

## Nombre del Grupo

django-valeiro-rodriguez-torres  
14/4/23

## Miembros del grupo

Jorge Rodríguez Rivera  
(jorge.rodriguez.rivera@udc.es)  
Xaquín Valeiro Pena  
(joaquin.valeiro.pena@udc.es)  
Juan Torres Oróns  
(juan.orons@udc.es)

## Resumen

Nuestra APP tratará de sacar datos de Reddit y analizarlos con una librería third-party TextBlob. Esos datos pueden ser datos de perfil de usuario, datos sobre una temática específica/subreddit o información sobre un árbol de comentarios de un subreddit. Podremos crear gráficos de toda esa información y además tendremos la posibilidad de añadir una publicación en instagram de la gráfica resultante o ver las publicaciones realizadas y los likes y comentarios de las mismas. A parte se podría acceder a un mapa interactivo en el que poder visitar con clicks los diferentes subreddits de los distintos países.

## Listado exhaustivo de las funcionalidades a implementar

1. Búsqueda de perfil usuario. La API de Reddit permite buscar perfiles de usuario mediante el uso de una solicitud GET a la URL <https://www.reddit.com/user/{username}/about.json>, donde `{username}` es el nombre de usuario de la cuenta que se desea buscar. Esta solicitud devuelve un objeto JSON que contiene información sobre el perfil de usuario, incluyendo su nombre de usuario, el tiempo que ha estado en Reddit, la cantidad de karma de publicaciones y comentarios que ha obtenido, la fecha en que se unió a Reddit y algunos otros datos.
2. Filtrar por temática y mirar la cantidad de comentarios positivos y negativos con la librería de procesamiento de lenguaje natural. La API de Reddit permite filtrar los resultados de una búsqueda por temática mediante el uso de un parámetro de consulta `q` en una solicitud GET a la URL <https://www.reddit.com/search.json>. El valor del parámetro `q` es una cadena de búsqueda que se utiliza para filtrar los resultados por palabra clave o frase.
3. Publicación en Instagram de las estadísticas. Para publicar una imagen con una breve descripción en Instagram mediante la API de Graph de Instagram, se requiere que el desarrollador tenga permiso para hacerlo y que la cuenta de Instagram en cuestión esté conectada a la aplicación.

Primero, se debe crear una solicitud POST a la siguiente URL: <https://graph.instagram.com/{user-id}/media>. En esta URL, `{user-id}` se refiere al ID de la cuenta de Instagram que se utilizará para publicar la imagen.

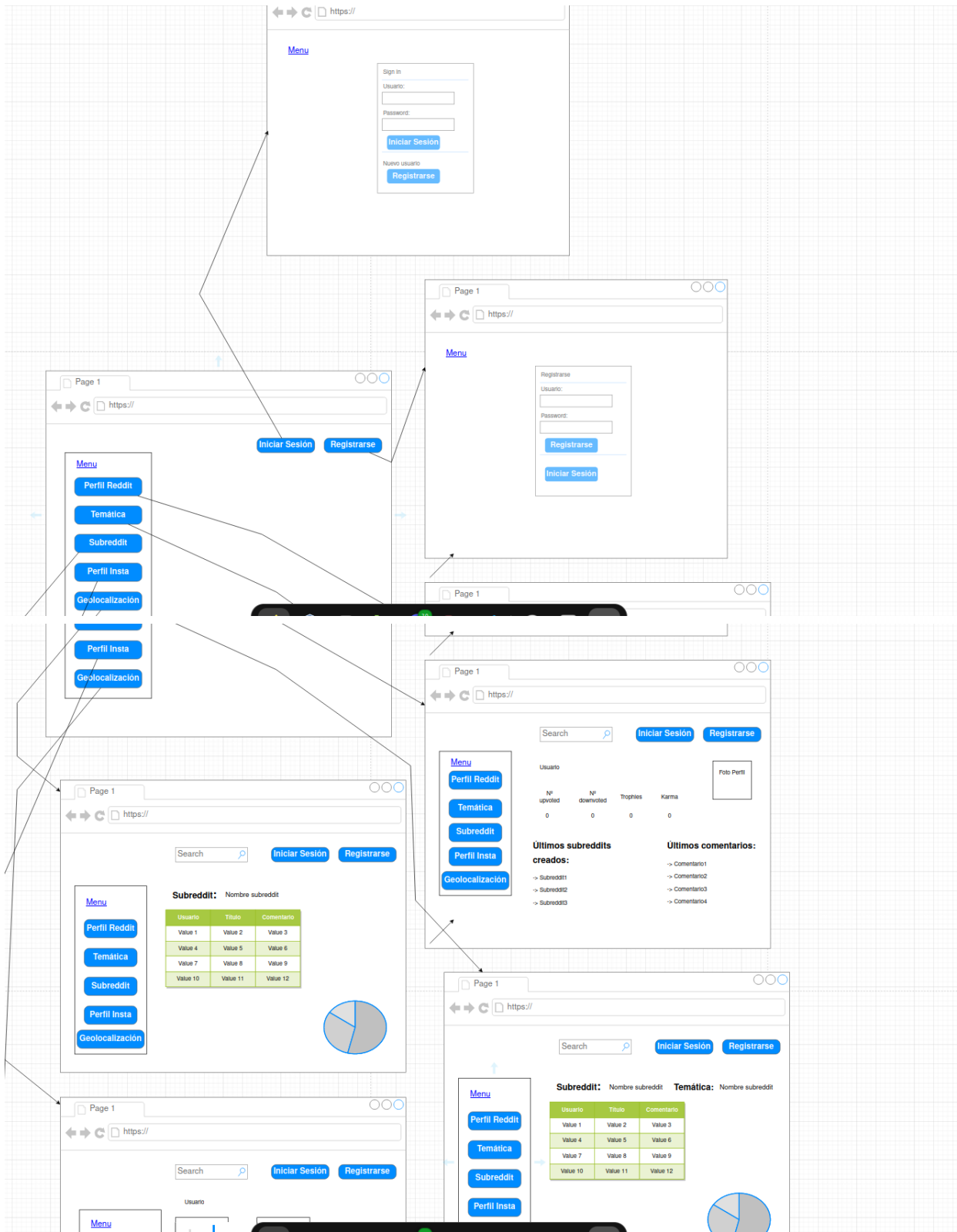
En el cuerpo de la solicitud POST, se debe incluir la siguiente información:

