# Pregunta 3: Datos No Estructurados

En esta sección se describe el servicio y el código utilizado para consumir datos no estructurados provenientes de redes sociales como Twitter y Reddit. Estos datos no siguen un formato fijo, contienen texto libre, publicaciones y comentarios, y requieren procesamiento adicional para su análisis.

## Servicio 1: Twitter API v2

Para consumir tweets recientes en español relacionados al hashtag #tecnología, se utilizó la API v2 de Twitter a través de la librería `tweepy`. Se accedió con una Bearer Token configurada desde un archivo .env. Los datos se almacenan en archivos JSON.

Código de ejemplo:

***from dotenv import load\_dotenv  
import tweepy  
import os  
  
load\_dotenv()  
bearer\_token = os.getenv("TWITTER\_BEARER\_TOKEN")  
client = tweepy.Client(bearer\_token=bearer\_token)  
  
query = "#tecnología -is:retweet lang:es"  
response = client.search\_recent\_tweets(query=query, max\_results=10, tweet\_fields=["created\_at", "text", "author\_id", "lang"])***

******

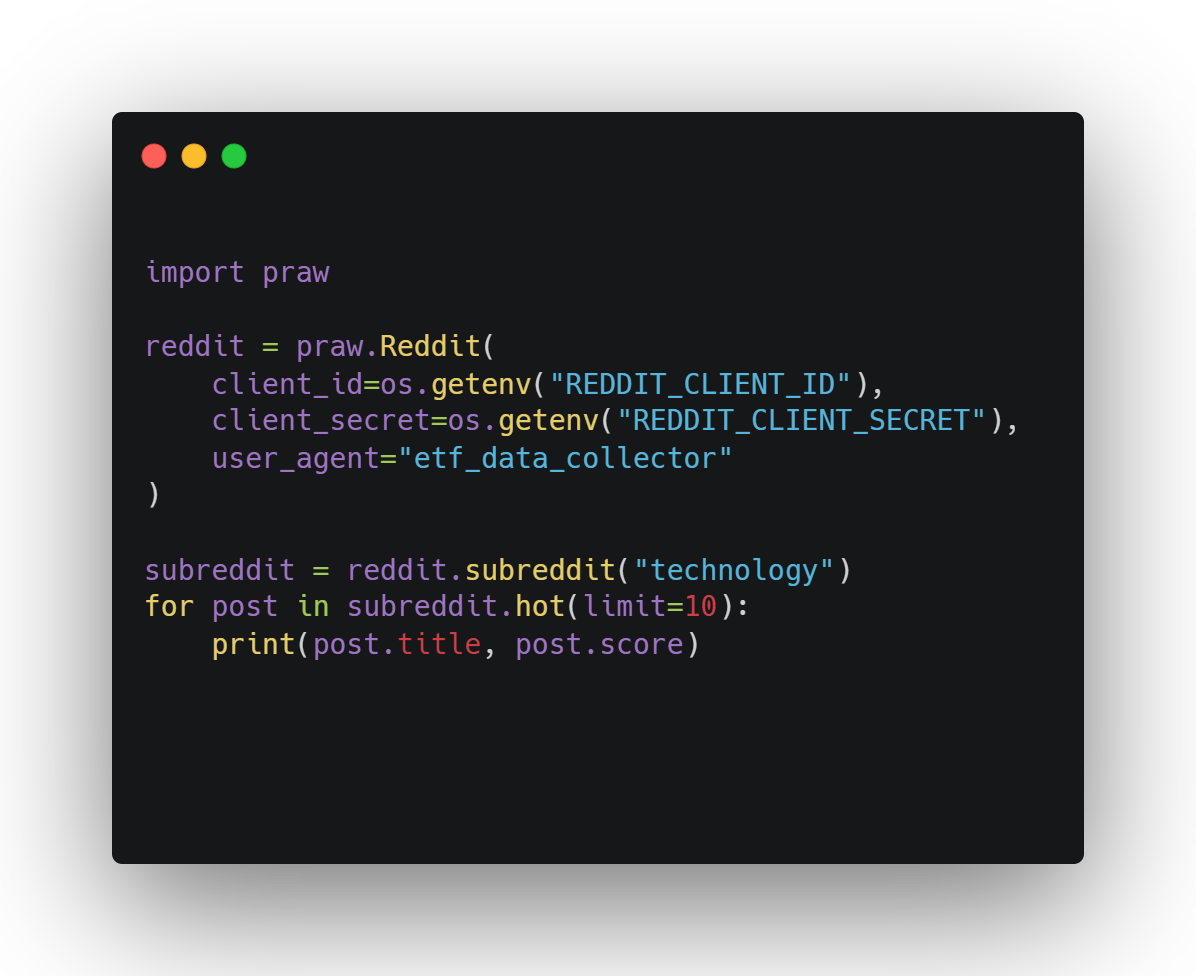
Se puede analizar

* **Tendencias** de uso del hashtag #tecnología.
* **Análisis de sentimiento** del texto (usando librerías como TextBlob, spaCy o transformers).
* **Frecuencia de menciones** de palabras clave.
* **Usuarios más influyentes** por cantidad de likes/retweets.
* **Hora del día con mayor actividad**.

## Servicio 2: Reddit API (PRAW)

Se utilizó la API de Reddit mediante la librería `praw` para extraer publicaciones recientes del subreddit r/technology. Las credenciales se cargaron desde variables de entorno. Los datos obtenidos incluyen título, puntuación, autor, número de comentarios y URL.

Se puede realizar análisis sobre

* **Temas populares** según títulos y texto de publicaciones.
* **Frecuencia de publicación y horario** (con created\_utc).
* **Engagement** (puntuación vs. número de comentarios).
* **Comparación de interacción** entre posts de texto y posts con enlaces.
* **Análisis de sentimiento** del selftext o de los comentarios (si también los extraes).  
  

Esto te permite:

* Hacer **minería de opiniones**.
* Estudiar el **tono y lenguaje** de la comunidad.
* Detectar **palabras clave** o conceptos recurrentes.

### 🧩 Ideas de uso para proyectos:

* Dashboards con visualización de tendencia de temas (#tecnología).
* Modelos de clasificación de sentimientos.