

## Tarea: POS tagging

Paso 1: descarga uno de los siguientes temas propuestos:

- [Amazon](#)
- [Meta](#)
- [Coca cola](#)
- [Bimbo](#)
- [Google](#)

Paso 2: a partir del tema seleccionado, responde las siguientes preguntas:

- ¿Cuántas palabras hay en el texto? (descarta signos de puntuación)
- ¿Cuántas palabras diferentes existen?
- Después de lematizar, ¿cuántas palabras diferentes existen?
- ¿Cuáles son las 20 palabras únicas más frecuentes - gráficas? (elimina stopwords)
- ¿Cuál es el número promedio de palabras por oración?
- ¿Cuál es la frecuencia de sustantivos, adjetivos y verbos en el texto?
- ¿Cuáles son las entidades nombradas más frecuentes? ¿de qué tipo son?
- ¿Cuáles fueron las 10 localizaciones más frecuentes en el texto?
- ¿Cuáles fueron las 10 personas nombradas más frecuentemente en el texto?
- Selecciona un párrafo de tu preferencia y visualiza las entidades nombradas encontradas.
- De las 20 palabras más frecuentes en el texto, selecciona 5 palabras y usa WordNet. Identifica los lemmas de esas palabras y su definición.

Para esta tarea, utiliza una de las siguientes herramientas:

- Freeling
- Stanza
- Tree tagger