate access and accommodations? Assign an "accessibility advocate" before the meeting to make the arrangements, inform participants of the protocols and details, and monitor live transcriptions, captioning, or other services.

- ☐ **Capitalize on Captions**
  If you plan on showing any videos during the meeting, make sure the videos are captioned. Research shows that video captions benefit everyone. To add captions, follow industry standards, check out DIY captioning resources, or contact a captioning vendor. For videos you produce yourself, be wary of apps or programs that provide auto-generated captions, which are not considered to provide equitable access due to their high error rate.

- ☐ **Make Accommodations**
  Though captions help provide equal access, they may not be enough. In addition to captioning, accommodations that can be provided include an in-person American Sign Language (ASL) interpreter, remote ASL interpreter, remote speech-to-text services, large-print materials, and presentation slides as handouts with presenter notes.

  If you are using remote interpreters or service providers, book them at least a week before your meeting, provide them with (and, if possible, test) access to your conferencing platform, and give them any meeting handouts or presentations in advance.

© 2020 National Deaf Center on Postsecondary Outcomes
Licensed under Creative Commons BY-NC-ND 4.0 International



HOW TO HOST EFFECTIVE AND ACCESSIBLE ONLINE MEETINGS WITH DEAF PARTICIPANTS

For your online meeting attendees who use hearing aids or cochlear implants and are struggling with access, these tips on using hearing assistive technology at home for online meetings and classes can be helpful.

- ☐ **Record and Transcribe Your Meeting**
  In case there are issues with technology or internet connections, record your online meeting so attendees can rewatch anything they missed.

  Some online conferencing platforms can also automatically transcribe the meeting, or you can have the recording transcribed later. Without editing for accuracy, these types of automatic transcriptions are not considered to provide equitable access because of their high error rate. Read NDC's Your Questions, Answered feature for more ideas on how to caption video recordings more efficiently.

- ☐ **Establish a Protocol**
  Just a few online meeting ground rules will reap major benefits. Establish a turn-taking and participation protocol, such as using the raise hand feature, brainstorming in the chat box, or identifying your name before commenting.

  If an ASL interpreter or captioner is in the meeting, participants should pause between questions and/or comments to give the interpreter time to catch up. Take care that participants do not talk over one another.

  For large meetings, instruct attendees to turn on their video only to ask a question. Limiting the number of participants on screen at the same time can increase video quality and size. The same is true for sound. To reduce background noise, tell attendees to stay in mute mode until they have something to say.

  However, if you're meeting with only a handful of people who are remote, you may want to require video. Knowing that everyone can see them helps people stay focused. Plus, it helps people identify more personally with remote colleagues and understand the need for remote interpreters or service providers.

- ☐ **Take a Break**
  Be mindful of screen fatigue. Attention spans vary greatly from person to person, especially in long presentations and meetings. ASL interpreters also need regularly scheduled breaks, typically every hour.

Share these tips with your fellow meeting hosts, managers, or colleagues. Before you know it, they will become standard meeting protocols that improve everyone's access and success. In need of further assistance? Connect with NDC or contact its Help Team.

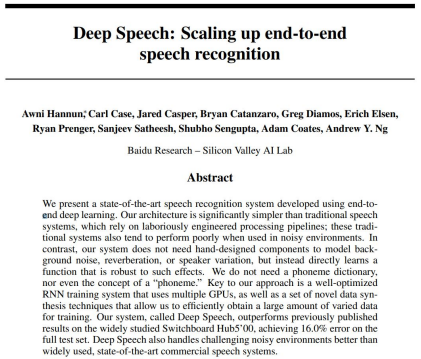
This document was developed under a jointly funded grant through the U.S. Department of Education's Office of Special Education Programs and the Rehabilitation Services Administration, #H326D160001. However, the contents do not necessarily represent the positions or policies of the federal government.

nationaldeafcenter.org

Transcriber - main features

- Transcriber allows you to convert your audio to text with **automatic transcription**.
- You don’t need to record your audio or to wait - Transcriber can provide you a **real time transcription** of your meeting.
- As the transcriptions are generated in real time, Transcriber doesn’t need to record and store your meetings - **we value your privacy**!
- The users also are able to communicate with each other using a built-in chat.
- You do not need to install special software - you can access the Transcriber in your browser!

