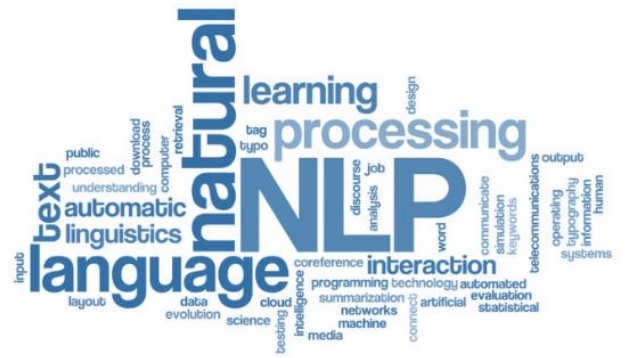
CENG 3526
Natural Language Processing

Lecture 0
Course Introduction

Instructor
Bekir Taner Dinçer

Teaching Assistant
Selahattin Aksoy



MUĞLA SITKI KOÇMAN ÜNİVERSİTESİ
COMPUTER ENGINEERING

1

Course Introduction

Logistics, Grading, Resources, etc.

2

Lecture Management Systems – LMS

- **dys.mu.edu.tr**, the official LMS site of MSKÜ
  - **Quizzes & Exams**
  - **Lecture Notes**
  - **Online Guides, Reading Assignments**,
  - Any materials related to the course.

Telegram Group

Telegram Group Name:
MSKU CENG 3526 NLP 2025-26 Fall

Evaluation & Grading

- Grading
  - **40% Midterm**
    - 20% - On site exam
    - 10% - Quizzes
    - 10% - Homework
    -
  - **60% Final**
    - 30% On site exam
    - 10% - Quizzes + Homework
    - 20% - Project work

Course Outline and Schedule

Part I: Foundations of NLP
Weeks 1-5

- **Week 1**: Introduction to NLP and Text Representation
- **Week 2**: Text Preprocessing and Representation
- **Week 3**: Language Models and Probability
- **Week 4**: Part-of-Speech Tagging and Named Entity Recognition
- **Week 5**: Syntax and Parsing

Part II: Machine Learning for NLP
Weeks 7-8:

- **Week 6**: Sentiment Analysis
- **Week 7**: Machine Learning for NLP I
  Supervised Learning, Feature Engineering, Hyperparameter Tuning
- **Week 8**: Machine Learning for NLP II
  Neural Networks, LSTM and GRU, Evaluation metrics, optimization techniques.

Part III: Advanced NLP Topics
Weeks 9-14:

- **Week 9**: Topic Detection
- **Week 10**: Unsupervised Learning in NLP
- **Week 11**: Transformers and Attention Mechanisms
- **Week 12**: Transfer Learning and Pre-trained Models
- **Week 13**: Machine Translation and Seq2Seq Models
- **Week 14**: Future Directions and Ethics in NLP

Course Objectives

By the end of the course, students will be able to:

1. Understand the fundamental concepts and techniques in NLP.
2. Preprocess and represent text data using various methods, including tokenization, stemming, lemmatization, and vectorization.
3. Apply machine learning algorithms to NLP tasks such as text classification, sentiment analysis, and topic detection.
4. Build and evaluate language models, including N-gram models and modern neural network-based models like RNNs and Transformers.
5. Implement and fine-tune advanced NLP models for tasks like machine translation, named entity recognition, and sentiment analysis.
6. Understand and address ethical issues related to bias, privacy, and fairness in NLP systems.
7. Develop and present a capstone project that demonstrates proficiency in multiple NLP techniques.

Academic Integrity & Ethics

- Don't **copy-paste** codes in projects from the internet. Don't **copy-paste** codes from AI
  - Doing so, **will result in a 0 score** and you'll immediately fail.

