

TRƯỜNG ĐẠI HỌC SÀI GÒN
KHOA CÔNG NGHỆ THÔNG TIN



PHÁT TRIỂN PHẦN MỀM MÃ NGUỒN MỞ

Xây Dựng Ứng Dụng Tải Video

trên PYTHON

GVHD: Từ Lăng Phiêu
SV: Nguyễn Thanh Thiên Tứ - 3121410570

TP. HỒ CHÍ MINH, THÁNG 4/2024

Mục lục

1	Phần giới thiệu	2
2	Nội Dung	2
2.1	Cơ sở lý thuyết	2
2.1.1	Yt-dlp	2
2.1.2	Flask	3
2.1.3	googleapiclient.discovery	3
2.1.4	json	4
2.2	Thiết kế ứng dụng	4
2.2.1	Tính năng tìm kiếm	4
2.2.2	Tính năng tải xuống	5
2.3	Hiện thực	7
2.3.1	Lấy API Youtube	7
2.3.2	Tính năng tìm kiếm	11
2.3.3	Tính năng download	14
2.4	Cách thức cài đặt ứng dụng	15
2.5	Nhiệm vụ, vai trò của từng thành viên trong nhóm	15
3	Tài Liệu Tham Khảo	15

1 Phần giới thiệu

Đây là một website Tải nhạc được xây dựng bằng Python, tận dụng sức mạnh của Youtube API để mang đến cho người dùng một phương tiện tải video dễ dàng và dễ tiếp cận. Dễ dàng tìm kiếm các video và tải về chúng một cách miễn phí. Dự án này không chỉ thể hiện khả năng kỹ thuật trong việc sử dụng Python, mà còn phản ánh sự tích cực của cộng đồng Python và sức mạnh của các công cụ API mở. Việc xây dựng một website như vậy không chỉ giúp người dùng tiện lợi trong việc tải nhạc mà còn thúc đẩy sự phát triển của ngành công nghệ thông tin. Nhờ vào Python, một ngôn ngữ lập trình đa năng và dễ học, chúng ta có thể tạo ra các ứng dụng mang lại giá trị cho người dùng một cách nhanh chóng và hiệu quả. Đồng thời, việc sử dụng API cũng giúp chúng ta kết nối với các nguồn dữ liệu lớn và cung cấp thông tin chính xác và đầy đủ. Tóm lại, dự án này là một minh chứng cho sức mạnh của Python và API trong việc sáng tạo và phát triển các ứng dụng web mang lại giá trị cho cộng đồng. Chúng ta cần tiếp tục khám phá và phát triển những dự án tương tự để nâng cao trình độ kỹ thuật và tạo ra những sản phẩm hữu ích cho xã hội.

2 Nội Dung

2.1 Cơ sở lý thuyết

Python là một ngôn ngữ lập trình mạnh mẽ và linh hoạt, được ưa chuộng trong việc phát triển web nhờ vào sự đơn giản và dễ đọc của nó. Một trong những ưu điểm lớn nhất của Python là thư viện phong phú, hỗ trợ mạnh mẽ cho việc xây dựng các ứng dụng web. Sau đây là các thư viện đã được sử dụng

2.1.1 Yt-dlp

Yt-dlp là một nhánh của dự án YouTube-DL, một công cụ mã nguồn mở được thiết kế để tải xuống video từ hàng trăm trang web phát trực tuyến khác nhau. So với YouTube-DL, yt-dlp được cập nhật thường xuyên hơn và có nhiều tính năng mở rộng hơn.

Với yt-dlp, bạn có thể tải video từ terminal và tùy chỉnh các thiết lập như loại file, độ phân giải, và nhiều hơn nữa. ọ.

Ưu điểm của yt-dlp:

- **Hỗ trợ nhiều nền tảng:** yt-dlp có thể tải video từ nhiều nguồn khác nhau, không chỉ giới hạn ở YouTube.
- **Cập nhật thường xuyên:** yt-dlp được cập nhật thường xuyên để hỗ trợ các thay đổi từ các trang web video.
- **Hỗ trợ tải xuống video và âm thanh:** yt-dlp cho phép tải cả video và âm thanh riêng lẻ, tùy chỉnh chất lượng và nhiều lựa chọn khác.
- **Tích hợp với YouTube API:** yt-dlp sử dụng API của YouTube một cách thông minh để tải video, giúp giảm thiểu việc bị chặn hoặc đổi đoạn mã.

Nhược điểm của Yt-dlp:

- **Đòi hỏi kiến thức kỹ thuật:** Việc tải và sử dụng yt-dlp có thể đòi hỏi một số kiến thức kỹ thuật cơ bản về lập trình.



- **Không có giao diện đồ họa:** yt-dlp hoạt động chủ yếu thông qua dòng lệnh, điều này có thể làm cho việc sử dụng không tiện lợi đối với một số người dùng.

Tóm lại, yt-dlp là một công cụ mạnh mẽ để tải video từ nhiều nguồn khác nhau, bao gồm YouTube, và đưa ra nhiều tùy chọn linh hoạt cho người dùng. Việc sử dụng yt-dlp cần phải được thực hiện cẩn thận và với kiến thức kỹ thuật nhất định.

2.1.2 Flask

Flask là một microframework web dành cho Python, nổi tiếng với sự nhẹ nhàng và linh hoạt. Nó cho phép lập trình viên xây dựng các ứng dụng web một cách nhanh chóng và dễ dàng.