- To provide the highest quality of your transcription, we use machine learning techniques based on the **Deep Speech research paper**.



Deep Speech: Scaling up end-to-end speech recognition

Awni Hannun, Carl Case, Jared Casper, Bryan Catanzaro, Greg Diamos, Erich Elsen, Ryan Prenger, Sanjeev Satheesh, Shubho Sengupta, Adam Coates, Andrew Y. Ng

Baidu Research – Silicon Valley AI Lab

Abstract

We present a state-of-the-art speech recognition system developed using end-to-end deep learning. Our architecture is significantly simpler than traditional speech systems, which rely on laboriously engineered processing pipelines; these traditional systems also tend to perform poorly when used in noisy environments. In contrast, our system does not need hand-designed components to model background noise, reverberation, or speaker variation, but instead directly learns a function that is robust to such effects. We do not need a phoneme dictionary, nor even the concept of a "phoneme." Key to our approach is a well-optimized RNN training system that uses multiple GPUs, as well as a set of novel data synthesis techniques that allow us to efficiently obtain a large amount of varied data for training. Our system, called Deep Speech, outperforms previously published results on the widely studied Switchboard Hub5'00, achieving 16.0% error on the full test set. Deep Speech also handles challenging noisy environments better than widely used, state-of-the-art commercial speech systems.

How does it work?

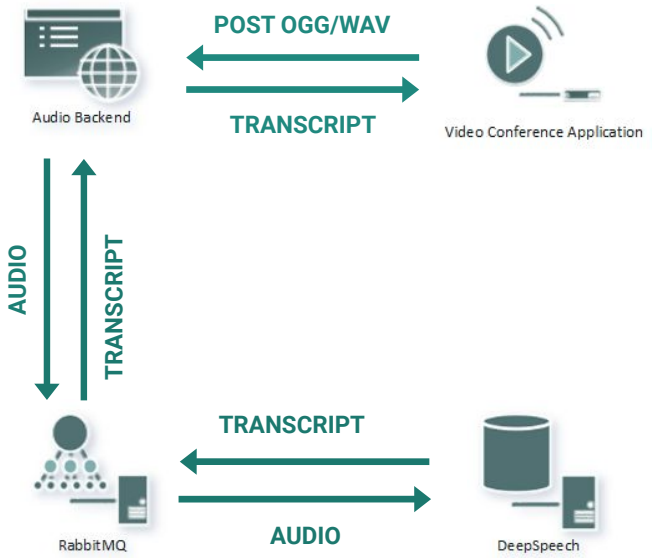
Several services come into play in the development of the application:

- a front-end using real time communication technologies
- a back-end that processes audio samples and generates transcripts
- a back-end that acts as a central hub for the real time text, video and audio communication.

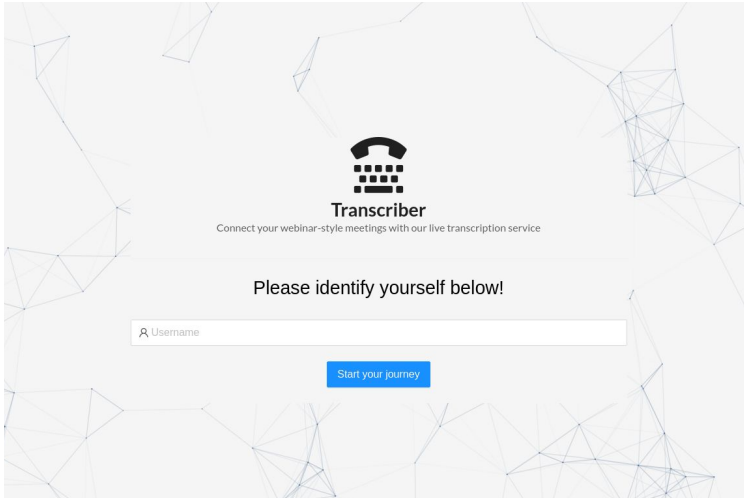
Use-case example:

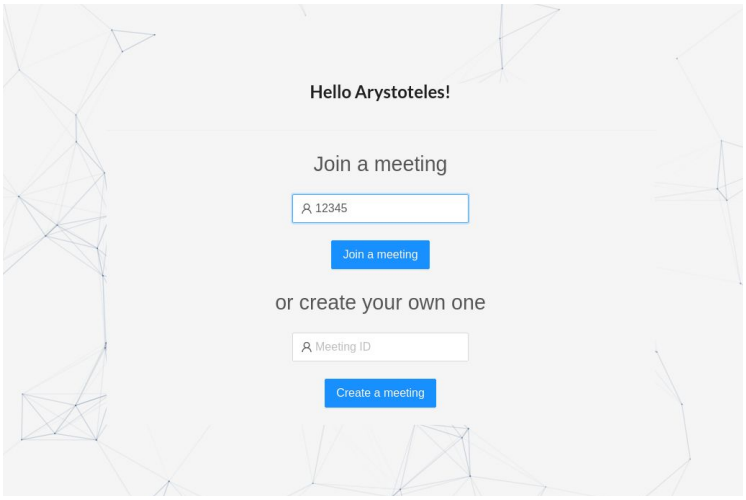
1. You host a webinar
2. You are presenting. As you present, automatic transcription is generated for the project.
3. You interact via real-time chat with your participants, that can see and hear you.
4. Everything is hosted on the cloud and nothing is stored after the meeting, all audio files are temporary.