- This class encourages the use of templates, use of AI and collaboration, as long as you clearly indicate (give reference to) when you use external resources

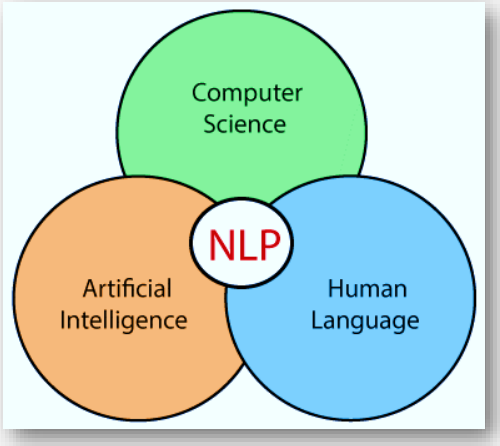
and

be sure the main part of the work is your own.

NLP in Brief

Definition, Goal, Use Cases, State-of-the-Art, NLP Pipeline.

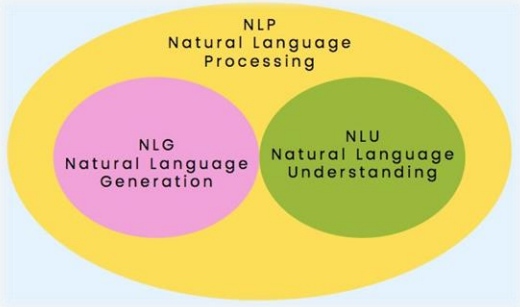
Natural Language Processing (NLP) : What is it?



Natural language processing is **a subfield of linguistics, computer science, and artificial intelligence (AI)**

concerned with the interactions between **computers** and **human language**

Natural Language Processing (NLP) : The Goal?



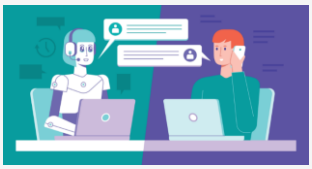
The goal

is

a computer capable of

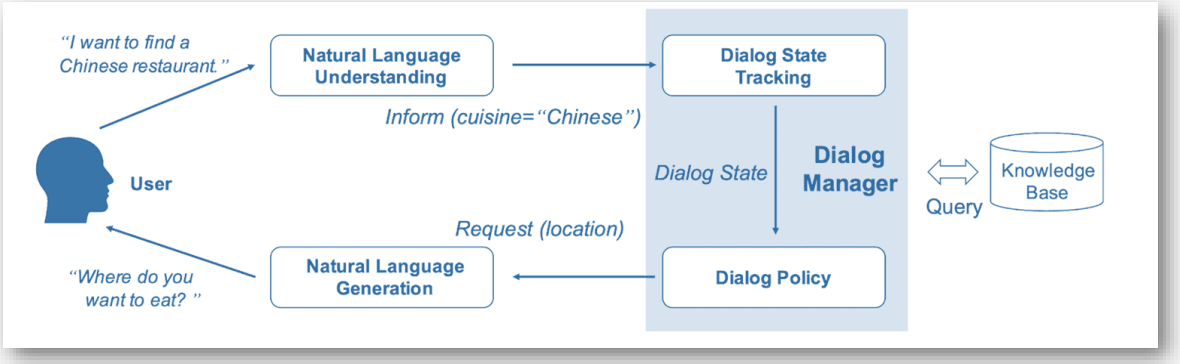
"understanding" & "generating"

the contents of
written texts, and speech, ...



Natural Language Processing (NLP) : The Goal?

General Model of NLP – Conversation Agent



Natural Language Processing (NLP) : Use Cases in Industries

Education
- Machine Translation, Spell Checking and Grammar, etc.

Healthcare
- Speech Recog./Synthesis, Language Gener./Under., Question-Answering

Marketing/Advertising
- Machine Translation, Document Classification, Sentiment analysis, etc.