Ưu điểm của Flask:

- **Đơn giản và dễ sử dụng:** Flask cung cấp một API đơn giản giúp lập trình viên dễ dàng bắt đầu và triển khai ứng dụng.
- **Linh hoạt:** Flask không áp đặt cấu trúc ứng dụng cụ thể, cho phép lập trình viên tự do tổ chức ứng dụng theo ý muốn.
- **Nhẹ nhàng:** Flask không bao gồm các thành phần không cần thiết, giúp giảm bớt tải trọng cho ứng dụng.

Nhược điểm của Flask:

- **Không phù hợp cho ứng dụng lớn:** Flask có thể không phải là lựa chọn tốt nhất cho các ứng dụng web quy mô lớn hoặc phức tạp do thiếu một số tính năng mà các framework lớn hơn cung cấp.
- **Cộng đồng hỗ trợ nhỏ hơn:** So với các framework lớn như Django, cộng đồng hỗ trợ cho Flask có thể nhỏ hơn.

Các hàm của Flask được sử dụng trong project:

- **render_template:** Hàm này được sử dụng để tạo ra một trang HTML từ một template. Nó cho phép lập trình viên truyền dữ liệu vào template và tạo ra trang web động.
- **request:** Đối tượng này cho phép xử lý các yêu cầu từ người dùng, bao gồm việc truy cập dữ liệu được gửi từ client.
- **jsonify:** Hàm này được sử dụng để gửi dữ liệu JSON như một phản hồi từ server, hỗ trợ việc tạo ra các API web.

2.1.3 googleapiclient.discovery

Thư viện googleapiclient.discovery là một công cụ mạnh mẽ cho phép các nhà phát triển xây dựng các thư viện khách hàng, plugin IDE và các công cụ khác tương tác với Google APIs. Nó cung cấp một danh sách các Google APIs và một “Discovery Document” có thể đọc được bằng máy cho mỗi API. Các tính năng chính bao gồm:

Ưu điểm:

- **Tự động hóa:** Tự động tạo ra các phương thức và tham số dựa trên “Discovery Document”, giảm thiểu công việc lập trình thủ công.

- **Hỗ trợ nhiều ngôn ngữ:** Có sẵn thư viện khách hàng cho nhiều ngôn ngữ lập trình, giúp dễ dàng tích hợp vào các dự án đa ngôn ngữ.
- **Cập nhật liên tục:** Các tài liệu khám phá được cập nhật định kỳ, đảm bảo tính nhất quán và cập nhật với các API mới nhất của Google.

Nhược điểm:

- **Phụ thuộc vào kết nối mạng:** Cần có kết nối internet để truy xuất “Discovery Document” từ Google, có thể gây trở ngại khi không có kết nối mạng.
- **Hạn chế tùy chỉnh:** Khi sử dụng các phương thức tự động sinh ra, việc tùy chỉnh có thể bị hạn chế so với việc viết code thủ công.
- **Hiệu suất:** Việc phân tích “Discovery Document” có thể ảnh hưởng đến hiệu suất, đặc biệt là khi tải lần đầu hoặc với các API lớn.

Nói chung, googleapiclient.discovery là một công cụ hữu ích cho việc phát triển ứng dụng sử dụng Google APIs, nhưng cũng cần cân nhắc kỹ lưỡng khi tích hợp vào các dự án có yêu cầu cao về hiệu suất và tùy chỉnh.

2.1.4 json

Thư viện json trong Python cung cấp một cách thuận tiện để chuyển đổi dữ liệu giữa định dạng JSON và các kiểu dữ liệu Python như dictionaries, lists, strings và numbers. Với các hàm như ‘json.load()’ và ‘json.loads()’ để đọc dữ liệu JSON, cùng với ‘json.dump()’ và ‘json.dumps()’ để ghi dữ liệu, thư viện này trở nên không thể thiếu trong việc xử lý dữ liệu định dạng JSON trong các ứng dụng hiện đại.

Ưu điểm:

- **Dễ sử dụng:** Cú pháp đơn giản và trực quan, dễ dàng tích hợp vào các dự án Python.
- **Hiệu suất cao:** Xử lý nhanh chóng, đặc biệt là với dữ liệu lớn.
- **Tính linh hoạt:** Hỗ trợ nhiều kiểu dữ liệu Python và có thể dễ dàng mở rộng.

Nhược điểm:

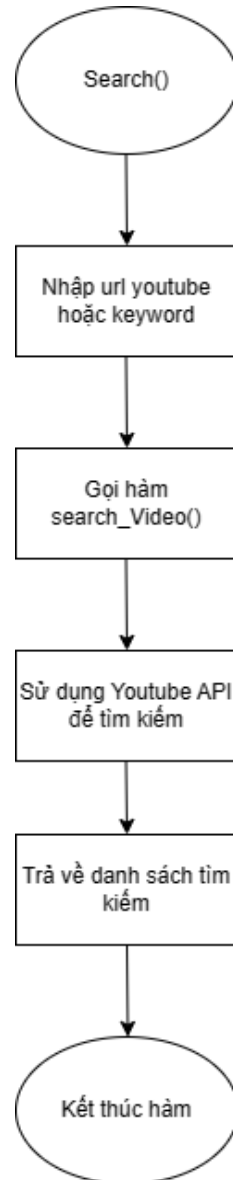
- **Không hỗ trợ tùy chỉnh mặc định:** Không thể serialize các kiểu dữ liệu không phải là built-in mà không cần tùy chỉnh.
- **Không hỗ trợ bình luận:** Định dạng JSON không hỗ trợ bình luận, điều này có thể gây khó khăn khi cần ghi chú trong dữ liệu.
- **Giới hạn về kiểu dữ liệu:** Một số kiểu dữ liệu như tuples không được hỗ trợ một cách tự nhiên và cần phải chuyển đổi thành list trước khi serialization.

