**Resources and tools for NLP**

**Text pre-processing**

**Input: raw text**

1. **Tokenize => tokens**
2. **Sentence split => sentences**
3. **POS => parts-of- speech**
4. **Lemmatize => lemmas**
5. **NER (name entity recognition) => entities**
6. **Chunking => constituents, phrases**
7. **Dependency parsing => binary grammatical relations**

**English language:**

* **Stanford parser:**

[**https://stanfordnlp.github.io/CoreNLP**](https://stanfordnlp.github.io/CoreNLP)

https://web.stanford.edu/~jurafsky/slp3/

* **CST Online Tools: <https://cst.dk/tools/index.php?lang=en>**

(tokenizer, POS-tagger, lemmatizer, name recognizer**)**

* **Allen NLP: <https://demo.allennlp.org/>**
* **Natural Language ToolKit (NLTK) <https://www.nltk.org/>**
* **Parsers: Stanford , RASP, Collins, Charniak, Bibel, Assert, MINIPAR**

[**https://opensource.adobe.com/NLP-Cube/index.html**](https://opensource.adobe.com/NLP-Cube/index.html)

* **Text compactor <https://www.textcompactor.com/> --- summarizer**
* **Tools4noobs –** [**https://www.tools4noobs.com/summarize/**](https://www.tools4noobs.com/summarize/)
* **START – Natural Language Question-Answering system**

[**http://start.csail.mit.edu/index.php**](http://start.csail.mit.edu/index.php)

* [**https://huggingface.co/coref/**](https://huggingface.co/coref/) **- coreference resolution**

- **WordNet (on-line)** **http://wordnetweb.princeton.edu/perl/webwn**

**– lexical database in English**

**- WordNet Similarity**: <http://ws4jdemo.appspot.com/>

**Sentiment analysis tools:**

<https://www.danielsoper.com/sentimentanalysis/>

[https://sentigem.com/#](https://sentigem.com/)!

<https://text2data.com/Demo>

**Romanian language**

* [**www.racai.ro**](http://www.racai.ro)

[**https://relate.racai.ro/index.php?path=teprolin/custom**](https://relate.racai.ro/index.php?path=teprolin/custom) **-- preprocesare texte in limba romana**

**-** [**https://www.racai.ro/tools/text/**](https://www.racai.ro/tools/text/)

* [**http://www.racai.ro/p/corola/**](http://www.racai.ro/p/corola/) **-- corpus limba romana**
* [**https://opensource.adobe.com/NLP-Cube/index.html**](https://opensource.adobe.com/NLP-Cube/index.html)
* [**https://spacy.io/models/ro**](https://spacy.io/models/ro)
* **RoWordNet:** [**https://relate.racai.ro/index.php?path=rown/query**](https://relate.racai.ro/index.php?path=rown/query)