

AI and People with Disabilities

Shaomei Wu
AlImpower.org

Guest Lecture
CS 377Q/ME214
Designing for Accessibility

May 23, 2024



The Promise

Can AI be the magic pill that will "solve" disabilities once for all?

Using Artificial Intelligence to Help Blind People 'See' Facebook

April 4, 2016



By [Shaomei Wu](#), Software Engineer and [Hermes Pique](#), Software Engineer on iOS and [Jeffrey Wieland](#), Head of Accessibility

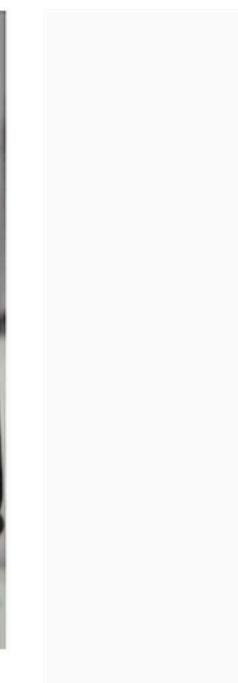
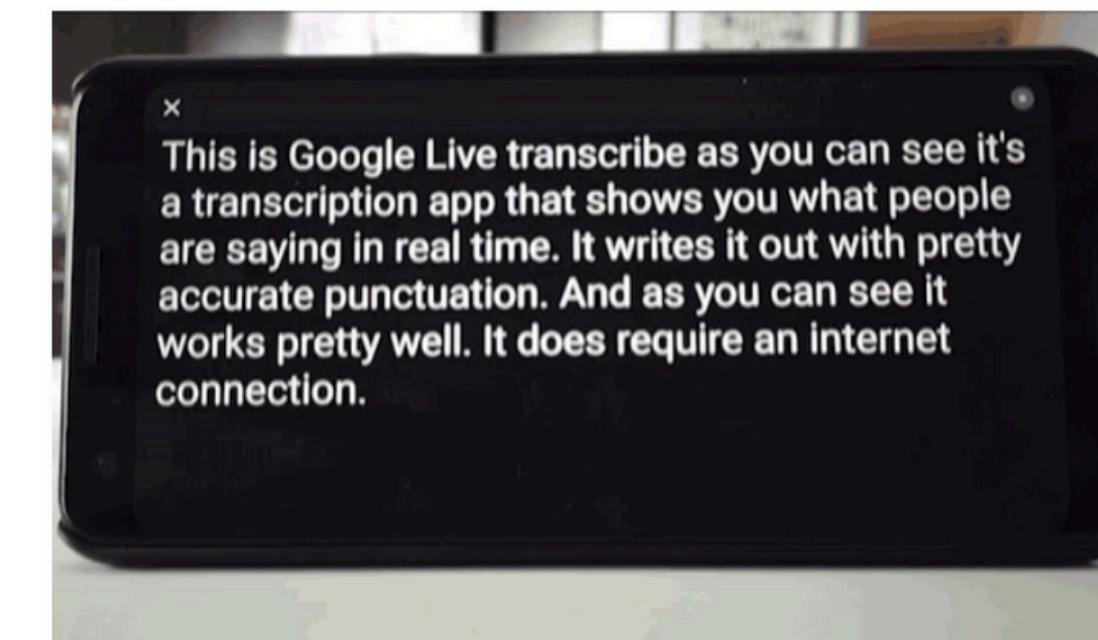
GOOGLE \ TECH \ ANDROID

Google Live Transcribe could be a big help for people who are deaf or hard of hearing