2.2 Thiết kế ứng dụng

2.2.1 Tính năng tìm kiếm

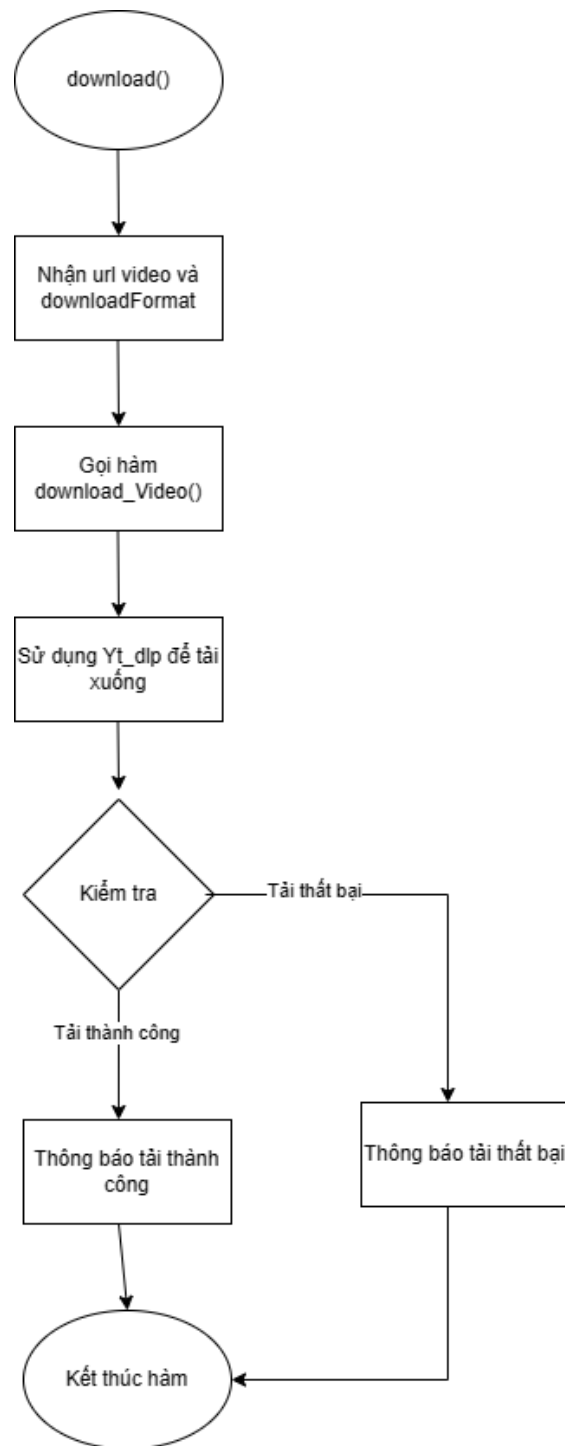
- Sử dụng YouTube Data API v3 để tìm kiếm.
- Người dùng có thể tìm kiếm bằng cách nhập tên video, url video

