

Automatická transkripčia archívneho fondu Oblastná úradovňa Osídľovacieho úradu a Fondu národnej obnovy v Košiciach, 1945 – 1951

APVV-24-0049 – Digitalizácia a
inteligentná obsahová analýza
historických dokumentov pre
odbornú archívnu činnosť a
historickú vedu

Workshop riešiteľov projektu

10. september 2026





Transkribus vs. veľké jazykové modely

- **Počiatočný predpoklad:**
automatická transkripcia
zvolených archívnych fondov
prostredníctvom platformy
Transkribus

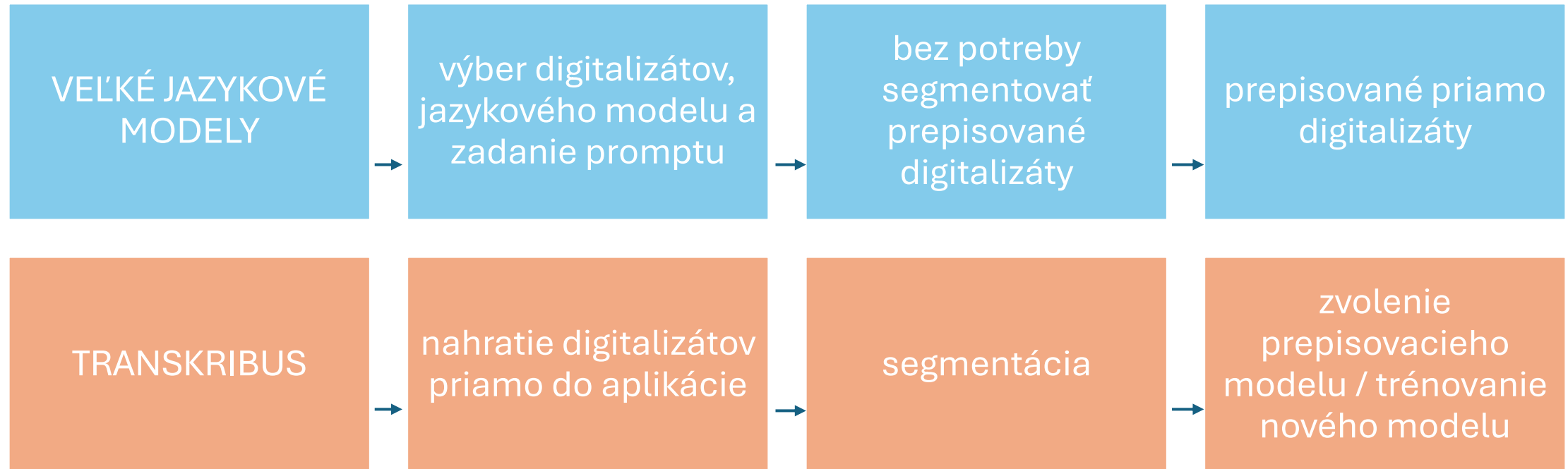
Experiment:

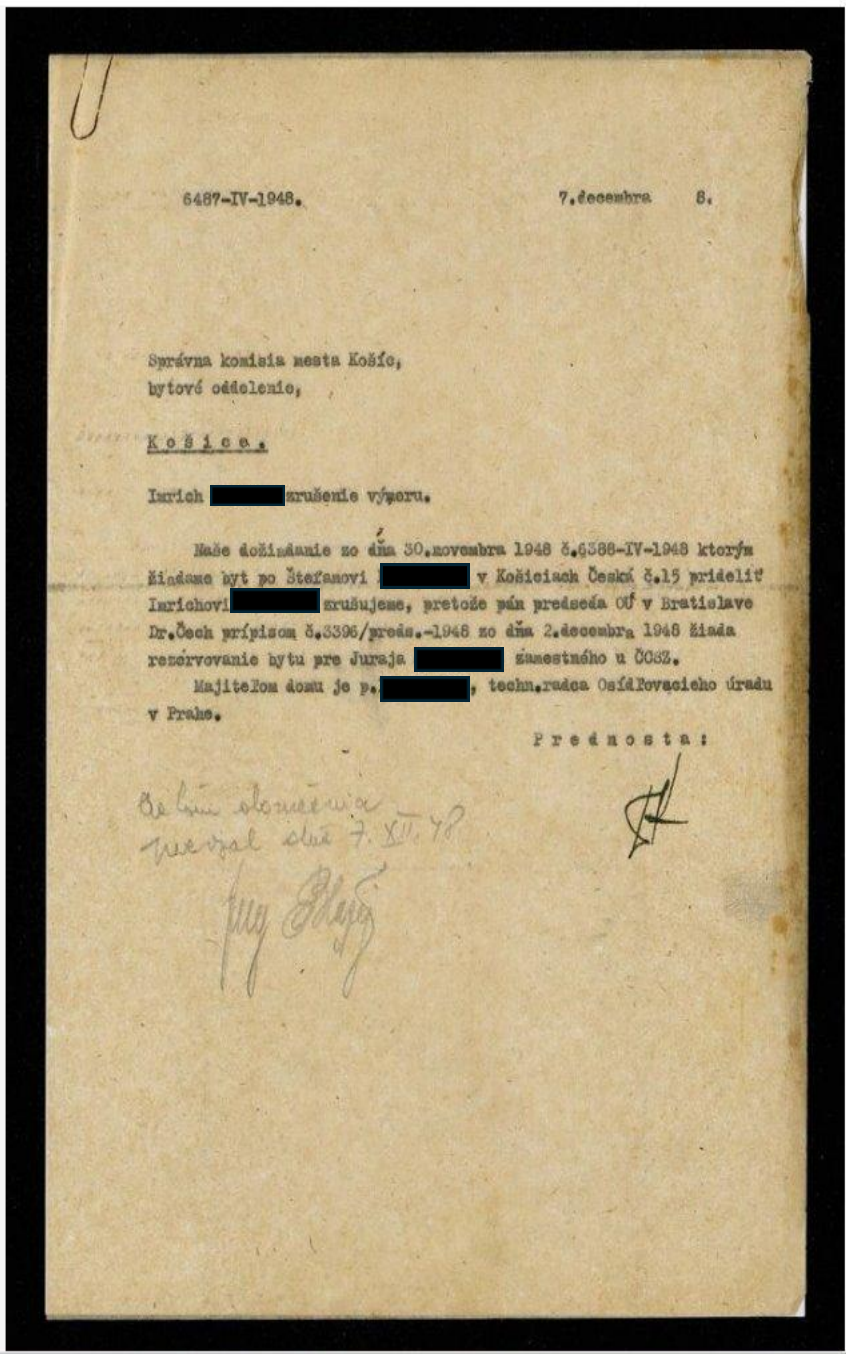
overiť a komparovať možnosti automatickej transkripcie historických prameňov prostredníctvom platformy Transkribus a veľkých jazykových modelov

- **náročnosť automatickej transkripcie a doba jej trvania**
- **chybovosť prepisov vyprodukovaných prostredníctvom automatickej transkripcie**



Proces automatickej transkripcie





Region 1

- 1 SslaN219486.
- 2 7. decembra 8.
- 3 Správna komisia mesta Košíc,
- 4 bytové oddelenie,
- 5 Košice:
- 6 Imrich [redacted] zrušenie výperu.
- 7 Naše dožiadanie zo dňa 30. novembra 1948 č. 6388-IV-1946 ktorým
- 8 žiadame byt po Štefanovi [redacted] v Košiciach Česká č. 15 pridelit'
- 9 Imrichovi Kozákovi zrušujeme, pretože pán predseda OÚ v Bratislave
- 10 Dr. Čech prípisom č. 3396/preds.-1948 zo dňa 2. decembra 1948 žiada
- 11 rezorvovanie bytu pre Juraja [redacted] zamestného u ČOSZ.
- 12 Majiteľom domu je [redacted], techn. radca Osídľovacieho úradu
- 13 v Prahe.
- 14 Prednosta:
- 15 Boka Sorstoice
- 16 Čpraso B4. 2. 0.7.
- 17 r. o.

Predmet : [REDAKTOVANÉ] Imrich, net.taj. pri SMÚ. Košice,
príkazanie bytu na Českej ul.15, presťahovanie.

Na výmer čís.:65001/1396/1948
zo dňa 2.dec.1948.

Titl.

Správna komisia mesta
bytové oddelenie

v K o š i c i a ch.

