CHAPTER 5

Semantic Analysis

Semantic Analysis is a subfield of NLP that attempts to understand the meaning of Natural Language

Semantic Analysis

- Parts of semantic analysis
- Semantic analysis approaches
- Main semantic tasks
- Basic semantic units
- Meaning representation

Semantics?

Semantic analysis purpose is understanding the meaning of words, phrases, sentences, and entire texts.

- **Understand** how words and combinations of words convey information, convey relationships, and express nuances
- **Syntax vs Semantics:**
  - Syntax: deals with the structure and rules governing the arrangement of words and phrases in a sentence
  - Semantics: interpretation and meaning derived from those structured words and phrases

Two parts of semantic analysis

- **Lexical Semantic Analysis:** understanding the meaning of each word of the text individually (e.g: using dictionaries)

- **Compositional Semantics Analysis:** understand the meaning of combinations of words (sentence, text).

  - Example: (same words, different meanings)

    Sentence 1: Students love NLP

    Sentence 2: NLP loves students

Semantic analysis approaches

- **Rule-Based:** involves POS tagging, Syntax and Dependency parsing
- **Word embeddings:** maps words into vectors and study semantic relationships (Word2Vec, GloVE)
- **Distributional semantics:** words with similar meanings tend to appear in similar contexts (LSA, LDA)
- **Learning-Based:** uses models trained on annotated corpora

Main semantic tasks

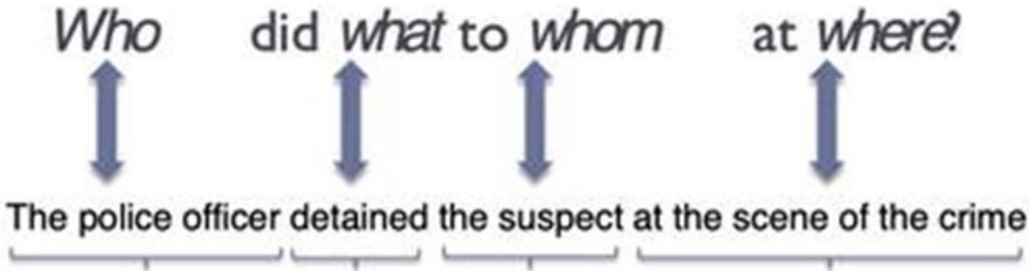
1- Word Sense Disambiguation (WSD): Interpreting the meaning of a word based on the context of its occurrence in a text

Example: He sat on the **bank** and watched the river flow.

- **Some Disambiguation Strategies:**
  - **Contextual Analysis:** Examining surrounding words and phrases to infer the correct meaning.
  - **Lexical Resources:** Utilizing dictionaries and thesauri to identify possible meanings
  - **Machine Learning:** Training algorithms on annotated corpora to predict word senses based on context
  - **Diacritization:** Add diacritical marks to Arabic words

Main semantic tasks

2- Semantic Role Labeling (SRL): identifying the roles that different words play in a sentence, such as identifying the **agent**, **patient**, or **location** in a given **action**

