

# **DATA SCIENCE**

## **NATURAL LANGUAGE PROCESSING**

**I. NATURAL LANGUAGE PROCESSING**

**II. NLP APPLICATIONS**

**III. NLP IN PYTHON**

# **I. NATURAL LANGUAGE PROCESSING**

**Natural language processing (NLP) is a field of computer science, artificial intelligence, and linguistics concerned with the interactions between computers and human (natural) languages. As such, NLP is related to the area of human–computer interaction. Many challenges in NLP involve natural language understanding -- that is, enabling computers to derive meaning from human or natural language input.**

**The interface between human and computer language**

**However, language is often ambiguous. Many words have different meanings based upon context.**

**Humans can often decipher the ambiguity; computers cannot.**

**How do we deal with that?**

**Create a complex set of rules?**

**That's too fragile as language constantly changes.**

**Semantic models work well but are computationally expensive.**

**Statistical models provide a good balance.**

# **II. NLP APPLICATIONS**

**Take five minutes and jot down any common, real-world examples of NLP that you can think of.**

**Let's talk about some common applications.**



## **Speech Recognition**

**Siri or Google Now**

**Speech recognition software incorporate language models  
along with audio signal.**

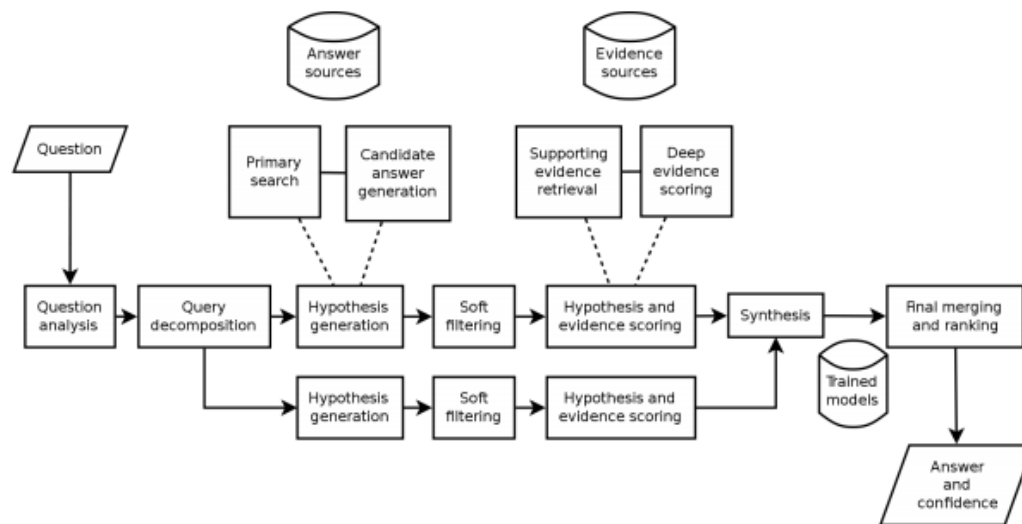
## **Machine Translation**

### **Google Translate**

**Can translate almost any word in any language to another language.**

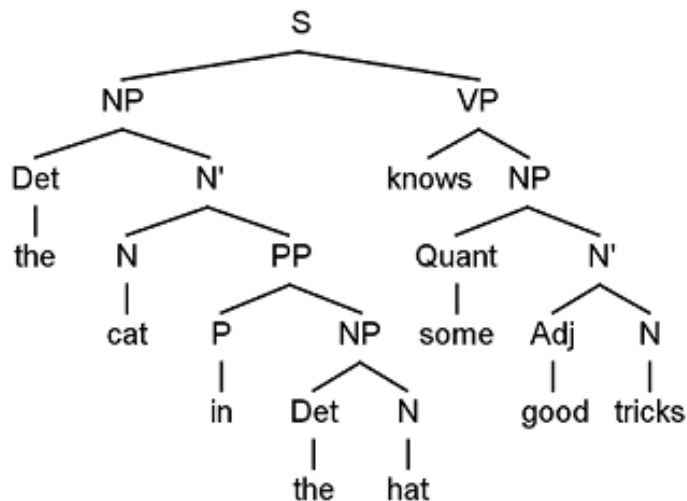
# Question Answering

## IBM Watson, Wolfram Alpha



## Part of Speech Tagging

**Aids in many other NLP tasks such as Named Entity Recognition**



## **Topic Modeling**

**Finding latent groupings of documents based upon the words in them.**

**Useful for news aggregators**

## **Sentiment Analysis**

**Determining the emotional content of a document.**

**Not only the emotion but the intensity of the emotion.**

**Not always straight forward**



CLARABRIDGE

# **III. NLP IN PYTHON**



## Libraries

NLTK

TextBlob

SpaCy

---

# DATA SCIENCE

---