

Pohled tvůrců, implementátorů
a provozovatelů



Ne náhrada, ale
nová příležitost

Přepis mluveného slova do textové podoby nápad z roku 2000



Největší komerční archiv plných textů v ČR

- *Unikátní patentovaný systém zpracování dat, (společnost založena v roce 1995)*
- *Multimediální archiv pro vydavatele a vysílatele (Mafra, Ringier, TV Nova, ČRo, PSP,...)*
- *Měsíčně zpracováváme 8000 hodin záznamu, 600 hodin doslovně přepisujeme*
- *Okamžitá úspora času (30-75%)*
- ***Spolupráce s Technickou univerzitou v Liberci (od roku 2005):***
 - *Propojení vývojových týmů, sdílení výsledků a zkušeností, zpětná vazba*
 - *Spolufinancování vývoje technologie **Voice to Text (V2T)***
 - *Samostatné nadstavbové moduly, nové aplikace, servis, komunikace s klienty*

**Komunikace řečí – pro člověka přirozená věc,
pro počítač velmi komplikovaná úloha**

Systém pro rozpoznávání hlasu se stále učí

1. HLAS – Variabilita hlasového projevu, každý člověk mluví jinak

(výška hlasu, podoba hlásek, intonace, výslovnost, volba slov, momentální stav, prostředí ...)

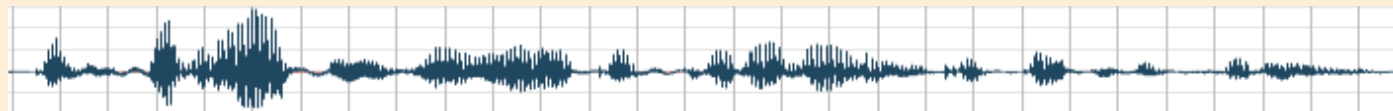
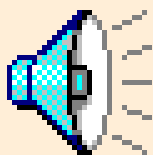
2. SLOVNÍK – Rozsáhlost slovní zásoby (Angličtina 50 tis. slov, Čeština > 1 milion)

Systém rozpoznává jednotlivá slova, mnoho stejně znějících slov („byly“, „bily“, „Billy“ ...),
250 000 příjmení v ČR, místopisné názvy apod.

V současnosti systém rozpozná 500.000 slov, slovo které není ve slovníku systém nikdy nepřepíše

3. KONTEXT - Volba nejpravděpodobnější sekvence po sobě jdoucích slov (odlišná témata)

chybí oddělovače (pauzy) mezi slovy, dochází ke koartikulaci, asimilaci,



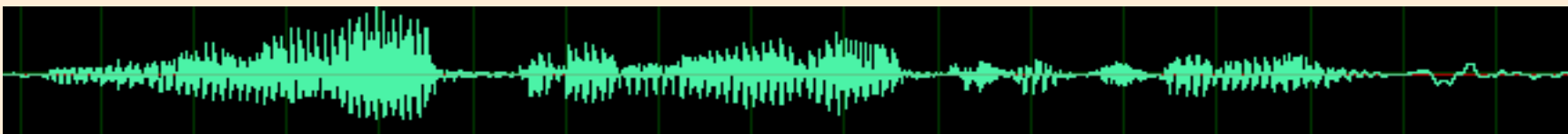
Text: *Až se vrátíš, zavolej jako včera na linku pět šest sedm.*

Text bez mezer: *ažsevrátíšzavolejjakovčeranalinkupětšestsedm*

Foneticky: *ašsevrátížzavolejakofčeranalinkupěčsedum*

4. PROSTŘEDÍ - Vliv prostředí a přenosové cesty – degradace signálu

hluk okolí, vlastnosti přenosové linky (mikrofon/diktafon), další mluvící osoby, ruchy, šumy...



Co bylo řečeno a v jakých časových okamžicích?

| - | zavolá | - | premiéra | předseda | - |

| - | za | vola | premiéra | předsedá | - |

| - | zvolá | prima | éra | před | sebou | - |

| - | zavolej | premiéra | předsedo | - |

| - | zavolá | - | premiér | a | předseda | - |

Při rozpoznávání spojitě řeči není apriori známo:

- Kolik slov bylo řečeno?
- Jaká sekvence slov byla řečena?
- Byla všechna vyřčená slova ze slovníku?
- Byla to skutečně jenom řeč nebo i další zvuky a hluky?
- V jakých časových okamžicích začínala jednotlivá slova?

