

Evidencia	Realizar un script que realice Web Scraping y búsqueda de información
Objetivo	Implementar los diversos temas vistos en la UA para realizar un proyecto en Python que permita analizar información y extraer datos significativos de forma automatizada.
Forma de trabajo	Individual o de forma colaborativa de hasta 4 personas
Factores a evaluar	<p>Escribe un script que resuelva el problema planteado y que incluya los siguientes tópicos:</p> <p>Manejo de módulos y uso de la función main.</p> <p>El script debe estar dividido en, al menos, 2 archivos .py (5%)</p> <p>Manejo de estructuras de datos (5%)</p> <p>Expresiones regulares (10%)</p> <p>Uso del módulo os (5%)</p> <p>Escritura y lectura de archivos .txt (10%)</p> <p>Escritura de libros de Excel (10%)</p> <p>Uso de requests y BeautifulSoup para hacer web scraping y descargar imágenes (15%)</p> <p>Consulta de API y uso de JSON (10%)</p>
	Presentación del trabajo (15%)
	Subir a Teams el proyecto completo en lenguaje Python, incluyendo el script principal (que debe contener el <code>"__main__"</code>) y los otros módulos (archivos .py); así como todos los entregables solicitados. El proyecto completo no debe tener ninguna falla (10%)
	Incluir un documento breve (word o txt o pdf) o video con las instrucciones de uso (5%)
Fecha límite de entrega	Domingo 15 de mayo a las 23:30 horas (sin prorrogas)

Descripción general

Se quiere contar con un script que realice investigación de un equipo (local o nacional) de algún deporte o deportista (local o nacional). La idea es que ese script se alimente de uno o más archivos de texto que contengan direcciones

URL de páginas web donde se pueda consultar información sobre el equipo deportivo o deportista que se está investigando.

Además, se consultará una API para conocer las predicciones meteorológicas de los 3 próximos partidos o encuentros, entrevistas o participación en algún evento importante (olimpiadas, torneo especial, copa, panamericanos, etc.) del equipo o deportista en cuestión.

Toda la información deberá almacenarse dentro de un Excel.

Funcionamiento específico

El script debe ser capaz de realizar las siguientes tareas:

- Buscar archivos con extensión txt en una carpeta para leerlos
- Leer y editar archivos txt, los cuales deben de tener las direcciones URL para hacer la investigación del equipo o deportista. Podemos agregar o eliminar fuentes de información.
- Definir al menos 3 patrones de expresiones regulares para búsqueda de información relevante del equipo o deportista. Puede ser información personal (fechas importantes, correo electrónico, redes sociales, teléfonos) o datos relevantes de su deporte como estadísticas, fechas relevantes o lugar obtenidos en los torneos.
- Aplicar web scraping (con los módulos requests y BeautifulSoup) para:
 - Encontrar al menos 3 titulares de noticias relacionados con el equipo o deportista (generalmente vienen en los Headers de las páginas HTML)
 - Buscar la información correspondiente apoyándose en las expresiones regulares
 - Descargar al menos 3 imágenes del equipo o deportista.
- Usar el módulo os para crear una carpeta y almacenar ahí las imágenes descargadas.
- Integrar, en una estructura de datos, la información relevante del equipo o deportista que, en su mayoría, debería ser obtenida con el web scraping a través de la búsqueda de información en los textos y, para datos no encontrados, se podrá optar por preguntar al usuario.
- Definir 3 fechas próximas en las que se efectuarán los juegos, presentaciones, entrevistas o eventos importantes donde

participara y consultar con la API de OpenWeather, la temperatura para dichos días.

- Almacenar toda la información obtenida en un Excel de forma automática.

Entregables

- Todos los archivos de Python generados
- El Excel generado de la demostración del video
- Las imágenes descargadas durante la demostración del video
- Los archivos txt que consulta el script
- Las instrucciones de uso

Tabla de Entregables

Insertar aquí todos los nombres de los entregables y la información solicitada. Se puede cambiar el tamaño de la tabla según se requiera. Los archivos se colocan en una carpeta y uno de los integrantes del equipo lo sube a Teams.

Nombre de los integrantes del equipo (máximo 4 personas por equipo)	
Nombre del equipo o deportista	
Deporte referido	
Nombre de los archivos .py que integran el proyecto. Colocar el que contiene al main al inicio	
Nombre del Libro de Excel	
Nombres de los archivos txt de consulta	
Nombres de las Imágenes	
Documento doc, docx o txt con las instrucciones de uso	