# Report: YouTube Metadata Scraper using yt-dlp

## 1. Objective

The purpose of this project is to extract metadata from a YouTube video without downloading the video itself. The script uses the yt-dlp library to fetch information such as title, channel name, description, publish date, and view count. The extracted data is stored in a structured format (Pandas DataFrame) for further analysis.

## 2. Tools & Libraries Used

- Python 3.10+

- yt-dlp (for extracting YouTube metadata)

- pandas (for tabular data storage and manipulation)

## 3. Code Explanation

def scrape\_youtube(video\_url):  
 ydl\_opts = {  
 "quiet": True,  
 "skip\_download": True, # we only want metadata  
 }  
  
 with yt\_dlp.YoutubeDL(ydl\_opts) as ydl:  
 info = ydl.extract\_info(video\_url, download=False)  
   
 video\_data = {  
 "Title": info.get("title"),  
 "Channel": info.get("uploader"),  
 "Description": info.get("description")[:500], # truncate for readability  
 "Publish Date": info.get("upload\_date"),  
 "Views": info.get("view\_count"),  
 }  
 return video\_data

The function scrape\_youtube() takes a YouTube video URL. It uses yt\_dlp.YoutubeDL to extract metadata and stores important fields in a Python dictionary. The description is truncated to 500 characters to avoid excess text.

## 4. Usage

video\_url = "https://www.youtube.com/watch?v=dQw4w9WgXcQ"  
datas = scrape\_youtube(video\_url)  
  
import pandas as pd  
df = pd.DataFrame([datas])  
  
# Convert Publish Date to datetime  
df['Publish Date'] = df['Publish Date'].astype('datetime64[ns]')

## 5. Sample Output

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Channel | Description | Publish Date | Views |
| Rick Astley - Never Gonna Give You Up (Official...) | Rick Astley | The official video for “Never Gonna Give You U…” | 2009-10-25 | 1695231113 |

## 6. Conclusion

This script successfully extracts key YouTube video metadata using yt-dlp and organizes it in a Pandas DataFrame. The approach is efficient for building datasets of video information for research, analytics, or content management without downloading the actual videos.