První AI produkty 2005-2008

MyVoice a MyDictate



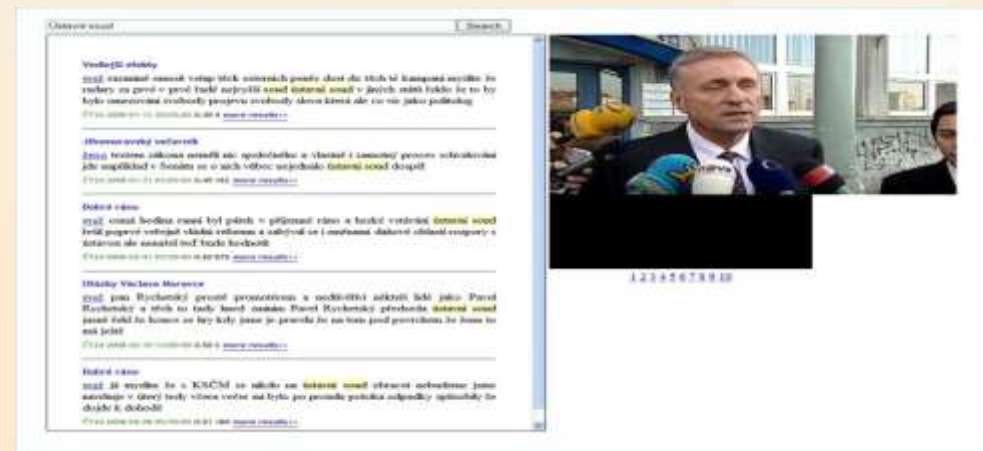
NewtonDictate – diktovací program



ATT – přepis a titulkování TV



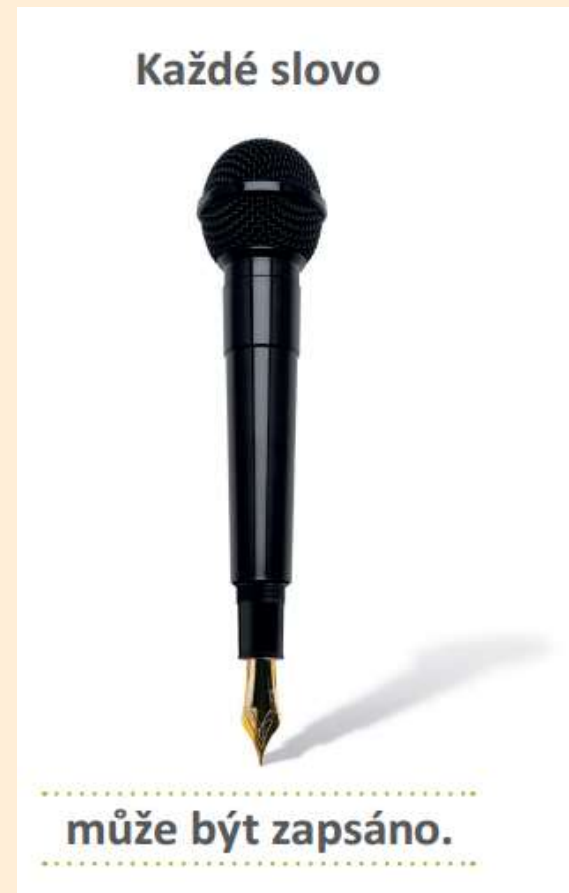
AudioSearch – vyhledávání v archivech



Projekt Ministerstva spravedlnosti ČR

„Racionalizace práce a výkonnosti soudů a státních zastupitelství prostřednictvím digitalizace administrativních procesů“