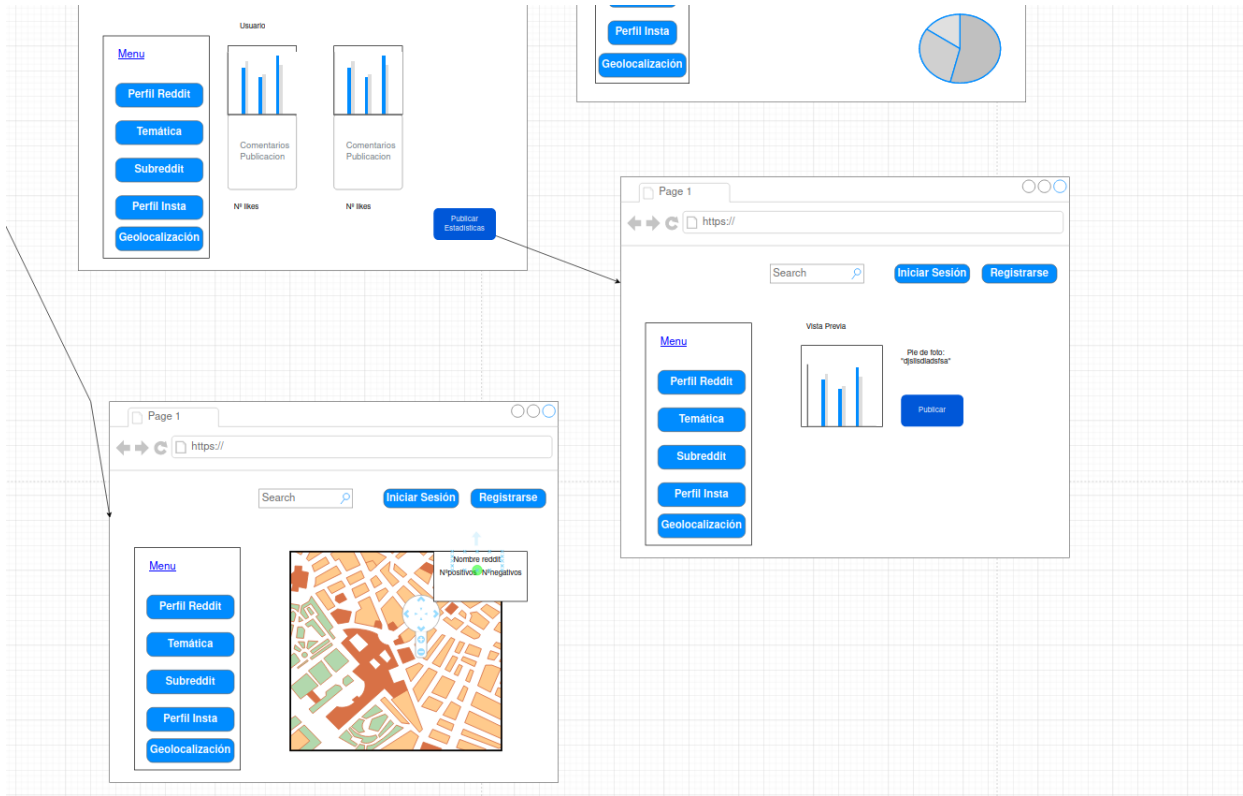
- `image_url` - una cadena que contiene la URL de la imagen que se va a publicar. La imagen debe estar alojada en un servidor externo.
- `caption` - una cadena que contiene la breve descripción que se publicará con la imagen.