It's also releasing a Sound Amplifier feature — both only on Android

By [Dieter Bohn](#) | [@backlon](#) | Feb 4, 2019, 9:00am EST

f t SHARE



The problems

How AI penalize People with disabilities



Denied services



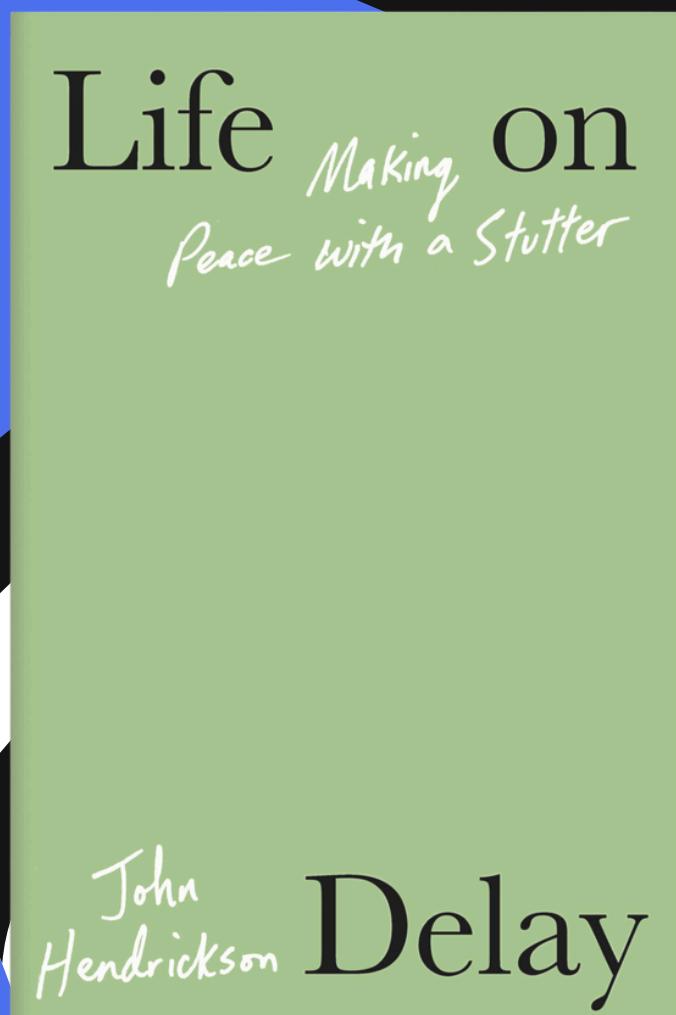
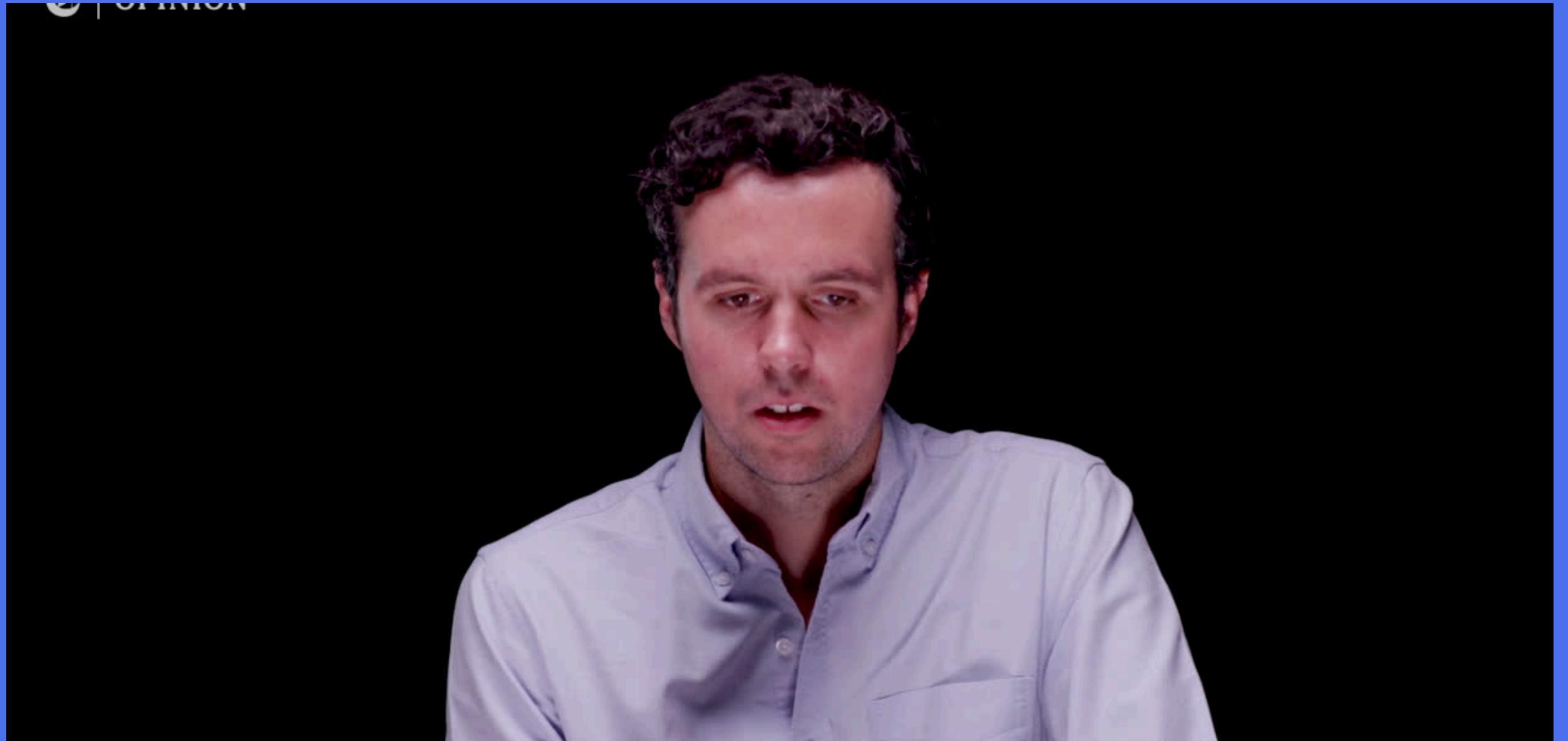
Materialized harm



