

Representación del conocimiento y razonamiento

SEMINARIO 1: SPACY



1. Introducción a NLP
 1. Concepto y aplicaciones
 2. Ambigüedad
 3. Etapas
2. Introducción a Spacy
 1. Instalación
 2. Etapas (en notebook)
3. Ejercicios



Un Sistema Inteligente

Sabiduría ≠
Inteligencia

Representación Conocimiento

Ontologías, reglas, acciones, procesos,...

Inferencia, Razonamiento (lógicos)

Inducción, Deducción, Abducción,...

Modelos Razonamiento

Temporalidad (Modelos Temporales)

Incertidumbre (Redes Bayesianas)

Imprecisión (Razonamiento Difuso)

Búsqueda de Soluciones

Heurística, Metaheurística, Métodos Bioinspirados, Inteligencia Social, etc.

Aprendizaje

Inductivo, Estadístico, ...

Sistemas Sociales

IA distribuida, MultiAgentes. Roles, Interacción y colaboración, Reputación, etc.

Respuesta, Decisión,
Solución Inteligente

Interacción, Interfaz Humano-Máquina

Visión, Lenguaje natural, Síntesis voz, Sensorización, Reconocimiento...

Aprendizaje, Deep Learning

Modelos conexionistas
Redes Neuronales

■ Concepto

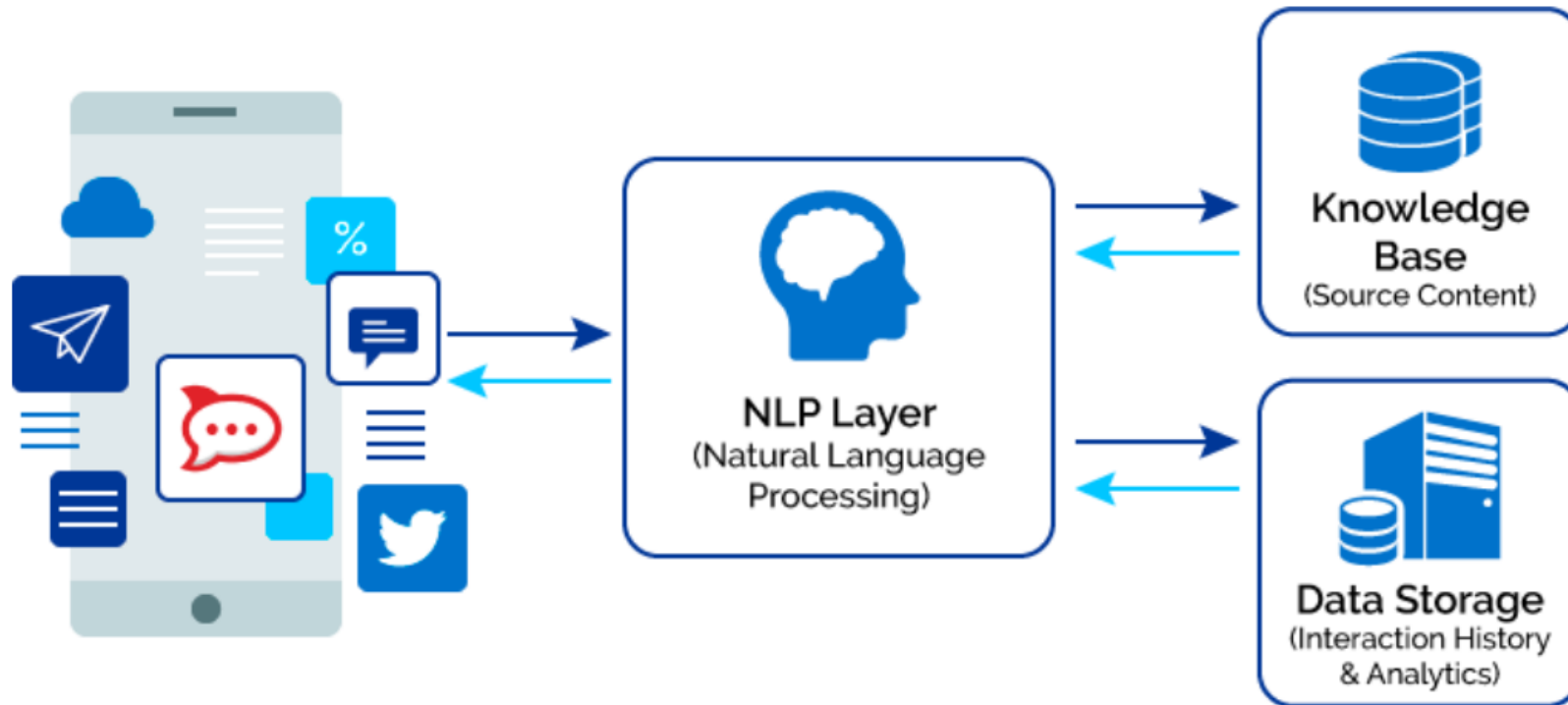
El NLP (Natural Language Processing) se ocupa de la formulación e investigación de mecanismos computacionalmente eficaces para la comunicación entre personas y máquinas por medio del lenguaje natural.



