Knowledge Graph

NLP and its challenges

Stop words

POS tagging

NER

co-reference

Lemmatization

Stemming

IE (Information Extraction)

Triplet Extraction, example, different rule based (openIE)

WordNet

Spark, RDD etc

Map reduce

IR (information Retrieval)

N-gram

Term Frequency

Inverse Document Frequency

Word2Vec (Skip Gram and CBOW), also how to extend it on multiple documents

Code to perform NLP technique, spark code basically on logic