Application architecture and technologies diagram

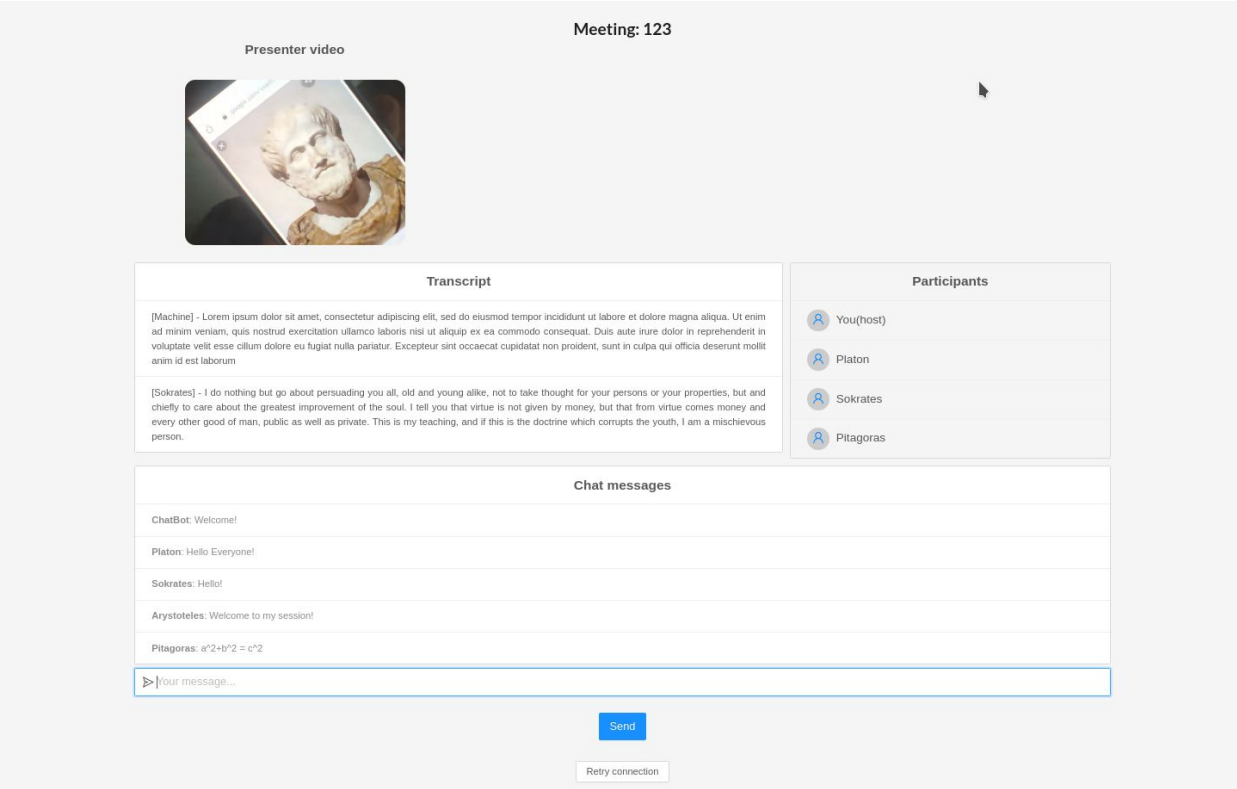


Application Design - Join / create meeting

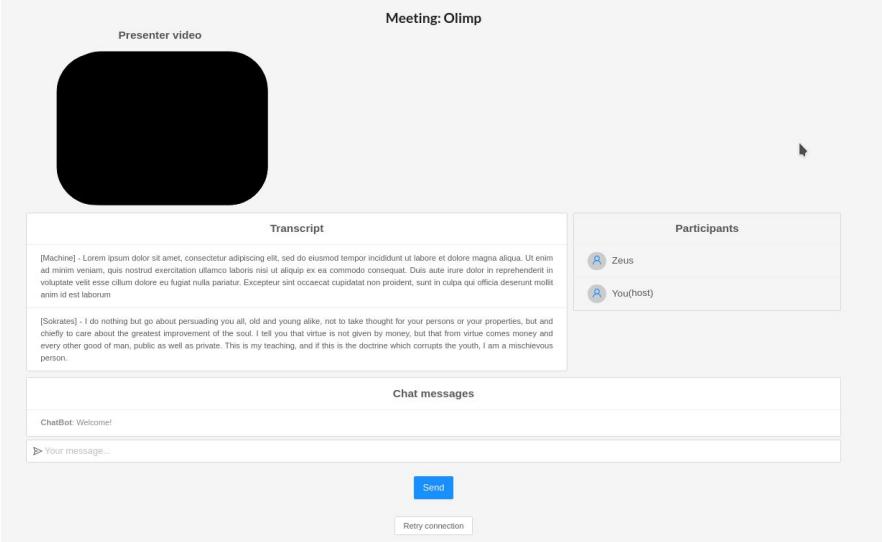


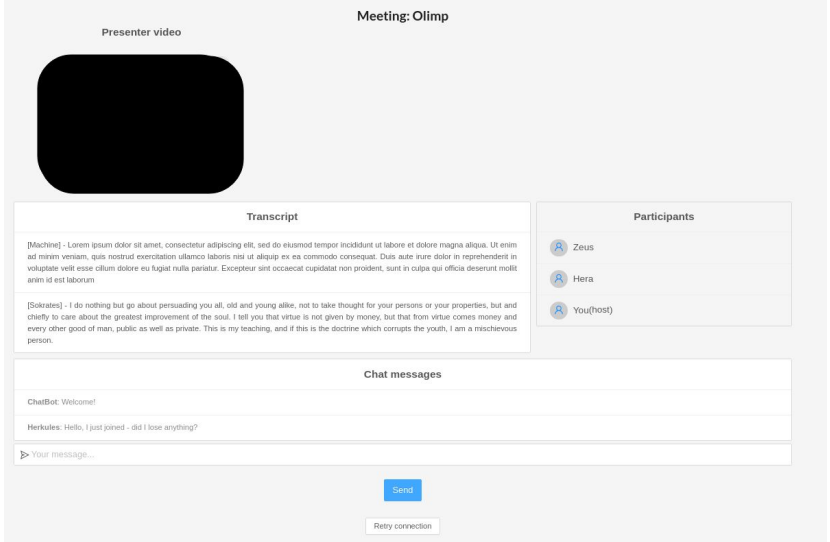


Chat Example - host view



New participants joining...





Our Team

Cristian Schuszter

Kamil Faber

Karolina Cynk