

Dr. Fériel BEN FRAJ



NLP:

Basic Concepts

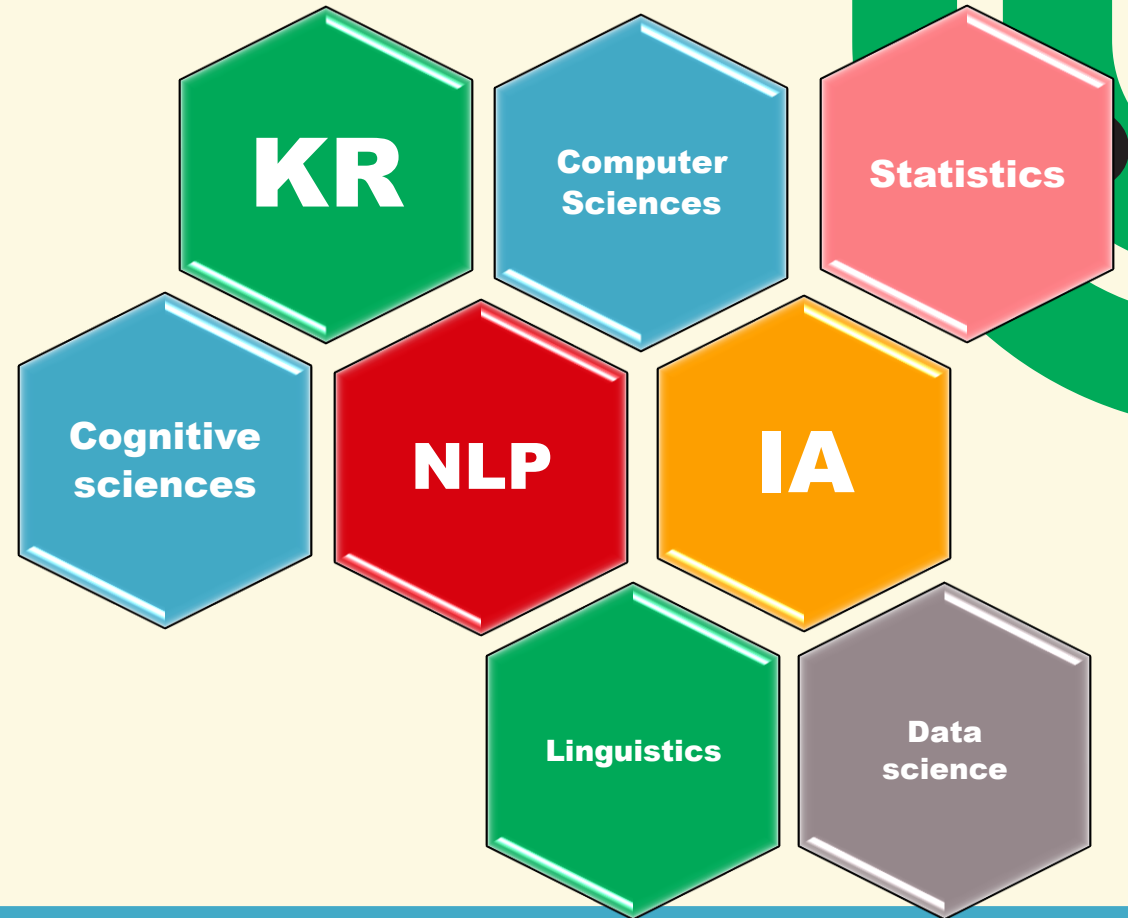


1st March 2023



Definition

Natural Language Processing (NLP) is the use of linguistic knowledge to create intelligent computer systems that can **recognize, understand, interpret,** and **replicate** different forms of human language..





Why NLP?

- Information availability and communication mediation → facilitate Human/Machine and Human/Human communication
- Getting intelligence through communication (Language as a reflective thinking of ideas, reasoning, and common sense)
- Huge pool of available texts

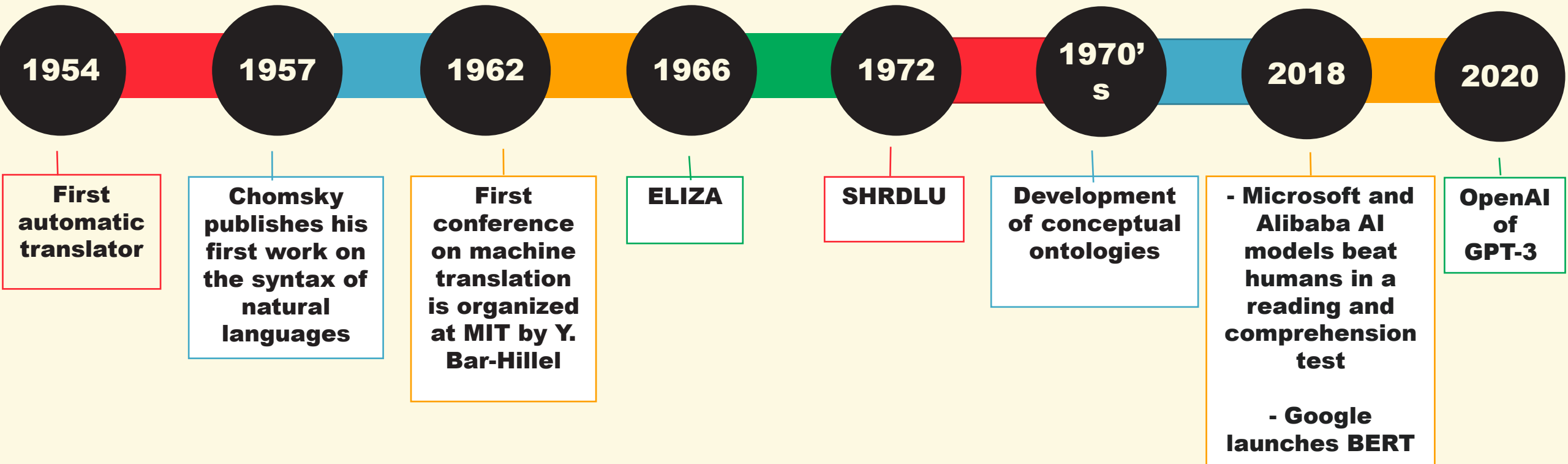


Fields of Interest

- **political, economic:** observation and translation of exchanges (from the cold war to PRISM), intelligence, "language industries"
- **legal:** tracing of sources and plagiarism, jurisprudence
- **cultural:** digitizing and exploiting the literary heritage, preserving the diversity of languages (CE), developing exchanges
- **pedagogical:** all languages, e-learning, automatic correction
- **sociological:** Web 2.0, opinion analysis, trends, e-reputation
- **social:** facilitate mobility, fight against handicaps
- **philosophical :** can machines speak, think, etc
- **scientific, technological:** multidisciplinary, innovation



Background





NLP Celebrities



Alain Turing

(1912-1954)



Noam Chomsky

(1928)



Karen Spärck Jones

(1935-2007)



Christopher Manning

(1965)



Hinrich Schütze

(1962)



Yushua Bengio

(1964)

1st March 2023



Regina Barzilay

(1970)



Nizar Habash

-



Speech Analysis

- From a speech input there are essentially two pieces of linguistic information to be extracted:
 - **Phonemes:** successive sounds that make up words
 - **Prosody:** intonation, rhythm and intensity allowing,
- Phonemes must be grouped together to form words





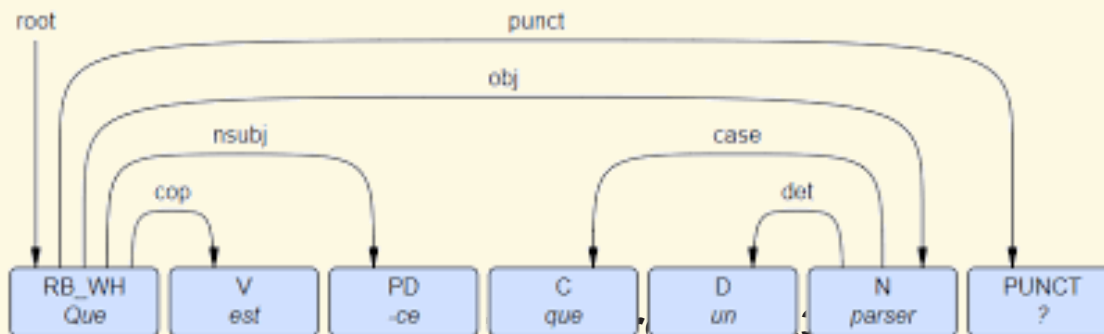
Lexical Analysis

- Study the formation of words from smaller units called morphemes
- Lexical analysis aims to :
 - Recognize the syntactic category and grammatical properties of words
 - Propose a lemmatization
 - Recognize named entities (names of people, organizations, companies, places, quantities, distances, values, dates, etc.)



Syntactic Analysis

- Study of the rules underlying the construction of sentences.
- The purpose of syntactic analysis may be to:
 - Break down the sentence into phrases (chunks or other constituents)
 - Recognize the functions of these phrases
 - Identify syntactic dependencies
 - Propose a hierarchical organization of the syntagms





Semantic Analysis

Semantics concerns the meaning of words and sentences.

- Semantic analysis consists in determining the meaning of an expression or a text by finely analyzing the word combinations and the context.



Pragmatic Analysis

- The purpose of pragmatics is to study the statement (one or more sentences) in an enunciative context.



Examples of Applications

Vocalic Interfaces

Hand-written recognition

Spelling correction

Information retrieval

Language translation

Opinion analysis

Cyber-security (fake news, cyberbullying, hate speech detection)



Main challenges

- 💡 Study of the contexts
- 💡 Word Sens Desambiguation (entre autres les homonymes)
- 💡 Ironic Expressions
- 💡 Error Detection (spelling, syntactic, etc)
- 💡 Dialects processing
- 💡 Low Ressource Languages
- 💡 Domain-specific language

THANK YOU

GRACIAS
ARIGATO
SHUKURIA
JUSPAXAR
DANKSCHEEN
TASHAKKUR ATU
SUKSAMA
GRAZIE
MEHRBANI
BOLZIN
MERCY
BIYAN
SHUKRIA
TINCHI
YAQHANYELAY
GOZAIMASHITA
EPCHABISTO
KOMAPSHANDA
KORAMASHITA
PALMIST

1st March 2023