**Plataforma para el análisis de datos de Instagram**

Alberto Macías Gutiérrez

Facultad de ciencias, Universidad de Salamanca

Trabajo de Fin de Grado

Lucía Martín Gómez

7 de julio de 2022

# Índice

# Introducción

Como bien se explicará más adelante, a finales de la década de los 90 empezaron a aparecer las primeras redes sociales. Al principio, las oportunidades de acceder a estas eran más bien escasas, únicamente se encontraban disponibles a aquellos privilegiados que podían tener acceso a internet.

Con el paso del tiempo, el desarrollo de dispositivos móviles y ordenadores, así como el acceso a internet por mayor cantidad de personas, ha impulsado la creación y uso de las redes sociales por parte de programadores y usuarios.

El desarrollo de las tecnologías en la red en la última década ha hecho que una de las formas más importarte de publicidad y contacto de cualquier negocio se realice mediante las redes sociales, haciendo así que estas sean de uso imprescindible para las empresas.

Sabiendo esto, y entrando en el contexto de una de las redes sociales más famosas como es Instagram, que en un principio fue creada con el fin de que las personas pudieran compartir fotos con sus seguidores y ponerse en contacto con estos, con el paso del tiempo esta red social también ha tenido un desarrollo profesional, donde muchos negocios promocionan sus servicios y contacto.

La elaboración de una aplicación web para el análisis de datos de Instagram ha supuesto la perfecta elección para poder observar y comparar la información que hay detrás de los distintos usuarios que forman parte de esta red social.

# Objetivos

El trabajo de fin de grado tiene como objetivo de realizar una aplicación web en la que usuario pueda obtener información a la hora de buscar cuentas de Instagram en concreto, así como la información a la hora de buscar hashtags.

Almacenar las cuentas de usuario de la aplicación y su respectiva información dentro de una base de datos relacional como MySQL.

Mostrar información de cada usuario o hashtag de Instagram a buscar, en conjunto con una serie de estadísticas y gráficos a crear mediante esta información.

Comparar distintas cuentas de Instagram para ver cual de ellas es la más popular o famosa y mostrar datos de estas cuentas en tablas y gráficos.

Crear dos ramas de la aplicación, una para los usuarios que quieran probar la aplicación localmente para poder cambiar la cuenta de Instagram usada para buscar la información, y otra rama para ser subida a un servidor y así poder usarla desde este.

# Antecedentes

Para poder meternos un poco en el contexto de las redes sociales veremos su historia (Hera, 2022).

**1997**

*SixDegrees*: Considerada la primera red social del mundo, esta permitía localizar a otros miembros de la red y crear listas de amigos. Esta desaparece en el año 2001.

**2002-2003**

*Friendster*: Se crea en 2002 como una red social para amantes de los videojuegos.

*Linkedln*: Se crea en 2003 como una red social más profesional orientada a empresas.

**2004**

*Facebook*: Considerada como la red social por excelencia, en un principio llamada Facemash, con el fin de que los estudiantes de la Universidad de Harvard pudieran tener un lugar virtual donde compartir acerca de quienes eran las personas más atractivas de esta.

**2005**

*YouTube*: Sitio web dedicado a compartir videos

**2006**

*Twitter*: Red social que permite la comunicación con otros usuarios mediante mensajes

con un tamaño máximo de 140 caracteres.

**2009**

*WhatsApp*: App de mensajería instantánea. Esta se vincula con la agenda de contactos del dispositivo del usuario.

**2010**

*Instagram*: Red social de fotografía por excelencia. Pionera junto con Twitter en el uso de hashtags, facilitando a los usuarios descubrir fotografías que los demás usuarios compartían sobre un mismo tema.

*Pinterest*: Red social que colecciona imágenes, que permite a los usuarios almacenarlas en tableros y dotarlas de pines.

**2011**

*Twitch*: Plataforma de live streaming más grande del mundo.

