Silvius: Training a Custom Speech System for Coding by Voice

David Williams-King[1], Yanbei Pang, Avijit Shah, Homayoon Beigi [1]dwk@voxhub.io Columbia University

Introduction

- Many computer users develop Repetitive Strain Injuries (RSIs) from heavy computer use; **incidence is about 50%** [1,2] over lifetime
- Recovery often requires 1+ years of near zero computer use
- Heavy computer users like programmers are forced to hire human typists, undergo surgery (57% failure rate), or switch careers (after surgery, 77%)
- Speech recognition (e.g. Dragon) provides an alternative input mechanism
- Rather than simply writing emails and documents, heavy computer users require complex and precise computer input to replace a keyboard

[1] http://www.rsi-therapy.com/statistics.htm, [2] http://www.rsi.org.uk/whatis/prevalence.html

Design

- Kaldi-based voice coding with the **lowest possible barrier to entry**
  => provide optional free cloud recognition service with default grammar
- Allow **user-customizable voice grammar** with arbitrary words and actions
  => requires language model that supports general English & commands
- Allow for high quality, per-user **custom speech models** for accents/etc
  => given audio data for a specific user, train a new acoustic model ……..

Voice Language

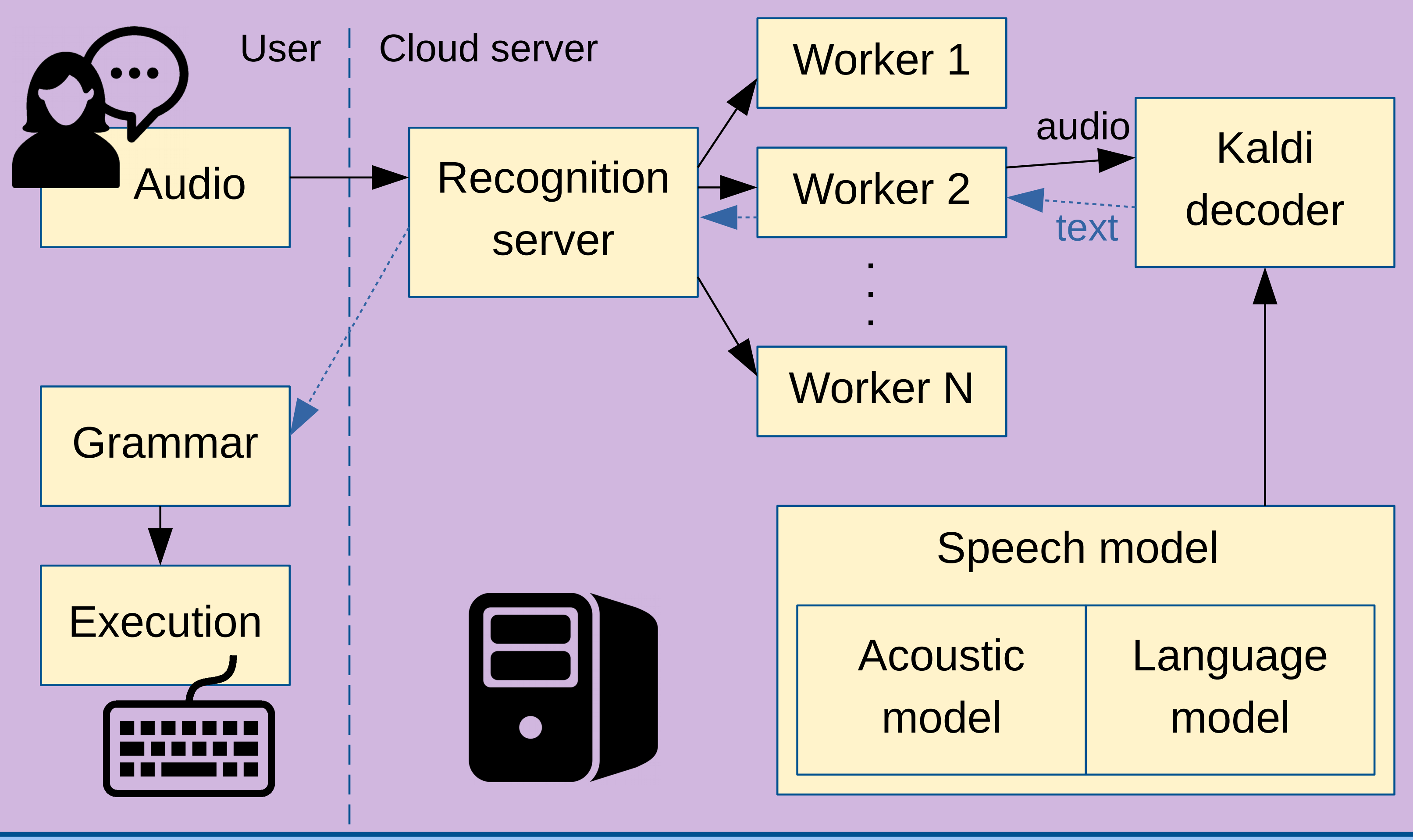
a b c	arch, bravo, char
A D	sky arch, sky delta
( ) { }	len ren (par**ens**), lace race (br**aces**)
[ ] < >	lack rack (br**ack**ets), langle rangle
<escape>, <tab>, !, #	escape/act, tab, bang, hash
1 42	number one, number forty two
<up><up><up>	up up up/up three
<ctrl><left> <pgup> <pgdn>	lope, gope, drop
<backspace> <delete>	scratch, chuck
<ctrl> <alt><tab>	control, alt tab/switch window
why hello there	phrase why hello there
whyHelloThere	camel ...
whyhellothere	jumble ...
why_hello_there	score ...
WHY_HELLO_THERE	upper score ...

