

AI-Powered Language Learning: Dual Subtitles with Whisper and ChatGPT 🤖 🐍

by Thibault Bétrémieux | 2024-06-04





Dual Subtitles: What's the Big Deal?

Different ways of experiencing subtitles:

- No Subtitles: for adventurers 🦖
- One Subtitle (Translation): learning by matching words you hear with the ones you read 🤔
- One Subtitle (Transcription): improving listening skills, trying to guess the meaning of words 🧐
- Dual Subtitles: best of both worlds? If you can read fast enough 🐢



А ти ким була перш ніж стати почварою?
And who were you before you became a monster?

*Ukrainian dub of
Disenchantment
S01E07*

*Dual-subtitles
generated by Whisper
and ChatGPT 4o*



Oh, non ricordo più qual è la manica per l'aria e quella delle scorie.
Oh, I no longer remember which is the air sleeve and which is the waste sleeve.

*Italian dub of Futurama
S07E01*

*Dual-subtitles
generated by Whisper
and ChatGPT 4o*

A new language learning experience



Curious about Ukrainian after learning Russian



How a personal interest sparked my motivation for learning Ukrainian



Duolingo has my back



Transitioning to more advanced Methods 🤔

- Leveling up with one-on-one lessons on [italki](#)
- Combining fun and learning: playing online video games with Ukrainians
- Facing the boss battles: slow progress and language barriers



Discovery of Whisper

- A suggestion from the Python User Group Freiburg 🙌
- Early experiments with YouTube videos: a mixed bag but seems promising
- Real-time gaming hiccups: microphones, dialects (including weird Russian/Ukrainian mixes), and "a few" accuracy problems 🐒

Technical Demo #1

Using Whisper as an interpreter

- + The library [whisper mic](#) simplifies this task a lot
- Without a powerful computer 🤖:
 - Bad results
 - Too much latency
- Requires running Whisper twice at the same time (translation and transcription must run as their own task) 🐢



Turning to Media for Learning

...after being super bored with modern online games 🐢

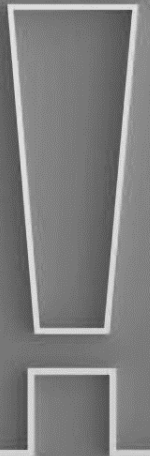
Discovering some Ukrainian gems:

- Mavka
- The Stolen Princess
- Victor_Robot



Turning to Media for Learning (cont.)

- Being a bit picky
- Lack of content in Ukrainian language
- Started watching dubs, such as:
 - Disenchantment: 50 episodes, ~30 minutes each
 - Futurama: over 150 episodes, ~22 minutes each



