alex: Bootstrapping a Spoken Dialogue System for a New Domain By Real Users

Ondřej Dušek, Ondřej Plátek, Lukáš Žilka and **Filip Jurčíček** {odusek,oplatek,zilka,jurcicek}@ufal.mff.cuni.cz
Institute of Formal and Applied Linguistics, Faculty of Mathematics and Physics, Charles University in Prague

Motivation: Bootstrapping a Dialogue System with No In-domain Data

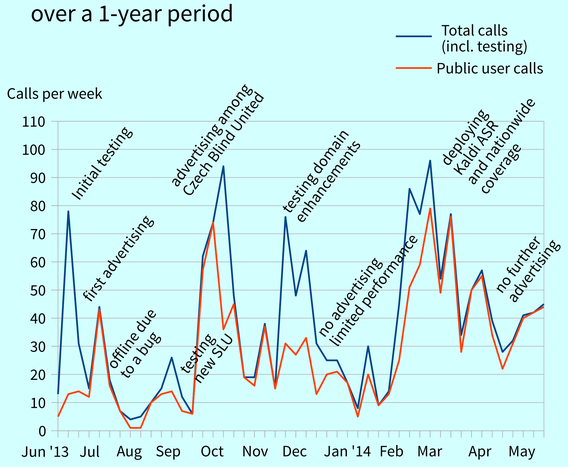
- Our approach: **use the public to bootstrap the system**
- Starting from a very minimal implementation
- **Iterative approach**
- Gradually introducing statistical components
- Minimal cost, alternative to Wizard-of-Oz

The Alex SDS Framework

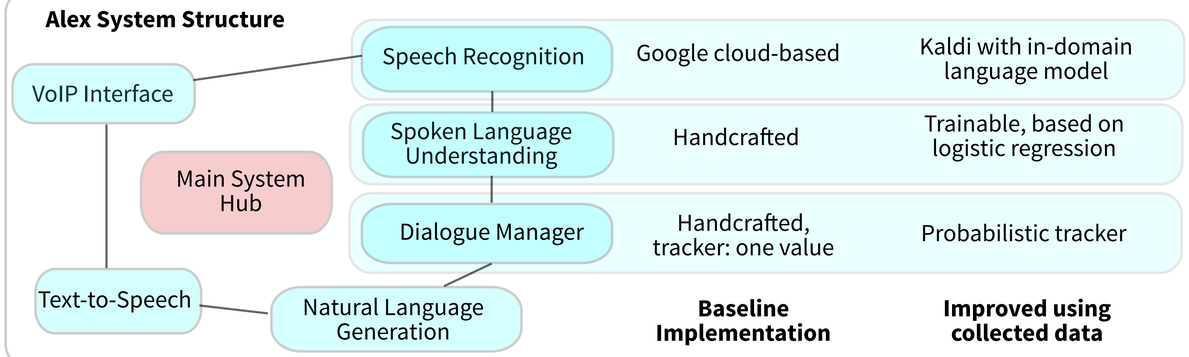
- **Open-source spoken dialogue systems framework**
- Available for download on Github
- Includes all usual SDS components
- **Easy to customize**, adapt for new domains and/or languages

Data Collection

- Domain: Czech public transport information
- System available at a toll-free phone number: **call 800-899-988** from the Czech Republic
- All calls recorded (including our testing)
  - consent required at the beginning of the call
- Small-scale (free) advertising
- More than 1,300 calls from the public over a 1-year period



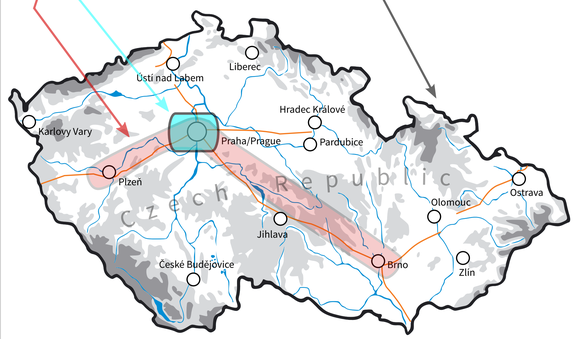
Alex System Structure



Czech Public Transport: Domain Extensions

Public transport network coverage:

1) Prague (ca. 2,500 bus/tram/subway stops)
   - using Google directions
2) Prague, Plzeň, Brno
3) Nationwide coverage (ca. 5,000 towns and cities, 45,000 stops)
   - using Czech Ministry of Transport service



Baseline abilities:

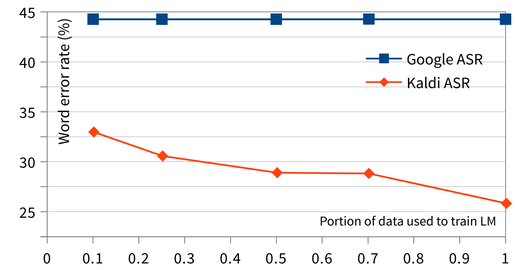
- **Next connection**(s)
- Repeat last statement

Improved:

- Specifying **departure/arrival time**: e.g., *in 15 minutes, at 6 p.m. tomorrow*
- **Connection details** (transfers, duration)
- Travel among different cities
- **Guessing city from stop name** (and vice versa)
- **Weather** information (for any city or town, for a given time)

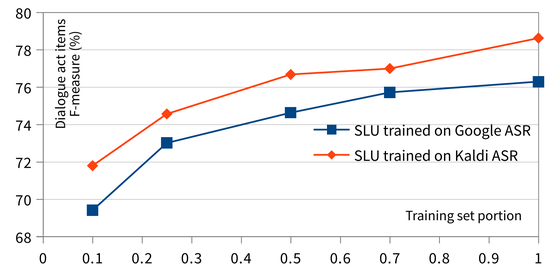
Speech Recognition Improvements

- Replacing Google ASR with **Kaldi** (see other poster!)
- In-domain LM trained on collected data
- **Substantial WER reduction** even for small data



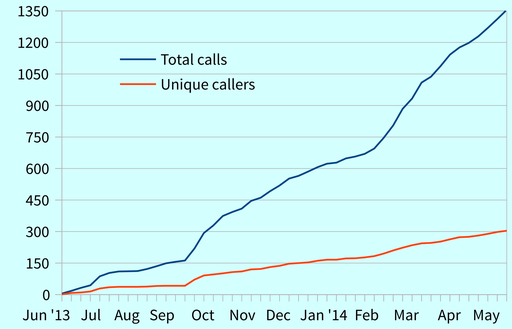
Spoken Language Understanding Improvements

- SLU based on **logistic regression**
- Trained on **annotation generated from transcripts** (fast to acquire, can be rerun for domain extensions)
- Learns to overcome frequent ASR errors



Call Tracking & Evaluation

- Google Analytics, detailed call logs
- Total numbers of calls and callers keep growing



- Information provided in vast majority of calls
- Question at the end of each call: "Did you receive the information you wanted?"
  - Answers mostly positive

Conclusions

- Even a very **limited system can be used to gather data** from the public
- Very **small amounts of in-domain data** greatly **improve ASR** accuracy
- **Generating semantic annotation** from dialogue transcripts **helps maintain statistical SLU**

Download / Fork Us!

- The Alex framework & the system for our domain are **available under Apache 2.0 license** at:

https://github.com/UFAL-DSG/alex

Presented at Sigdial, Pennsylvania, PA, USA, June 2014. Supported by the Ministry of Education, Youth and Sports of the Czech Republic grant LK11221 and core research funding, SVV project 260 104, and grants GAUK 2058214 and 2076214 of Charles University in Prague. Using language resources distributed by the LINDAT/CLARIN project of the Ministry of Education, Youth and Sports of the Czech Republic (project LM2010013).