This document is an additional description of the codes I submitted.

Detailed information about each function and python file are written in each file.

Twelve Labs – Technical Code Assignment

Due to security concerns, I have removed sensitive information such as my AWS key and MongoDB passwords from the code.

As a result, the code provided in Github is for reference only and cannot be executed successfully without the necessary credentials.

Therefore, I captured screenshots of the relevant results and attached them to this document.
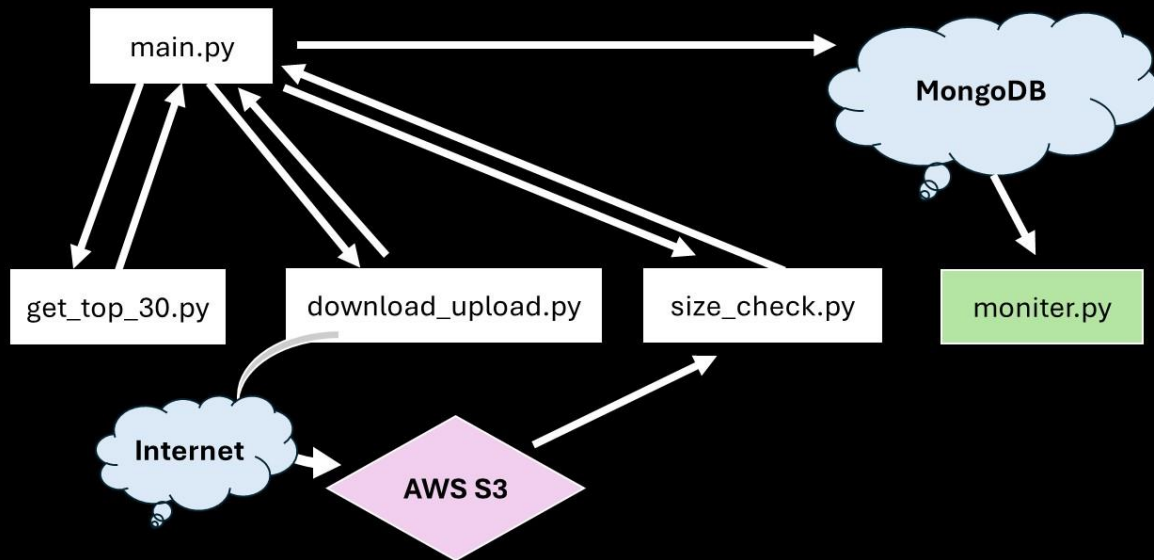
# Part 1. Configuration