Empathy machine

Denied Services

The deployment of AI systems that do not function for PwD create new accessibility barriers.



Materialized Harm: Micro Aggression

AI generated image with prompt
“a person with cerebral palsy, photo”

“These images have a lot of anguish, [the] last one looks like [a] scene out of [a] horror film.”

(P41)

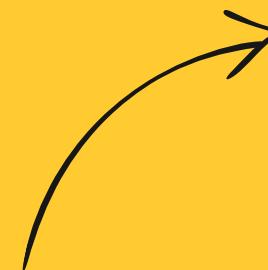


Source: Avery Mack, Rida Qadri, Remi Denton, Shaun K Kane, and Cynthia L Bennett. “**They only care to show us the wheelchair**”: disability representation in text-to-image AI models. CHI 2024

Materialized Harm: Targeted Erasure

Risks when PwD are identified by AI

How marginalized communities, including people with disabilities, are algorithmically suppressed in the name of "protection".



Alice Wong
@DisVisibility



"TikTok Admits It Suppressed Videos by Disabled, Queer, and Fat Creators

TikTok says that as part of an anti-bullying policy, it restricted videos that featured people whose bodies had “problems.””

[slate.com/technology/201... via @slate](https://slate.com/technology/2019/12/tiktok-suppresses-disabled-fat-queer-creators.html)

#TikTok #SuckItAbleism

9:56 AM · Dec 5, 2019 · TweetDeck

144 Retweets 53 Quote Tweets 133 Likes

Empathy Machine

TECHNOLOGY

Disabled people don't need so many fancy new gadgets. We just need more ramps.

Technology isn't always the answer.

by **s.e. smith**

Apr 30, 2019, 4:30 AM PDT



A woman in a wheelchair uses a ramp. Constantinis/Getty Images/iStockphoto

Liz Jackson
@elizejackson · [Follow](#)

X

Disability Dongle: A well intended elegant, yet useless solution to a problem we never knew we had. Disability Dongles are most often conceived of and created in design schools and at IDEO.