Pharmaceuticals/BioTech
- Document Classification, NER,Entity-Linking / Knowledge Graphs

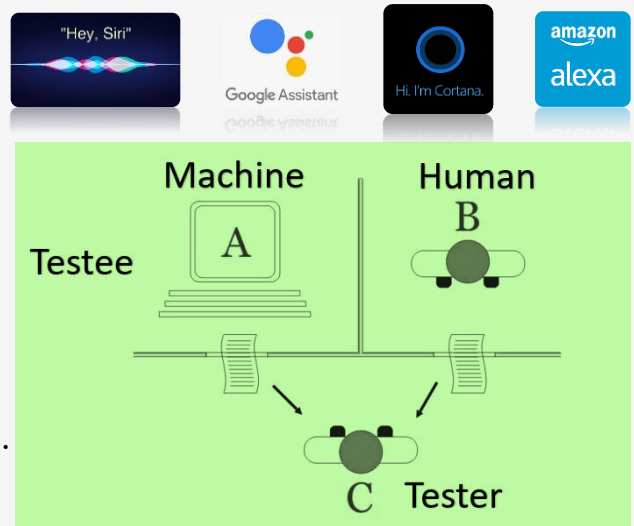
Banking/Finance
- Information Extraction, Text Summarization, NER etc.

Miscellaneous
- Automotive, Defense & National Security, Food, Tourism, Law, etc.

Natural Language Processing (NLP) : Where are we now?

Turing Test

A method of inquiry in **artificial intelligence** (AI) for determining whether or not **a computer is capable** of **thinking like a human being**.



Natural Language Processing (NLP) : Where we are now?

OpenAI GPT-3

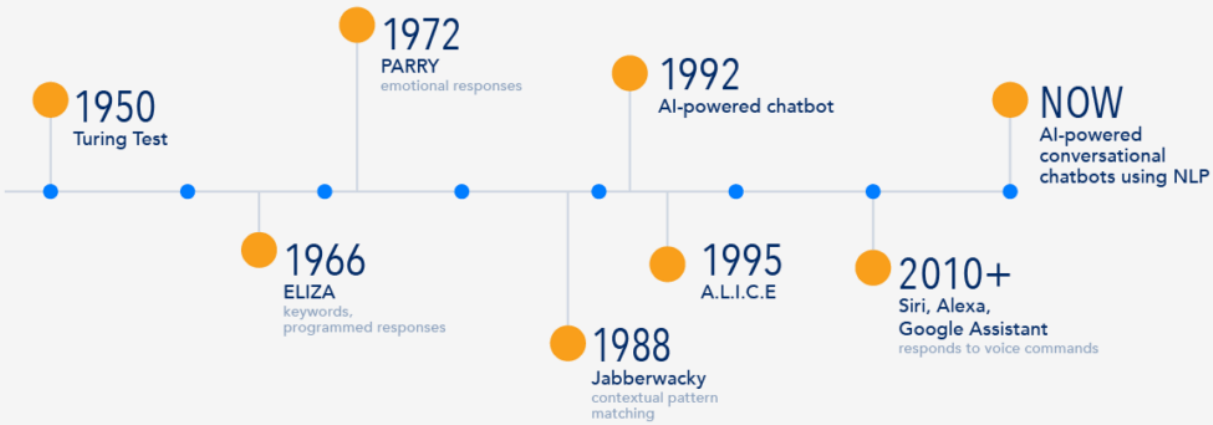
```
Q: What is your favorite animal?
A: My favorite animal is a dog.

Q: Why?
A: Because dogs are loyal and friendly.
```

```
Q: How many rainbows does it take to jump from Hawaii to seventeen?
A: It takes two rainbows to jump from Hawaii to seventeen.
```