■ Aplicaciones



Interpretación de un texto



De <https://medium.com/trang-nguyen/an-introduction-to-natural-language-processing-b3de96cff299>

El lenguaje natural es extremadamente rico en forma y estructura.

Y es muy ambiguo. Hay diferentes niveles de ambigüedad:

- Ambigüedad léxica: a nivel de palabra

Can you help me with the can?

Figure 3: NLP example sentence with the text: "Can you help me with the can?"

In the sentence above, we can see that there are two “can” words, but both of them have different meanings. Here the first “can” word is used for question formation. The second “can” word at the end of the sentence is used to represent a container that holds food or liquid.



- Ambigüedad sintáctica: una frase puede interpretarse de formas distintas

I saw a man on a hill with a telescope.

Figure 2: NLP example sentence with the text: "I saw a man on a hill with a telescope."

These are some interpretations of the sentence shown above.

- There is a man on the hill, and I watched him with my telescope.
- There is a man on the hill, and he has a telescope.
- I'm on a hill, and I saw a man using my telescope.
- I'm on a hill, and I saw a man who has a telescope.



- Ambigüedad referencial: cuando se usan pronombres

Por ejemplo: “Él está cansado”

¿Quién exactamente está cansado?

- Un texto de entrada puede tener diferentes significados.
- Varias entradas pueden significar lo mismo.



Análisis léxico

- Se divide el texto en párrafos, oraciones y palabras. Consiste en identificar y analizar la estructura de las palabras.

Análisis sintáctico

- Consiste en analizar las palabras en una oración para extraer la relación entre ellas.

Análisis semántico

- Se estudia el significado exacto de cada palabra y se analiza el significado de la oración.

Integración del discurso

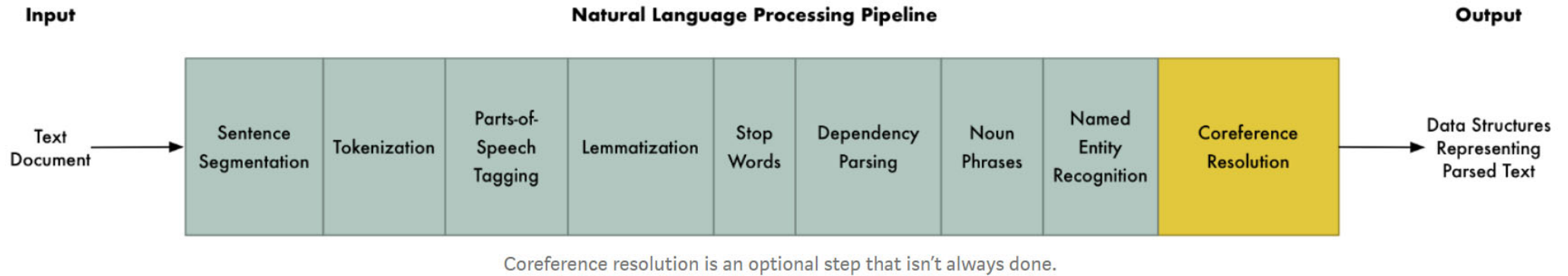
- Se considera el contexto del texto, analizando el significado de cada oración con respecto a la oración anterior y a la oración siguiente.