12:49 PM · Mar 26, 2019

ⓘ

367 [Reply](#) [Copy link](#)

What went wrong?

Technical factors that contribute to AI-powered marginalization of PwD

01

Data

Under and misrepresentation of PwD in AI data

02

Model Design

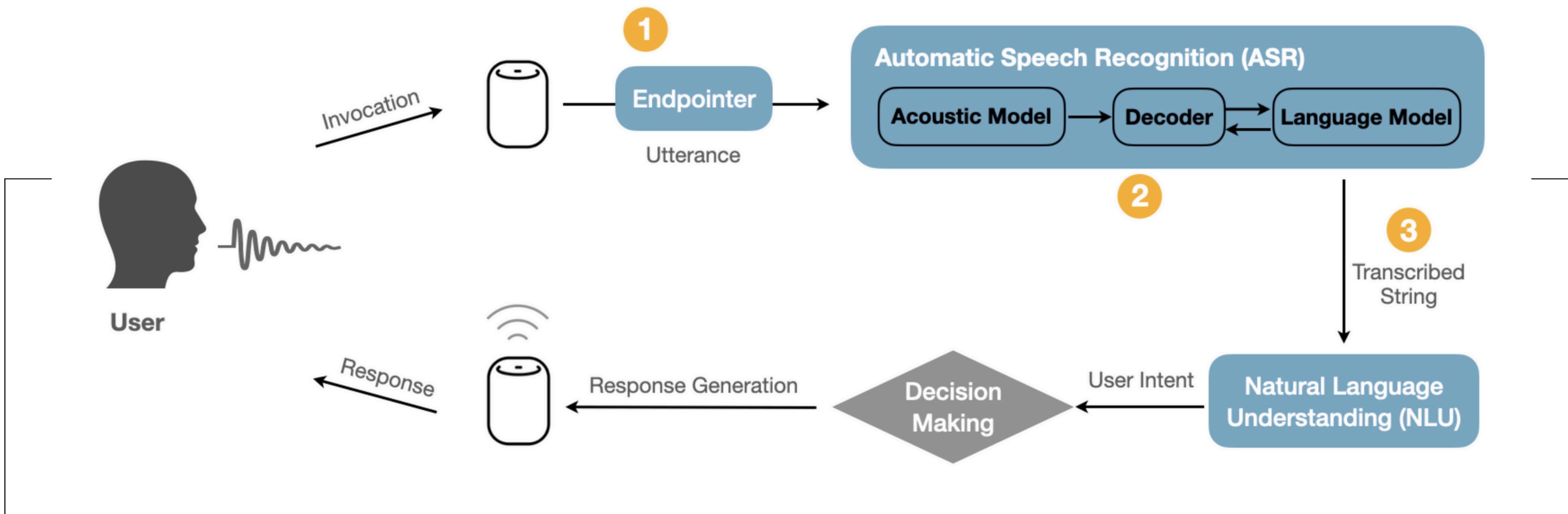
Hardcoded ableist assumptions

03

Metrics

Misalignment between standarized evaluation metrics and PwD's user experience of AI models

