



Predstavitev govornih tehnologij za Slovenščino

Marko Bajec

27. 5. 2022



Univerza v Ljubljani
Fakulteta *za računalništvo
in informatiko*

VITASIS



**Laboratorij za
podatkovne tehnologije**
Laboratory for
data technologies



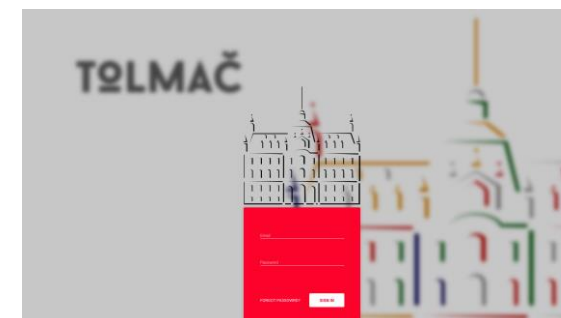
Govorne tehnologije

- Razpoznavalnik slovenskega govora (off-line in on-line)
- Prevajalnik Sl-En, En-Sl
- Sintetizator slovenskega govora



Reference

- Vitabits, **Vitasis**
SISTEM ZA ODDALJENO SPREMLJANJE IN TELEMEDICINSKO
OBRAVNAVO PACIENTOV, 2018
- Neo, **Telekom Slovenije**
TV SET TOP BOX
2019
- ON/Online Notes, **UL**
SISTEM ZA AVTOMATSKO PREVAJANJE V REALNEM ČASU
2019



Reference

- Samodejno podnaslavljanje TV programov, **Ministrstvo za kulturo, RTV**

PILOTNI PROJEKT SAMODEJNEGA PODNASLAVLJA TV PROGRAMOV, 2021

- Podnaslavljanje konference, **Državno odvetništvo**

ZAGOTAVLJANJE PODNAPISOV ZA MEDNARODNO KONFERENCO, 2021

- Uvedba strojne razpoznavne govora, **Državni zbor**

PILOTNA UVEDBA RAZPOZNAVALNIKA ZA POTREBE DRŽAVNEGA ZBORA, 2022-

- Razvoj Slovenščine v digitalnem okolju, **Ministrstvo za kulturo**

PRIPRAVA UČNE GOVORNE BAZE ZA RAZVOJ



© 2022 UL, FRU & VITASIS
verzijski: 1.0 | Hvalite v okviru MDO

Evaluator

Način primerjave: brez ločil brez velikih začetnic | Število znakov v vrstici: 60

Referenčna datoteka | Št. pojavnih: 280
Choose file: DZR_15.txt

Transkript | Št. pojavnih: 309
Choose file: DZR_15_transcript_nemo.txt

[Izračunaj napake](#) [Ponovno](#)

Metrika	Vrednost	Ločila	N	Natančnost	Prilike	F1
WER	0.99	plika	-	-	-	-
Št. zadržkov	256	veljica	-	-	-	-
Št. zametov	22	vpradej	-	-	-	-
Št. brisanj	1	klicaj	-	-	-	-
Št. duševanj	31	začetnica	-	-	-	-

Za izračun natančnosti ločil in velikih začetnic se mora upamati število pojavnih v referenčnem besedilu in transkriptu

[Napake - transkript](#) [Napake - ločila in začetnice](#)

Napake transkripta	Transkript
1 ne bvala predsednik splošni premijer @da	1 in našlepe hvala predsednik splošni premijer @da
2 koš dobro veste vai je minimalna plača eno od temeljnih	2 ko dobro veste vai je minimalna plača eno od temeljnih
3 temeljnih vprašanj a katerim se ukvarjamo in to	3 @@ temeljnih vprašanj a katerim se ukvarjamo in to
4 temeljno vprašanje je tudi kako nabojno se pravi dostojno	4 temeljno vprašanje je tudi kako nabojno se pravi dostojno
5 plačilo za opravljeno delo da stanja ki trenutno	5 plačilo za opravljeno delo da stanja ki trenutno
6 vlada v naši državi je pa popolnoma obratno vemo kakšno	6 vlada v naši državi je pa popolnoma obratno vemo kakšno
7 kakšno so popili v katerih ljudje ki prejmejo minimalno	7 kakšno so popili v katerih ljudje ki prejmejo minimalno
8 plačo preživljajo ali pa skušajo preživeti pa	8 plačo preživljajo ali pa skušajo preživeti pa
9 preživeti a takimi prejemki je zelo oteženo obhajajo	9 preživeti a takimi prejemki je zelo oteženo obhajajo
10 rene anomalije kot je razred j da tudi tistih	10 rene anomalije kot je razred j da tudi tistih
11 nekaj nad razredom ki ki so v bistvu pod minimalno plačo	11 nekaj nad razredom ki ki so v bistvu pod minimalno plačo
12 in to vse skupaj povzroča neko reno socialno stisko	12 in to vse skupaj povzroča neko reno socialno stisko
13 pri ljudeh ki so zaposleni in prejemajo nek nov fenomen	13 pri ljudeh ki so zaposleni in prejemajo nek nov fenomen
14 ki se mu reče revidni zaposleni a kot rešitev	14 ki se mu reče revidni zaposleni da kot rešitev
15 temu stanju smo v levici predlagali novoletno zakona o	15 temu stanju smo v levici predlagali novoletno zakona o
16 minimalni plači a treni javni rešitveni smo je dris	16 minimalni plači a treni javni rešitveni smo je dris
17 na sedemsto evrov neto do leta 2020 izločitev	17 na sedemsto evrov neto do leta dva tisoč dvajset izločitev
18 vseh dodatkov in minimalne plače in uvedbo do leta	18 vseh dodatkov in minimalne plače in uvedbo do leta
19 2021 tako izločitev formula ki bi	19 dva tisoč dvajsetev tako izločitev formula ki bi

