In [10]:
```
sample_document = "Natural language processing is a subfield of artificia
l intelligence that deals with the interaction between computers and huma
ns using natural language. It involves tasks such as text analysis, langu
age translation, and sentiment analysis."
```

Tokenization:

In [4]:
```
import nltk
nltk.download('punkt')
from nltk.tokenize import word_tokenize

tokens = word_tokenize(sample_document)
print("Tokenization:")
print(tokens)
```

```
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\Bhaskar\AppData\Roaming\nltk_data...
[nltk_data]   Unzipping tokenizers\punkt.zip.

Tokenization:
['Natural', 'language', 'processing', 'is', 'a', 'subfield', 'of', 'arti
ficial', 'intelligence', 'that', 'deals', 'with', 'the', 'interaction',
'between', 'computers', 'and', 'humans', 'using', 'natural', 'language',
'.', 'It', 'involves', 'tasks', 'such', 'as', 'text', 'analysis', ',', '
language', 'translation', ',', 'and', 'sentiment', 'analysis', '.']
```

POS Tagging (Part of Speech Tagging):

In [5]:
```
nltk.download('averaged_perceptron_tagger')
pos_tags = nltk.pos_tag(tokens)
print("\nPOS Tagging:")
print(pos_tags)
```

```
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]     C:\Users\Bhaskar\AppData\Roaming\nltk_data...

POS Tagging:
[('Natural', 'JJ'), ('language', 'NN'), ('processing', 'NN'), ('is', 'VB
Z'), ('a', 'DT'), ('subfield', 'NN'), ('of', 'IN'), ('artificial', 'J
J'), ('intelligence', 'NN'), ('that', 'IN'), ('deals', 'NNS'), ('with',
'IN'), ('the', 'DT'), ('interaction', 'NN'), ('between', 'IN'), ('comput
ers', 'NNS'), ('and', 'CC'), ('humans', 'NNS'), ('using', 'VBG'), ('natu
ral', 'JJ'), ('language', 'NN'), ('.', '.'), ('It', 'PRP'), ('involves',
'VBZ'), ('tasks', 'NNS'), ('such', 'JJ'), ('as', 'IN'), ('text', 'JJ'),
('analysis', 'NN'), (',', ','), ('language', 'NN'), ('translation', 'N
N'), (',', ','), ('and', 'CC'), ('sentiment', 'NN'), ('analysis', 'NN'),
('.', '.')]

[nltk_data]   Unzipping taggers\averaged_perceptron_tagger.zip.
```

Stop Words Removal:

In [6]:
```
from nltk.corpus import stopwords

nltk.download('stopwords')
stop_words = set(stopwords.words('english'))
filtered_tokens = [token for token in tokens if token.lower() not in stop
_words]
print("\nStop Words Removal:")
print(filtered_tokens)
```

```
Stop Words Removal:
['Natural', 'language', 'processing', 'subfield', 'artificial', 'intelli
gence', 'deals', 'interaction', 'computers', 'humans', 'using', 'natural
', 'language', '.', 'involves', 'tasks', 'text', 'analysis', ',', 'langu
age', 'translation', ',', 'sentiment', 'analysis', '.']

[nltk_data] Downloading package stopwords to
[nltk_data]     C:\Users\Bhaskar\AppData\Roaming\nltk_data...
[nltk_data]   Unzipping corpora\stopwords.zip.
```

Stemming:

In [7]:
```
from nltk.stem import PorterStemmer

stemmer = PorterStemmer()
stemmed_tokens = [stemmer.stem(token) for token in filtered_tokens]
print("\nStemming:")
print(stemmed_tokens)
```

```
Stemming:
['natur', 'languag', 'process', 'subfield', 'artifici', 'intellig', 'dea
l', 'interact', 'comput', 'human', 'use', 'natur', 'languag', '.', 'invo
lv', 'task', 'text', 'analysi', ',', 'languag', 'translat', ',', 'sentim
ent', 'analysi', '.']
```

Lemmatization:

In [8]:
```
from nltk.stem import WordNetLemmatizer

nltk.download('wordnet')
lemmatizer = WordNetLemmatizer()
lemmatized_tokens = [lemmatizer.lemmatize(token) for token in filtered_to
kens]
print("\nLemmatization:")
print(lemmatized_tokens)
```

```
[nltk_data] Downloading package wordnet to
[nltk_data]     C:\Users\Bhaskar\AppData\Roaming\nltk_data...

Lemmatization:
['Natural', 'language', 'processing', 'subfield', 'artificial', 'intelli
gence', 'deal', 'interaction', 'computer', 'human', 'using', 'natural',
'language', '.', 'involves', 'task', 'text', 'analysis', ',', 'language'
, 'translation', ',', 'sentiment', 'analysis', '.']
```