19.12.2008	<i>Smlouva o zahájení testovacího provozu na 6 měsíců</i>
10.12. 2010	<i>Žádost MSp na realizaci projektu schválena OSF MV ČR</i>
15.3. 2011	<i>Začátek soutěže na dodavatele systému</i>
3.5.2013	<i>Podepsaná smlouva</i>
10.5.2013	<i>První školení v Praze (2. účastníci)</i>
28.5.2013	<i>100. uživatel</i>
8.8.2013	<i>600. uživatel</i>
17.10.2013	<i>860 uživatelů - Závěrečná konference</i>



End to End modely

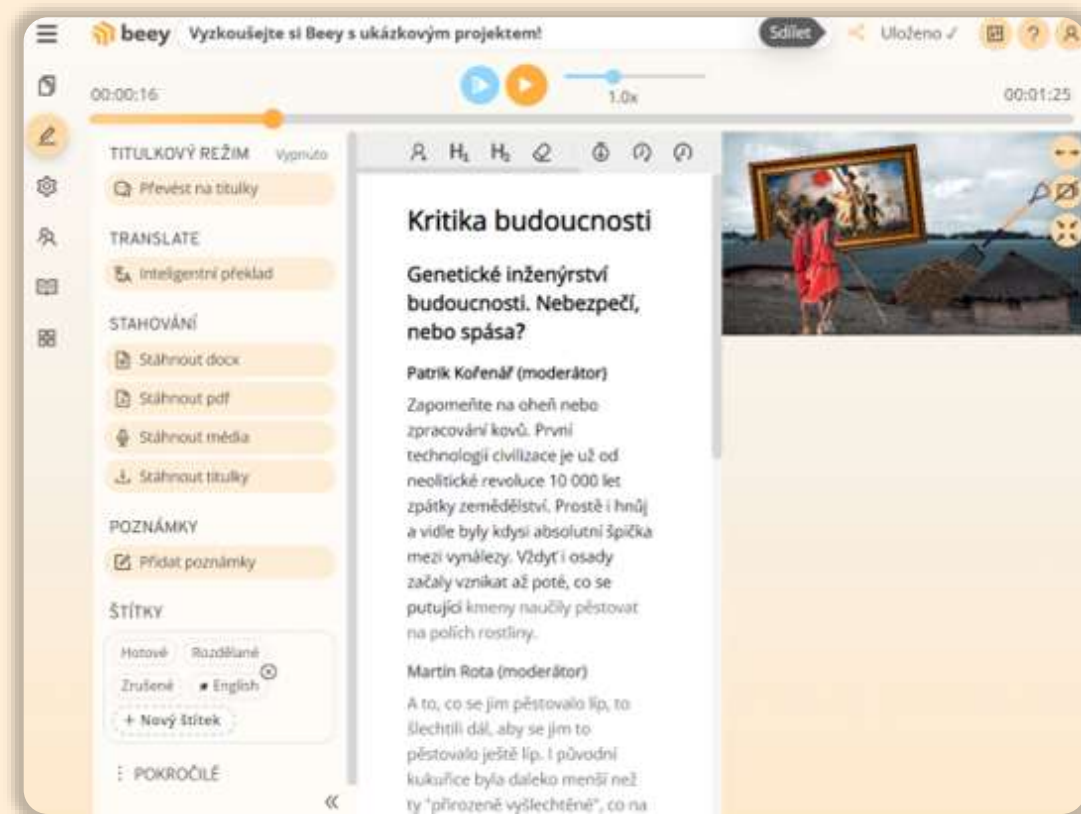
- 2015–2017 revoluce ve strojovém učení, přechod na neuronové sítě, úspěšnost se skokově zlepšila.
- Hluboká neuronová síť je mechanismus pro lenochy, kteří nechtějí složitě počítat matematické modely, ale chtějí, aby to za ně udělal počítač předložením spousty příkladů.
- Přístup sequence-to-sequence – je velmi univerzální, dá se nasadit na jakoukoliv kombinaci dat (nejen text-text, ale i zvuk-text, video-text, atd.).

Komplexní systém pro převod řeči do textu, rychlou editaci,
fulltextové vyhledávání, práci v týmu, titulkování, překlady, anotace,
sumarizace, další (nejen) AI funkce.