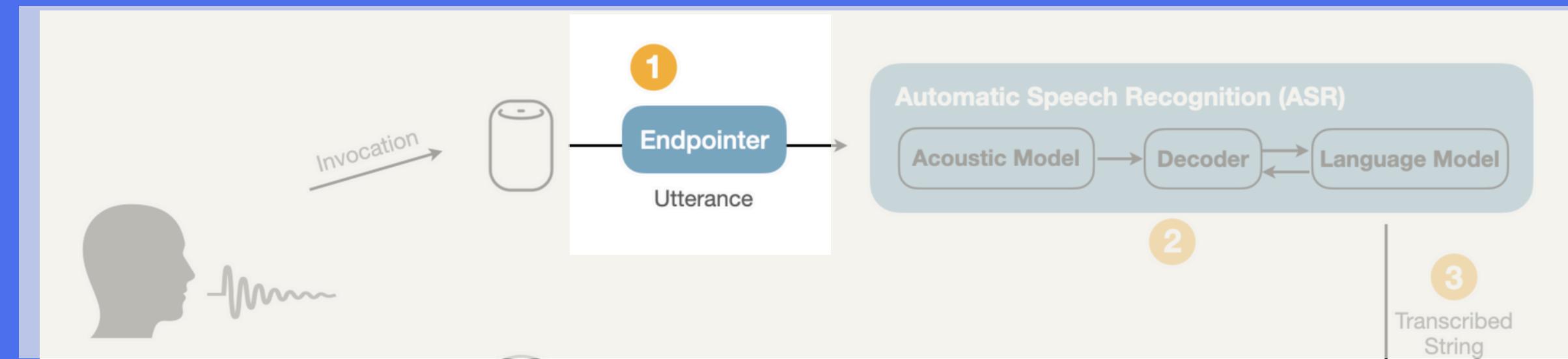
Architecture of Voice AI



Source: Lea, C.S., Huang, Z., Narain, J., Tooley, L., Yee, D., Tran, D.T., Georgiou, P.G., Bigham, J.P., & Findlater, L. (2023). From User Perceptions to Technical Improvement: Enabling People Who Stutter to Better Use Speech Recognition. CHI '23

Model Design

Ableist assumptions are often hardcoded in the model



Endpointer (aka end-of-speech-detector)

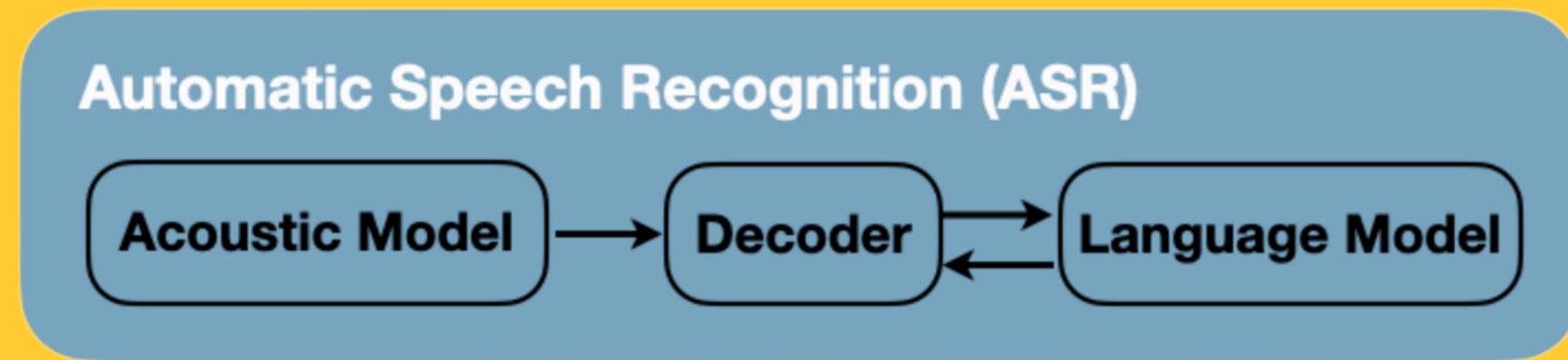
Detect whether the user has finished speaking, often tuned and optimized for low interaction latency, and tend to cut off people who stutter [1][2].



- [1] Anna Bleakley et al. 2022. Exploring Smart Speaker User Experience for People Who Stammer. ASSETS '22.
- [2] Lea, C.S. et al. (2023). From User Perceptions to Technical Improvement: Enabling People Who Stutter to Better Use Speech Recognition. CHI '23