Cách thức hoạt động



2.2.2 Tính năng tải xuống

- Sử dụng yt-dlp để tải
- Lựa chọn định dạng file trước khi tải xuống



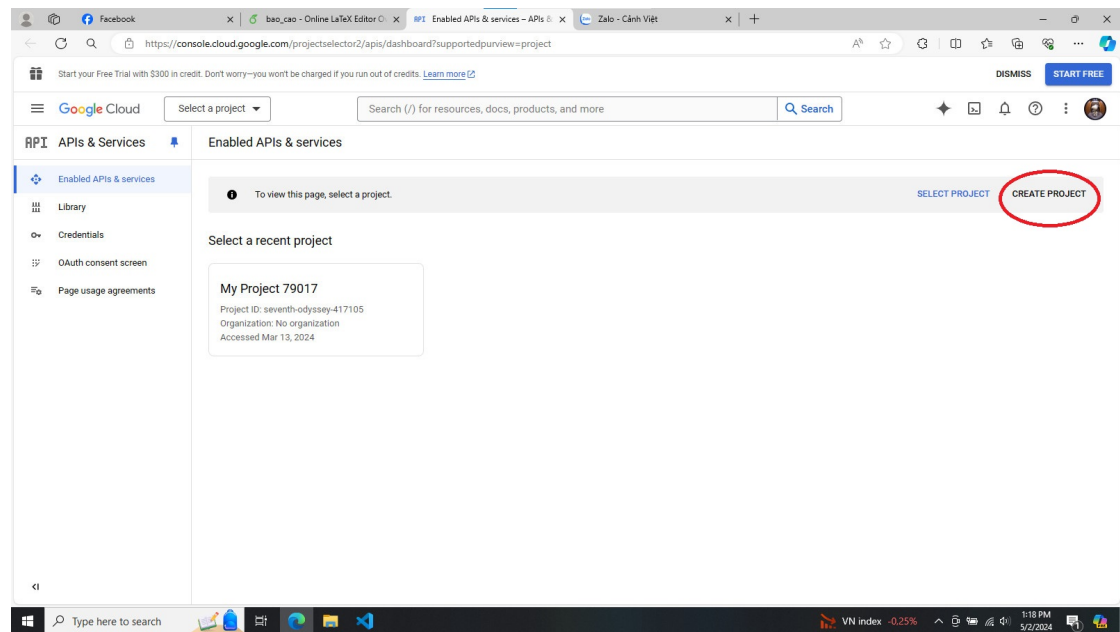


2.3 Hiện thực

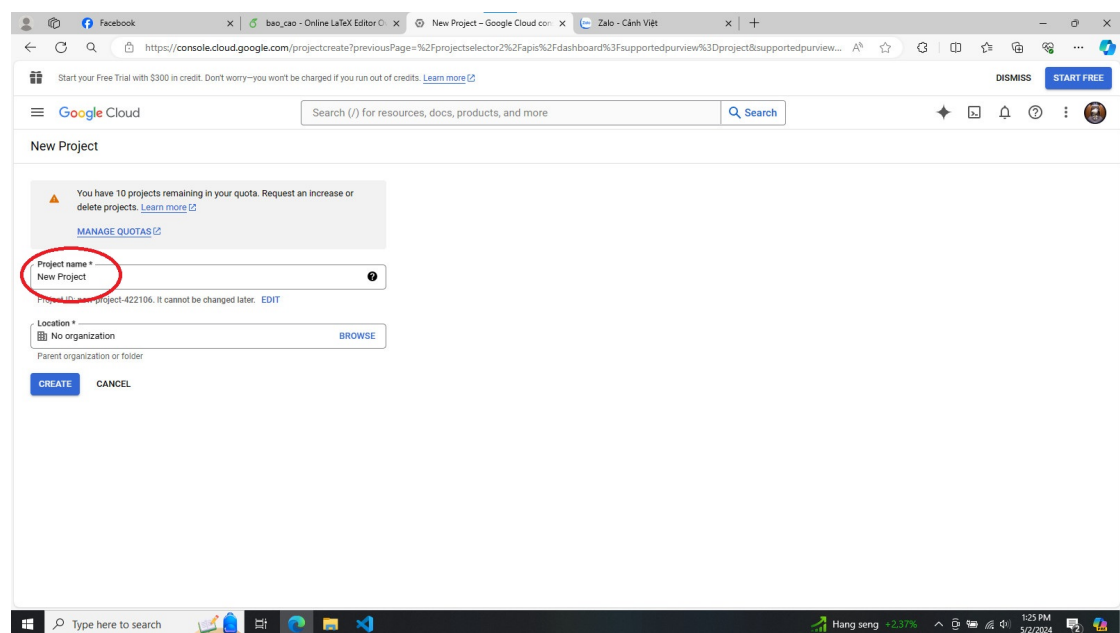
2.3.1 Lấy API Youtube

Bước 1: Truy cập đường dẫn sau: [link](#)

Bước 2: Create project:

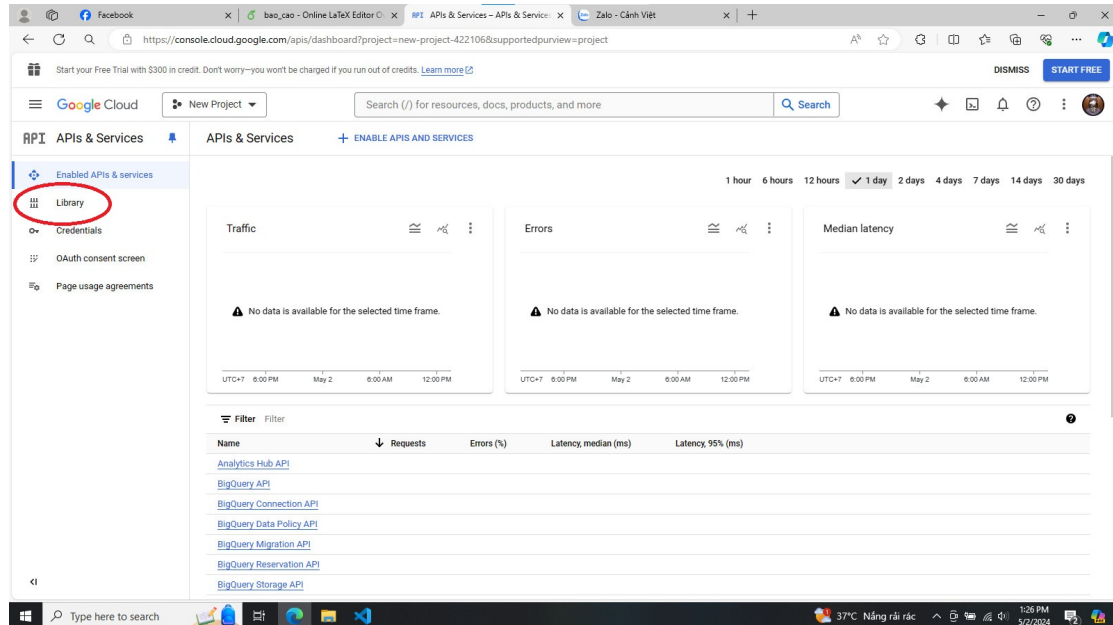


Bước 3: Đặt tên project

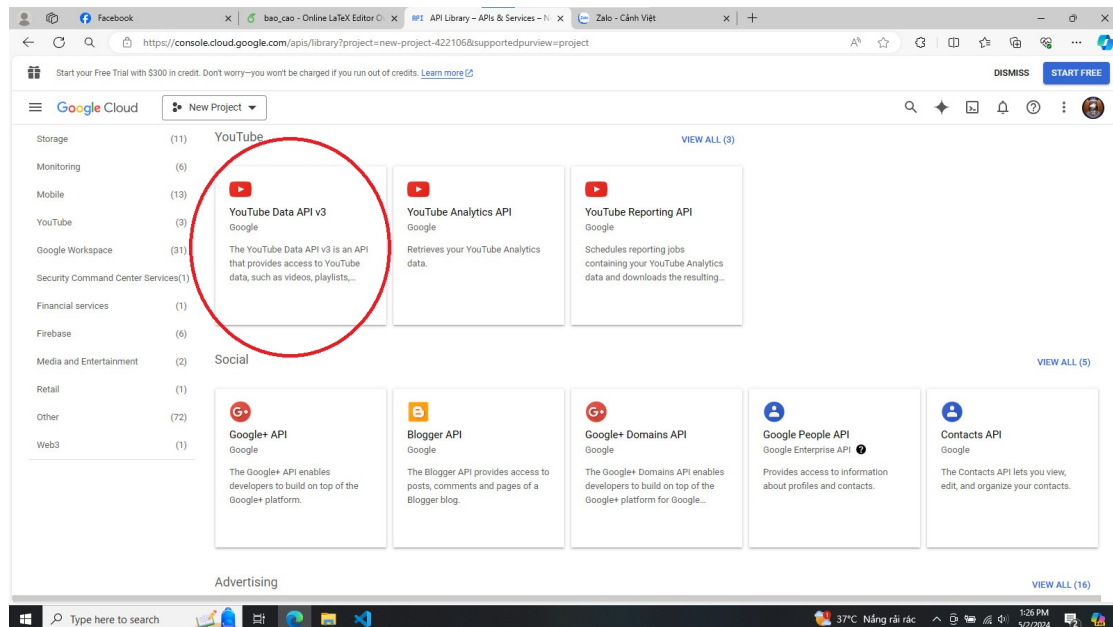




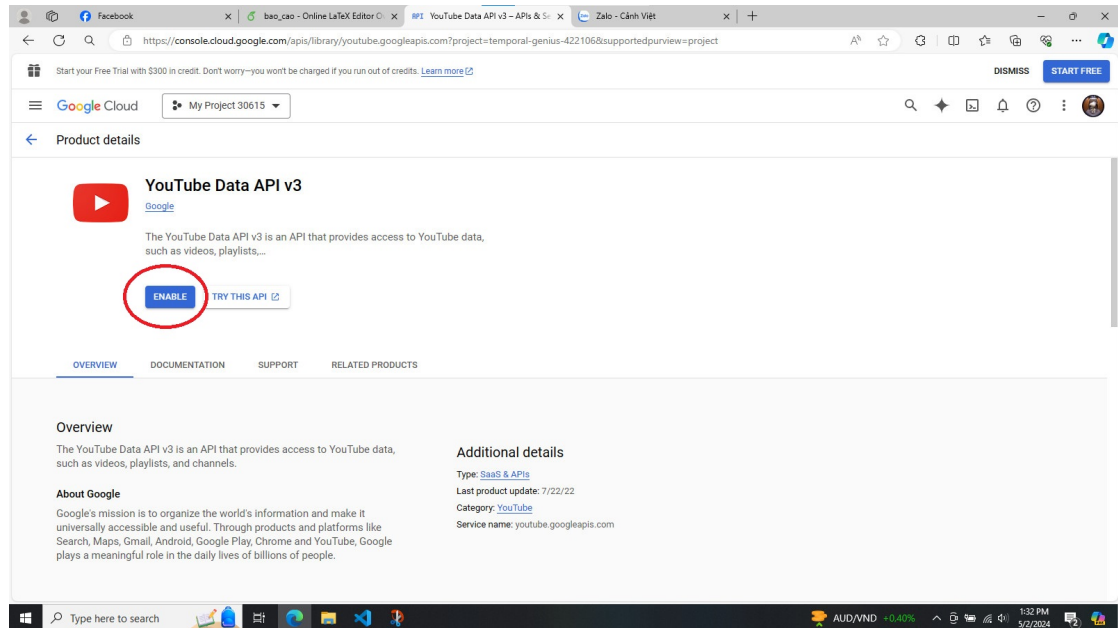
