

# Project Resources



**CS 6320**

# Overview

---

- Data Resources
- Software Resources
  - Search Index Software
  - WordNet Data and API
  - NLP Tools
    - Sentence Segmentation
    - Tokenizer
    - Part-of-Speech Tagger
    - Syntactic or Dependency Parser
    - Semantic Parser

# Data Resources

---

- NLTK: [http://www.nltk.org/nltk\\_data/](http://www.nltk.org/nltk_data/)
- <https://github.com/niderhoff/nlp-datasets>
- <https://nlp.stanford.edu/links/statnlp.html#Corpora>

# Software Resources

---

- Search Index Software
  - SOLR
    - <http://lucene.apache.org/solr/>
  - Lucene
    - <http://lucene.apache.org/index.html>

# Software Resources

---

- WordNet Data and API
  - Data and C++ API:
    - <https://wordnet.princeton.edu/wordnet/download/>
  - Java API:
    - <http://yawni.sourceforge.net/wiki/index.php>
  - Python API:
    - <http://www.nltk.org/howto/wordnet.html>

# Software Resources

---

- NLP Features
  - Sentence Segmentation
  - Tokenizer
  - Part-of-Speech Tagger
  - Syntactic or Dependency Parser
  - Semantic Parser
- NLP Tools
  - <http://www.nltk.org/>
  - <http://nlp.stanford.edu/software/>
  - <https://opennlp.apache.org/>