Course Presentation & Logistics

Natural Language Processing
Master in Applied Artificial Ingelligence

Jerónimo Arenas-García, Lorena Calvo-Bartolomé, Jesús Cid-Sueiro

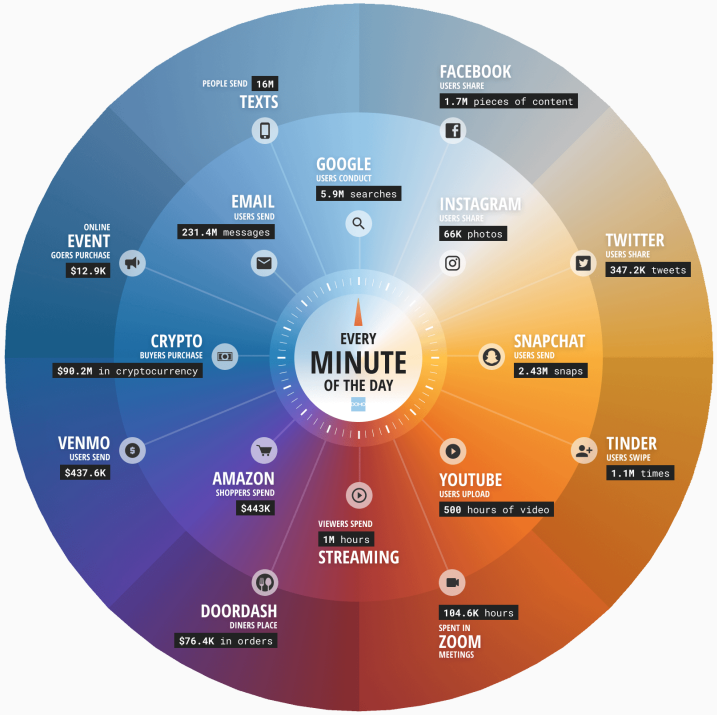
November, 2023

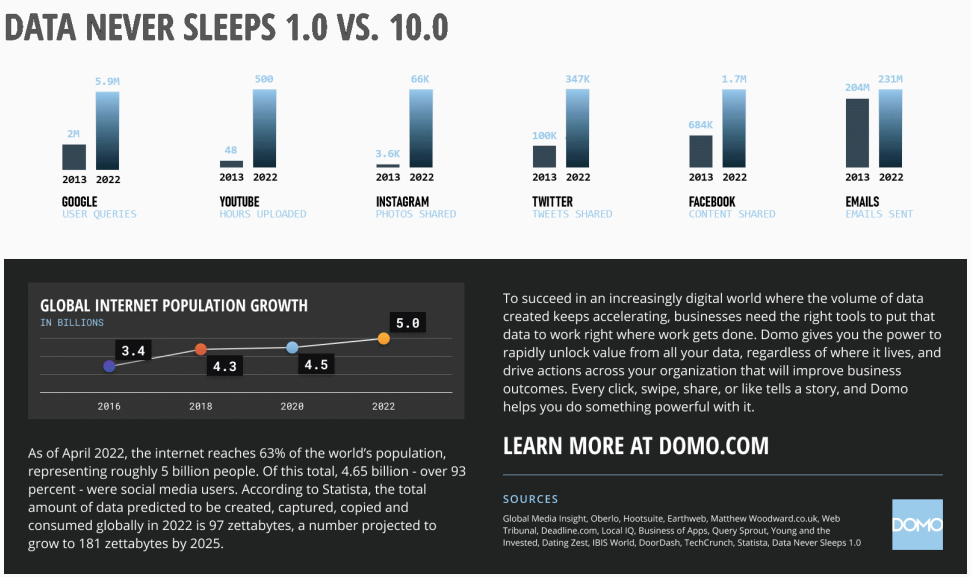
Department of Signal Theory and Communications
Universidad Carlos III de Madrid

Part I

NLP Introduction

Motivation I





https://www.domo.com/data-never-sleeps

Motivation II

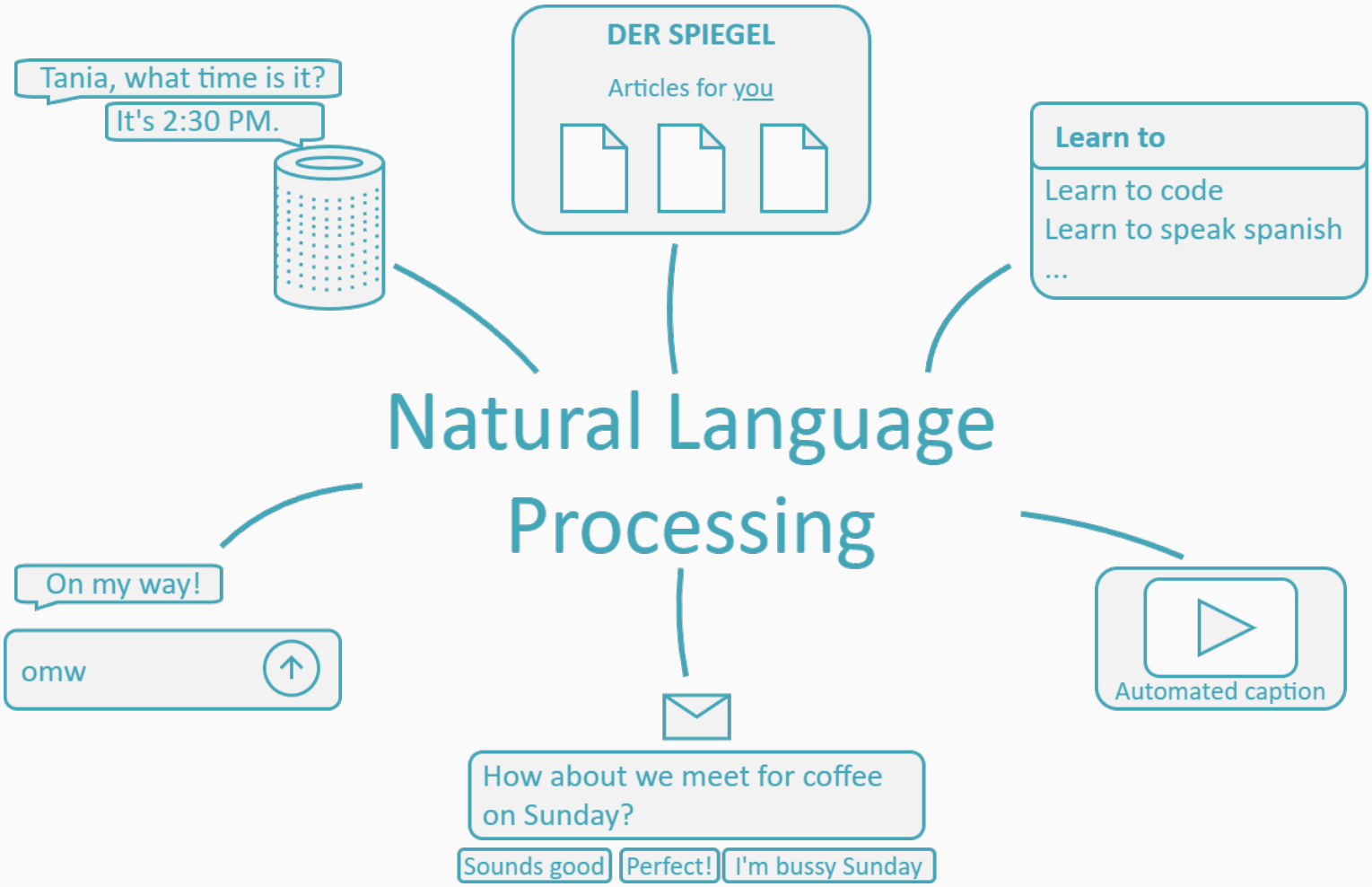
Large amounts of unstructured text data are generated every second

» **We can no longer use the common approach to understand the text and this is where NLP comes in**

What is NLP?

