# Agenda

- Command-line usage of vw-l2s for canonical tasks:
  - Sequence labeling
  - Sequence span labeling
  - Graph labeling

- Intro to pyvw (vw in python interface)

- Learning to search in pyvw
  - Part of speech tagging walk-through
  - Named entity recognition exercise

# Part of speech tagging on one slide

```
wget http://bit.ly/1FVkLEK
unzip 1FVkLEK

vw --search          45                          \
   --search_task     sequence                    \
   --search_rollin   learn                       \
   --search_rollout  none                        \
   --affix           -2w,+2w                     \
   --spelling        w                           \
   --search_history_length    2                  \
   --search_neighbor_features  -1:p,1:p,-1:w,1:w \
    -b 26                                        \
    -f wsj.train.model                           \
    -d wsj.train.vw

... patience ...

vw  -i wsj.train.model        \
    -p wsj.test.pred          \
    -d wsj.test.vw            \
    -t
```

# Sequence span labeling

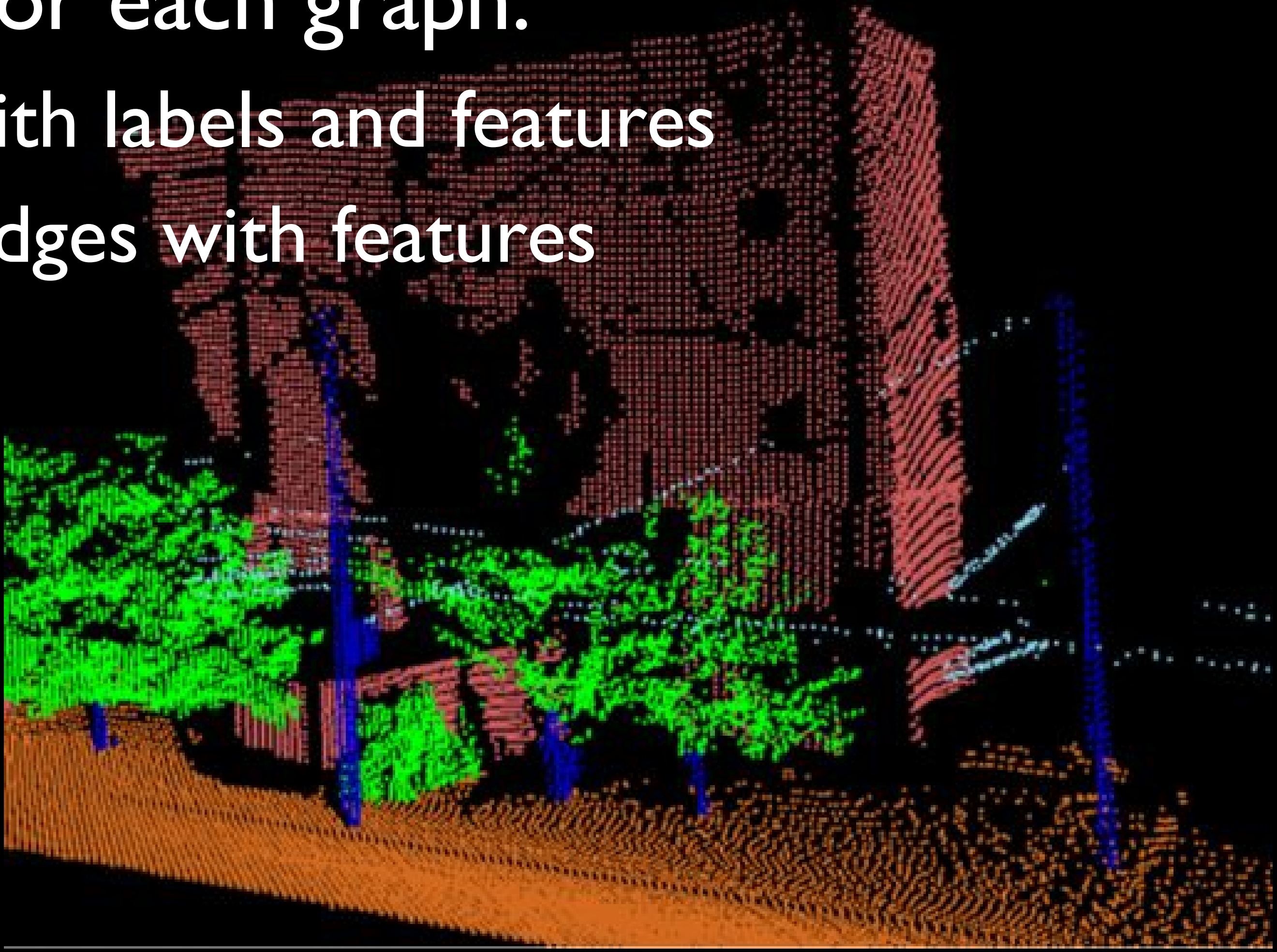
```
               --search_task sequencespan
  optional:    --search_span_bilou
```

- Plus special BIO encoding of labels:
  - “Out” = 1
  - “Begin-X” = any even # at least 2
  - “In-X” = “Begin-X” + 1

# Graph labeling

--search_task graph

- Data encoding; for each graph:
  - List of nodes with labels and features
  - List of (hyper)edges with features



- See search_graph.cc for more docs

# Intro to pyvw

- From vowpal_wabbit directory, run:

```
      cd python
      make
      python test.py
```

  If that doesn't work, look on with your neighbor

- If you have iPython installed, run:

```
      ipython notebook VW_in_Python.ipynb
```

- Or view at: `http://tinyurl.com/pyvwintro`

# Pythonic part of speech tagging

- Open notebook

  Learning_to_Search.ipynb

  or view at

  http://tinyurl.com/pyvwsearch

# Your homework assignment

- Download:

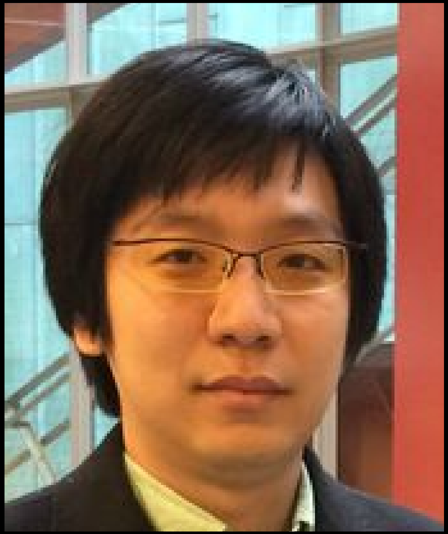
  `http://hal3.name/ner.zip`

- Let's build a named entity recognizer!

- Files:

| | |
|---|---|
| `ner.py` | basic scaffolding |
| `ner_assignment.txt` | your homework |
| `ner_solution.py` | my solution to your homework |
| `moredata.py` | a larger dataset to play with |

We're here to help!



Kai-Wei

Hal

He

John

Sudha