Análisis pragmático

- Se ocupa de la comunicación como un todo, de forma que se pueda derivar un uso significativo del lenguaje en diferentes situaciones (por ejemplo, ironía)

<https://medium.com/towards-artificial-intelligence/natural-language-processing-nlp-with-python-tutorial-for-beginners-1f54e610a1a0>



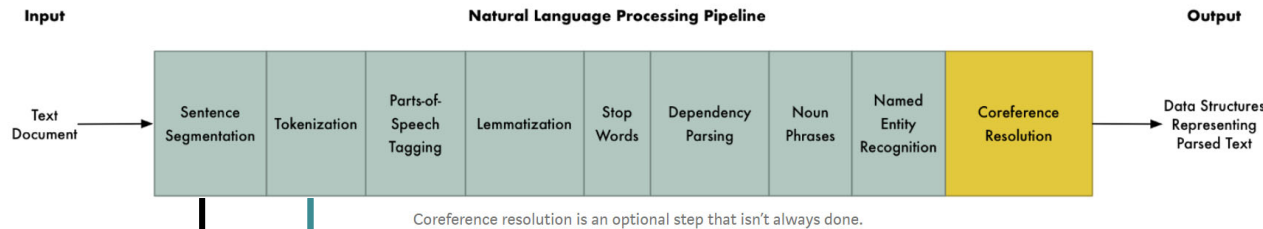


London is the capital and most populous city of England and the United Kingdom. Standing on the River Thames in the south east of the island of Great Britain, London has been a major settlement for two millennia. It was founded by the Romans, who named it Londinium.

(Source: Wikipedia article "London")