Úctivo oznamujem, že na základe Vášho horecit.výmeru jako
aj v smysle dispozície OOÚ. v Košiciach zo dňa 30.nov.1948 číslo :
6388/-IV-1948 dňom 5.decembra 1948 som sa natrvalo presťahoval z do
terajšieho bytu- domu: Košice-Hviezdoslavová u.10 do domu resp.bytu
na Č e s k e j ul.15.

Ráčte toto oznámenie láskave v známost vziať už aj preto,
že mám osobnú vedomosť o pozmenení dispozícií zo strany OOÚ.Košice
pozatýmne vydaných a o ktorej som doposiaľ nebol vyrozumieť.Žiadam,
aby pri prípadnom meritornom vybavovaní pozmeňujúcej dispozícií bol
som vypočutý poťažne skutkový stav v známost vzatý.

V Košiciach, dňa 7.decembra 1948.

[REDAKTOVANÉ]
podpis zainteresovaného.



SK_1540_11913_12_0022.txt

Na výmer čís.:65001/1396/1948

zo dňa 2.dec.1948.

Titl.

Správna komisia mesta

bytové oddelenie

v K o š i c i a ch.

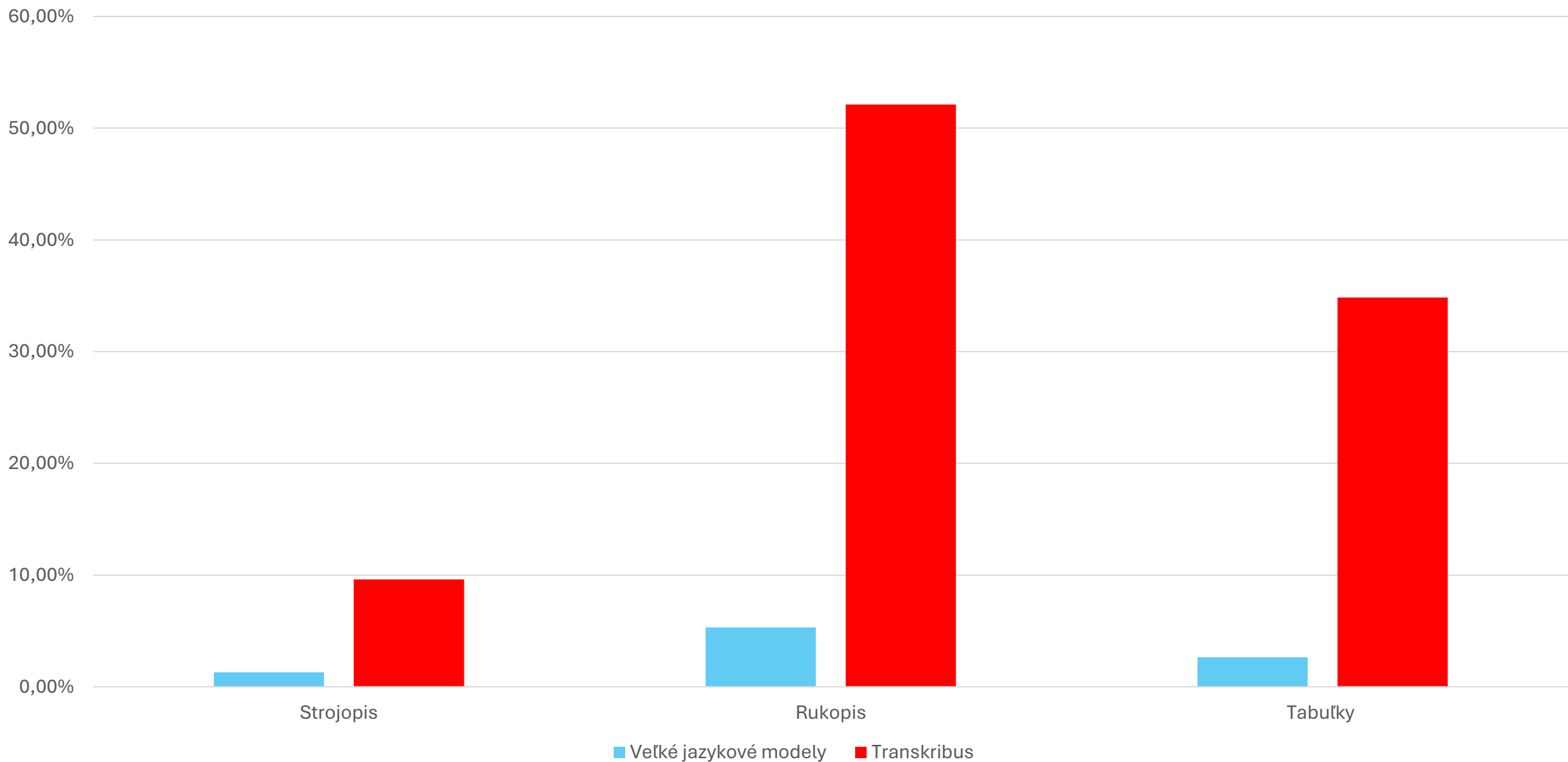
Úctivo oznamujem, že na základe Vášho horecit.výmeru jako
aj v smysle dispozície OOÚ. v Košiciach zo dňa 30.nov.1948 číslo :
6388/-IV-1948 dňom 5.decembra 1948 som sa natrvalo presťahoval z do
terajšieho bytu- domu: Košice-Hviezdoslavová u.10 do domu resp.bytu
na Č e s k e j ul.15.