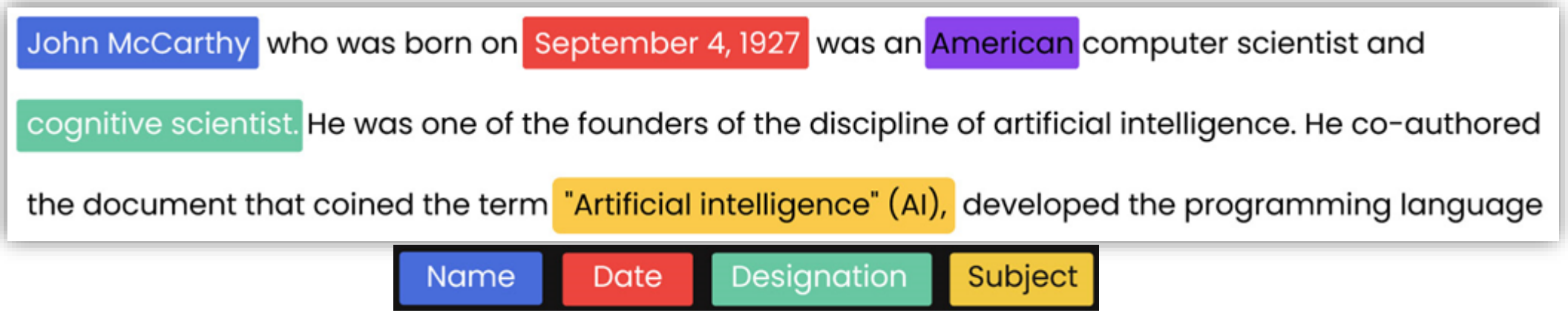


SRL key concepts

Agent: The entity that performs the action	***John*** kicked the ball
Patient: The entity that is affected by the action	John kicked the ***ball***
Instrument: The entity used to perform the action	She cut the bread with a ***knife***
Experiencer: The entity that experiences or perceives something	***Mary*** heard a strange noise
Theme: The entity that is moved or the topic of the action	She gave the ***book*** to him
Location: The place where the action occurs	He stayed in the ***house***
Source: The starting point of the action	She came from the ***village***
Goal: The endpoint of the action	He walked to the ***park***

Main semantic tasks

3- Named Entity Recognition (NER): Identifying and categorizing entities such as **names of people**, **organizations**, **products, locations, events, date and time**



- **NER goes through:**
  - POS Tagging
  - Dependency parsing

Main semantic tasks

3- Named Entity Recognition (NER):

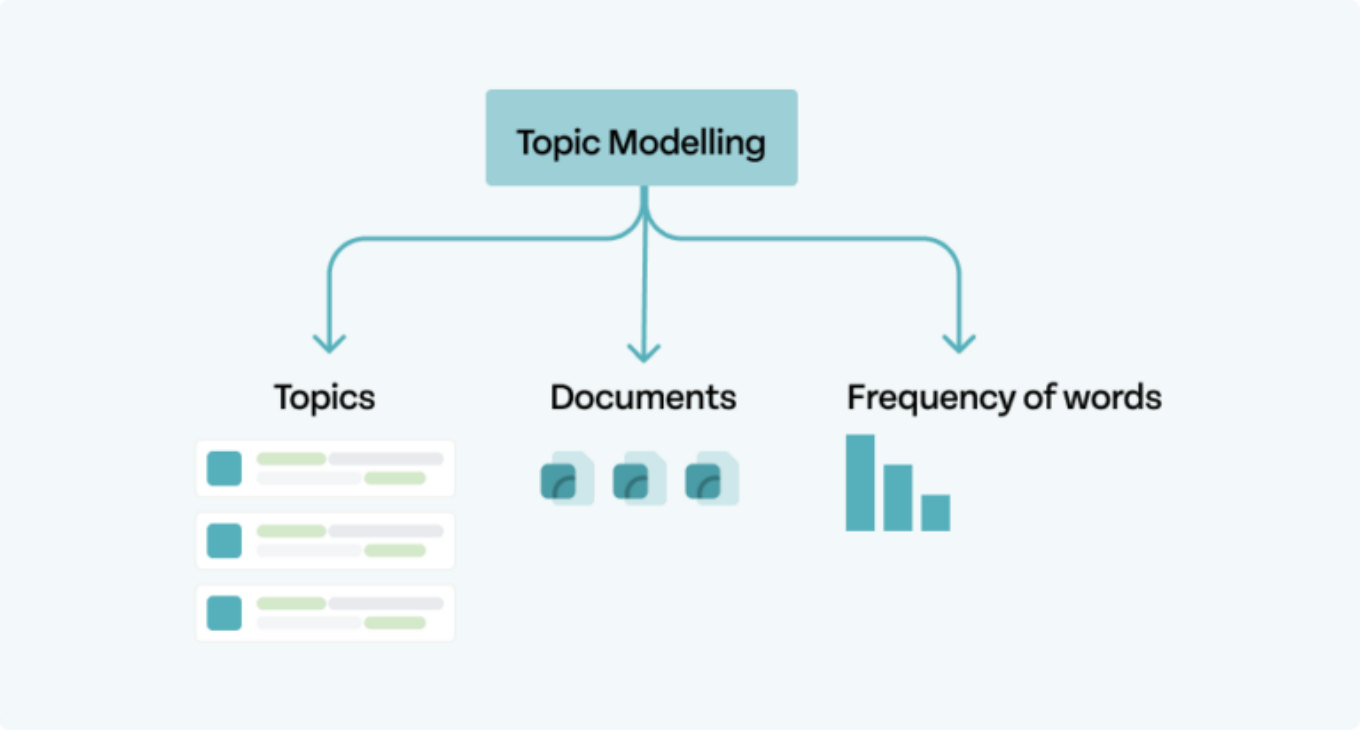
- **Main NER labels**

PERSON	People, including fictional
NORP	Nationalities or religious or political groups
FAC	Buildings, airports, highways, bridges, etc.
ORG	Companies, agencies, institutions, etc.
GPE	Countries, cities, states
LOC	Non-GPE locations, mountain ranges, bodies of water
PRODUCT	Objects, vehicles, foods, etc. (Not services.)
EVENT	Named hurricanes, battles, wars, sports events, etc.
WORK_OF_ART	Titles of books, songs, etc.
LAW	Named documents made into laws
LANGUAGE	Any named language
DATE	Absolute or relative dates or periods
TIME	Times smaller than a day
PERCENT	Percentage, including "%"
MONEY	Monetary values, including unit
QUANTITY	Measurements, as of weight or distance
ORDINAL	"first", "second", etc.
CARDINAL	Numerals that do not fall under another type

Main semantic tasks

4- Other semantic tasks:

- Sentiment analysis
- Topic modeling
- Document embeddings



Semantic elements

Hyponymy (النوع والجنس) Refers to a term that is an instance of a generic term	'*Color*' is a hypernymy while '*grey*', '*blue*', '*red*', etc, are its hyponyms
Homonymy (المشترك اللفظي) Refers to two or more lexical terms with the same spellings but completely distinct in meaning	'*Rose*' might mean '*the past form of rise*' or '*a flower*'
Synonymy (الترادف) Two or more lexical terms that might be spelt distinctly but have the same or similar meaning	*(Job, Occupation), (Large, Big), (Stop, Halt)*
Antonymy (التضاد) Refers to a pair of lexical terms that have contrasting meanings	*(Day, Night), (Hot, Cold), (Large, Small)*
Polysemy (التعدد الدلالي) Refers to lexical terms that have the same spelling but multiple closely related meanings	'*man*' may mean '*the human species*' or '*a male human*' or '*an adult male human*'
Meronomy (الجزء والكل) Refers to a relationship wherein one lexical term is a constituent of some larger entity	'*Wheel*' is a meronym of '*Automobile*' (which is holonym)
Collocations (التلازم اللفظي) Words that frequently appears together	Fully aware, Low cost, ..

Meaning representation

- **Building blocks**

- **Entity:** A particular unit or individual in specific such as a person or a location (Mohamed, Algiers,..)
- **Concept:** A generalization of entities (Country, Student,..)
- **Relation:** Relationships between various entities and concepts
  - Example: [Semantic analysis] **is_subtopic_of** [NLP] **is_subfield_of** [AI]
- **Predicate:** The verb or relation that expresses an action, state, or property
  - Example: *eats* in the sentence "John eats an apple.“
- **Arguments:** Entities participating in the action or state described by the predicate
  - Example: In "John eats an apple“, *John* is the subject (agent), *an apple* is the object (patient)

Meaning representation

- **Approaches**
  - First-Order Predicate Logic
  - Semantic nets
  - Conceptual graphs & Ontologies