NLP – Keyword searching, Information Extraction,

Components of NLP:

Natural language Understanding

Natural Language Generation – meaningful

Steps in NLP:

Tokenization:

Breaking strings into tokens.

Stemming:

Normalizing words into base or root form.

Lemmatization:

Grouping words into one word

Pos Tags:

What kind of word is it. Verb, Noun

Named Entity Recognition:

Organization, Person, location

Chunking:

Grouping tokens into chunks

NLTK Natural language toolkit library runs all of these functions in python