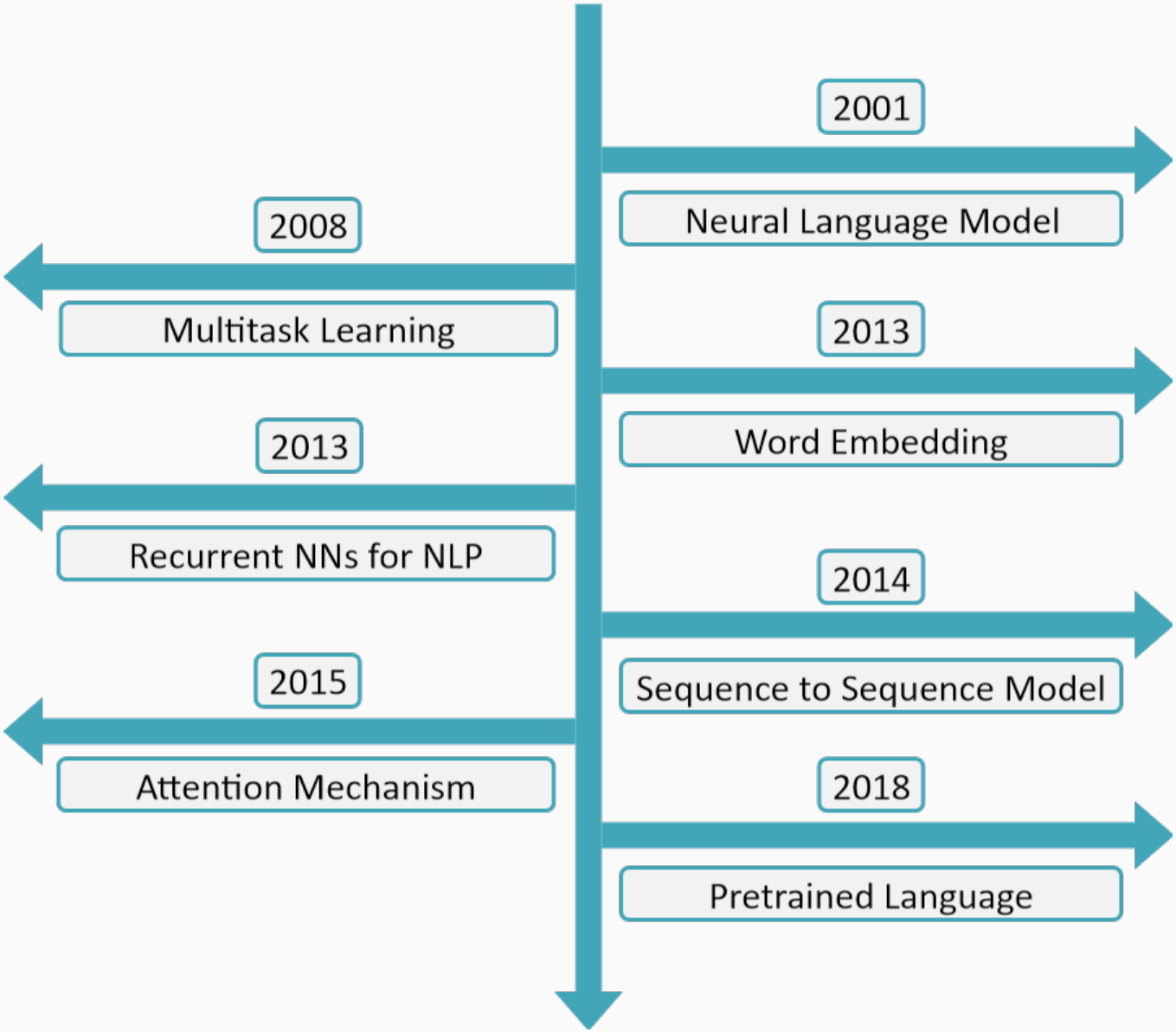
AI's discipline concerned with giving computers the ability to understand written and spoken human languages in the same manner humans do

hoaaaa



Modified from

A walk-through of recent developments in NLP



Some NLP applications

Treffen wir uns am Sonntag zum Kaffee?

Machine translation

The application of computers to the task of translating texts from one natural language to another

¿Nos vemos el domingo para tomar un café?

