

实验三

要求：利用 Chinese.txt 和 English.txt 的中英文句子，在实验二的基础上，继续利用以下给定的中英文工具进行词性标注和命名实体识别。并对不同工具产生的结果进行简要对比分析，将实验过程与结果写成实验报告，实验课结束后提交。

英文工具：

Nltk:

<http://www.nltk.org/index.html>

<https://github.com/nltk/nltk>

<https://www.jianshu.com/p/9d232e4a3c28>

Spacy:

<https://blog.csdn.net/u012436149/article/details/79321112>

<https://spacy.io/usage/linguistic-features#section-tokenization>

Stanfordnlp:

<https://github.com/Lynten/stanford-corenlp>

中文工具：（部分工具命名实体识别没有直接调用的函数，可以根据词性标注的结果自己实现）

Jieba:

<https://github.com/fxsjy/jieba>

StanfordCoreNLP:

<https://github.com/Lynten/stanford-corenlp>

SnowNLP:

<https://github.com/isnowfy/snownlp>

<https://www.jianshu.com/p/4692d1b5364d>

THULAC:

<https://github.com/thunlp/THULAC-Python>

NLPIR:

<https://github.com/tsroten/pynlpir>

https://blog.csdn.net/weixin_34613450/article/details/78695166

HanLP（选做，需要 Microsoft Visual C++ 14.0）

<https://github.com/hankcs/pyhanlp>