Subtitles in Ukrainian media 🙄

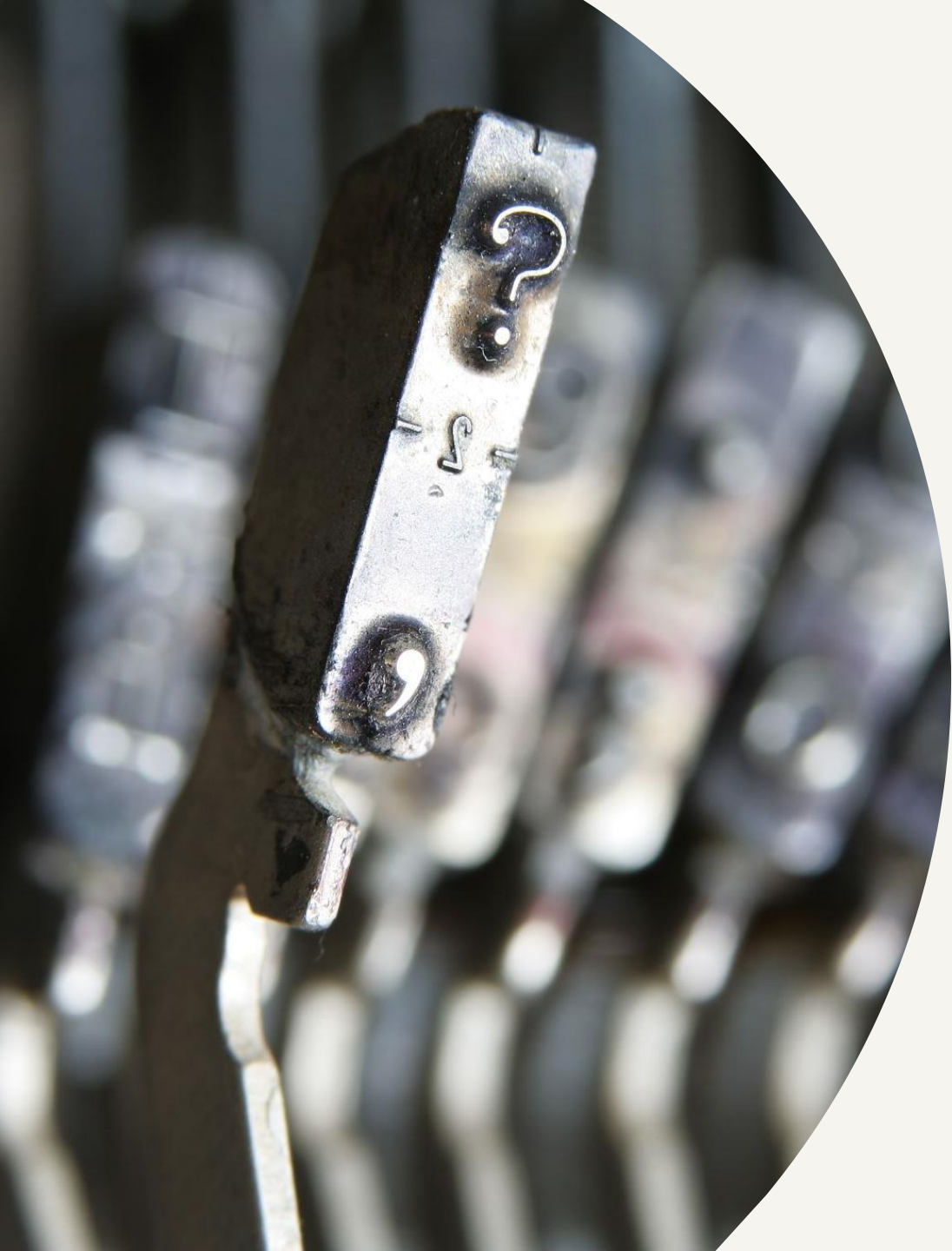


Actually, you'll
be lucky to have
any subtitle...



...and the subtitles rarely transcribe exactly what's being said





Generating dual subtitles using "ancient technology"



-
- [BingLingGroup/autosub](https://binglinggroup.github.io/autosub/) uses **A**utomatic **S**peech **R**ecognition APIs, by default Google Cloud Speech-to-Text V2
 - [PotPlayer](#) 🐱 - a media player that can display multiple subtitles at once and translate them in real-time (with Deepl by default)

Note: PotPlayer does not work on Linux 🐱



The future is now, old man.