Data

How the data from PwD are missed from popular ML datasets



LibriSpeech: 1000+ hours
audio books



WikiText: verified Good
and Featured Wikipedia
articles



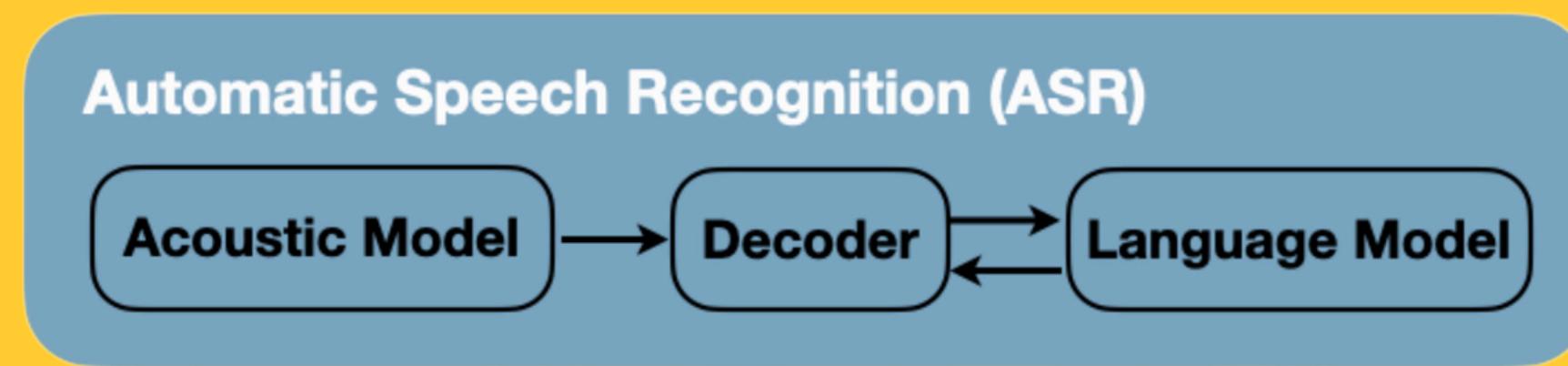
7000+ hours of
crowdsourced reading



Over 250 billion pages
crawled from the web

Data

How the data from PwD are missed from popular ML datasets



LibriSpeech: 1000+ hours
audio books



WikiText: verified Good
and Featured Wikipedia
articles



7000+ hours of
crowdsourced reading



Over 250 billion pages
crawled from the web

?

*No speech or text
with disfluencies is
represented in the
training of ASR!*

Metrics

How common metrics might not fully align with experiences

— WER (word error rate): common metrics for ASR

Measures the ratio of errors in the transcript.

Label: Hi my name is Shaomei Wu and I, um, worked at Facebook.

Wav2Vec: Hi my name is Charmaru and I am worked at Facebook.

Google Cloud: Hi my name is Joe and I [] worked at Facebook.

— Same WER score, different user experience

The two outputs elicit very different emotional responses in me as a PWS.

What can we do

Meaningful involvement in every step of AI development



Source: International Disability Alliance (IDA)
@IDA_CRPD_Forum

01

Representation in data collection

Collect data about PwD in a respectful and transparent way.

02

Representation in model development

Better surface outliers and avoid the predominance of the majority.