Una vez que se envía la solicitud POST, la API de Graph de Instagram devuelve un objeto JSON que contiene información sobre la imagen publicada, incluyendo su ID, la URL de la imagen, la descripción y algunos otros datos.

4. Obtener árbol de comentarios de un subreddit y analizarlo: La API de Reddit permite obtener el árbol de comentarios de un subreddit mediante el uso de una solicitud GET a la URL `https://www.reddit.com/r/{subreddit}/comments/{post_id}.json`, donde `{subreddit}` es el nombre del subreddit que se desea buscar y `{post_id}` es el ID de la publicación que se desea buscar. Esta solicitud devuelve un objeto JSON que contiene información sobre la publicación y sus comentarios, incluyendo el autor de cada comentario, la fecha en que se publicó y el contenido del comentario. Como en la funcionalidad de la temática, podremos analizar tanto la cantidad de comentarios positivos, negativos o neutros de un subreddit, como hacer una petición a la API sobre la predicción de la categoría del comentario.
5. Registro e inicio de sesión: El usuario podrá registrarse e iniciar sesión.
6. Observar información sobre las distintas publicaciones que se han hecho con la propia aplicación, es decir, comentarios y likes de las mismas.
7. Mostrar un mapa del mundo que sea interactivo (con la API de google Maps) con diferentes puntos de interés que representen los países, y al hacer clic en cada punto de interés se podría mostrar al usuario una etiqueta en la que haya información sacada del modelo de python que calcule con la librería TextBlob la cantidad de comentarios positivos y negativos de un muestreo de comentarios del subreddit del país.

## Bocetos de las pantallas de la aplicación





## Flujo de datos de aplicación

- Registro e inicio de sesión: al registrarse o iniciar sesión en la aplicación, el usuario proporciona sus credenciales y la aplicación verifica si son correctas. Si lo son, el usuario tendrá acceso a las funcionalidades de la aplicación.
- Búsqueda de perfil de usuario: cuando el usuario realiza una búsqueda de perfil de usuario, la aplicación envía una solicitud GET a la API de Reddit con el nombre de usuario proporcionado. La API de Reddit devuelve un objeto JSON con la información del perfil de usuario, que la aplicación utiliza para mostrarla en la interfaz de usuario.
- Filtrar por temática: cuando el usuario realiza una búsqueda filtrada por temática, la aplicación envía una solicitud GET a la API de Reddit con el valor del parámetro q correspondiente a la palabra clave o frase proporcionada por el usuario. La API de Reddit devuelve un objeto JSON con los resultados de la búsqueda, que la aplicación utiliza para mostrarlos en la interfaz de usuario.
- Convertir datos recogidos a gráfica con la librería third-party matplotlib(con scipy): la aplicación utiliza los datos recogidos en las búsquedas de perfiles de usuario y filtrado por temática para generar una gráfica utilizando la librería third-party matplotlib(con scipy). La gráfica se muestra en la interfaz de usuario.
- Publicación en Instagram de las estadísticas: cuando el usuario desea publicar estadísticas en Instagram, la aplicación envía una solicitud POST a la API de Graph de Instagram con la imagen y la breve descripción que el usuario ha proporcionado. La API de Graph de Instagram publica la imagen y devuelve un objeto JSON con la información de la publicación, que la aplicación utiliza para mostrarla en la interfaz de usuario.
- Obtener árbol de comentarios de un subreddit y analizarlo: cuando el usuario desea obtener el árbol de comentarios de un subreddit y analizarlo, la aplicación envía una solicitud GET a la API de Reddit con el nombre del subreddit y el ID de la publicación correspondientes. La API de Reddit devuelve un objeto JSON con la información de la publicación y sus comentarios, que la aplicación utiliza para analizar y mostrar la información en la interfaz de usuario.
- Acceder a los datos de una publicación en la API de Graph de Instagram: se inicia con la creación de una solicitud POST. En el cuerpo de la solicitud, se debe incluir la URL de la imagen a publicar y la breve descripción que se desea publicar con la imagen. Luego, la API de Graph de Instagram devuelve un objeto JSON que contiene información sobre la imagen publicada, incluyendo su ID, la URL de la imagen, la descripción y algunos otros datos. Para acceder a los datos de una publicación en particular, es necesario conocer el ID de la publicación y realizar una solicitud GET. Finalmente, la API de Graph de Instagram devuelve un objeto JSON que contiene la información solicitada.