Results / Future Work

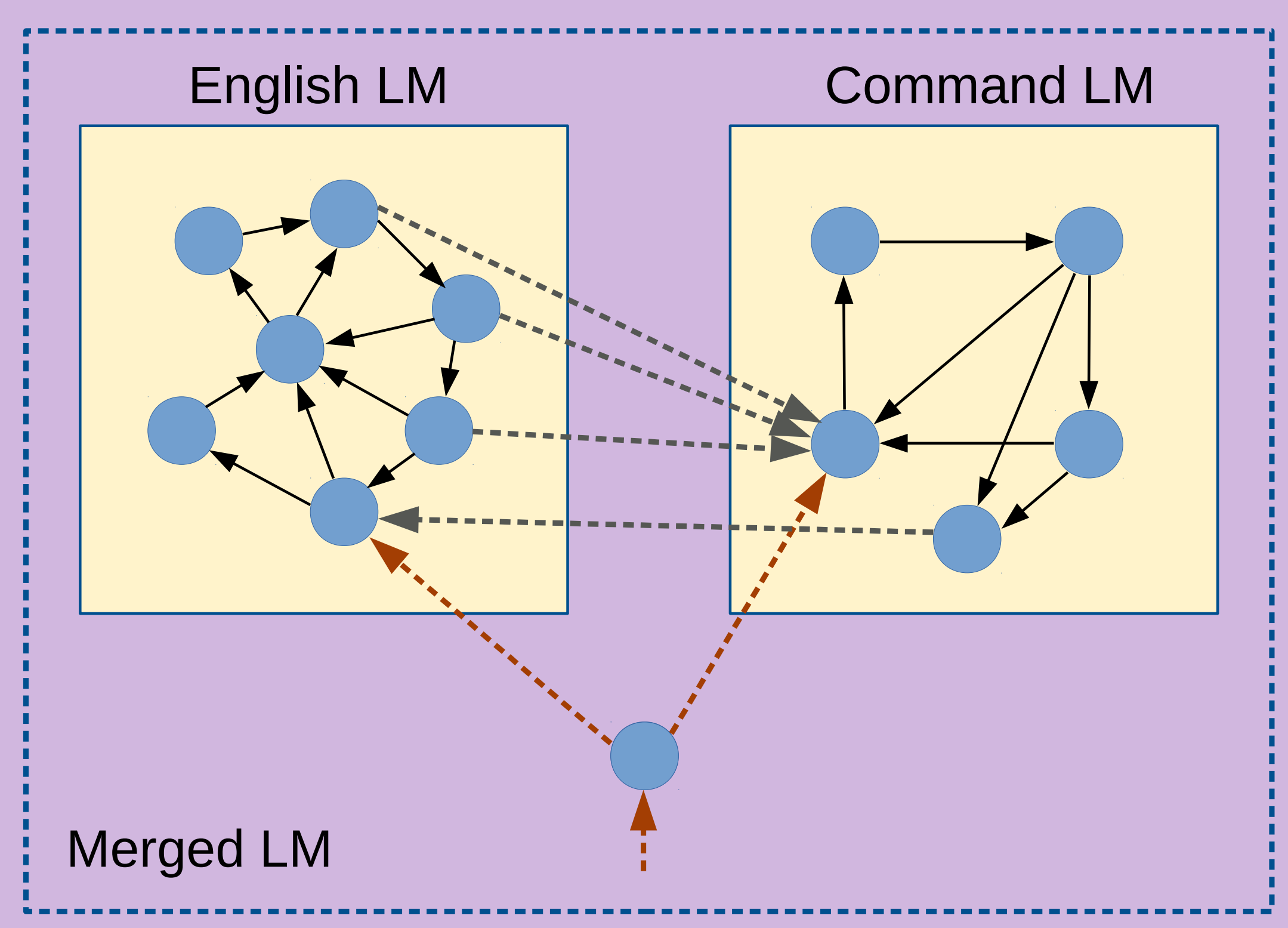
- Developed Silvius grammar format; aim to support Dragonfly format later
- Recognition server has been online for 2 years, many casual users
- Simple boosted language model implemented, accuracy is not good
- Trying union language models, but transition probabilities need tweaking
- Successfully filtering audio transcripts from Dragon and human typist sources
- Training new acoustic models is ongoing (significant GPU time required)

- Decoding speed still slower than Dragon, will try newer Kaldi chain models
- Will develop website that creates custom language models for grammars
- **Seeking interested collaborators** in accessibility or speech