Whisper for Transcription

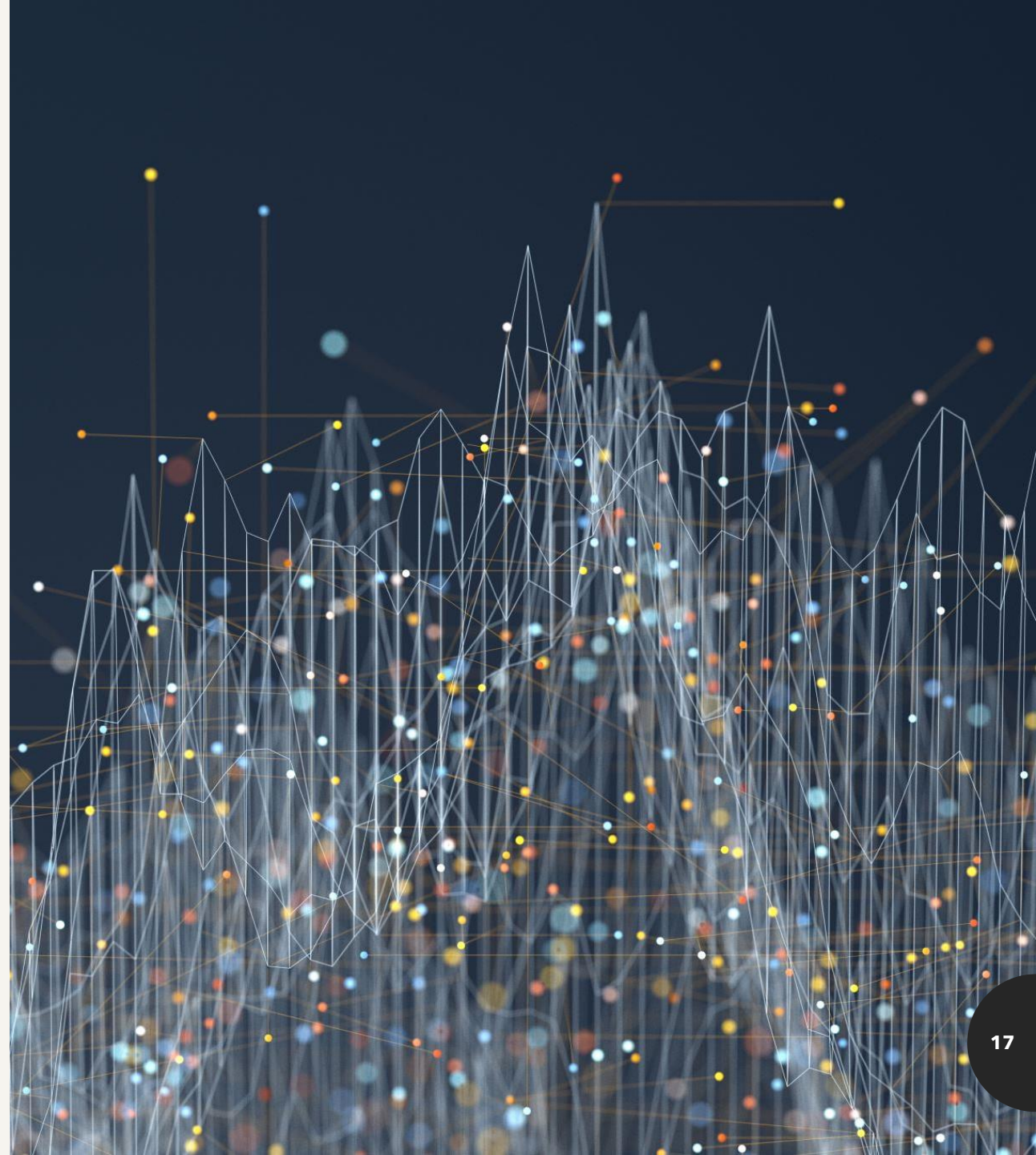
- Challenges with [Whisper](#):
 - memory requirements for large models
 - setup on Windows
 - big media files
 - hallucinations 🌀
- [faster-whisper](#) 🐱💖 - reduces memory requirements, runs faster, has no problems with big media files

Technical Demo #2

Transcriptions and measurements

Transcription and accuracy measurements of free ASR tools:

- Google Cloud Speech-to-Text V2 (via autosub)
- OpenAI's Whisper large-v3 model (via faster-whisper)



Whisper for subtitle translation

Problem:

- Subtitles **out-of-sync** with the transcription, since Whisper adjusts them for English

Solutions? Using data from the transcription we could:

- x Split medias in one file per subtitle line then put all texts back together 🐒
- x Aligning subtitles to the transcription 🤪

Note: Whisper can only translate into English!