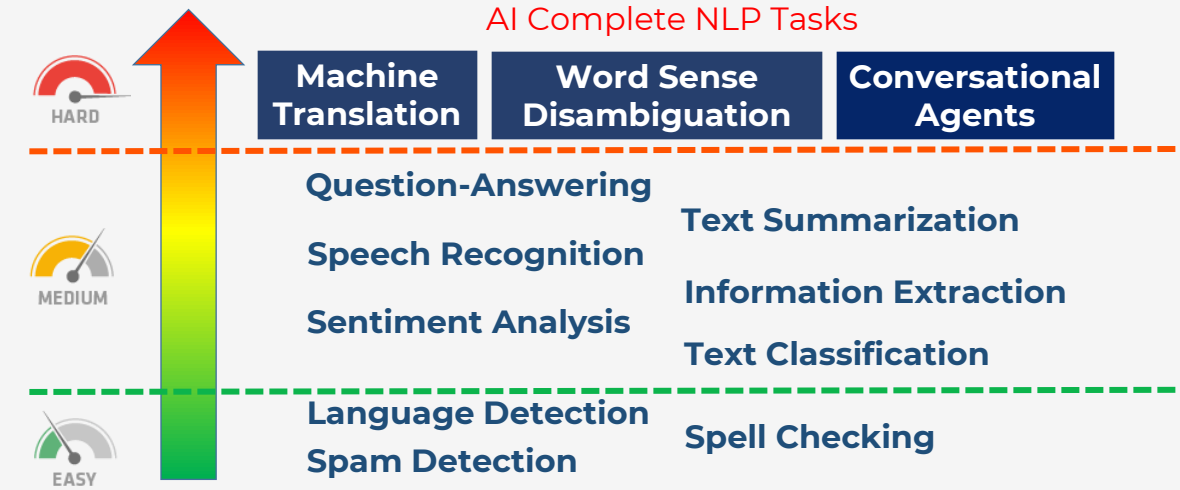
Giving GPT-3 a Turing Test (lacker.io)

https://lacker.io/ai/2020/07/06/giving-gpt-3-a-turing-test.html

Natural Language Processing (NLP) : Where we are now?



Natural Language Processing (NLP) : NLP is Hard!



18

Natural Language Processing (NLP) : NLP is Hard, Why?

Ambiguity

Uncertainty in Meaning

Ayşe and Fatma are **sisters**.

Ayşe and Fatma are **mothers**.

Metaphors

My lawyer is a shark.

Idioms

He is as good as John Doe.

Common Sense/Knowledge

The facts that all humans are aware of

Dog bit man. ✓

Man bit dog. ✗

Creativity / Open Vocabulary

Poems, Genres (Literature) / Languages are generative.

Levesque, Hector, Ernest Davis, and Leora Morgenstern. "The Winograd Schema Challenge." *The Thirteenth International Conference on the Principles of Knowledge Representation and Reasoning* (2012)

19

Natural Language Processing (NLP) : NLP is Hard, Why?

More on Later !

Ambiguity

Common Sense/Knowledge

Creativity / Open Vocabulary

Semantic-Information

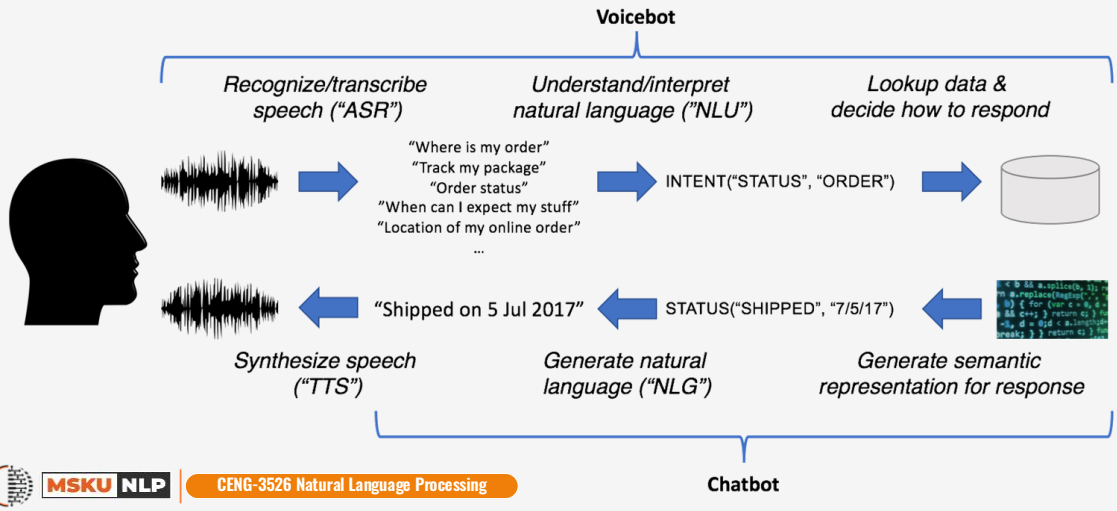
Reducible ?