Editor usnadňuje zpracování a editaci textu

- Nejlepší automatický přepis
- Identifikace mluvčích
- Detekce konce vět
- Rychlá editace
- Možnost použití klávesových zkratk
- Fulltextové vyhledávání
- Úprava a časování titulků
- Snadná implementace (embedding)



Co ještě systém umí

- 🔗 **Link** přepíše videa a nahrávky přímo z webu
- 🔗 **Publish** publikuje audio/video přepis (open source knihovna)
- 🔗 **Splitter** rozdělí a zpracuje i velmi dlouhé nahrávky
- 🔗 **Stream** přepisuje a titulkuje živé vysílání z jednání nebo z televize
- 🔗 **Voice** nahraje a přepíše diktovaný text
- 🔗 **ChannelMerger** přepisuje více kanálové nahrávky do jednoho výstupu
- 🔗 **Summarize** automaticky z přepisu vytvoří krátké shrnutí
- 🔗 **DSS Converter** přepíše soubor ve formátu DSS
- 🔗 **ChannelMixer** Mono/Stereo poskytne inteligentní přepis mono/stereo nahrávek
- 🔗 **Live** titulkuje projev v reálném čase





**Multimediální komunikační
platforma pro znevýhodněné**

Základní informace



Webová platforma využívající nástroje na bázi umělé inteligence (artificial intelligence, dále jen AI) pro rychlou a nezkreslenou distribuci informací týkajících se COVID-19 mezi občany znevýhodněných skupin obyvatel ČR (např. sluchově postižených), seniorů a cizinců.