<https://medium.com/@ageitgey/natural-language-processing-is-fun-gaobff37854e>

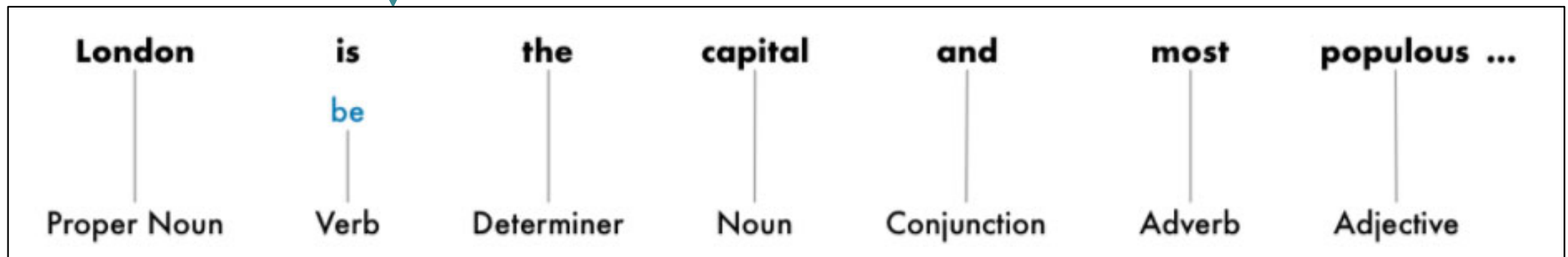
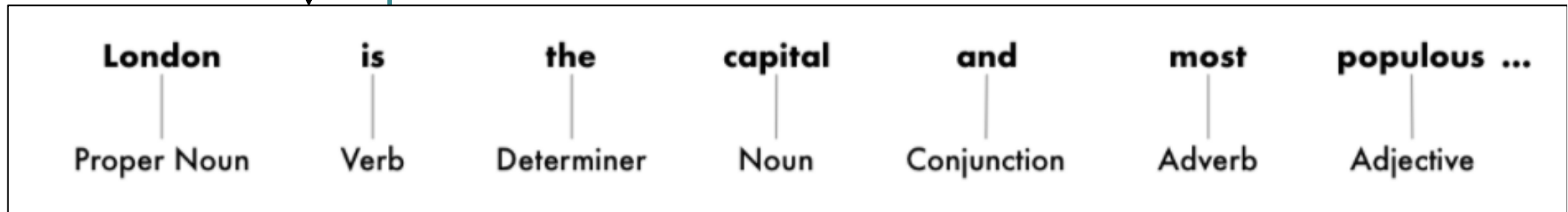
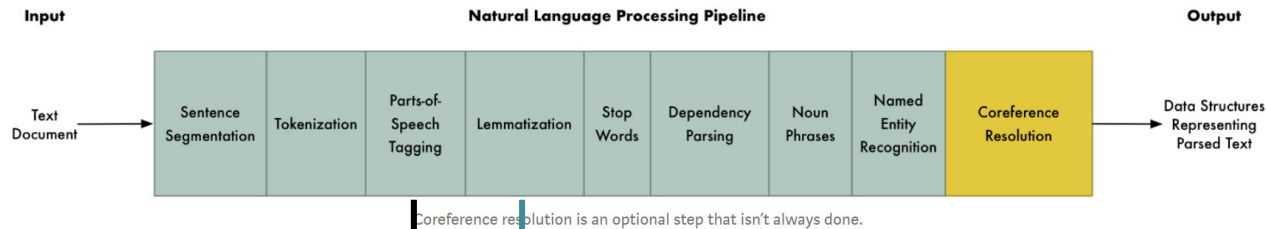


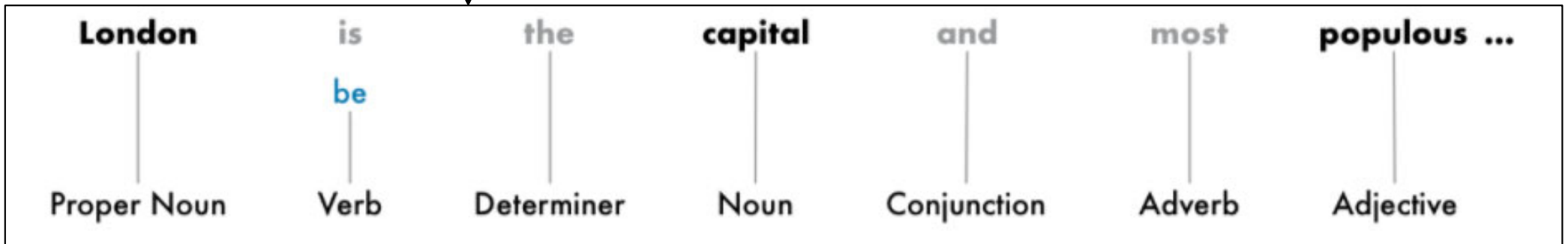
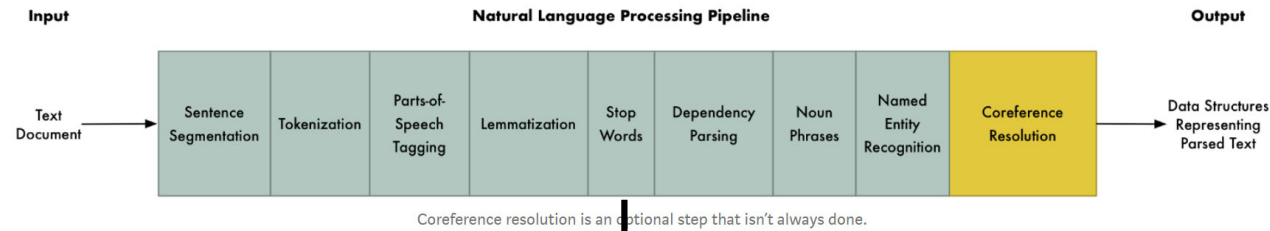