**2016**

*Tik Tok*: Red social que permite compartir vídeos cortos.

# Conceptos Teóricos

## Introducción

Diagrama

Descripción generada automáticamenteGran parte de las aplicaciones web usan la arquitectura Cliente Servidor (Schiaffarino, 2019), la cual es una de las principales usadas en muchos servicios y protocolos de internet. Esta arquitectura tiene dos partes bien diferenciadas, por un lado, la parte del servidor, que suele ser una máquina bastante potente la cual actúa como depósito de datos y ejecuta la aplicación web, y por otro la parte del cliente o clientes, que son dispositivos o máquinas que demandan servicios al servidor.

**Figura 1**

*Modelo Cliente-Servidor*

En la figura 1 se observa cómo funciona el Modelo Cliente Servidor dentro de internet, por el que los clientes o usuarios envían peticiones y el servidor responde a esas peticiones.

Un ejemplo de la arquitectura Cliente Servidor es la red de internet donde existen ordenadores de diferentes personas conectados alrededor del mundo, las cuales se conectan a través de los servidores de su proveedor de internet por ISP donde son redirigidos a los servidores de las páginas que desean visualizar.

## Desarrollo Web

El desarrollo web ha evolucionado exponencialmente desde su inicio. En la actualidad es uno de los servicios más solicitados, que nos brinda una gran cantidad de funciones distintas.

**Figura 2**

*Historia del Desarrollo Web*

Escala de tiempo

Descripción generada automáticamente

En la figura 2 (Gima, 2019) podemos observar, como hemos pasado de páginas HTML estáticas, a la creación de aplicaciones web mediante la suma de diferentes leguajes de programación y estilo, gracias a el uso de JavaScript para dar dinamismo a las páginas web, las hojas de estilo CSS para separar del HTML el formato y la apariencia y el diseño responsivo o adaptable para que las páginas web se puedan representar en pantallas con tamaños y resoluciones muy diferentes.

## Tecnologías Web

### Python

Fue creado por Guido van Rossum (Visus, 2020) a principios de los 90, su nombre se debe a la afición de este al grupo Monty Python. Python es un lenguaje de programación interpretado cuya filosofía es que sea legible por cualquier persona con conocimientos básicos de programación.

Características:

* Es gratis, ya que es un lenguaje de código abierto
* Respaldado por una comunidad enorme, por el carácter gratuito hace que continuamente se estén desarrollando nuevas librerías y aplicaciones.
* Es un lenguaje multiparadigma, es decir, combina propiedades de diferentes paradigmas de programación
* Sus aplicaciones tienen áreas tan dispares como el diseño de aplicaciones web o la inteligencia artificial.
* Es acto para todas las plataformas, es decir, se puede ejecutar en cualquier sistema operativo.

El principal inconveniente del lenguaje de programación Python es que se trata de un lenguaje de programación interpretado, es decir, que no se compila, sino que se interpreta en tiempo de ejecución, dando como resultado a que sea más lento que otros lenguajes de programación como Java, C o C++. Pero este inconveniente no es un gran problema actualmente ya que en la actualidad las máquinas son muy potentes.

### Django

Es un framework web gratuito (MDN contributors, 2021) de alto nivel que permite el desarrollo rápido de sitios web seguros y mantenibles. Django se encarga de una gran parte de las complicaciones del desarrollo web, por lo que permite al programador centrarse únicamente en escribir las funcionalidades de la aplicación. Este posee una comunidad próspera y activa, mucha documentación y soporte gratuito. Django se encuentra escrito en Python.

Django funciona con patrón de diseño organizando la arquitectura en Modelo Vista Plantilla “Model View Template (MVT)”, este tiene una estructura muy similar a la arquitectura Model View Controller, con la diferencia que la vista es el controlador intermediario entre el modelo, que se comunica con la base de datos, y las plantillas en formato HTML que reproducen la información y los formularios para pedir o cambiar información en la base de datos.