Přístupnost

Zpřístupnění informací pro znevýhodněné skupiny



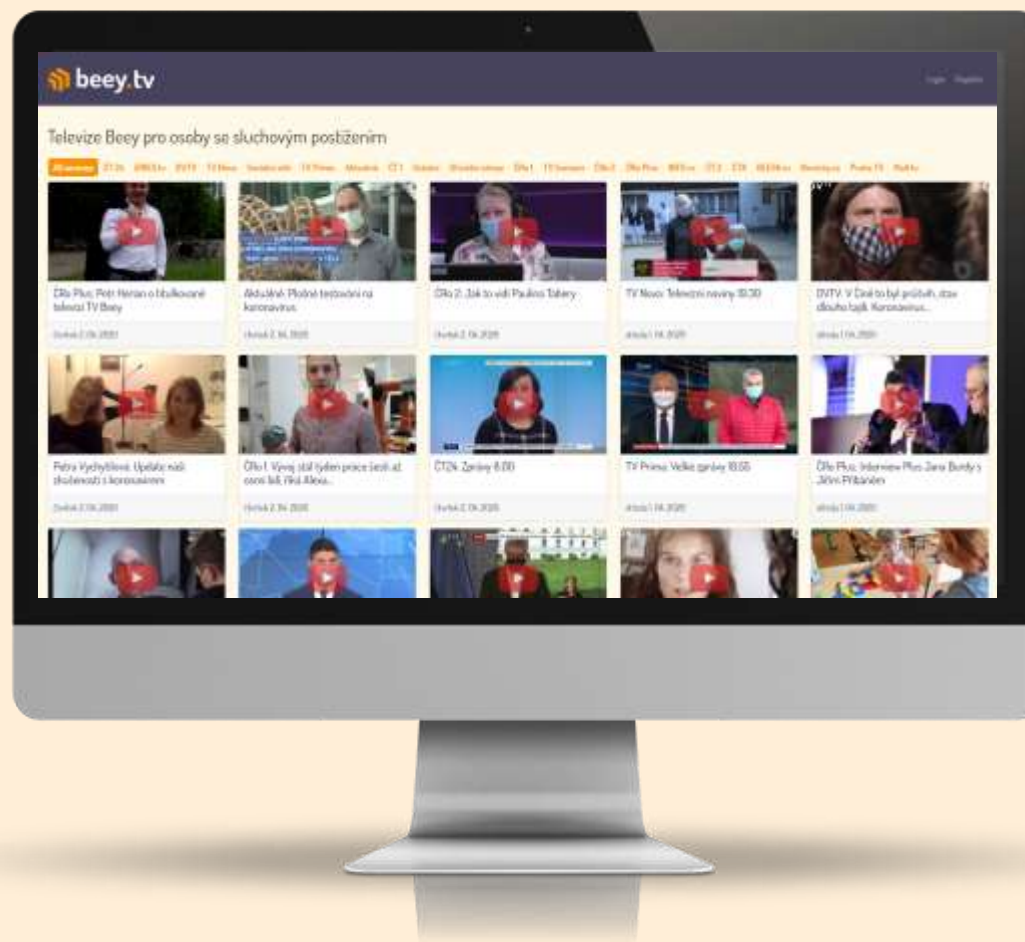
Nejlepší Kvalita

tuzemská technologie pro rozpoznání řeči



Dostupnost

Rozhraní optimalizované i pro **starší zařízení**.



Dosah

Velký dosah díky distribuci na **sociálních platformách**



Univerzálnost

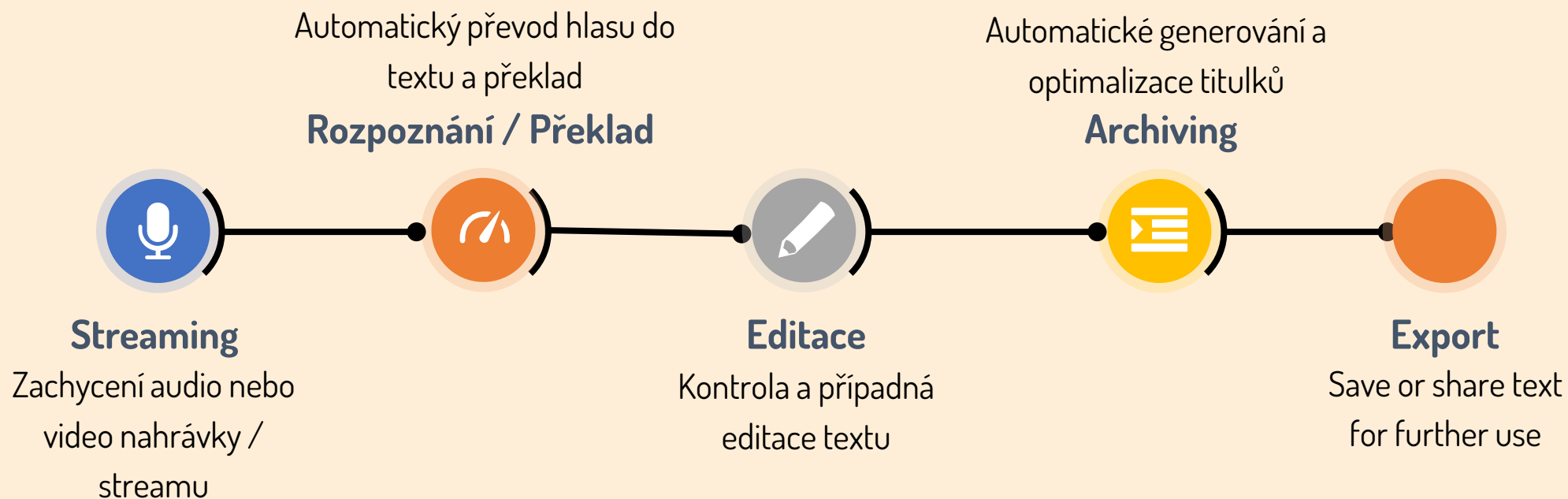
Přestože je řešení vyvíjeno pro **COVID-19**, je použitelné univerzálně



Rychlost

Okamžité automatické zpracování a publikování klíčových informací

Beey.tv
Workflow



Využívání platformy v prvních dnech po spuštění

Beey.tv



320 pořadů s titulky

Od startu platformy v neděli 15/3 jsme otitulkovali více než 300 pořadů.



7 150 aktivních uživatelů

Máme neustále rostoucí komunitu aktivních uživatelů.



590 hodin sledování

Už v rámci testovacího provozu naši diváci shlédli stovky hodiny otitulkovaných pořadů a videí.



25 zdrojů zpráv

Pravidelně publikujeme otitulkované pořady z více než dvou desítek různých zdrojů.