Bước 4: Vào Library



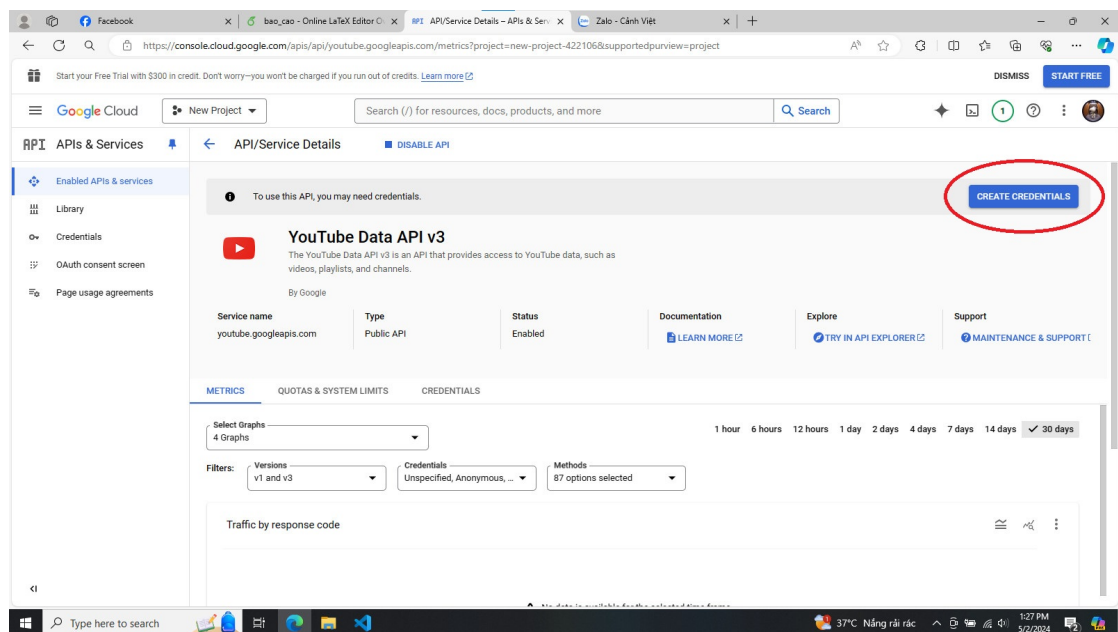
Bước 5: Chọn YouTube Data API v3



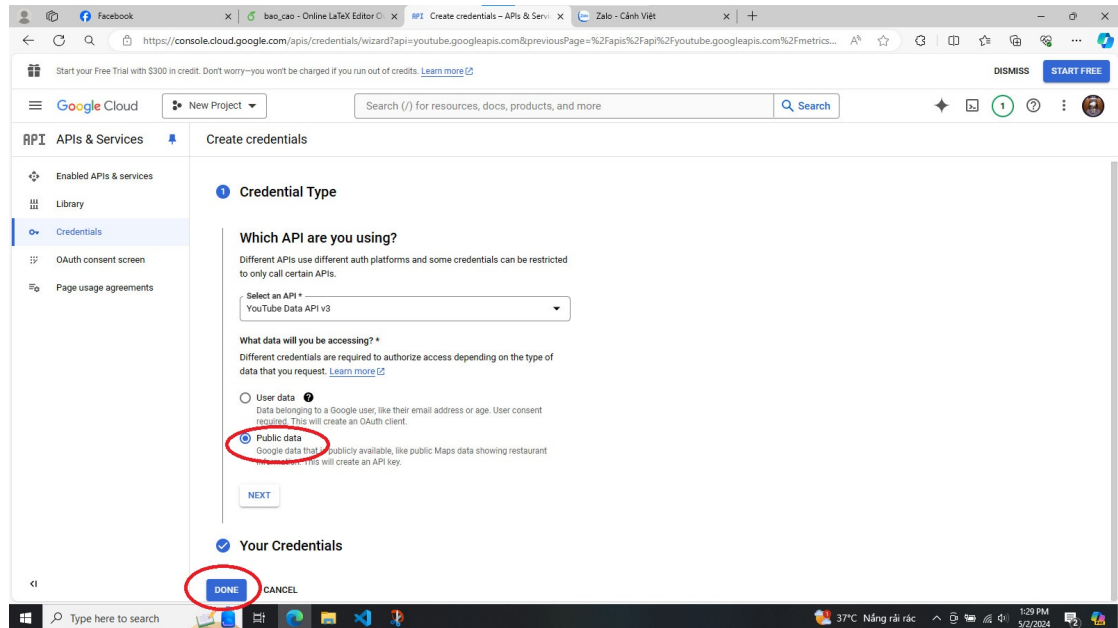
Bước 6: Chọn Enable:



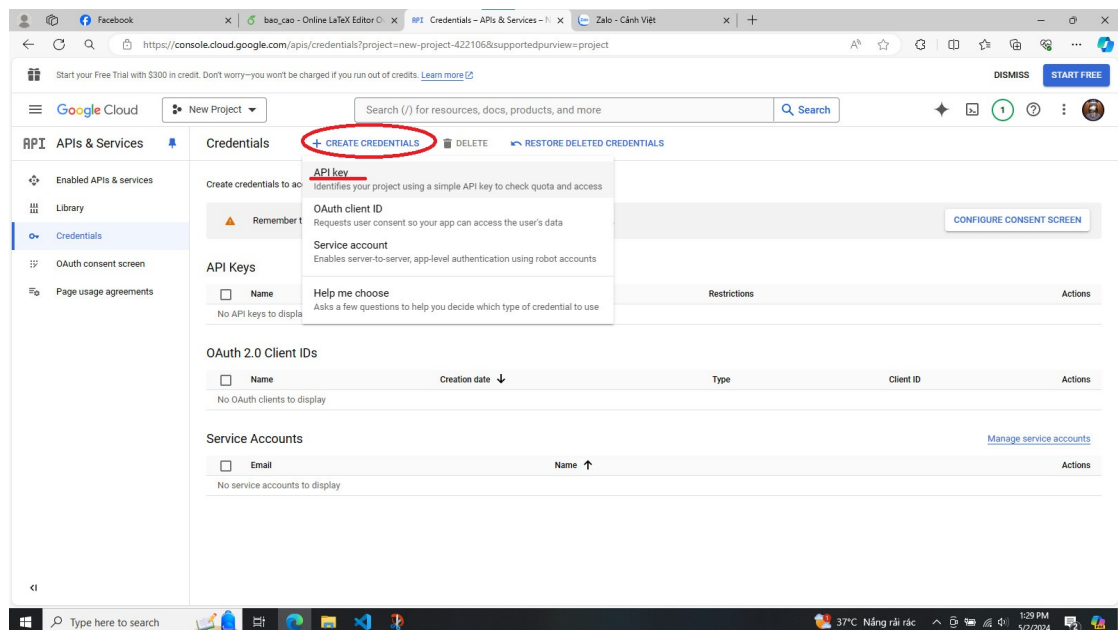
Bước 7: Chọn Create Credentials



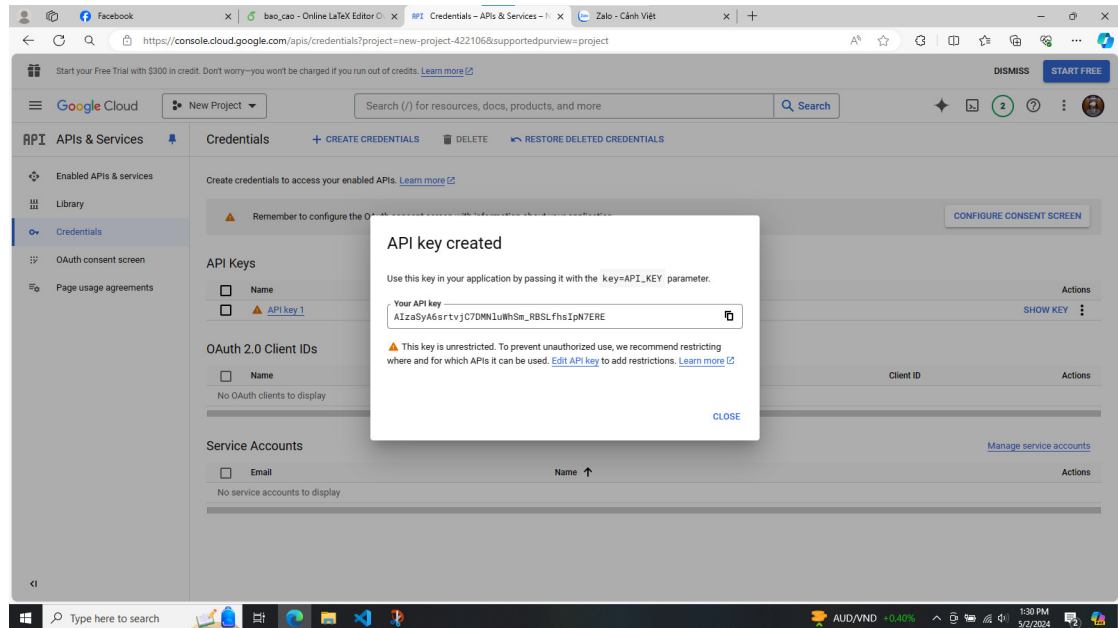
Bước 8: Chọn public data sau đó nhấn Done



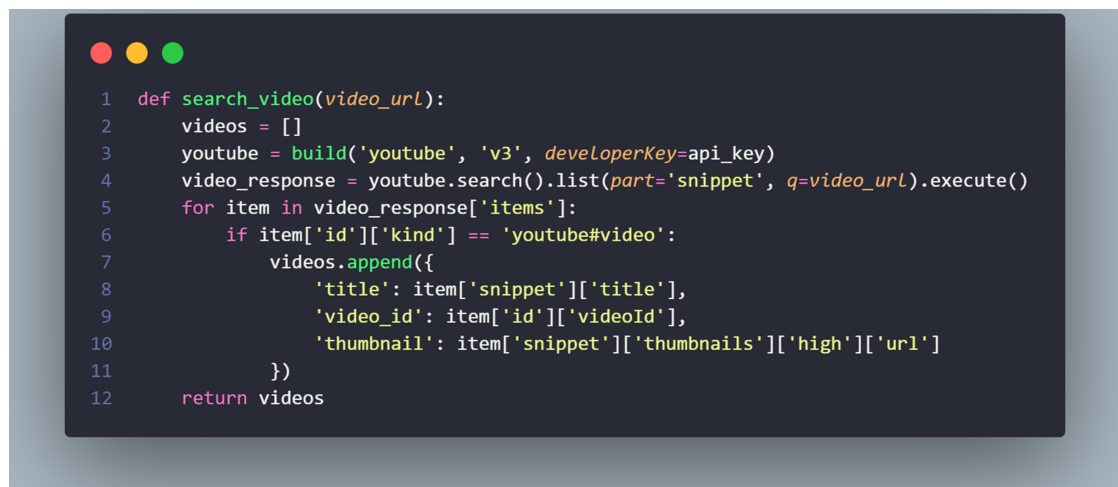
Bước 9: Chọn Create Credentials sau đó chọn API Key