RSDO - cilji

- Izdelati **govorno bazo** in **podporna orodja**, ki so ključna za razvoj razpoznavalnika slovenskega govora.
- Razviti en **splošen** in dva **specializirana** razpoznavalnika govora, ki bodo uporabnikom omogočali brezplačno izdelavo transkripcij za naložene zvočne datoteke.
- Opisati **postopek izdelave** splošnih in domenskih razpoznavalnikov.

Komercialne rešitve

- Truebar™ narekovalnik

DIKTIRANJE, RAZPOZNAVA V ŽIVO, GLASOVNO UPRAVLJANJE

- Truebar™ razpoznavalnik in urejevalnik

OFFLINE RAZPOZNAVA IN UREJANJE PREPISOV

- Truebar™ TV

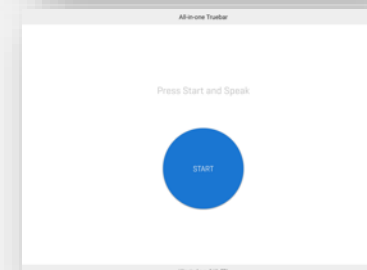
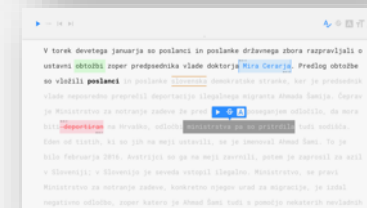
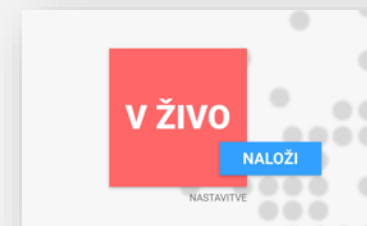
RAZPOZNAVA IN FORMIRANJE PODNAPISOV V ŽIVO

- Truebar™ klepetalnik

RAZPOZNAVA, FORMIRANJE ODGOVORA, SINTEZA ODGOVORA V REALNEM ČASU

- Truebar™ Zoom

PODNASLAVLJANJE V VIDEOKONFERENČNEM SISTEMU ZOOM



Ogrodja in tehnologije

- Razpoznavalnik:
 - Ogrodje Kaldi (NNet3, časovno zakasnjene NN)
 - Ogrodje NeMo (Jasper, QuartzNet, Citrinet, CTC-Conformer, Conformer-Transducer...)
 - ...
- Sintatizator:
 - Tacotron + WaveGlow
 - FastPitch + HifiGan
 - ...

Izzivi

- Omejeni učni viri: 100/1.000 vs 60.000
- Pristop:
 - Preskušanje različnih ogrodij (Kaldi, NeMo, Fairseq...) in pristopov (AM+LM, e2e, besedne podenote ipd.)
 - Iskanje tehnologije/pristopa, ki je najmanj odvisen od količine učnega gradiva
 - Uporaba tehnike prenosa znanja (angl. [Transfer Learning](#))

Demo

- Uporaba splošnega razpoznavalnika za [samodejno podnaslavljanje TV programov](#)

Pilotni projekt, Ministrstvo za kulturo, RTV Slovenije, 2021, 2022

- [ON/Online Notes](#)

Univerza v Ljubljani, CJVT, 2020-

- Sintetizator govora:

- [Jurček](#) (primer odraslega moškega glasu; glas Jure Longyka)
- [Ajda](#) (primer odraslega ženskega glasu; glas Ajda Sokler)