10 000

Aktivní uživatelé platformy

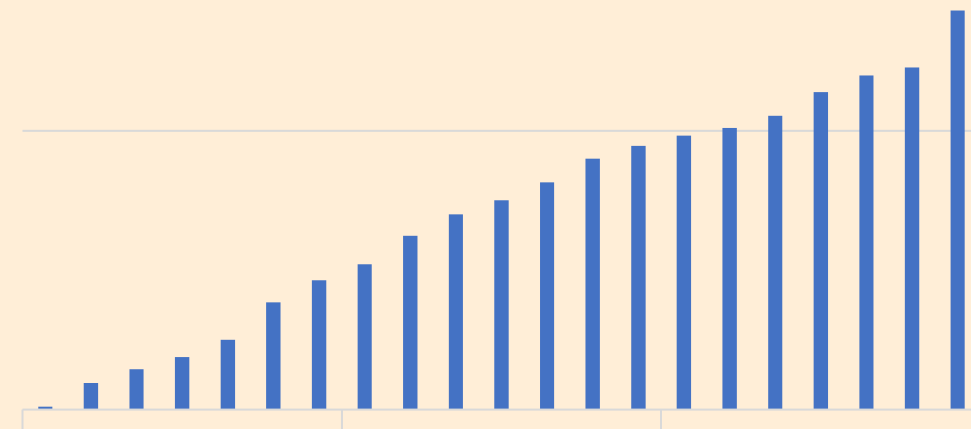
5 000

0

15.3

22.3

29.3



Co řekli o projektu



“

Chci moc poděkovat za titulkovanou internetovou televizi. Hodně lidí z naší komunity si ji chválí. Opravdu nám titulkování pomáhá.

Ing. Martin Novák

Ředitel České unie neslyšících

”



“

Konečně se mohou lidé se sluchovým postižením seznámit i s obsahem, který nám byl až do teď nepřístupný. Snad se zapojí i další média.

Mgr. Věra Strnadová

Komise pro skryté titulky
ASNEP

”



“

Pro půl milionu osob s postižením sluchu má sledování videa bez titulků stejnou hodnotu, jako kdyby se slyšící diváci dívali na televizi bez zvuku.

Jaroslav Winter

expert na informační
technologie pro
handicapované

”

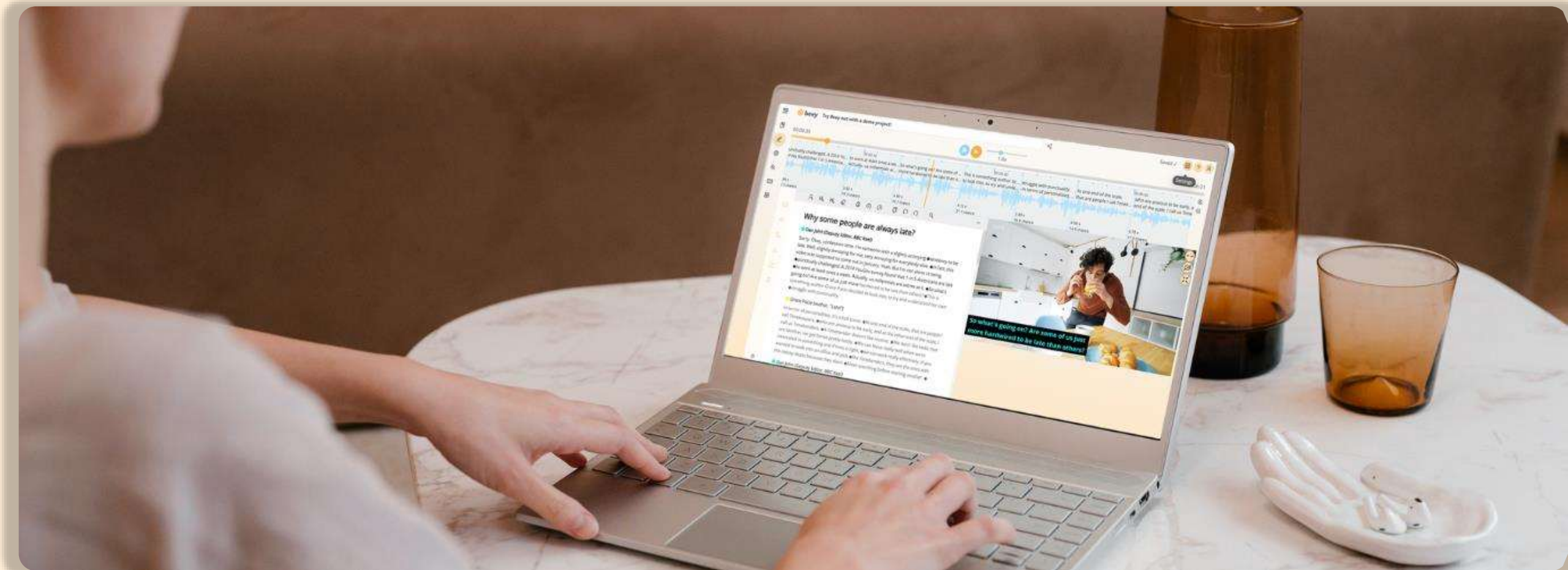
Skryté Titulky pro televize a webové platformy