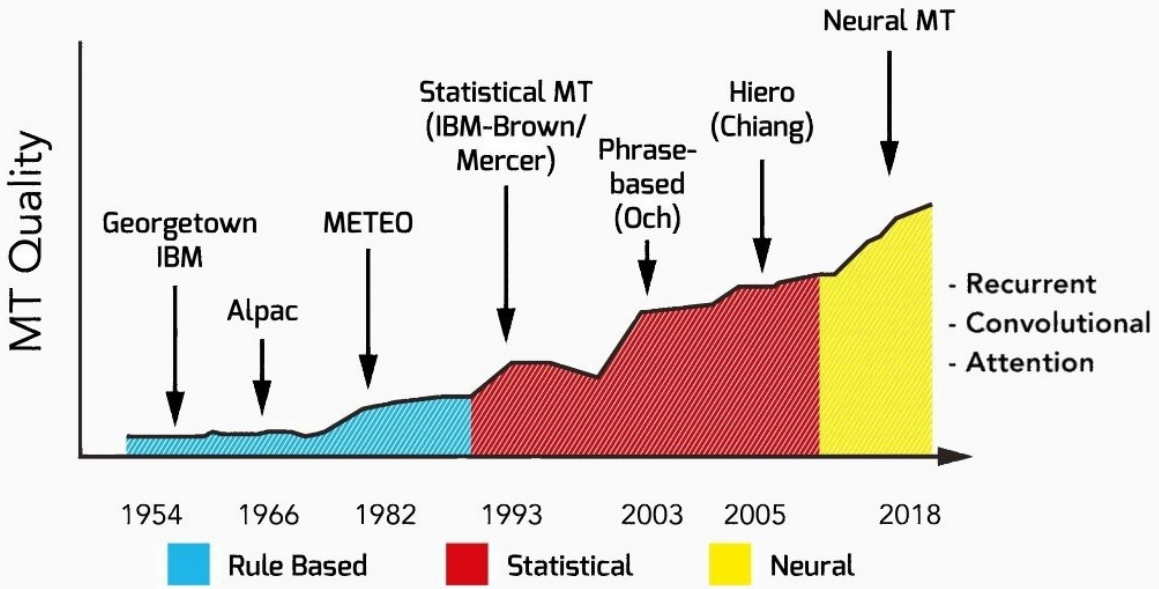
Google Translate

Amazon Translate

DeepL

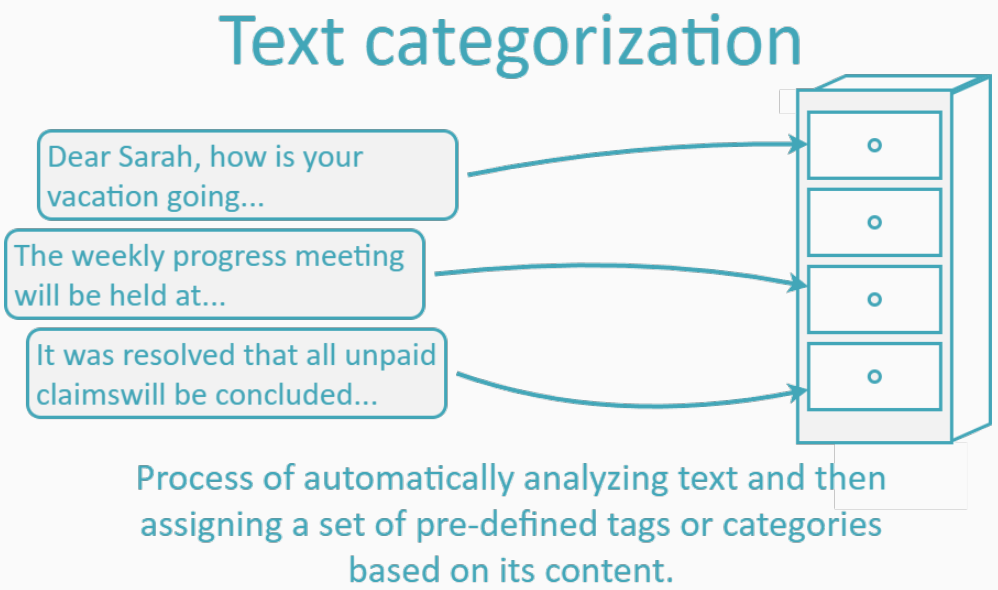
Microsoft Translator

Baidu Translate

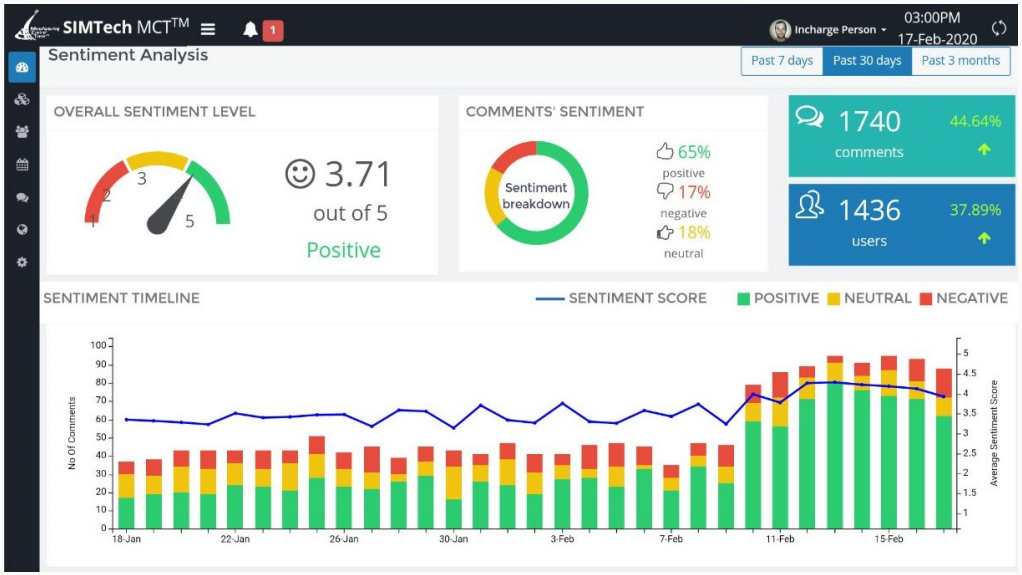


Source: IconicTranslation, modified from

Some NLP applications



