

Speech to Text

*Language models you need,
built and customized for you*

- *A unique machine learning approach that generates high accuracy models for different use cases at a rapid rate.*
- *Integrated in Video Indexer to combine Zoom Media Speech Recognition with Microsoft Cognitive Services.*
- *Easy to use API powered by Azure Cloud.*

What is Speech to Text?

Automatic Speech Recognition can turn the spoken word into written text automatically, enabling customers to search through their audiovisual content in an easy fashion.



Why customers use Zoom Media

- Index large amounts of content in real-time.
- We assure high accuracy language models.
- We put privacy first and delete all data after processing.
- We can customize our models to better fit our clients needs.
- We believe in building strong partnerships to deliver end users the best solutions.

High Accuracy

- **A model is ready when it reaches a 90%+ accuracy with clear speech**
- Monthly iterations to update all our languages
- Customization of engines for specific verticals

Average Word Error Rate of 14%

Realtime & Cloud based

- **Realtime algorithms enabling live processing**
- No restrictions on amount of content processed
- Security and scalability, we put privacy first.

Secure, Scalable and Cloud based

Video Indexer Integration

- **Zoom Media Speech to Text and Microsoft Cognitive Services combined**
- Easy integration in customer workflows
- Add all relevant additional metadata

Apply all available AI

"We use Zoom Media speech to text to add automatically generated subtitles to clips we post on social media. In doing so we can reach a wider audience, for example people that are hearing impaired or who want to watch our clips with the sound turned off."

- Wiljan Hobbelink, Media Management, AT5/NH