AUDIO PODCAST GENEROVANÝ AI

MIROSLAV SUCHÝ

@mirek@rodina-sucha.cz

Cíl: 5 minutový podcast

Cíl: 5 minutový podcast

každý den

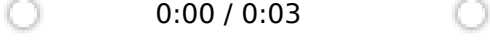
Cíl: 5 minutový podcast

každý den

za méně než minutu.

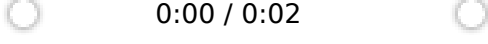
```
$ dnf install festival
```

```
$ echo Zdravim ucastniky OpenAltu | festival --tts
```



```
$ dnf install espeak-ng
```

```
$ espeak-ng -v cs+f1 'Zdravím účastníky OpenAltu'
```



ElevenLabs.io

0:00 / 0:02

```python
from elevenlabs.client import ElevenLabs
from elevenlabs.play import play


elevenlabs = ElevenLabs()

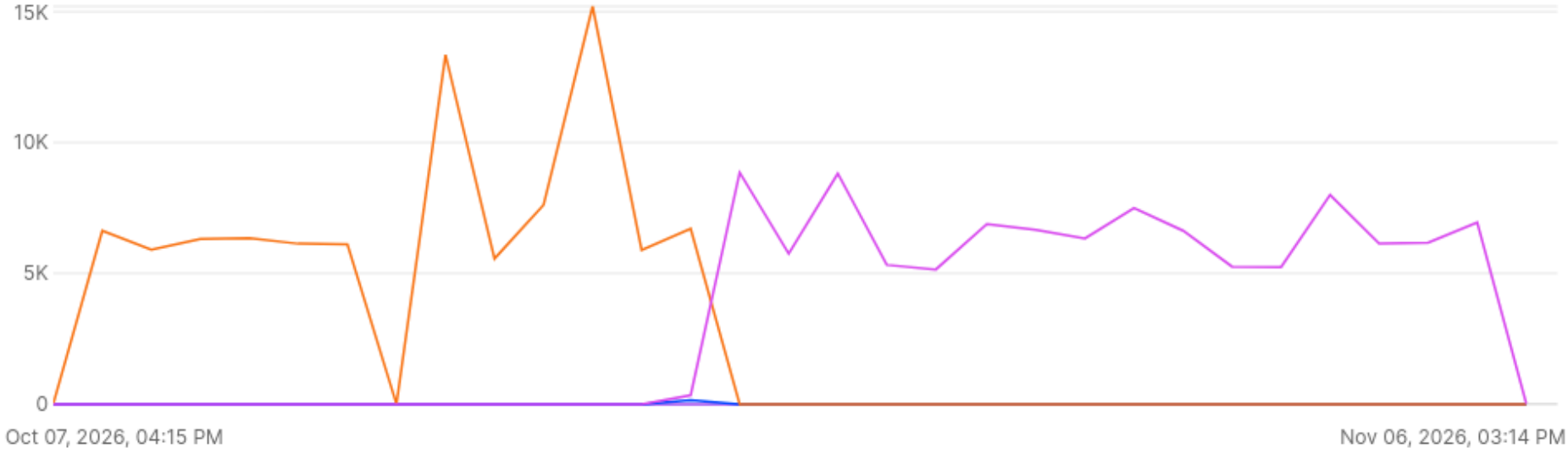
audio = elevenlabs.text_to_speech.convert(
    text="Zdravím účastníky OpenAltu",
    voice_id="JBFqnCBsd6RMkjVDRZzb",
    model_id="eleven_v4",
    output_format="mp3_44100_128",
)

play(audio)
```

$0.08 / 1k characters (cca minuta)

(některé hlasy mají příplatek)

Usage	Character Usage	Total Duration	Average Cost per Request	Billable Request Count
198K credits	198K	4.1 h	394 credits	502



Plusy:

- rychlost - 5 minut za 30 sekund!
- divná jména nejsou problém (demo)
- Hodně hlasů

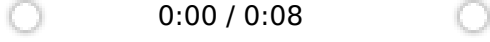
Mínusy:

- Cena
- Hodně hlasů
- Dříve číslovky a zkratky - ale záleží na hlasui (demo)
- zkratky: DPMB, např.

- divná jména

0:00 / 0:11

Mezi mrtvými jsou dva čtyřiačtyřicetiletí hasiči Denys Konovalenko a Oleksij Suščevskyj a třiapadesátiletý vojenský medik Oleksandr Zarjaščuk ze Žytomyru.

- výslovnost jmén 

Zelenskyj si ve středu telefonoval devadesát minut s americkými vyjednavači Stevem Witkoffem a Jaredem Kushnerem.

- číslovky (dnes už ne) 0:00 / 0:05

Celkem vyslalo v noci Rusko 465 dronů a 30 raket.

Registrace do podcastových služeb (by H. Javorek)

SSML TAGY (SPEECH SYNTHESIS MARKUP LANGUAGE)

XML značky obalující text, které říkají modelu, jak ho vyslovit

```
<speak>
  Hmm, give me a few seconds. <break time="2.0s
</speak>
```

Název značky
co udělat: vložit pauzu

Samouzavírací
nebo uzavřete každou otevřenou značku

STANDARD, ALE NE JEDNOTNÝ

Standard W3C, který každý poskytovatel implementuje jinak

Skript fungující v jednom SSML enginu nemusí fungovat stejně v jiném

```
<speak>
  Please welcome
  <phoneme alphabet="ipa" ph="ʃɪˈvɔːn">Siobhan</phoneme>
  from
  <phoneme alphabet="cmu-arpabet" ph="M AE1 D IH0 S AH0 N">Madiso
</speak>
```

Vlastní model? TTS Arena

Kokoro

Kokoro

jenže nemá čestinu

Kokoro

Jenže nemá čestinu.

Jazyky jsou řešené projektem Misaki

Konveze textu na fonémy (např. [Opənˈɑlt])

Podpora češtiny PR#91

Shrnutí:

- Generovat podcast jde.
- ElevenLabs je jednoduché, ale drahé
- Kokoro potřebuje dotáhnout podporu češtiny.

miroslav.suchy.cz/presentations/OpenAlt-2026-podcast