The dataset includes

* 3 folders:
  + train: data.json
  + dev: data.json
  + test: data.json
* dictionary.json
* gold.json

data.json file is a collection contains records where each record is stored as follows:

{

“**text**”: // The string sentence in draid record that has one acronym

“**start\_char\_idx**” : // The index of the ambiguous, i.e., acronym in text

“**length\_acronym**”: // Length of acronym in text

“**expansion**”: // The correct meaning, i.e., expansion of the acronym in the sample. These expansions are selected from dictionary.json files.

}

dictionary.json: A dictionary of the acronyms and their possible meanings, i.e., expansion. All predictions should use this dictionary to expand the ambiguous acronyms in the text

gold.json: A prediction file created from test/data.json to test the scoring script.

Statistics:

