



Spóznawanje rěče Staw a wuhlad

Daniel Zoba, 04.12.2021



Nastork

- Za čo poprawom spóznawanje rěče?
- Nałoženja:
 - rěčny asistent / smart home (prototyp)
 - hrajki (produkt)
 - diktowanska funkcija (słužba)
 - simultane přełožowanje / awtomatiske podtitle (słužba)
 - dalše?



Započatk: Digitalizaciska konferenca 2019

- Mało rěčnych resursow → „klasiski“ system
 - akustiski model (nahrawana rěč → rječaz fonemow)
 - słownik (fonemy → rječaz słowow)
 - rěčny model (słowa → dowolene słowne skupiny)
- Potrjeby:
 - transkrirowane natočenja za akustiski model
 - pisane teksty
 - fonetiska / rěčespytna wěda



Resursy a pomoc

- Nahrawanja we „Mozilla Common Voice“
- Kooperacija z Fraunhofer IKTS a BTU Choćebuz za hornjoserbšćinu:
 - statistiske wuhódnoćenje tekstoweho korpusa
 - nastajenje fonetiskeje lisćiny a wurjekowanskich prawidłow
 - wuběr nahrawanjow z „Common voice“ / Nowe nahrawanja
- Wuspěšne nałożenje wužiwajo němski akustiski model we studiji
 - http://www.essv.de/pdf/2021_43_50.pdf



Wuslědki studije

- „Přiměrjeny“ akustiski model derje funguje za wobmjezowanu domenu (smart home)
- Wozjewjenje demonstratora (open source):
 - https://github.com/ZalozbaDev/web_demo_spoznawanje
- Wozjewjenje prototypa „digidom“ (open source):
 - https://github.com/ZalozbaDev/raspberry_pi_demo_spoznawanje
- Dalše dźěło na robustnosći a wukonje trěbne



Pokročowanje 2021-dźensa

- Akustiski model:
 - přechod na ryzy hornjoserbski model
 - přiměrjenje na rěčnika a akustiske wuměnenja
- Słownik:
 - zlepšíšenja wurjekowanskich prawidłow
- Rěčny model:
 - zlepšíšenje technologije za natwar wjetšich modelow
- Nahrawanja:
 - wobšěrny nawod za wuzwolenje sadow a wuměnenja za nahrawanje



Wuhlrad

- Wozjewjenje wuslědkow z 2021
- Delnjoserbščina
- Planowane kooperacije
 - Serbski Institut: fonetika / linguistika za hornjo- a delnjoserbščinu
 - RCW: kombinacija spóznawanja ze sotra.app
 - FHG a BTU: nałoženja z wulkim rěčnym modelom
- Pomoc při wosobinskim wužiwanju a přiměrjenju prototypa „digidom“
- Dalše nahrawanja



Kónc / Prašenja

- Přeća?