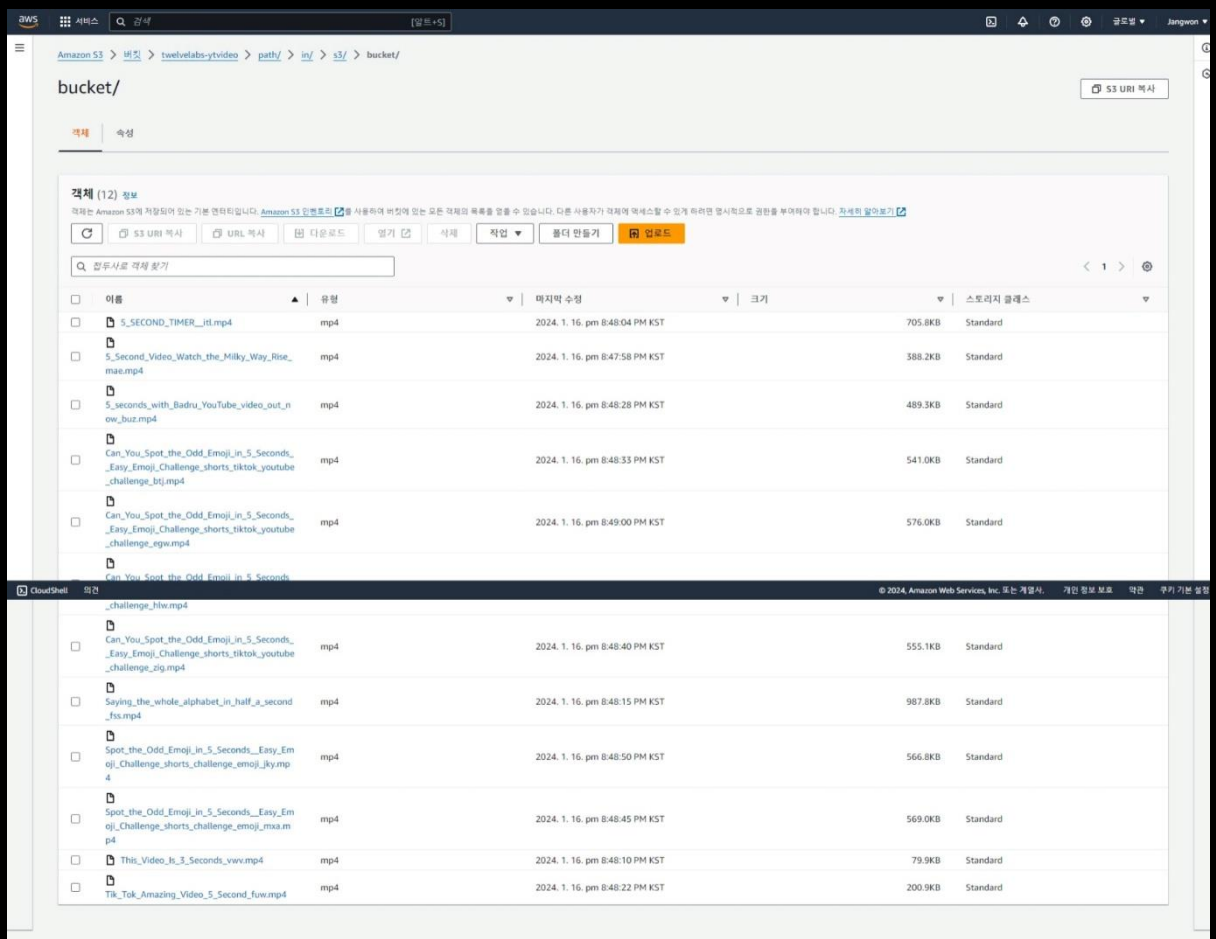**Figure 1. Flow Chart**

**There is total 5 Python files:**

**1. main.py**

**2. get_top_30.py : This function return lists of videos' name and URL, number of videos after filtering, and number of videos retrieved by query search**

**3. download_upload.py : This function download the video, and upload it to S3 bucket.**

**4. size_check.py : This function return the size of the video file in S3 bucket.**

**5. montier.py : This function get log data from MongoDB and print the following information.**

**Part 2. Screenshot of the Result**

To validate my code, I gave a query as "5seconds videos". To save AWS S3 resources, I refine time constraints from 8 minutes~12 minutes to shorter than 15 seconds to reduce the size of videos.

**① After run main.py**

- **AWS S3 Bucket**



- ■ We can verify whether the videos retrieved through the query have been successfully uploaded.

- **MongoDB screenshot**



# mydatabase.mycollection

STORAGE SIZE: 36KB    LOGICAL DATA SIZE: 4.13KB    TOTAL DOCUMENTS: 17    INDEXES TOTAL SIZE: 36KB

**Find**        Indexes        Schema Anti-Patterns **0**        Aggregation        Search Indexes

Filter⬚                Type a query: { field: 'value' }

**QUERY RESULTS: 1-17 OF 17**

```
_id: ObjectId('65a68161af0be4a489386a77')
Item: "Query_Base"
Query: "5seconds videos"
Number of Videos after filtering: 10
Number of Videos retrieved by Query search: 30
Worker_Name: "LAPTOP-HU5R4S9U_240116_5"
Number of Videos Successfully Upload in S3: 7
Time_for_Download: 27.359541177749634
```

```
_id: ObjectId('65a68161af0be4a489386a78')
Item: "Video_Base"
Video_Name: "5_Second_Video_Watch_the_Milky_Way_Rise_nag"
Video_URL: "https://www.youtube.com/watch?v=QC8iQqtG0hg"
Worker_Name: "LAPTOP-HU5R4S9U_240116_5"
Video_size: 397567
```

- ■ **Relevant logs are presented in MongoDB**
- ■ **In the "Query_Base" object, keys such as Query, Number of Videos after search, filtering, and upload, worker, as well as time for download and upload, are configured to facilitate organized data management.**
- ■ **In the "Video_Base" object, keys such as Video's name and URL, worker, as well as size of the video, are configured to enhance structured data management.**