En un sitio web tradicional, una aplicación web espera peticiones HTTP del cliente. Cuando se recibe una petición la aplicación elabora la respuesta que se necesita basándose en la URL y a veces, en una información incluidas en los datos de una petición POST o GET. La aplicación también devolverá una respuesta al explorador web, normalmente creando dinámicamente una página HTML con los datos correspondientes a mostrar.

Sin embargo, las aplicaciones web de Django agrupan el código que gestiona cada uno de estos pasos en ficheros separados.

**Figura 3**

*Funcionamiento Framework Django*

Diagrama

Descripción generada automáticamente

URLs: Como se observa en la figura 3, cuando se recibe una petición por parte de un cliente, esta se procesa en el archivo urls.py, de forma que funciona como un mapeador URL para redirigir las peticiones HTTP a la vista apropiada.

View: El mapeador URL redirige la petición a una vista en concreto, la cual accederá a los datos que se necesitan para la petición por medio del modelo y formateará la respuesta al cliente por medio de las plantillas.

Models: Los modelos definen la estructura de los datos de la aplicación y proporcionan mecanismos para gestionar esta (añadir, modificar y borrar) y consultar registros en la base de datos.

Templates: Las plantillas son ficheros de texto que definen la estructura o diagrama de otro fichero, con marcadores de posición para representar contenido real. La vista creará dinámicamente una página usando una plantilla, rellenando esta con los datos del modelo. Normalmente estos ficheros suelen estar codificados en lenguaje HTML pero no tiene porque ser así, pueden ser ficheros de cualquier tipo.

### HTML (HyperText Markup Language)

HTML (Discovery Formación, 2021) es el lenguaje con el que se crea y define el contenido que se ve en las páginas web. Es el lenguaje encargado de mantener la estructura de cada elemento dentro de una página web. Sus siglas significan “Lenguaje de Marcas de hipertexto”, este permite definir donde comienza y acaba un elemento.

HTML se basa en un conjunto de etiquetas que sirven para marcar los elementos que se usan a la hora de hacer páginas web, dando igual la herramienta que se esté utilizando.

### CSS (Cascading Style Sheets)

CSS (Santos, s.f.) es un lenguaje que se ocupa del diseño y la presentación de las páginas web, funciona junto con el lenguaje HTML. Sus siglas significan “hojas de estilo en cascada”. Con este puedes crear reglas para indicar a la página web como mostrar la información y el estilo de cada elemento de la página.

### JavaScript

JavaScript (Ramos, s.f.) es el lenguaje de programación encargado de dar más interactividad y dinamismo a las páginas web. Este no necesita ser compilado, sino que el navegador lee directamente el código. Se trata de uno de los tres lenguajes nativos de las páginas web junto a HTML (contenido y estructura) y a CSS (diseño del contenido y estructura).

Este lenguaje de programación se ejecuta en la máquina de los clientes, el cual permite crear efectos y animaciones o responder a eventos causados por el usuario como pulsar botones o modificar el DOM (document object model). El código JavaScript se ejecuta en todos los navegadores, tanto de escritorio como móviles.

### Bootstrap

Bootstrap (Invitado, 2020) es un framework CSS desarrollado por Twitter en 2010, para estandarizar las herramientas de la compañía. Este framework combina CSS y JavaScript para estilizar los elementos de las páginas web HTML, además proporciona interactividad en las páginas incluyendo una serie de componentes que facilitan la comunicación con el usuario, como menús de navegación, controles de página y barras de progreso.

El principal objetivo de Bootstrap es permitir la construcción de sitios web responsive (los elementos de la página escalan en función de la resolución de la pantalla) para dispositivos móviles, ordenadores y tablets, de una manera simple y organizada.

## Base de datos

Una base de datos (Oracle, s.f.) es una recopilación organizada de información o datos estructurados, que normalmente se almacenan de forma electrónica en un sistema informático. Normalmente, una base de datos se controla por un sistema de gestión de bases de datos (DBMS). El conjunto de los datos, el DBMS y las aplicaciones asociadas a ellos, reciben el nombre de sistema de bases de datos.