- 🔧 Vysoká kvalita přepisu televizních pořadů
- 🔧 Podpora SLT (skryté a cizojazyčné titulky)
- 🔧 Integrace do webových stránek
- 🔧 Fulltextové vyhledávání, zvýšení SEO

Česká televize, TV NOVA, CNN Prima NEWS, DTV, Seznam.cz, přepisuje nahrávky pro **Český rozhlas, Polský rozhlas, Mafra, CN Center, rádio TOK FM**, a další.



Titulkování videí na sociálních sítích



Proč titulkovat videa na sociálních sítích

- 🟡 Titulky zvyšují dostupnost a zapojení uživatelů, zejména pro osoby se sluchovým postižením.
- 🟡 Titulky zajišťují, že všichni uživatelé mají stejný přístup k obsahu bez ohledu na jejich schopnosti.
- 🟡 Přístupný obsah je přehlednější a srozumitelnější pro všechny uživatele, což vede k lepšímu zapojení a delšímu sledování videí.



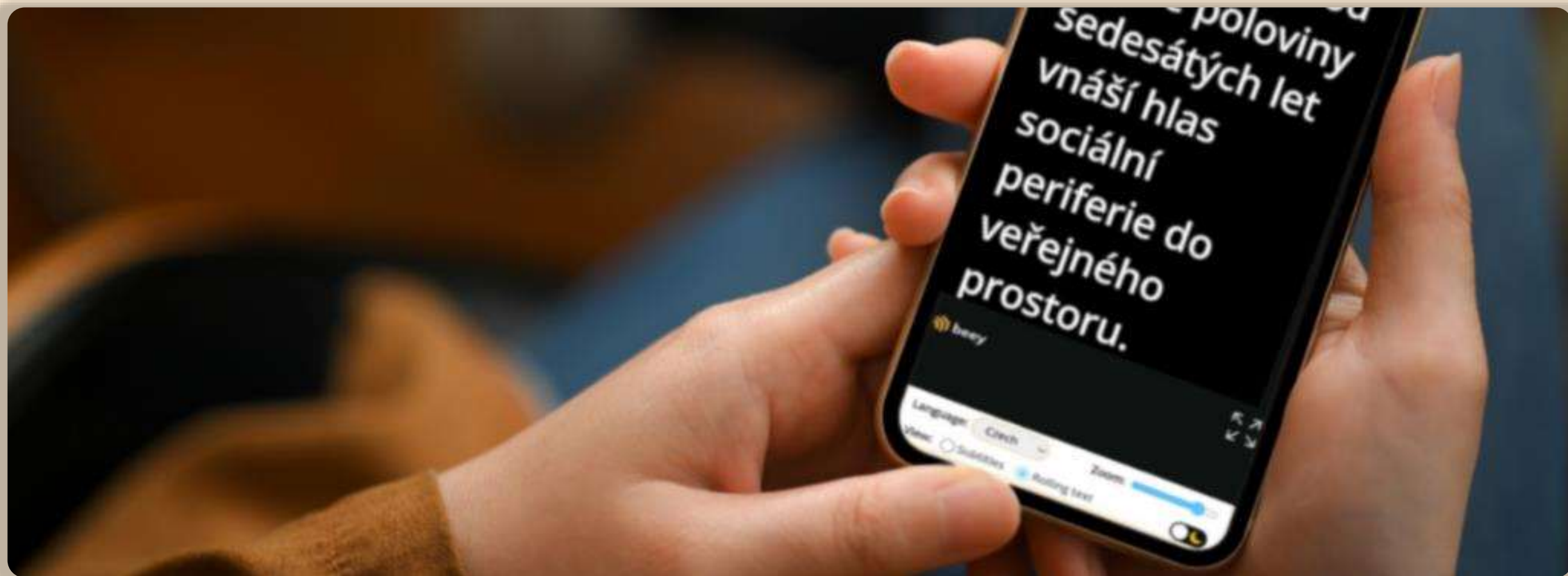
Statistiky a fakta

- 📊 **85 %** uživatelů sleduje videa na Facebooku bez zvuku.
- 📊 **93 %** videí na Facebooku je sledováno na mobilních zařízeních
- 📊 Titulky mohou zvýšit průměrný čas sledování videa až o **12 %**.
- 📊 Videa s titulky mají tendenci získávat více zhlédnutí a interakcí, což je klíčové pro úspěch na sociálních sítích.

Zdroje: Why America is obsessed with subtitles (preply.com), mediaguru.cz



AI asistent pro živé titulkování



Živě titulujeme mluvený projev

- System průběžně zaznamenává vše, co na vaší akci řeknete do mikrofону.
- Živý záznam můžete okamžitě sledovat na mobilu a kompletní text je uložen k dalšímu použití.
- Konference, přednášky, školení, veřejné slavnosti nebo tiskové konference se okamžitě přepisují do textu a generují titulky do mobilních zařízení účastníků.



Kde vám **živý přepis** pomůže

Konference