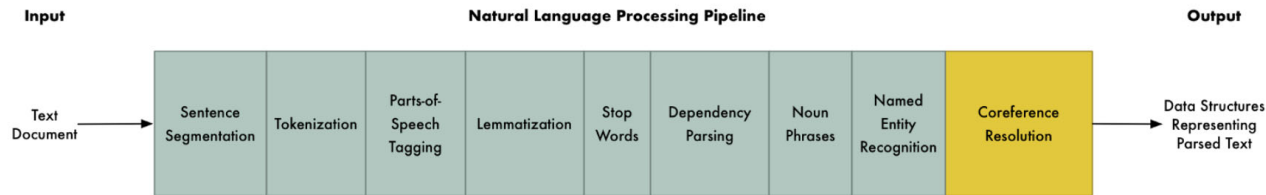
1. "London is the capital and most populous city of England and the United Kingdom."
2. "Standing on the River Thames in the south east of the island of Great Britain, London has been a major settlement for two millennia."
3. "It was founded by the Romans, who named it Londinium."

"London", "is", "the", "capital", "and", "most", "populous", "city", "of",
"England", "and", "the", "United", "Kingdom", "."

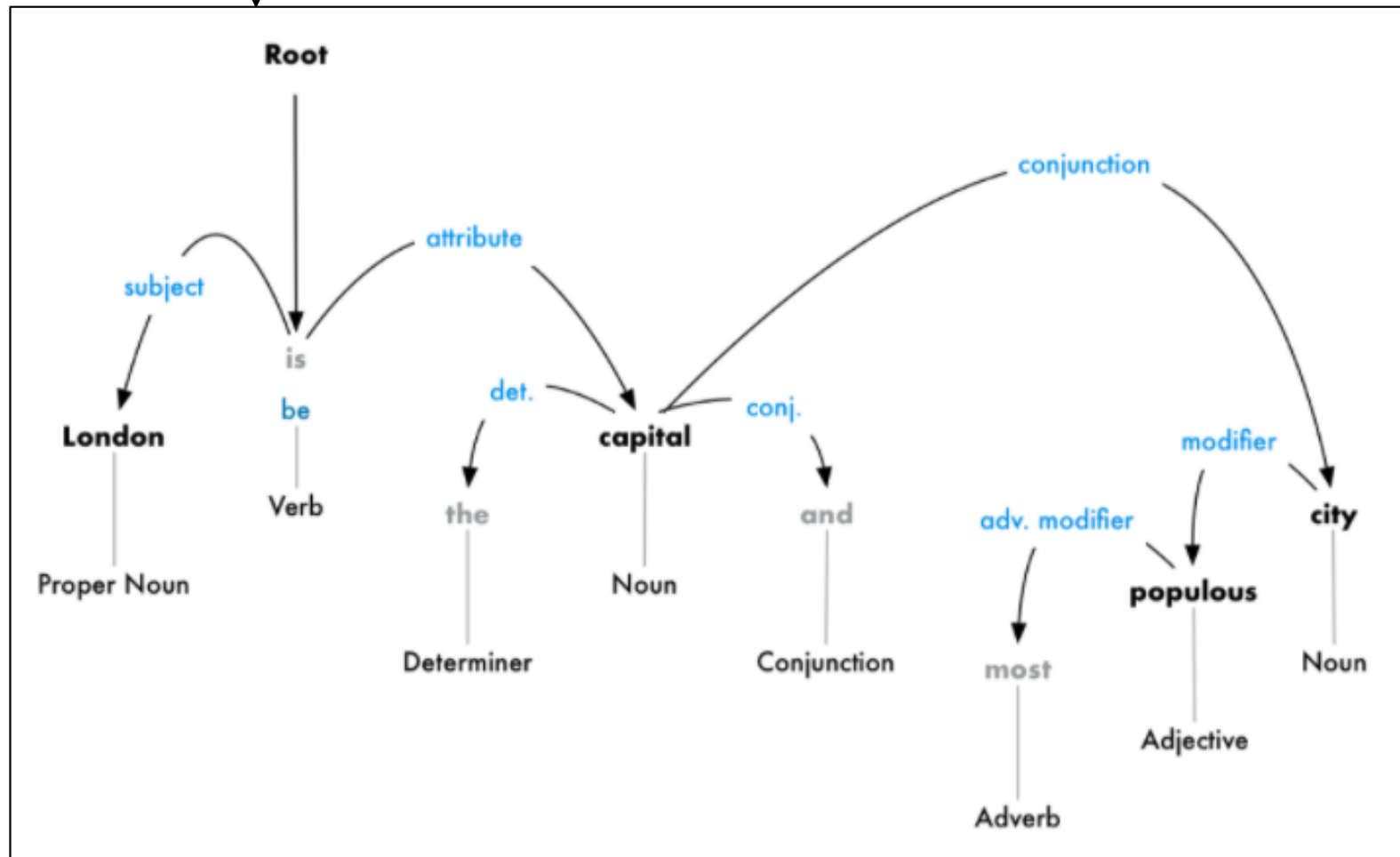








Coreference resolution is an optional step that isn't always done.

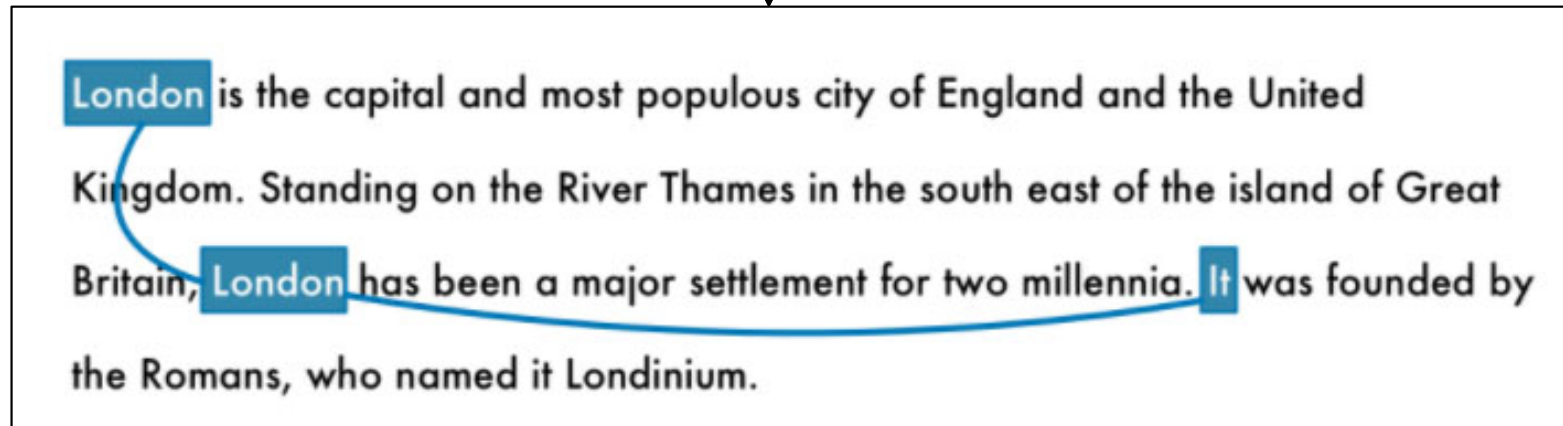
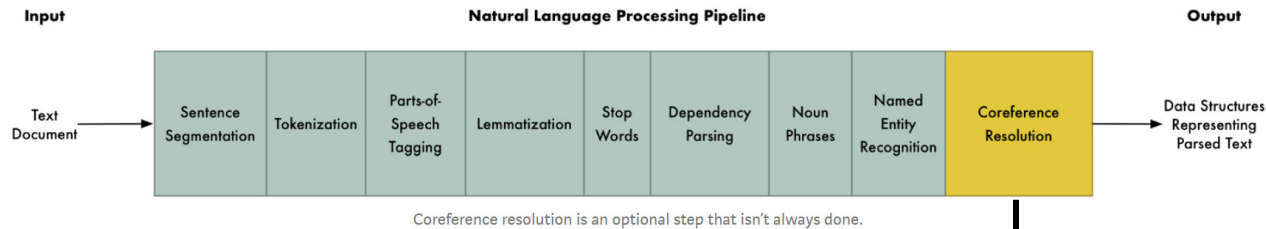