## APIs utilizadas

1. API de Reddit
2. API Google Maps
3. API de Instagram

### Información a extraer del API 1

Información sobre perfiles de usuario, como su nombre de usuario, el tiempo que ha estado en Reddit, la cantidad de karma de publicaciones y comentarios que ha obtenido, la fecha en que se unió a Reddit y algunos otros datos.

También es posible filtrar los resultados de una búsqueda por temática a través de la API de Reddit mediante el uso de un parámetro de consulta.

Además, la API de Reddit permite obtener el árbol de comentarios de un subreddit y su análisis, lo que incluye información sobre el autor de cada comentario, la fecha en que se publicó y el contenido del comentario.

### Información a extraer del API 2

En la API de Graph de instagram, se pueden extraer distintos tipos de información y realizar acciones:

- Publicación de contenido multimedia: se pueden publicar imágenes o videos en la cuenta de Instagram conectada a la aplicación, incluyendo la descripción y los hashtags.
- Acceso a la información de las publicaciones: se pueden obtener datos de las publicaciones de una cuenta, como su ID, la descripción, la fecha de publicación, el número de me gusta y comentarios, entre otros.

### Información a extraer del API 3

La API de Google Maps proporciona una amplia variedad de mapas, incluyendo mapas de calles, mapas de terreno, mapas de relieve y mapas de satélite. Los desarrolladores pueden personalizar los mapas para que se ajusten a sus necesidades y mostrar información específica, como las carreteras, los edificios, los puntos de interés y más. Lo que haríamos sería utilizar la API de Maps para mostrar un mapa del mundo con diferentes puntos de interés que representen los países, y al hacer clic en cada punto de interés se podría mostrar al usuario una etiqueta que muestre el nombre del subreddit del país junto con el porcentaje de comentarios positivos y negativos de una cantidad "x" de comentarios. Así podremos ver las diferencias entre los países que queramos.

## Uso de Pandas dentro de la práctica

Para la primera funcionalidad (Búsqueda de perfil usuario), se utilizará Pandas para unificar los Dataframes obtenidos de las peticiones correspondientes con la API. Para ello, necesitamos que las contemos con el mismo índice para unificar los Dataframes, en este caso el nombre de usuario y a partir de ahí unificar las diferentes tablas correspondientes a sus trofeos, datos de votos, karma, fechas, etc. Para esto, usamos la función `'merge()'` de Pandas.

```
merged_df = pd.merge(DataFrame1, DataFrame2, on='nombre_user')
```

En la segunda funcionalidad, un uso interesante de Pandas sería en cuanto a la ordenación de comentarios según sus votos positivos y obtener así un estudio de la temática filtrada más visual. A partir de esto, realizaríamos nuestro estudio. Podríamos utilizar la función `sort_values()`:

```
df_sorted = df.sort_values('columnaVotos')
```

En resumen, todas las funcionalidades donde tratemos con Dataframes, haremos uso de las funciones de Pandas para tratar con ellos pero estos casos son los más notorios y significativos en nuestro proyecto.

Las gráficas que obtengamos de ordenar con pandas van a usarse posteriormente en el frontend para enseñar toda la información.

## Funcionalidades a implementar en la primera iteración de la práctica

En la primera iteración de la práctica se implementarían las funcionalidades 1,4,5 y 7.

## Uso de librerías externas (opcional)

La librería externa que vamos a usar va a ser TextBlob para procesar análisis de sentimiento en los comentarios que obtengamos de la API de reddit. Además vamos a analizar la subjetividad en los comentarios de un subreddit o en un subreddit de una temática.

## Integración lenguajes distintos a python

Para la integración de otros lenguajes, vamos a integrar bootstrap(con javascript) e integrarlo con peticiones AJAX.