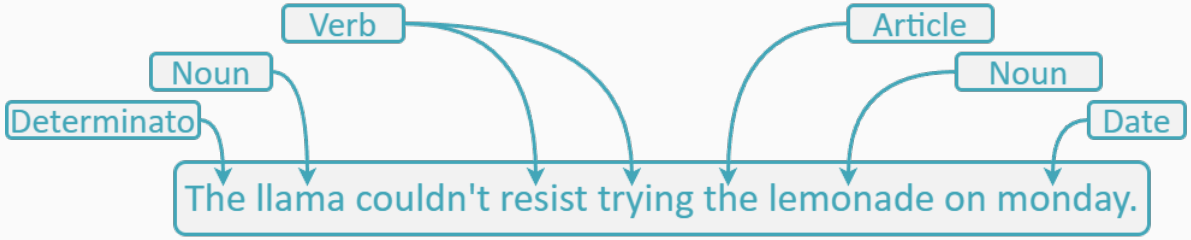
- Spam Filtering
- Sentiment Analysis
- Alert detection
- Automating customer walkthroughs in support systems
- ...



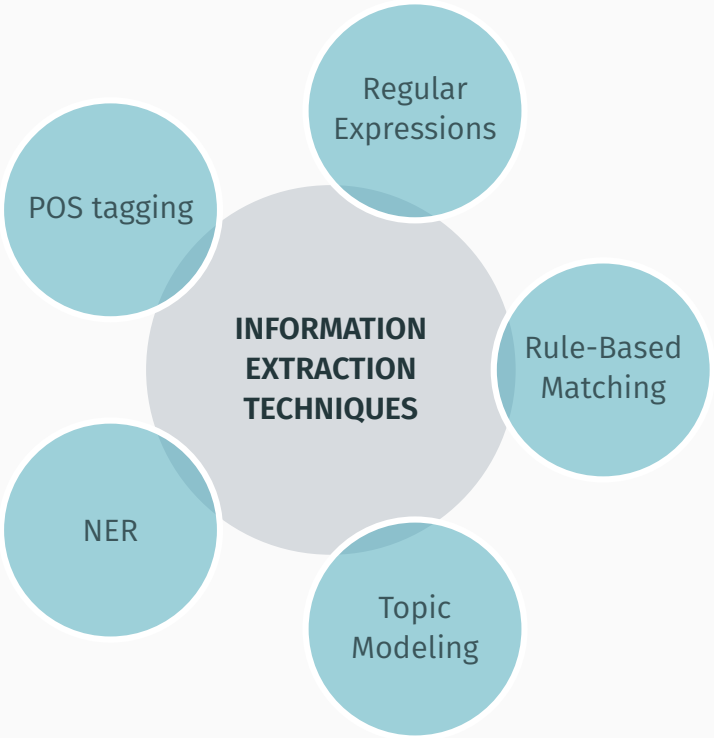
Source: Singapore Institute of Manufacturing Technology (SIMTech). Free demo

