Natural Language Processing

Orevaoghene Ahia

Lecture 1: Introduction and History of NLP

Meeeee !

I am a research engineer interested in Machine Translation

and Dialogue Systems for Low-resourced Languages.

I am a songwriter and I play the guitar at leisure .

Co-instructor

Kelechi Ogueji

We are still open to volunteer tutors for this track too .

Private Slack Channel

Assignments

Mini projects

Final Projects

What we hope to teach ?

- The NLP Pipeline
  - Key Components for understanding text
- NLP Systems / Applications
  - Current techniques and Limitations

How important is Language to you ?

What is NLP ?

Ideas ?

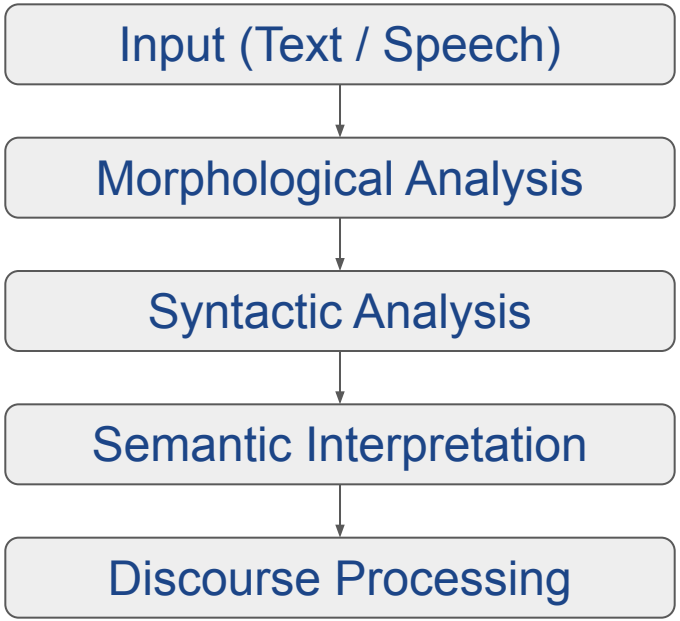
What is Natural Language Processing (NLP) ?

- Natural Language Processing is a field that intersects:
  - **Computer Science**
  - **Artificial Intelligence**
  - **Linguistics**
- Goal: Correctly determining semantic content of natural language in order to perform useful tasks. For example
  - Machine Translation; *Google translate*
  - Dialogue systems; *Siri, Alexa*
- Fully **understanding** and **representing** the meaning of language is a difficult goal .

History of NLP

- NLP has been through at least 3 major eras:
  - 1950s - 1980s: Linguistic Models and HandWritten Rules
    - Too complex to maintain, can’t scale, can’t generalize .
  - 1980s - Now: Corpus/Statistical Methods
    - Using statistical learning over huge datasets of unstructured texts.
  - Now - ???: Deep Learning
    - Deep learning provides a very flexible, (almost?) universal, learnable framework for representing world, visual and linguistic information
    - Major state of the art breakthroughs in NLP

Levels of NLP



So what are some applications of Natural Language Processing ?

Some NLP Applications

Range from Simple to Complex

- Spell checking, keyword search
- Information extraction from websites
  - Finding product prices, dates, location, people
- Classifying text, e.g positive and negative sentiments in documents.

- Automatic Summarization
- Dialogue Systems
- Machine Translation
- Question Answering