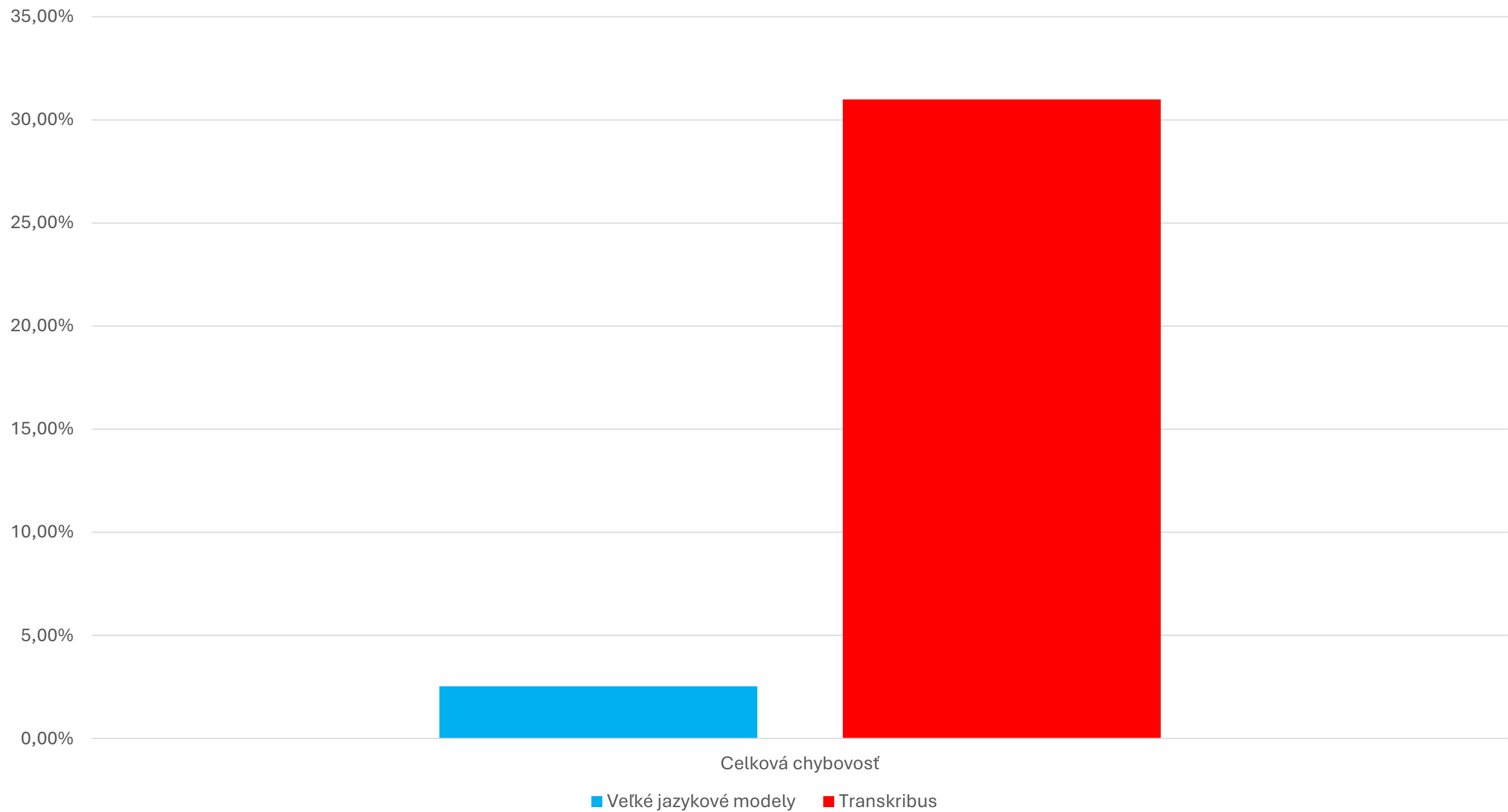
Ráčte toto oznámenie láskave v známost vziať už aj preto,
že mám osobnú vedomosť o pozmenení dispozícií zo strany OOÚ.Košice
pozatýmne vydaných a o ktorej som doposiaľ nebol vyrozumieť.Žiadam,
aby pri prípadnom meritornom vybavovaní pozmeňujúcej dispozícií bol
som vypočutý poťažne skutkový stav v známost vzatý.

