

People want to listen to and watch content in their native language

Traditionally achieved through **dubbing** - a post-production process where the original language of recording is swapped with audio recorded by human in a different language



Expensive

~\$100/min

→ Approximate dubbing cost including voice actors fee, post-production, and studio cost

Long Process

>2 weeks

→ 10 minute video takes at least 2 weeks to dub. Involves multiple functions. Longer ones can take months!

There are no affordable tools to make **content watchable in any language** with **high quality**.

Human quality **automated dubbing** as a SaaS

Human Quality

Preserving voice features

Automated dubbing based on thousands of hours of professional dubbing - keeping the original emotions, intonation & speakers performance

Personalized

Dubbing with your own voice

For the first time training a deep-learning model that preserves your own voice across languages

Simple & Quick

Accessible through an E2E solution

SaaS that takes an input audio or video, and enables with a click of a button to do full dubbing - human-in-the-loop is supported for improving quality even further

We have already built a prototype with **state-of-the-art research for dubbing**

1. Any movie or audio input in English
2. Subtitles generation - either automatic speech recognition or metadata extraction
3. Translation from language A to B
4. Background noise + dialogue separation
5. **Automatic dubbing - voice generation in another language - core technology**
6. Dubbed video ready for download



Demo video

Quick (10 minute video dub time)

2 minutes

We have studied, lived and worked together. We are best friends since high-school.



Piotr Dabkowski | CTO
[ML Researcher](#)

Previously Machine Learning @ Google
Computer Science at Cambridge & Oxford
University

Deep-learning researcher - published a [paper](#)
at NeurIPS with >300 citations

Open-source work - created [Js2Py](#) with >250k
downloads / month and other projects



Mati Staniszewski | CEO
[Deploying Products at Scale](#)

Deployment Strategist @ Palantir
Mathematics at Imperial College London

Experience at BlackRock & Opera Software –
modelling usage and risk metrics

Founder of new communities - created
[Mathscon](#) – first Mathematics student led
conference with >1000 students over 3 years

Eleven's **automatic dubbing** will power **seamless** communication and content across any language.

Eleven Expansion



Use-cases

Real-time dubbing

automatic language conversion across
online video and audio conversations

Real-time voice conversion

online privacy protection,
call-centers improvements,
metaverse

Professional Dubbing

full control of voice modification for
highest quality automated dubbing
in the feature movies

Localization & Advertising

advertisements
language embedding in core tools

Offline voice generation

game development,
audiobooks,
podcasts

Automatic dubbing for creators

content creators,
audio & video
editing software

Users



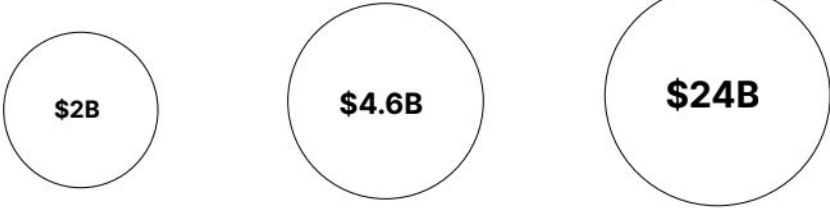
NETFLIX



LIONBRIDGE

GemiusGroup





\$2B

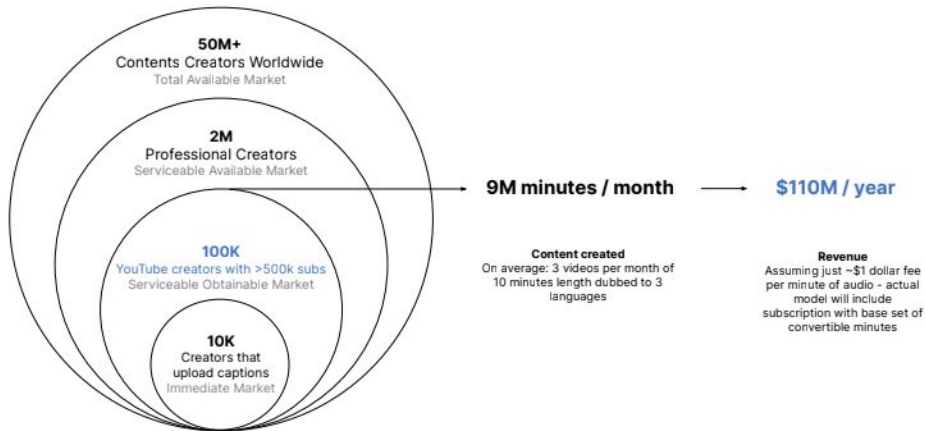
Estimate for yearly TAM in for all professional content creators across podcasts and videos

\$4.6B

Current yearly spent on game localization and movie dubbing - industry will disrupt

\$24B

Localization, translation, interpreting total market



MrBeast English channel subscribers 

96M

→ MrBeast is one of top 5 YouTube creators by subscribers, starting his career in early 2012

MrBeast Spanish channel subscribers 

19M

→ New channel started in 2021 with content dubbed professionally to Spanish. **One video generates ~\$50k!**

Key insights

- Creators will explore the same model to reach **more viewers & revenue**
- **Quick** dubbing process requirement but a lower quality bar
- High volume data allows to **improve speech & text datasets** to build long term defensibility



New way to automatically dub - preserves speakers voice, emotion, intonation

- Instead of traditional Text-to-Speech approach we take both **Speech and Text as an input to generate Speech** in a new language - with state-of-the-art results.
- **Novel speech representation** as a combination of:
 - **prosody (emotions, intonation)** - a sequence of per-phoneme, speaker independent annotations - based on professional dubbing
 - **speaker's voice** - separate speaker embedding - based on thousands of voices
- **Quick, affordable, generalizable** - easy to scale to new languages, where the end dubbing takes minutes instead of weeks