User Workflow



Language Model Multiplexing



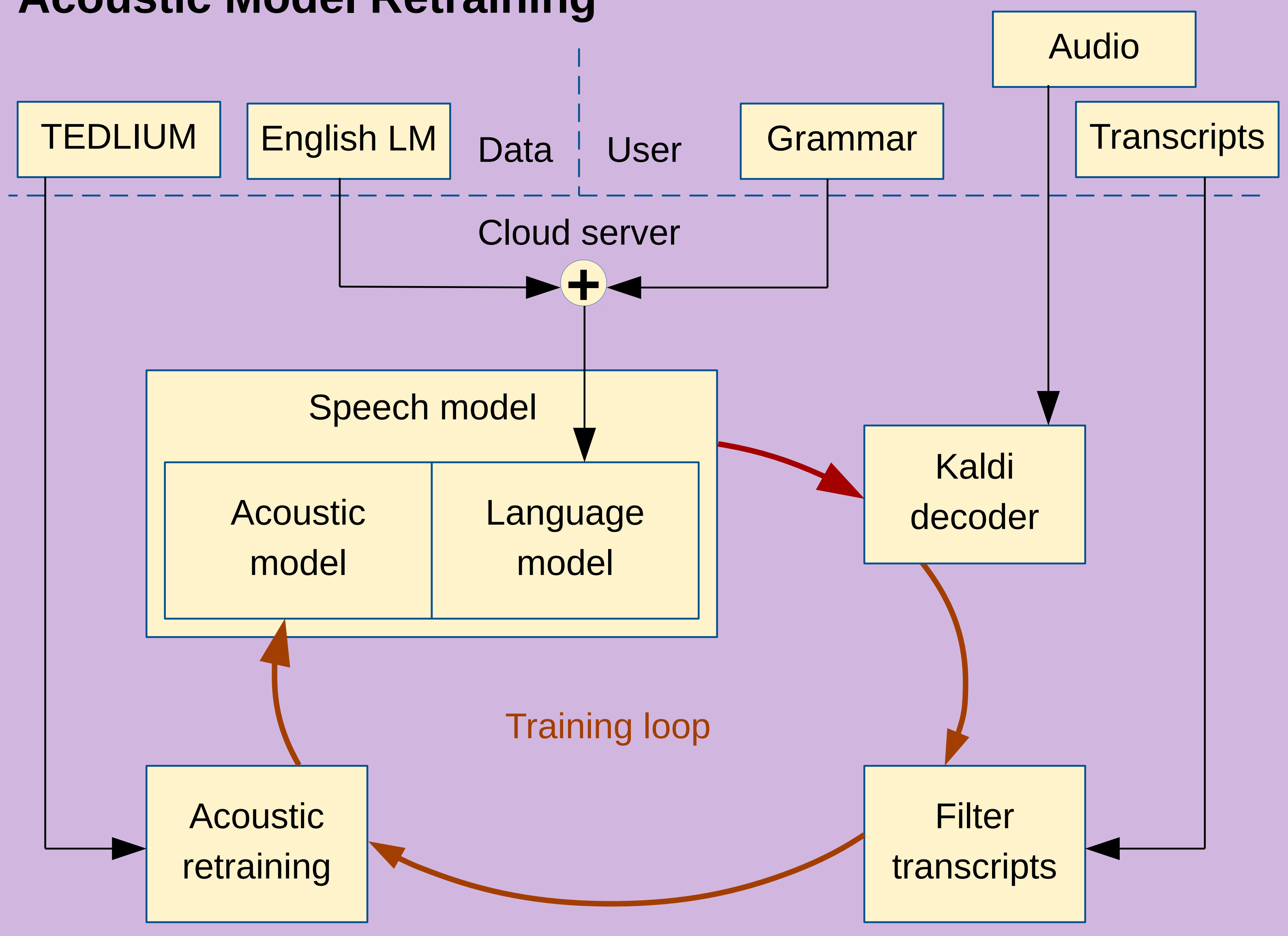
- Create union of two language models, for English and for commands (from the grammar)

