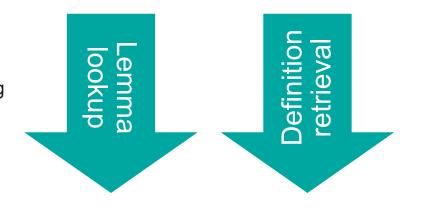
## Exercise: Dictionary Advanced Search

cognitive

adjective
: of, relating to, or being conscious mental activities (as thinking, reasoning, remembering, imagining, learning words, and using language)



relates to reasoning and learning

adjective: cognitive

## Classify POS with Natural Language Classifier

- Get dictionary data
  - Es. WordNet (http://wordnetcode.princeton.edu/wn3.1.dict.tar.gz)
- Train the classifier
  - Extract training set (classification pairs) from dictionary
    - CSV: POS, defintion
  - Launch classifier (takes some time)
- Evaluation:
  - Extract some definition (not it training set) and check:
    - definition > POS



## Search definitions with Retrieve and Rank

- Get dictionary data
  - Es. WordNet (http://wordnetcode.princeton.edu/wn3.1.dict.tar.gz)
- Build «ground truth»:
  - Sense := <lemma, defintion>
  - This can be directly extracted from Wordnet and serialized in a Json doc
- Feed Solr service instance with Json dictionary doc
- Train the ranker
  - Create a training set:
    - Pick up a number of senses from the «ground truth» and provide a number of definition queries (could be fragments or paraphrases)
    - <sense id, definition fragment or paraphrase>
  - Launch training (takes some time)
- Evaluation:
  - Pick up a number of senses not in the training set and check:
    - definition > sense



## Integrate Classifier for re-ranking

- Classify the query
  - query > POS
- Retrieve sense
  - query > sense
- Rerank senses based on POS in your app
  - E.g. filter
- Better method: extend R&R with a POS feature
  - score the query POS
  - score the pair: <sense POS, query POS>
    - trivially, less trivially
  - This way, the POS contribution enters the learning process
- Evaluate
  - Measure improvement in both cases

