

UWebASR - Web-base ASR engine for Czech and Slovak



Jan Švec, Martin Bulín, Aleš Pražák, Pavel Ircing



NTIS- New Technologies for the Information Society and
Dept. of Cybernetics, University of West Bohemia,
Plzeň, Czech Republic



Introduction

We present a first version of a user-friendly Web-based ASR engine for Czech and Slovak that is free to use for research purposes and does not require any background knowledge about the inner workings of the ASR engine or the API usage. The uploaded audio recording is automatically transcribed and stored in a structured XML format that allows efficient manual post-processing.

Acoustic model

- 3-state HMM with output probabilities modeled by DNN
- trained on 1 990 hours of diverse speech from 15k different speakers

Language model

- simple and robust trigram model
- diverse sources:
 - newspapers (480 mil. tokens)
 - web news (535 mil. tokens)
 - subtitles (225 mil. tokens)
 - transcriptions of TV programs (210 mil. tokens)

Recognize from File

100% of 442452.19 kB sent to the engine

speech signal nonspeech signal

Processing file Bělohávková Olga R09_0001.wav...

Partial result Confidence: <0.0, 0.3> | <0.3, 0.6> | <0.6, 0.9> | <0.9, 1.0>

ono se toho nebude tolik pamatovat která je tak dobrá mimo vliv vy jste na toho tatínka jsem hrál já jen chci říct tenkrát ale na to se Petr čtyřicet devět nebo po byli jsme tedy k babičce do Průhonice Jonathan Terra jsme jako ty babičky byly do školy a a mám takový dojem jsou to zase moudra od maminky pevně mně bylo osm nebo prostě žádala jsem malá holka takže vedle sousedka rozvedená paní no tak jako se s tím a tam byla posádka Sky takže se tam krásně s ním a tak nějak jako stýkala nebo bavila se s ním mám takový dojem ze musela někomu říct že víte o tom že vedle prosím máte také jako Rusa kolegů z atd. no tak nevím to si jako myslíme že to bylo zrovna nějaký udání takového protože mohl jako se list od sebe potom poslouchal jinde tak oni si je dovedli najít tělo potřebovali tak nevím jestli to bylo jako tzv. na top ten z o o tatínkovi řekl ano no vím že tak jako maminka já tak jako říkala že chtěl tatínek té už jsem měl nějaký takový jako zprávy z jeho pera se toto děje jak říkal že odjede do Plzně a naleziště mamince vlasti byl dřívě i já z Brna večer vyjel večer prostě pro něj přišel to se pamatuju že tam snad i pokojíčku vedle na půdě křesel která někdo hodil že by se s ním promluvil na chvíli chvíle bylo by to pak hlídali pět knížka nebo tak já jsem si tenkrát měla okna dívat dnes vyjít ven a vtáhnout rád synu hrál sraz před toho to je jediný co by udělal špatně a osm samozřejmě maminka hned začal tak nějak aby ho dostala zpátky a ministerstvo navýšil den a nikdo si nic nedělal a

Underlying ML models

Back-end

- *SpeechCloud* - platform for remote real-time speech recognition and synthesis
- able to handle the lexicon exceeding 1 mil. word forms
- dynamic modification of the lexicon during the recognition is possible
- transcoding of various audio formats into the raw PCM data
- session-based allocation of recognition engines
- each client opens 2 connections to the platform:
 - interaction with the speech engine - JSON messages over WebSocket protocol
 - transfer of the audio data to be recognized - SIP/RTP

GUI

- implemented in JavaScript
- 4 basic states:
 - waiting for a file
 - processing a file
 - file processed successfully
 - file processing error

Uploaded file name:	test_1.wav
Uploaded file size:	1828.08 kB
Link to the file:	5bb1f82d342d7e0178f29874
Recording timestamp:	2018-10-01T10:34:00Z

Raw recognition result

Confidence: <0.0, 0.3> | <0.3, 0.6> | <0.6, 0.9> | <0.9, 1.0>

dobrý den toto je prostě rozpoznal kg Lugana Západočeské_univerzité v Plzni děkujeme za pozornost

CLARIN CENTRE B 2014 DSA 2017 CLARIN CENTRE K

Partners, Coordination, Funding

- Dept. of Cybernetics, Univ. of West Bohemia
- Institute of Formal and Applied Linguistics (Prague)
- Institute of Czech Language (Prague)
- NLP Centre, Masaryk University (Brno)
- Ministry of Education, Sports and Youth of the Czech Republic

Repository

- Main page
- Contact
- Submission
- Lifecycle
- FAQ
- About and Policies

More

- CLARIN Knowledge Centre | INESS
- CLARIN
- META-Net
- Service Status
- How to Sign in
- Terms of Use

LINDAT CLARIN

MINISTRY OF EDUCATION, YOUTH AND SPORTS

<https://lindat.mff.cuni.cz/services/uwebasr/>

Download transcription Upload another file

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE Trans SYSTEM "trans-14.dtd" >
<Trans audio_filename="test_1.wav">
  <Episode>
    <Section type="report" startTime="0" endTime="7.93">
      <Turn startTime="0.39" endTime="7.93">
        <Sync time="0.39"/>
        dobrý
        <Sync time="0.98"/>
        den
        <Sync time="1.21"/>
        toto
        <Sync time="1.69"/>
        je
        <Sync time="1.8"/>
        prostě
        <Sync time="2.26"/>
        rozpoznal
        <Sync time="2.98"/>
        kg
        <Sync time="3.64"/>
        Lugana
        <Sync time="4.25"/>
        Západočeské_univerzité
        <Sync time="5.68"/>
```