London is **the capital** and **most populous city ...**

Proper Noun Verb Noun Conjunction Noun

Grado en Ciencia de Datos





- spaCy is a free, open-source library for advanced Natural Language Processing (NLP) in Python.
- spaCy is designed specifically for production use and helps you build applications that process and “understand” large volumes of text. It can be used to build information extraction or natural language understanding systems, or to pre-process text for deep learning.

spacy.io



- spaCy proporciona modelos pre-entrenados en más de 15 idiomas distintos: chino, alemán, inglés, español, griego, francés, italiano, japonés, portugués, etc.
- Instalación

```
$ pip install -U spacy
```

```
$ python -m spacy download en_core_web_sm
```

```
$ python -m spacy download es_core_news_sm
```

Modelo en inglés

Modelo en español

- Puede ser útil:

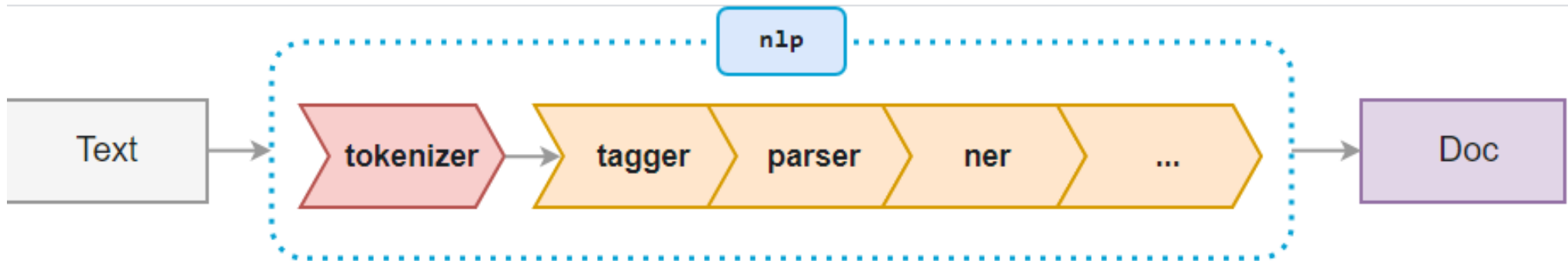
<https://www.datacamp.com/cheat-sheet/spacy-cheat-sheet-advanced-nlp-in-python>



Cargar el modelo, en este caso, en español

```
import spacy
nlp = spacy.load("en_core_web_md")
```

Pipeline en spaCy



```
text = """The Lonja is an emblematic building of the city and one of the most famous civ  
It was declared a National Historic and Artistic Monument and was made a World  
doc = nlp(text)
```



- Ejemplos de uso para cada fase del proceso
 - En el notebook Seminario 1