Využijte AI asistenta k otitulkování profesionálních událostí a vyjděte vstříc všem hostům

Školení

Zaznamenejte každé slovo a nic vám neuteče, ani během pauzy

Slavnostní a veřejné akce

Zajistěte, že vašemu programu bude rozumět úplně každý – i neslyšící

Tiskové konference

Umožněte novinářům získat textový přepis ihned po skončení akce

Přednášky

Usnadněte návštěvníkům porozumět termínům a orientovat se v debatách

Přístupnost v CAMPu

Centrum **A**rchitektury a **M**ěstského **P**lánování

Vybrané akce jsou simultánně přepisované do českého jazyka, tlumočené do českého znakového jazyka, případně opatřené českými titulky.

K dispozici je vám k zapůjčení mechanický vozík. Můžete využít přebalovací pult. Vítáni jsou i vaši asistenční psi.



Zpřístupnění bez bariér

Přístupnost díky AI = bez limitu počtu jazyků,
bez závislosti na lidech, bez bariér

Důležitá není jen dostupnost, ale také
dostupnost dostupnosti,

Lidský tlumočník má omezenou kapacitu
– a i **dobré úmysly někdy ztroskotají na rozpočtu.**

Zpřístupňujeme obsah i tam, kde by jinak
prostě nebyl.



LLM modely – zajímavosti

- LLM není vyhledávač! Nehledá weby, hledá pravděpodobné následující slovo. Nejdříve vygeneruje výsledek a pak hledá, kde to vzal.
- LLM nemá model sebe – neví, co ví, neví, jaké má schopnosti.
- LLM přemýšlí, jakým následujícím slovem nás nejlépe uspokojí.
- OpenAI upravila trénovací data do formy otázka + odpověď, proto je model tak dobrý v konverzacích
- Když model přizná, že něco neví, znamená to, že je tam ex-post cenzorní zásah.
- Nemáme v dnešní době dost dat, aby se modely ještě zlepšovaly – na světě prostě není dost dat.
- Už teď jsou LLM trénované z celého internetu.

Děkuji za pozornost



petr.herian@newtontech.cz