Some NLP applications

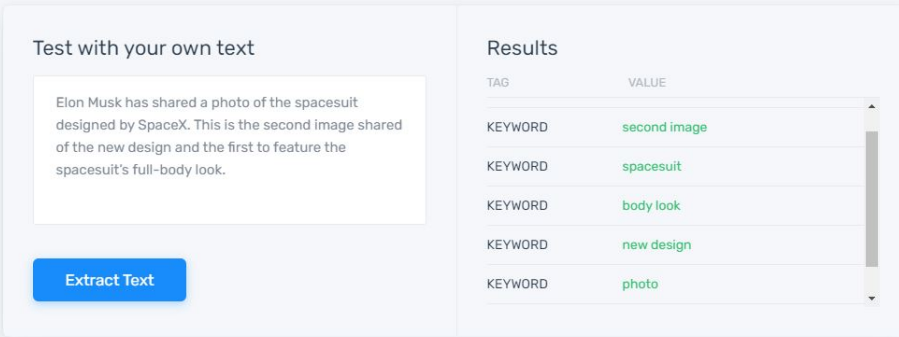
Information extraction



Process of extracting information from unstructured textual sources to enable finding entities (names, places, events, dates, etc.) that can be applied for a variety of purposes, e.g., identify keywords, classifying text items, etc.



Automate tasks with keyword extraction:



Free demo

Some NLP applications

Summarization

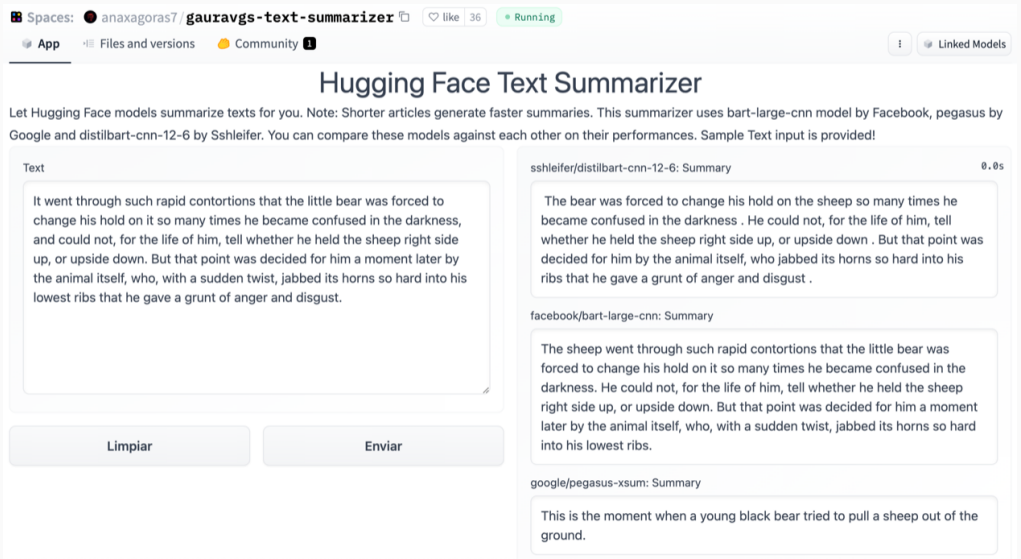
Pink ponies and purple giraffes roamed the field. Cotton candy grew from the ground as a chocolate river meandered off to the side. What looked like stones in the pasture were actually rock candy. Everything in her dream seemed to be perfect except for the fact that she had no mouth.

A little girl had a dream that she was in a candy factory.

Extractive
From sentences within the text.

Abstractive
Possible to contain words not explicitly present in the text.

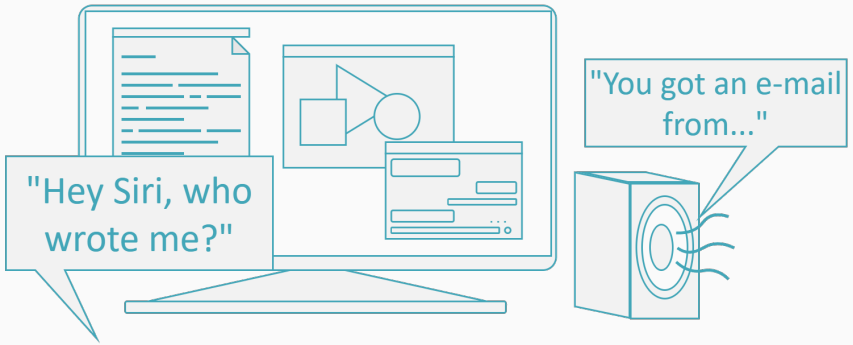
Text summarization is the problem of reducing the number of sentences and words of a document without changing its meaning.