Los datos de los tipos más comunes de bases de datos en funcionamiento actualmente se suelen utilizar como estructuras de filas y columnas en una serie de tablas para aumentar la eficacia del procesamiento y consulta de datos. La mayoría de las bases de datos utilizan un lenguaje de consulta estructurada denominado SQL para escribir y consultar datos.

### Tipos de bases de datos

* Bases de datos relacionales: Los elementos de una base de datos relacional se organizan como un conjunto de tablas con filas y columnas. La tecnología de estas proporciona la forma mas eficiente y flexible de acceder a la información estructurada.
* Bases de datos orientadas a objetos: La información de una base de datos orientada a objetos se representa en objetos, como en la programación orientada a objetos.
* Bases de datos distribuidas: Una base de datos distribuida consta de dos o más archivos que se encuentran en sitios diferentes.
* Almacenes de datos: Un repositorio central de datos o un data warehouse es un tipo de base de datos diseñado específicamente para consultas y análisis rápidos.
* Bases de datos NoSQL: También denominada base de datos no relacional permite almacenar y manipular datos no estructurados y semiestructurados, a diferencia de una base de datos relaciona, que define cómo se deben componer todos los datos insertados en la base de datos. Estas se han ido haciendo más populares cada vez que las aplicaciones web se hacen más comunes y complejas.
* Bases de datos orientadas a grafos: Almacena los datos relacionados con entidades y las relaciones entre entidades.
* Bases de datos OLTP: Se tratan de bases de datos rápidas y analíticas diseñadas para que muchos usuarios realicen una gran cantidad de transacciones.

# Desarrollo de la Propuesta

## Definición de la Propuesta

## Diseño de la Base de Datos

## Proceso de Obtención de Datos de Instagram

## Desarrollo de la Aplicación Web

# Resultados

# Conclusiones y Líneas de Trabajo Futuras

# Bibliografía

Discovery Formación. (21 de Junio de 2021). *HTML: Qué es y significado*. Discovery formacion: https://discoveryformacion.com/html-que-es-y-significado/

Gima, N. (30 de Septiembre de 2019). *La evolución del diseño web*. HostGator: https://www.hostgator.mx/blog/evolucion-del-diseno-web/

Hera, C. d. (2 de Junio de 2022). *Historia de las redes sociales: cómo nacieron y cuál fue su evolución*. Marketing4ecommerce: https://marketing4ecommerce.net/historia-de-las-redes-sociales-evolucion/

Invitado, A. (12 de Abril de 2020). *Bootstrap: guía para principiantes de qué es, por qué y cómo usarlo*. Rockcontent: https://rockcontent.com/es/blog/bootstrap/

MDN contributors. (11 de Febrero de 2021). *Introducción a Django*. MDN Web Docs: https://developer.mozilla.org/es/docs/Learn/Server-side/Django/Introduction

Oracle. (s.f.). *¿Qué es una base de datos?* Oracle: https://www.oracle.com/es/database/what-is-database/

Ramos, R. (s.f.). *¿Qué es JavaScript y para qué sirve?* Soyrafaramos: https://soyrafaramos.com/que-es-javascript-para-que-sirve/

Santos, D. (s.f.). *Introducción al CSS: qué es, para qué sirve y otras 10 preguntas frecuentes*. HubSpot: https://blog.hubspot.es/website/que-es-css

Schiaffarino, A. (12 de Marzo de 2019). *Modelo cliente servidor*. Infranetworking: https://blog.infranetworking.com/modelo-cliente-servidor/

Visus, A. (Octubre de 2020). *¿Para qué sirve Python? Razones para utilizar este lenguaje de programación*. Esic: https://www.esic.edu/rethink/tecnologia/para-que-sirve-python