Signal-Information

20

Natural Language Processing (NLP) : An NLP Walkthrough

General Model of NLP – Conversation Agent

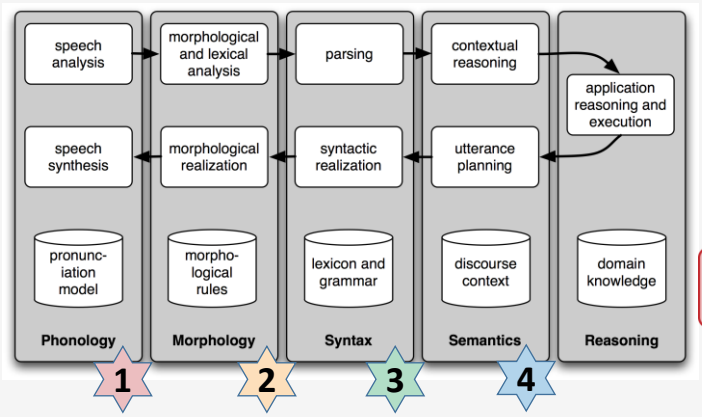


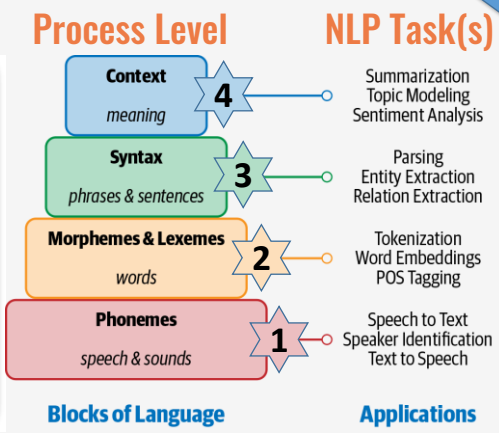
21

Natural Language Processing (NLP) : An NLP Walkthrough

More on Later !

