

10.2 Graph

El segundo de los plugins que aporta X-pack nos ayuda a representar y relacionar sistemas de Grafos a través de una nueva ventana en Kibana.

Descarga de los datos de prueba

Al igual que en el caso de *Machine Learning* visto en el tema anterior, el primer paso para montar esta prueba de concepto será la descarga de los datos necesarios. En nuestro caso la base de datos que se va a descargar e instalar son valoraciones de miles de usuarios a distintas películas y se verá cómo **Graph** es capaz de relacionar dichos datos buscando comportamientos o en este caso "gustos" parecidos.

Para la descarga de los datos se clonará directamente el directorio de pruebas y se accederá posteriormente al apartado de Graph:

```
git clone https://github.com/elastic/examples.git
cd examples/Graph/movie_recommendations
```

Una vez ahí, comprobamos que este proyecto requiere de Python 3 y algunos módulos extras, por lo que lo más fácil será instalar un entorno virtual donde poder instalar todo lo necesario y hacer la carga de datos de la manera más rápida y fácil posible:

```
apt install virtualenv
virtualenv -p /usr/bin/python3 py3env
source py3env/bin/activate
pip install requests
pip install -r requirements.txt
python3 download_data.py
```

Antes de realizar la subida a Elasticsearch de estos datos, será necesario realizar un par de modificaciones ya que la versión de ELK con la que estamos trabajando no es 100% compatible con estos scripts y tampoco tenemos activada la autenticación. Las dos modificaciones serán:

- En `index_users.py` comentar la parte de autenticación:

```
#es = elasticsearch.Elasticsearch(http_auth=('elastic', 'changeme'))
es = elasticsearch.Elasticsearch(["192.168.1.31"])
```

- En `movie_lens.json` eliminar las siguientes líneas:

```
"_all": {  
  "enabled": true  
}
```

Y ya si se puede ejecutar el indexado de eventos:

```
python3 index_users.py
```

En el caso de tener problema con la carga de datos por errores con Python, puede deberse a la versión de la librería de Elasticsearch que se está utilizando. Para solucionar este problema, deberá modificarse el archivo `requirements.txt` cambiando la versión de Elasticsearch por la 6.0.0 antes de ejecutar `pip install -r requirements.txt`.

Al ejecutar `index_users.py` puede que todavía arroje errores pero mientras se están indexando los datos. Se puede comprobar si efectivamente están indexándose esos datos desde DevTools, esto puede deberse a una no total compatibilidad entre las librerías instaladas, pero para lo que esta prueba de concepto refiere, lo importante es que al final termine indexando todos los datos.