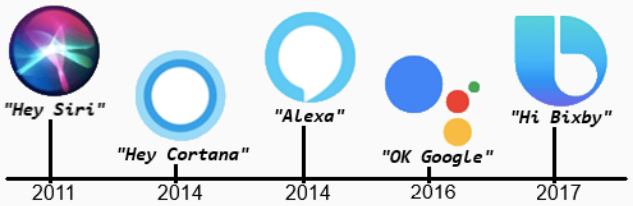
Free demo

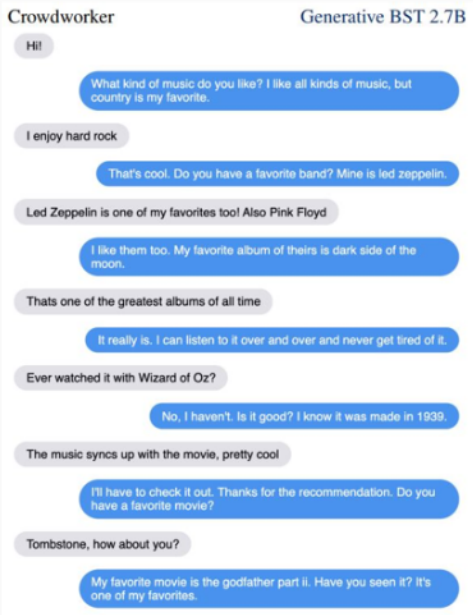
Some NLP applications

Dialogue Systems



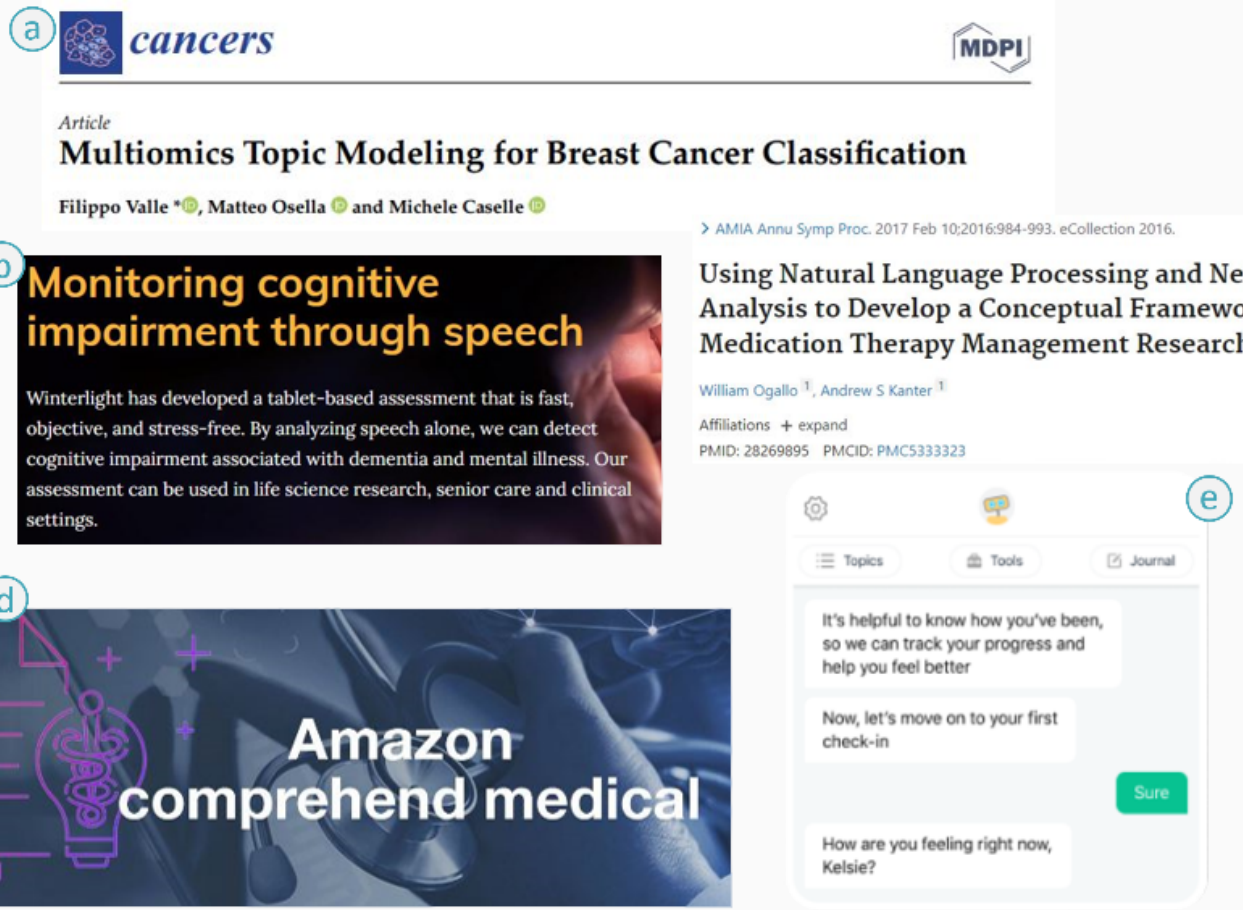
Computer systems intended to converse with a human, employing one or more of text, speech, graphics, haptics, gestures, and other modes for communication on both the input and output channel.