Machine Translation

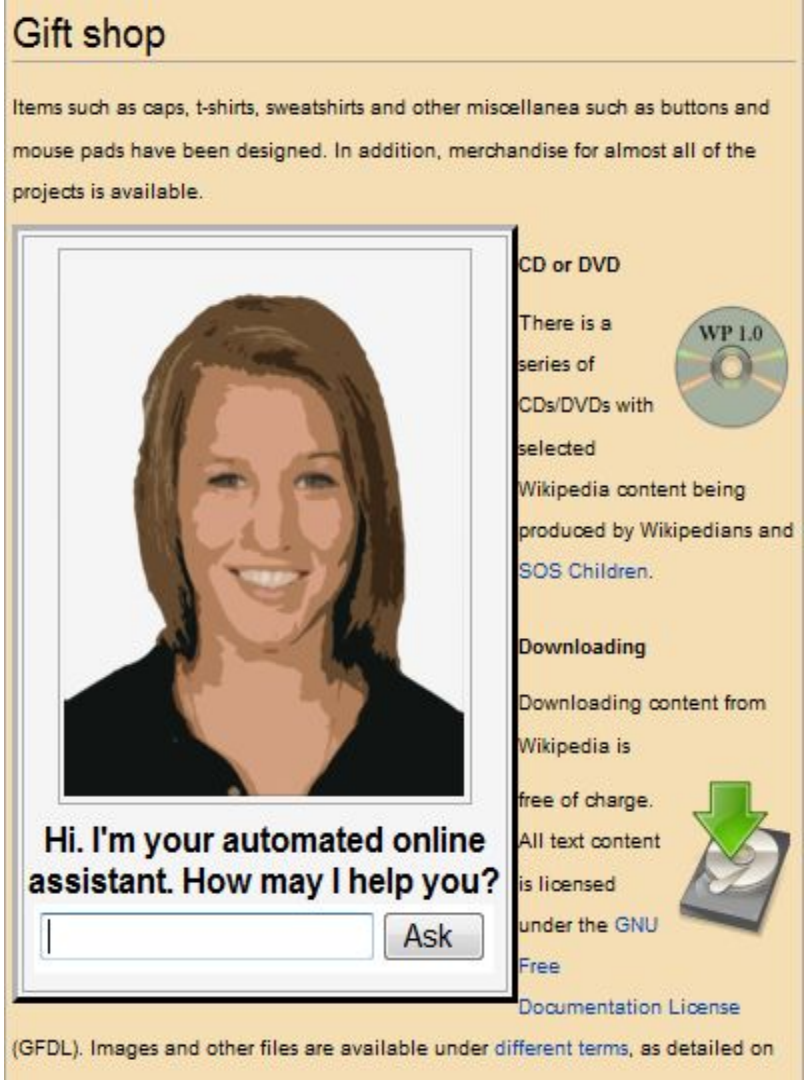
The task of translating a sentence **x** from one language (the source language) to a sentence **y** in another language (then target language) .

X: Quand rien ne va bien, allez á gauche.

Y: When nothing goes right, go left.