V Košiciach, dňa 7.decembra 1948.

Imrich [REDAKTOVANÉ]
podpis zainteresovaného.

Percentuálne vyjadrenie chybovosti automatickej transkripcie







Transkribus vs. veľké jazykové modely



Transkribus:

- niekoľkonásobne časovo náročnejší postup
- nutná segmentácia všetkých digitalizátov (problém s
- tabuľkami)
- potreba trénovania prepisovacích modelov
- niekoľkonásobne vyššia chybovosť prepisov
- textové rozloženie dokumentu nie je po prepise zachované

Veľké jazykové modely:

- časová úspora
 - nie je potrebná segmentácia dokumentov (ani pri komplikovanejšom layoute)
 - automaticky kopíruje pôvodné textové rozloženie dokumentu
 - bez namáhavého trénovania prepisovacích modelov
 - diametrálne nižšia chybovosť
 - bezproblémová transkripcia aj v prípade rukopisov
- 

Atlas

Vyskúšajte prvý predtrénovaný model Smart Extract

DEMO · Atlas je najsilnejší v ťačených a štruktúrovaných materiáloch

SlovenčinaZmiešané

Skryté tagyZrušiť všetkoNové

Stanica SNB Stratená, okres Rožňava.

Stratená, 13. júl 1949.

Císlo: 630 / 1949.

Vec: Evidencia presídencov -
hlásenie.

Odpoveď na číslo: 5251-IS/1949
zo dňa ./.

Prílohy: ./.

K o š i c e .

Pokiaľ sa týka presídencov, týto sa v tunajšom príkázanom služob-
nom obvode stanice SNB Stratená, nenachádzajú vôbec.
Súčasne v tomto smere ani soznam nepredkladám iba len negatívne
hlásenie.

Voliteľ stanice noust.
J. Martišovič

15. 7. 49.

40/49

16. 7. 49

EA

▼ BLOK (0)

▼ ODSTAVEC (1)

Ďakujem

▼ BLOK (0)

▼ ODSTAVEC (10)

Stanica SNB Stratená, okres Rožňava.

Stratená , 13. júl 1949 .

Číslo: 630 / 1949.

Vec: Evidencia presídencov -
Hlásenie.

Odpoveď na číslo: 5251-IS/1949

zo dňa ./.

Prílohy: ./.

▼ BLOK (0)

▼ ODSTAVEC (4)

Oblasťná úradovná osídlovacieho
úradu pre Slovensko .

K o š i c e .

▼ BLOK (0)

▼ ODSTAVEC (2)

Pokiaľ sa týka presídencov, týto sa v tunajšom príkázanom služob-
nom obvode stanice SNB Stratená , nenachádzajú vôbec.

▼ ODSTAVEC (2)

Súčasne v tomto smere ani soznam nepredkladám iba len negatívne
Hlásenie.

+ -

DOKUMENT 30 riadkov · 105 slov Zmiešané Rozloženie SEBAVEDOMIE Celkovo 71% Štruktúra 81% Obsah 67% ŠTRUKTÚRA 6 blokov 7 odsekov 2 Podpisové značky

SUBJEKTY Organizácia Miesto Dátum/Čas Osoba