Natural Language Chatbot “Eno” & “BlenderBot”

NLP is booming in the healthcare industry



(a) Multiomics Topic Modeling for Breast Cancer Classification
(b) Winterlight Labs
(c) Using Natural Language Processing and Network Analysis to Develop a Conceptual Framework for Medication Therapy Management Research
(d) Amazon comprehend medical
(e) Woebot Health

Part II

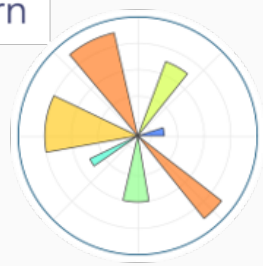
Course logistics

What is this course about?

Technologies

plotly | Dash

Hugging Face



Spark NLP

Table of Contents

Instructors

- **Jerónimo Arenas García (coordinator)**
  - Email: jarenas@ing.uc3m.es
  - Office: 4.2.C06 (Leganés)
  - Tutor Hours: Wednesday and Friday, 16 - 17 (request by email)
- **Lorena Calvo Bartolomé**
  - Email: lcalvo@pa.uc3m.es
  - Office: 4.2.C03 (Leganés)
  - Tutor Hours: Tuesday and Thursday, 12 - 13 (request by email)
- **Jesús Cid Sueiro**
  - Email: jcid@ing.uc3m.es
  - Office: 4.2.D03 (Leganés)
  - Tutor Hours: Wednesday and Thursday, 12 -13 (request by email)

Methodology

- Mixed theory + practical sessions
- Personal work
- Intensive use of Python Notebooks
- Lab exercises will be proposed during each block
- Mid writing quiz + graphs notebook + final open project

Assessment (Continuous evaluation)

1. **Mid-Writing Quiz on December 13th (30%)**
   This quiz will assess all covered material up until December 1st, excluding Graphs.
2. **Solve Python Notebook about Graphs (20%)**
   You'll need to work through the provided Python notebook and hand it in.
3. **Open Final Project (40%)**
   - For the final project, you'll choose a topic related to the subject and submit your proposal by November 29th.
   - On December 13th, there will be a dedicated session for project work and Q&A.
   - The project report is due on January 15th.
4. **Seminary Summary Hand-In on January 10th (10%)**
   Attend the external professor's seminar on January 10th, create a summary of the seminar, and hand it in before the session ends.

Assessment (II)

- **January call**: Blocks 1, 2 and 4 will be discarded, and a new grade (60%) will be based on the performance in a written + practical exam.
- **May call**: In addition to the previous exam (60%) , students will be able to hand in a new final project (40%). They can opt for just taking the exam, carrying out a new project, or both. Any existing grade will be replaced by that achieved in the new assessmnent items taken during the May call.

Final Project

Goals:

- To work on a real application scenario involving NLP
- To design experiments/assess results to validate or reject hypotheses
- To use visualization tools
- Prepare a report, introducing the used models/algorithms, presenting the achieved results, analyzing these results and drawing conclusions

Rules:

- Students will work in teams and hand in their code and a short report

Assessment:

- **Methodology:** 2 points
- **Report Quality + Visualization:** 1,5 points
- **Code:** 0,5 points

In any case, specific conditions will be published together with the final project statement.

Information Resources

- Notebooks provided for the lessons
- Other resources as published in the course description
- Plenty of high-quality material available on the web