NLP Task Pipeline for Conversation Agent

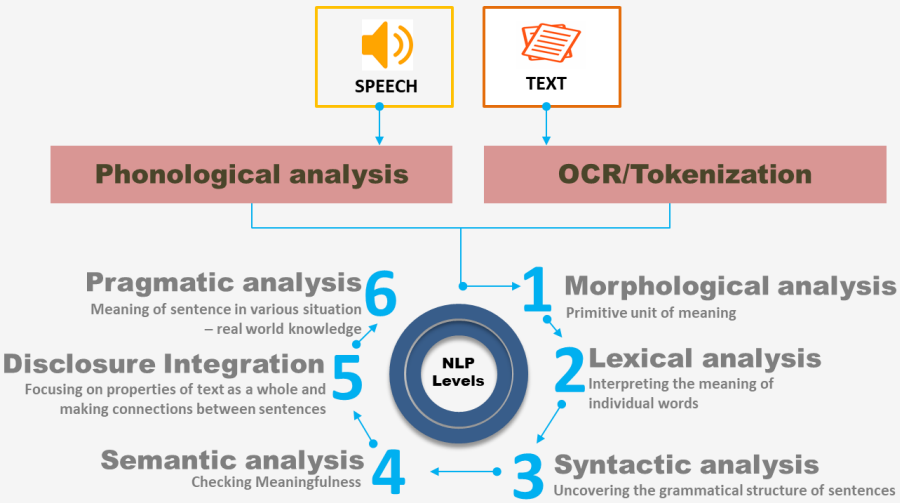




22

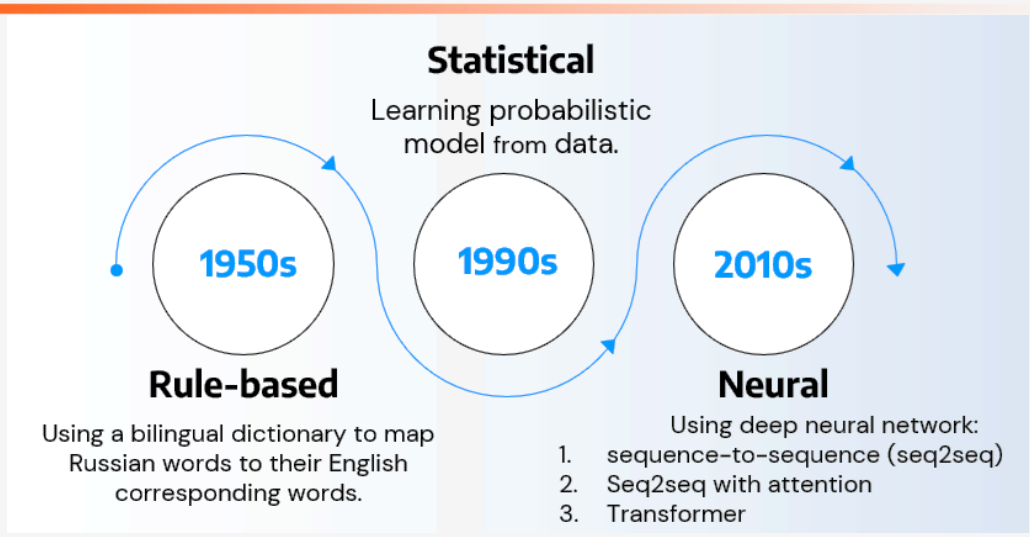
Natural Language Processing (NLP) : Classical NLP Pipeline

More on Later !



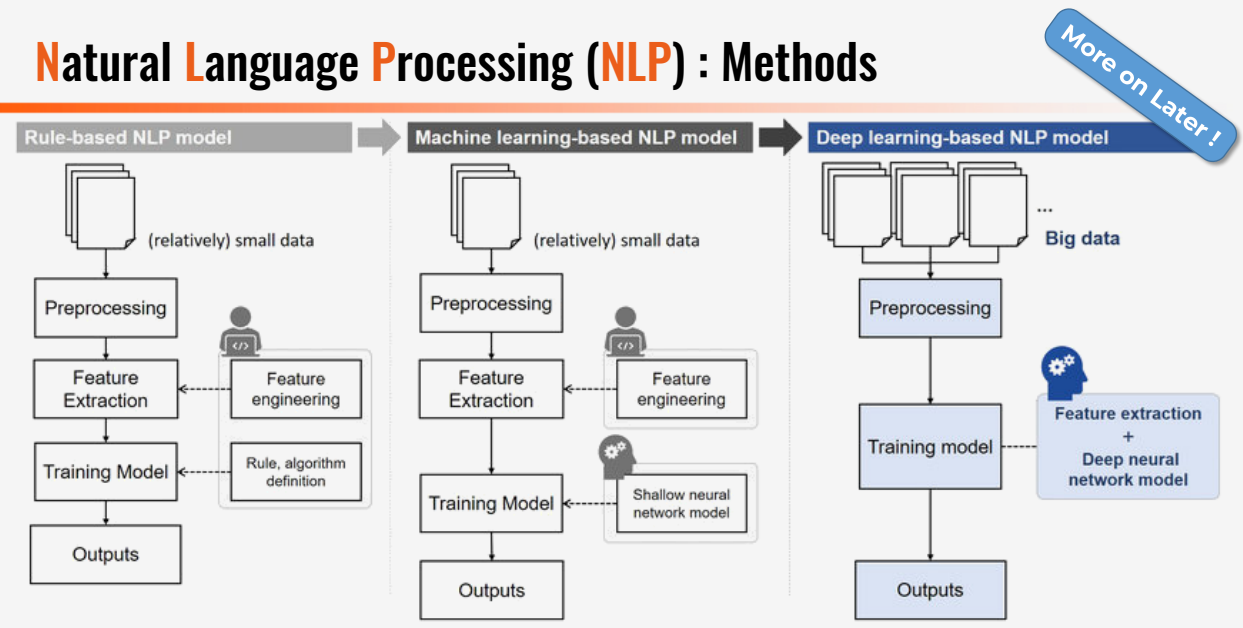
23

Natural Language Processing (NLP) : Methods



24

Natural Language Processing (NLP) : Methods



25