03

Representation in evaluation

Define new evaluation methods that represent the needs and interests of PwD

Representation in Data

Community-led Grassroots Stuttered Speech Data Collection

1 year

2023/1 - 2023/11

50 hours

speech collected

38k

*Stuttering events
annotated*

422k

*characters in verbatim
transcriptions*

70 PWS

participated

100%

positive experience

2 PWS

volunteers

2 tasks

*conversations &
command recitation*

stammer 口吃

Talk 说



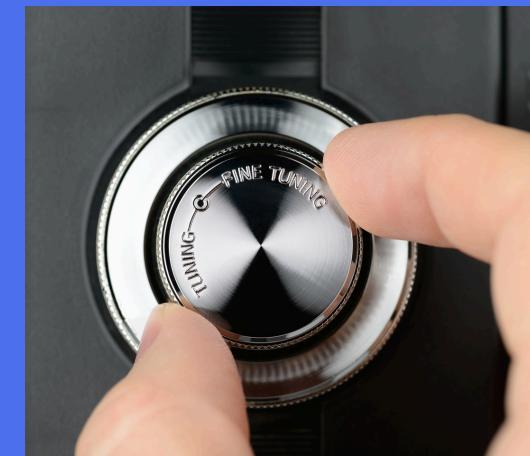
AImpower

- **Towards Fair and Inclusive Speech Recognition for Stuttering: Community-led Chinese Stuttered Speech Dataset Creation and Benchmarking.** Qisheng Li, Shaomei Wu. *CHI 2024*
- **AS-70: A Mandarin stuttered speech dataset for automatic speech recognition and stuttering event detection.** Rong Gong, Hongfei Xue, Lezhi Wang, Xin Xu, Qisheng Li, Lei Xie, Hui Bu, Shaomei Wu, Bin Jia, Jiaming Zhou, Yong Qin, Binbin Zhang, Jun Du. *InterSpeech 2024*.
- **“I Want to Publicize My Stutter”: Community-led Collection and Curation of Chinese Stuttered Speech Data.** Qisheng Li, Shaomei Wu. *CSCW 2024* (under review).

Representation in Model Development

Model	Pre-trained Dataset	Fine-tuned Dataset	mild	moderate	severe	conversation	command	all
Conformer	WenetSpeech	-	11.41	17.81	33.21	11.82	20.40	15.00
	WenetSpeech	AS-70	5.20	8.00	9.25	7.96	3.32	6.24
Hubert-large	WenetSpeech (unlabel)	AISHELL-1	16.28	22.86	26.76	16.80	23.79	19.06
	WenetSpeech (unlabel)	AS-70	6.25	9.67	7.85	9.75	2.03	7.25
Whisper-large-v2	Whisper	-	14.50	28.50	95.33	17.83	46.85	27.20
	Whisper	AS-70	5.18	13.46	18.92	10.19	5.74	8.75

— Fine tuning models with smaller datasets from marginalized groups [1]



[1] AS-70: A Mandarin stuttered speech dataset for automatic speech recognition and stuttering event detection. **Rong Gong**, Hongfei Xue, **Lezhi Wang**, Xin Xu, Qisheng Li, Lei Xie, Hui Bu, **Shaomei Wu**, **Bin Jia**, Jiaming Zhou, Yong Qin, Binbin Zhang, Jun Du. InterSpeech 2024.

Representation in evaluation

Develop human-centered eval metrics and methods with PwD

— The option for **verbatim transcription** by ASR

Respect and normalize disfluencies in human speech.

“How do we **normalize the stuttered speech**? The caption can transcribe my stuttered speech, as it is, like, **transcribe ‘my’ five times**. [...] Instead of making people who stutter feel like **the technology is waiting for a moment to correct my speech**, just **transcribe as it is**.”

— Adapt WER for verbatim transcription

Compare model transcription with verbatim transcription

- Jane (PWS participant) [1]

[1] Re-envisioning Remote Meetings: Co-designing Inclusive and Empowering Videoconferencing with People Who Stutter. Jingjin Li, Shaomei Wu, and Gilly Leshed. DIS 2024

