

Guía para reproducir trabajo

- Tener instaladas las librerías utilizadas en el notebook.
- Cambiar el rango de fechas a las que se quiera hacer el web scraping.

```
start_year = 2023
start_month = 11
start_day = 1
end_year = today.year
end_month = today.month
end_day = today.day
```

ElasticSearch:

- ```
-client = Elasticsearch('http://localhost:9200')
```
- ElasticSearch abierto en ese puerto y el jupyter en local o cambiar esa línea de código si se abren ambos en docker.

## **GraphDB:**

- Hay que tener el archivo boe.ttl en el mismo directorio de archivos donde se ejecute el notebook ya que es la ontología base con la que después se genera el RDF.
- ```
-sparql = SPARQLWrapper("http://localhost:8000/repositories/boe_rdf")
```
- Tener el GraphDB abierto en el docker en ese puerto.
- Hay que tener un repositorio creado en graphDB con nombre *boe_rdf* y subir el RDF generado por el código cuyo nombre es *boe_ejemplo.ttl*.

Usamos las versiones del github de gestbd tanto en ElasticSearch como en GraphDB.