## ※ Verification of the functioning of multiprocessing

```
Checkpoint1 : Download the video from YouTube
Checkpoint1 : Download the video from YouTube
Checkpoint1 : Download the video from YouTube
Checkpoint1 : Download the video from YouTube
Error: NNoKCHdpKho is age restricted, and can't be accessed without logging in.
Checkpoint1 : Download the video from YouTube
Error: UoKpp1ewyRc is age restricted, and can't be accessed without logging in.
Checkpoint1 : Download the video from YouTube
Checkpoint2 : Upload the file to S3 bucket
Checkpoint2 : Upload the file to S3 bucket
Checkpoint2 : Upload the file to S3 bucket
Checkpoint2 : Upload the file to S3 bucket
File uploaded successfully to twelvelabs-ytvideo/path/in/s3/bucket/5_Second_Video_Watch_the_Milky_Way_Rise_zoh.mp4
Checkpoint3 : Delete the video file from the computer
Video file '5_Second_Video_Watch_the_Milky_Way_Rise_zoh.mp4' deleted
Checkpoint1 : Download the video from YouTube
File uploaded successfully to twelvelabs-ytvideo/path/in/s3/bucket/5_SECOND_TIMER__sha.mp4
Checkpoint3 : Delete the video file from the computer
Video file '5_SECOND_TIMER__sha.mp4' deleted
Checkpoint1 : Download the video from YouTube
File uploaded successfully to twelvelabs-ytvideo/path/in/s3/bucket/The_Judges_Life_Flashed_Before_Her_Eyes_wrl.mp4
Checkpoint3 : Delete the video file from the computer
Video file 'The_Judges_Life_Flashed_Before_Her_Eyes_wrl.mp4' deleted
Checkpoint1 : Download the video from YouTube
File uploaded successfully to twelvelabs-ytvideo/path/in/s3/bucket/Can_You_Spot_the_Odd_Emoji_in_5_Seconds__Easy_Emoji_Challenge_shorts_ti
ktok_youtube_challenge_zuw.mp4
Checkpoint3 : Delete the video file from the computer
Video file 'Can_You_Spot_the_Odd_Emoji_in_5_Seconds__Easy_Emoji_Challenge_shorts_tiktok_youtube_challenge_zuw.mp4' deleted
Checkpoint1 : Download the video from YouTube
Error: hUhhHSlIKBc is age restricted, and can't be accessed without logging in.
Checkpoint1 : Download the video from YouTube
Checkpoint2 : Upload the file to S3 bucket
```

- Given that I executed the code on a 4-core CPU, the figure shows that four videos were downloaded concurrently.

## ② After run moniter.py

```
PS C:\Users\jy121\Desktop\twelvelabs> c:; cd 'c:\Users\jy121\Desktop\twelvelabs'; & 'C:\Users\jy121\AppData\Local\Programs\Python\Python
312\python.exe' 'c:\Users\jy121\.vscode\extensions\ms-python.python-2023.22.1\pythonFiles\lib\python\debugpy\adapter/../..\debugpy\launch
er' '55908' '--' 'C:\Users\jy121\Desktop\twelvelabs\moniter.py'
The number of active workers is  2                                                          ①
                                                       0                           1
Worker                            LAPTOP-HU5R4S9U_240116_5   LAPTOP-HU5R4S9U_240116_1
Query                                   5seconds videos          5seconds videos
# of Videos retrieved by Query search              30                       30        ②
# of Videos after filtering                        10                       11
# of Videos Successfully Upload in S3               7                        8
Average download and upload speed for LAPTOP-HU5R4S9U_240116_5 is 0.15584366286514242 MB/s   ③
Average download and upload speed for LAPTOP-HU5R4S9U_240116_1 is 0.2186737656241222 MB/s
                             Video                    Worker
0       5_Second_Video_Watch_the_Milky_Way_Rise_nag  LAPTOP-HU5R4S9U_240116_5
1                      5_SECOND_TIMER__vks           LAPTOP-HU5R4S9U_240116_5
2       The_Judges_Life_Flashed_Before_Her_Eyes_gwo  LAPTOP-HU5R4S9U_240116_5
3    Can_You_Spot_the_Odd_Emoji_in_5_Seconds__Easy_... LAPTOP-HU5R4S9U_240116_5
4    Can_You_Spot_the_Odd_Emoji_in_5_Seconds__Easy_... LAPTOP-HU5R4S9U_240116_5
5    Spot_the_Odd_Emoji_in_5_Seconds__Easy_Emoji_Ch... LAPTOP-HU5R4S9U_240116_5
6    Can_You_Spot_the_Odd_Emoji_in_5_Seconds__Easy_... LAPTOP-HU5R4S9U_240116_5
7        5_Second_Video_Watch_the_Milky_Way_Rise_zoh  LAPTOP-HU5R4S9U_240116_1   ④
8                      5_SECOND_TIMER__sha           LAPTOP-HU5R4S9U_240116_1
9       The_Judges_Life_Flashed_Before_Her_Eyes_wrl  LAPTOP-HU5R4S9U_240116_1
10   Can_You_Spot_the_Odd_Emoji_in_5_Seconds__Easy_... LAPTOP-HU5R4S9U_240116_1
11   Spot_the_Odd_Emoji_in_5_Seconds__Easy_Emoji_Ch... LAPTOP-HU5R4S9U_240116_1
12   Can_You_Spot_the_Odd_Emoji_in_5_Seconds__Easy_... LAPTOP-HU5R4S9U_240116_1
13   Spot_the_Odd_Emoji_in_5_Seconds__Easy_Emoji_Ch... LAPTOP-HU5R4S9U_240116_1
14   Can_You_Spot_the_Odd_Emoji_in_5_Seconds__Easy_... LAPTOP-HU5R4S9U_240116_1
PS C:\Users\jy121\Desktop\twelvelabs> []
```

1. The number of active workers
2. Total number of videos after searching, filtering and uploading for each query
3. Average download and upload speed (MB/s) for each worker
4. Worker who downloaded each video