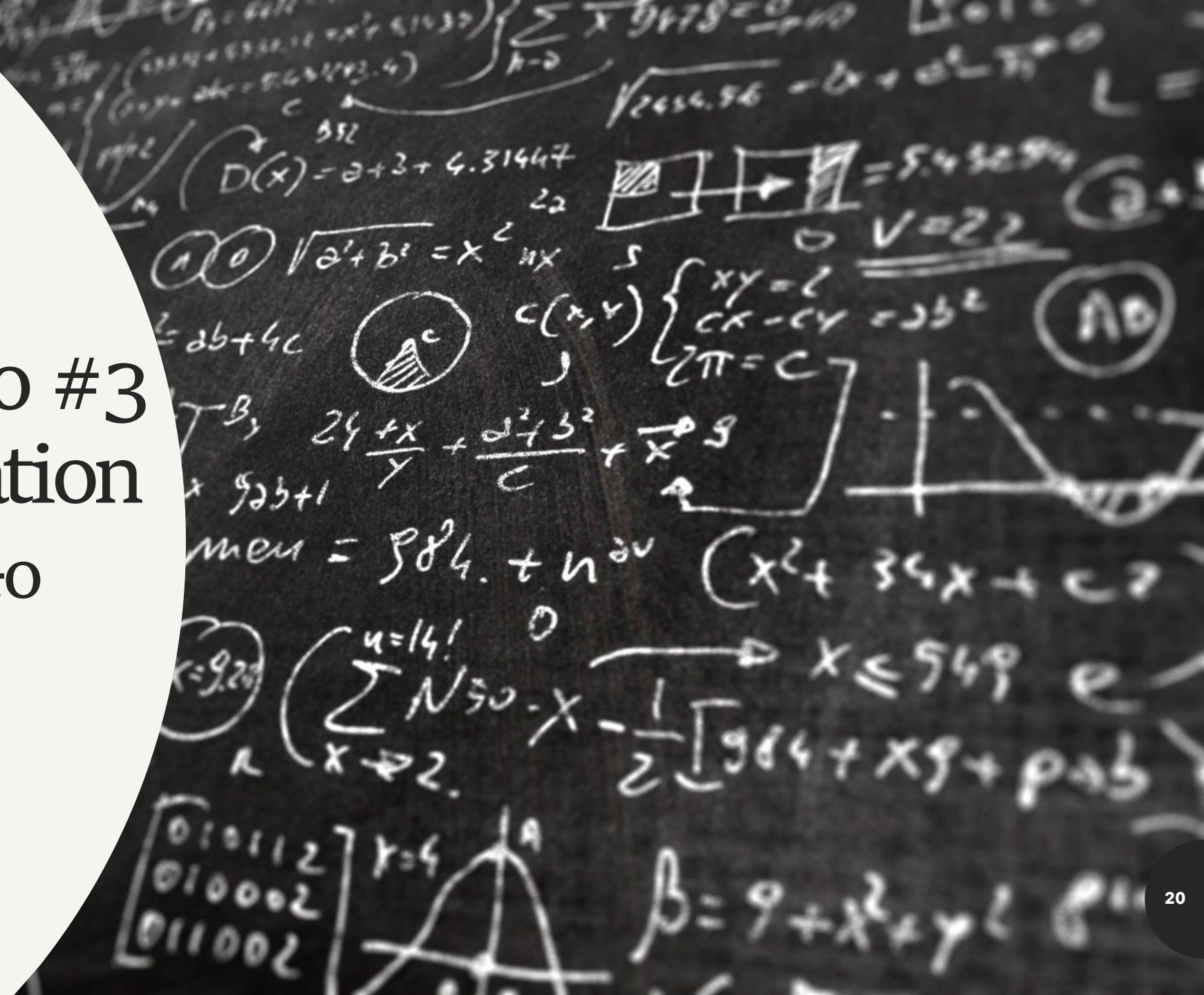
ChatGPT for subtitle translation

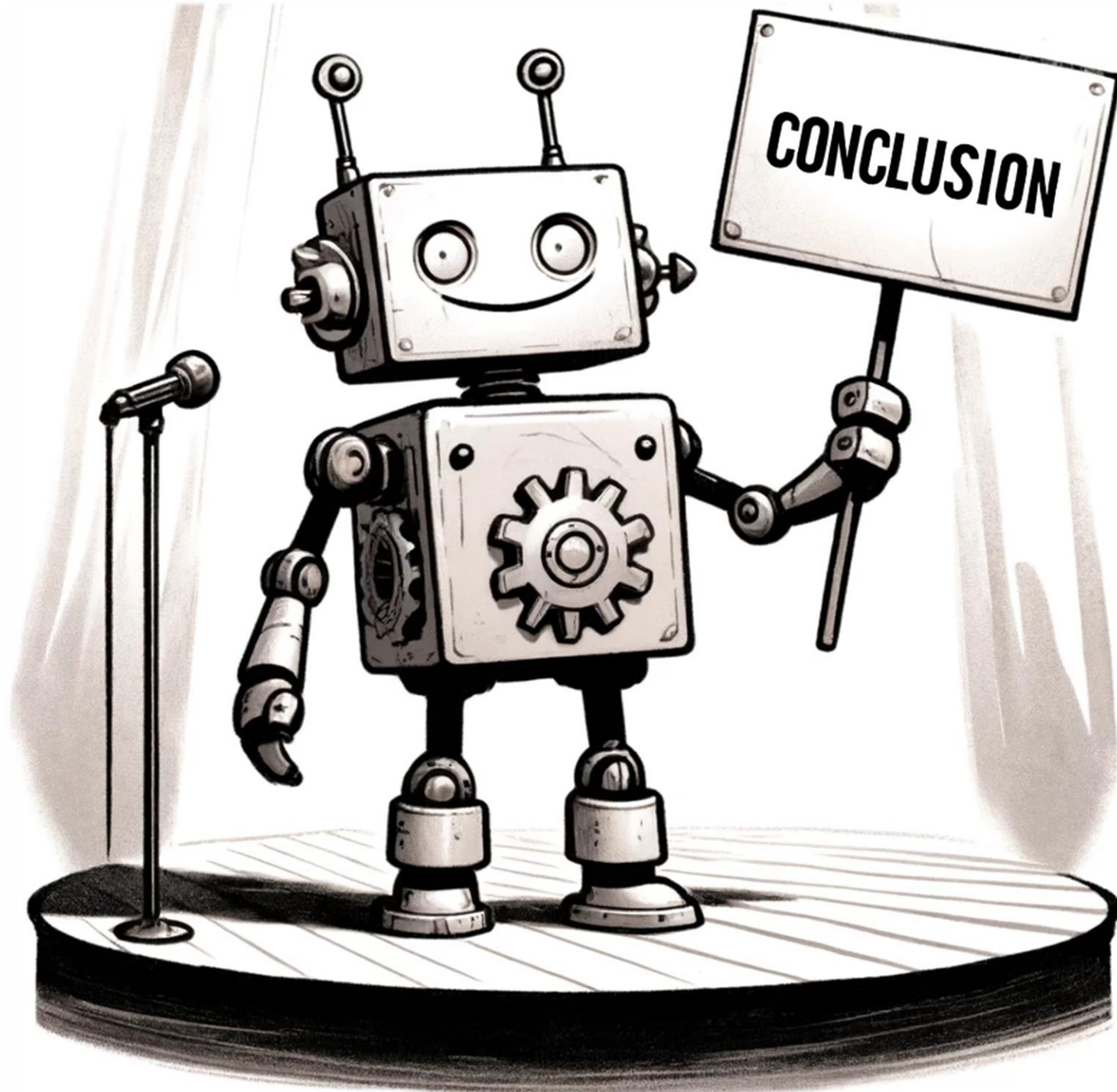
- It costs a bit of money since we have to use the API
- ChatGPT is a little disobedient 🐒 when asking specific output formats, but retrying calls can fix that
- + We can translate into any language we want