→ Intra-model edges

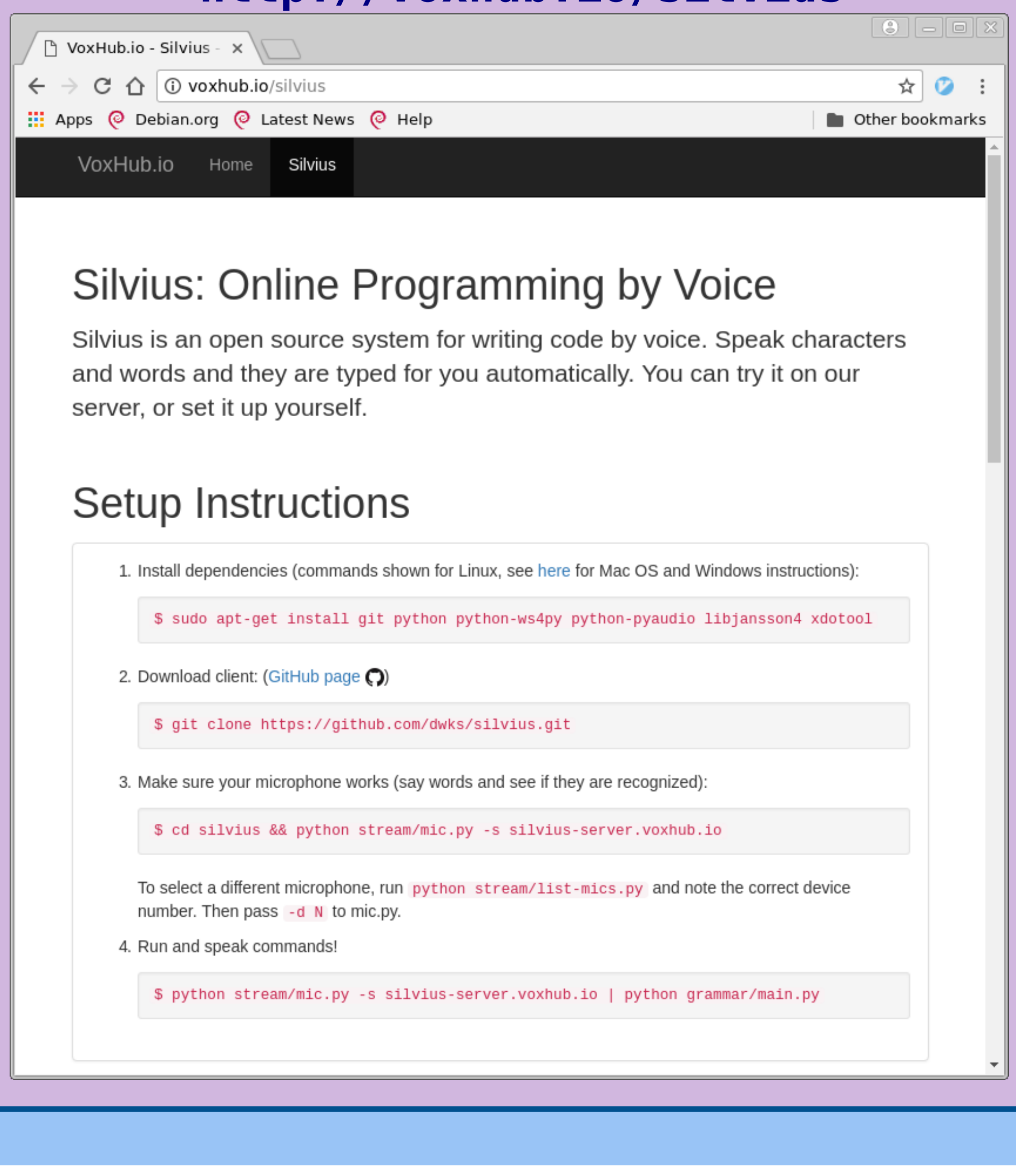
⇢ Links between one model and start node of the other

⇢ Union links

Acoustic Model Retraining



http://voxhub.io/silvius



Silvius: Online Programming by Voice

Silvius is an open source system for writing code by voice. Speak characters and words and they are typed for you automatically. You can try it on our server, or set it up yourself.

Setup Instructions

1. Install dependencies (commands shown for Linux, see here for Mac OS and Windows instructions):

```
$ sudo apt-get install git python python-ws4py python-pyaudio libjansson4 xdotool
```

2. Download client: (GitHub page)

```
$ git clone https://github.com/dwks/silvius.git
```

3. Make sure your microphone works (say words and see if they are recognized):

```
$ cd silvius && python stream/mic.py -s silvius-server.voxhub.io
```

To select a different microphone, run `python stream/list-mics.py` and note the correct device number. Then pass `-d N` to mic.py.

4. Run and speak commands!

```
$ python stream/mic.py -s silvius-server.voxhub.io | python grammar/main.py
```

Quotes

"Thank you for making available a code-by-voice system that's easy to install and easy to work on! The on-line server is a life-saver for getting started quickly."

"Thank you for your work on Silvius.
I like how it is all open source."

"A wonderful project and presentation. Please add a bitcoin address to the page so I can tip you :)"

"This is a life saver."

"I'm glad that there is a way that I can give up typing but still being able to code."

Personal Experience

- Still primarily use Dragon-based system, but the flexibility of Silvius is important once finished
- Continue programming & pair programming