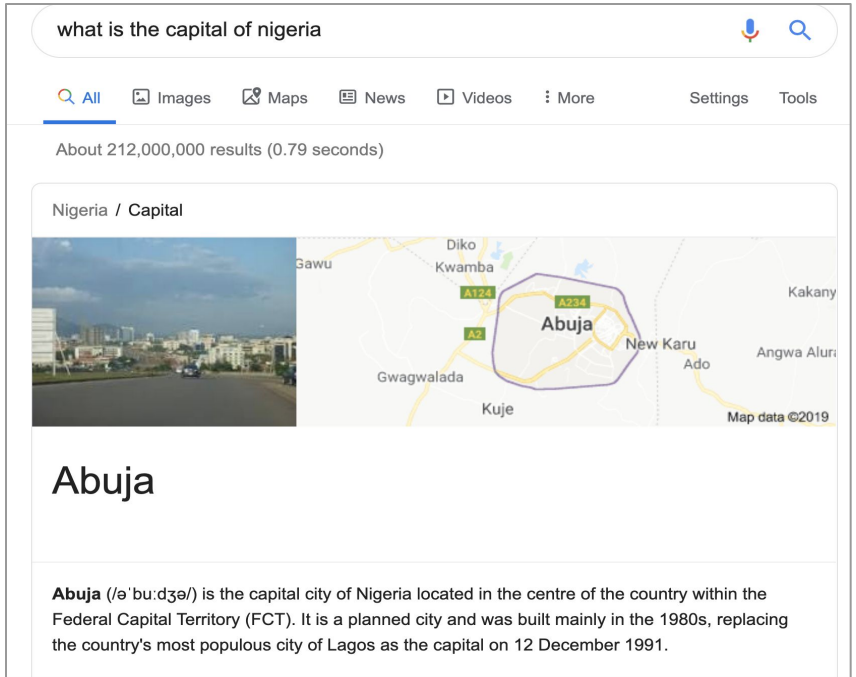
Dialogue Systems

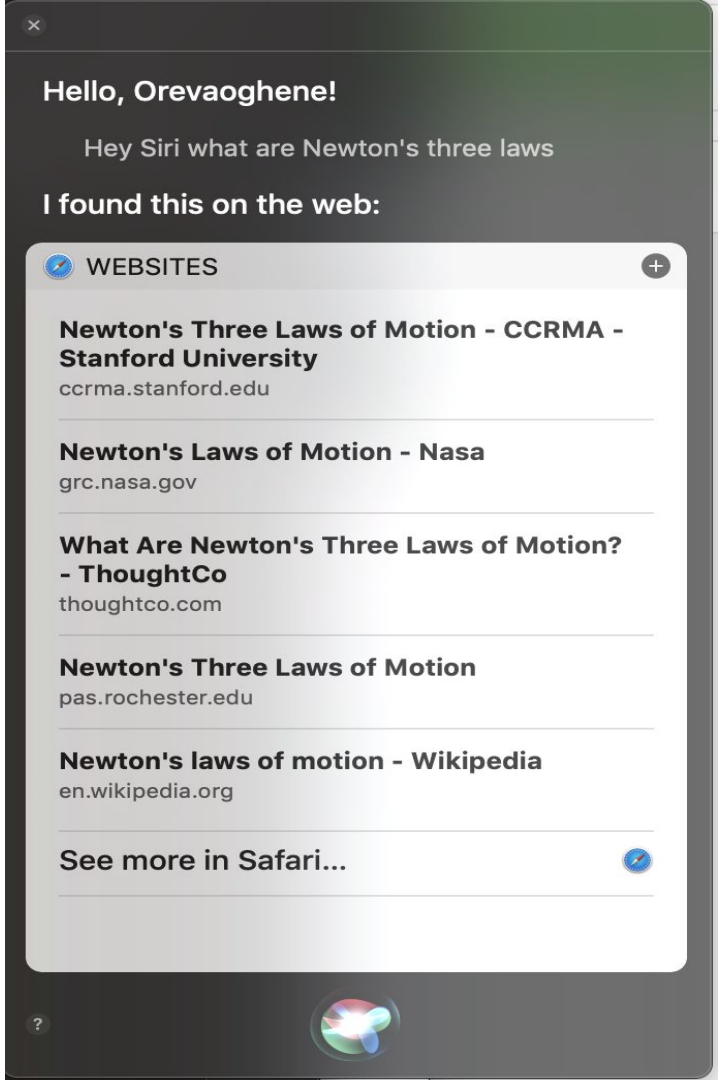
- Computer Systems intended to converse with humans .
  - Task Oriented
  - Non-task Oriented



Question Answering

- Question Answering (QA) systems enable users to retrieve exact answers for questions posed in natural language.





NLP in Industry

- Automated Translation
- Speech Recognition
- Chatbots / Dialogue agents
- Sentiment Analysis for customer reviews
- Automated Hiring and Recruitment
- Online Advertisement
- Text Mining in Healthcare

Why is NLP Complex ?

- Complexity in language representation.
- Ambiguity
  - *For example, Rima went to Gauri. She said, “I am tired.” – Exactly who is tired?*
- Human Language Interpretation Variation
  - Depends on real world, common sense and contextual knowledge.

Challenges

- Bias
  - Eg. Bias in Word Vector and Text Generation
- NLP for low-resourced scenarios
- Model Explainability
- Cross-lingual Learning
- Evaluation
  - BLEU Score; It admits too many variations
- Multi-task Learning
  - DecaNLP; Salesforce

Prerequisites

- Proficiency in python
- Multivariate Calculus , Linear Algebra
- Basic Probability and Statistics
- Fundamentals of Machine Learning and Neural Networks
  - Loss functions
  - Taking simple derivatives
  - Performing optimization with gradient descent.

Lecture Plan for the Cohort

https://docs.google.com/document/d/18Tz5rQsW3l7rkGk9pPHHJvYwPcR_lFwGWzeLbpSjrYs/edit?ts=5d7669e8#