Representation in evaluation

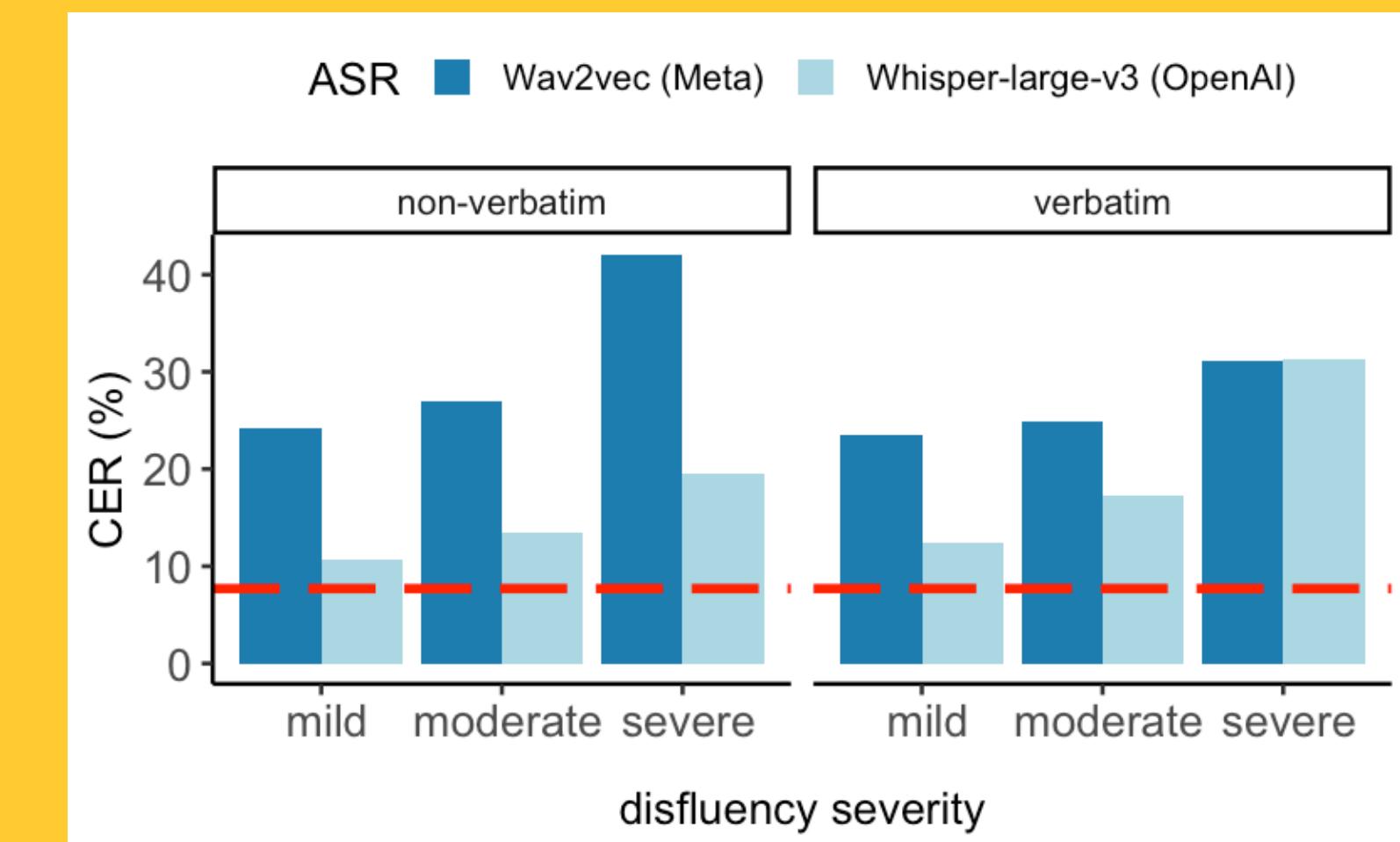
Develop human-centered eval metrics and methods with PwD

Performance disparities

40-760% increase in CER (character-error-rate) from non-stuttered speech (7.7%).

Challenges with verbatim transcription

Whisper tends to smooth over sentences, showing strong fluency biases.



Advocacy

Foster socio-structural changes through public awareness and policy frameworks

— Crip Technoscience for AI [1]

I don't just want us to get a seat at someone else's table, I want us to be able to build something more magnificent than a table, together with our accomplices.

Mia Mingus, keynote at 2018 Disability Intersectionality Summit

Advocacy

Foster socio-structural changes
through public awareness and
policy frameworks

—
Crip Technoscience for AI

—
Policy enforcement

Parity of performance to stuttered vs.
non-stuttered speech as a requirement
for commercial ASR models.



AImpower.org

Research and co-create empowering
technologies ***for, with, and by***
marginalized communities

Learn more about our ongoing work

contact@aimpower.org