Technical Demo #3

Subtitle translation with ChatGPT4o

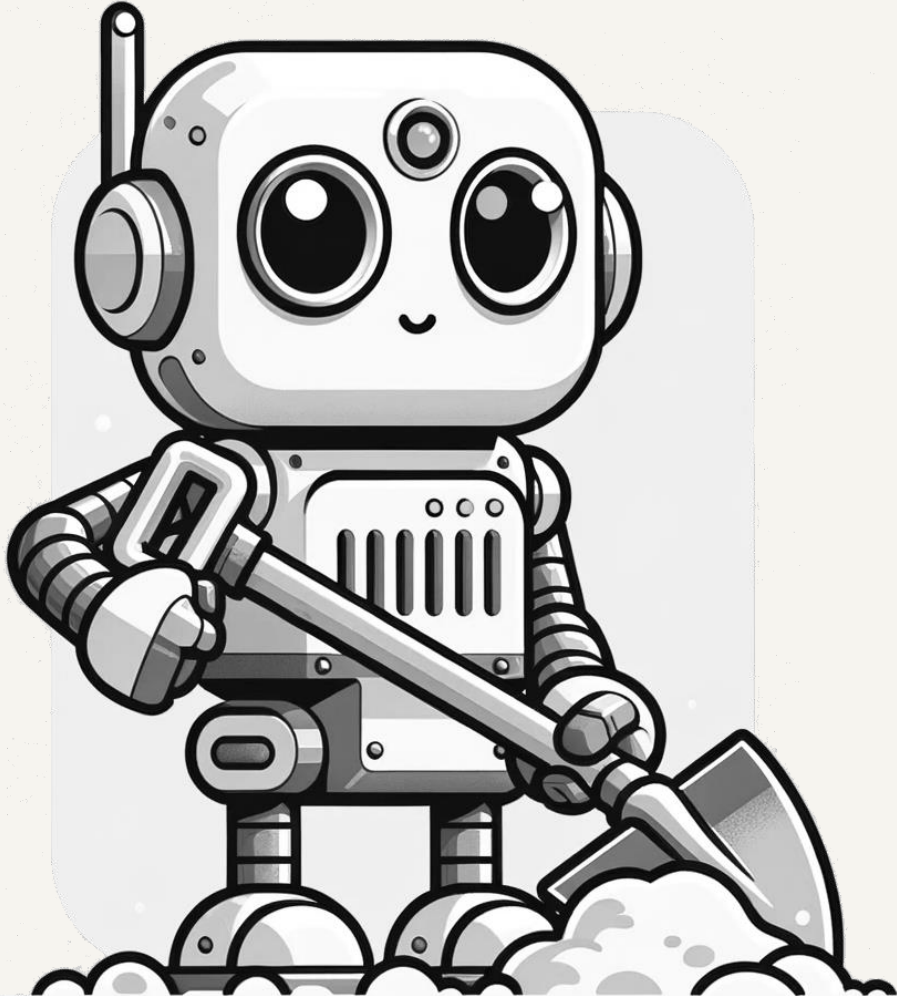


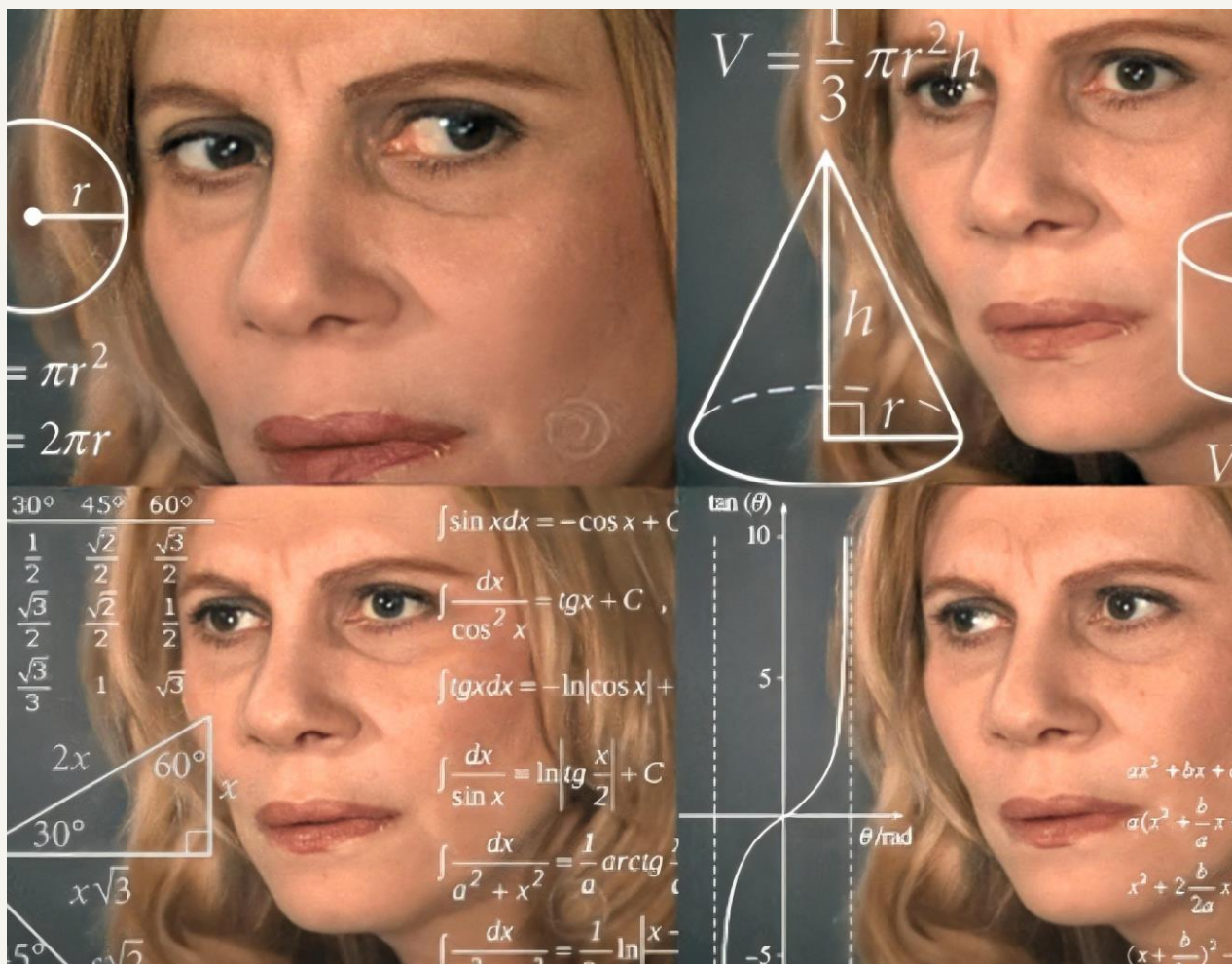


Scratching the surface

There is so much we did not talk about! AI offers so much more when it comes to language learning:

- Live conversations with **AI assistants**
- **Translation** of comic **books**, texts etc. in real time with **AR**
- **Dubs** generation with services like [ElevenLabs](#)
- **Idiomatic** and customized translation from and to many languages
- Thorough and clear **linguistic explanations** from language models like ChatGPT (e.g. German cases)





Q & A

The End

Thanks for your attention 🙏!