Bước 10: Copy link API vừa tạo



2.3.2 Tính năng tìm kiếm



Giải thích đoạn code trên:

1. Input: video-url: URL hoặc từ khóa tìm kiếm trên YouTube. 2. Logic: Hàm này sử dụng build từ googleapiclient.discovery để tạo một đối tượng YouTube API với khóa API đã cung cấp. Sau đó, nó sử dụng phương thức search().list() của YouTube API để tìm kiếm video dựa trên từ khóa q, được truyền vào từ biến video-url. Kết quả trả về bao gồm danh sách các video phù hợp được trả về dưới dạng một đối tượng JSON. Hàm này lặp qua danh sách kết quả và chọn ra các video có loại là 'youtube#video', và sau đó trích xuất các thuộc tính như tiêu đề, ID và hình ảnh thu nhỏ của video. Cuối cùng, danh sách các video được trả về. 3. Output: Một danh sách các video được tìm thấy trên YouTube, mỗi video là một dictionary chứa các thông tin như tiêu đề,



ID và hình ảnh thu nhỏ-

Ngoài ra còn có các part khác có thể tham khảo thêm ở <https://developers.google.com/youtube/v3/docs/videos>
videoCategoryId:

Ở đây chỉ định danh mục của các video muốn lấy ở đây videoCategoryId = 10 tức là chỉ lấy các video có danh mục là music.

Dưới đây là các danh mục bạn có thể dùng:

- 1 - Film & Animation
- 2 - Autos & Vehicles
- 10 - Music
- 15 - Pets & Animals
- 17 - Sports
- 18 - Short Movies
- 19 - Travel & Events
- 20 - Gaming
- 21 - Videoblogging
- 22 - People & Blogs
- 23 - Comedy
- 24 - Entertainment
- 25 - News & Politics
- 26 - Howto & Style
- 27 - Education
- 28 - Science & Technology
- 29 - Nonprofits & Activism
- 30 - Movies
- 31 - Anime/Animation
- 32 - Action/Adventure
- 33 - Classics
- 34 - Comedy
- 35 - Documentary
- 36 - Drama
- 37 - Family
- 38 - Foreign
- 39 - Horror
- 40 - Sci-Fi/Fantasy
- 41 - Thriller
- 42 - Shorts
- 43 - Shows
- 44 - Trailers

maxResult: Kết quả tối đa lấy được vd 9 thì sẽ lấy được 9 video khi search.

Ví dụ về kết quả trả về khi search



YouTube Downloaded App

tam quoc

SEARCH

Search Results



Toàn bộ Tam Quốc Diễn Nghĩa trong 49 phút

MP3

DOWNLOAD



Khổng Minh Dịch Thân Ra Trận Đêm Theo Bộ Ba Hồ Tướng Quan Trương Triệu Công Chiếm Nam Quận | 555TV

MP3

DOWNLOAD



[FULL] Tóm tắt: Tam Quốc (190 - 280) | Chi tiết & đầy đủ trong 98 phút | Tóm Tắt Lịch Sử

MP3

DOWNLOAD

Ở đây ta có thể lấy các thông tin như videoId, thumbnail, title, ...

2.3.3 Tính năng download

```
1 def download_video(video_url, download_format):
2     message = ''
3     errorType = 0
4     ydl_opts = {
5         'format': 'm4a/bestaudio/best',
6         'postprocessors': [{
7             'key': 'FFmpegExtractAudio',
8         }],
9     }
10    if download_format == 'mp3':
11        ydl_opts['format'] = 'bestvideo+bestaudio/best'
12        ydl_opts['postprocessors'][0]['preferredcodec'] = 'mp3'
13    elif download_format == 'mp4':
14        ydl_opts['format'] = 'bestvideo+bestaudio/best'
15        ydl_opts['postprocessors'][0]['preferredcodec'] = 'mp4'
16    elif download_format == 'webm':
17        ydl_opts['format'] = 'webm/best'
18        ydl_opts['postprocessors'][0]['preferredcodec'] = 'opus'
19    elif download_format == 'ogg':
20        ydl_opts['format'] = 'vorbis/best'
21        ydl_opts['postprocessors'][0]['preferredcodec'] = 'vorbis'
22    elif download_format == 'wav':
23        ydl_opts['format'] = 'wav/best'
24        ydl_opts['postprocessors'][0]['preferredcodec'] = 'wav'
25
26    with yt_dlp.YoutubeDL(ydl_opts) as ydl:
27        try:
28            ydl.download([video_url])
29            message = 'Downloaded successfully!'
30        except yt_dlp.utils.DownloadError:
31            errorType = 1
32            message = 'Invalid video URL!'
33            logging.error('Invalid Video URL: %s', video_url)
34        except Exception as e:
35            errorType = 1
36            message = 'An error occurred!'
37            logging.exception('An Error Occurred: %s', str(e))
38
39    return message, errorType
```



Giải thích đoạn code trên:

Input: video-url: URL của video trên YouTube mà bạn muốn tải về. download-format: Định dạng bạn muốn tải về (ví dụ: 'mp3', 'mp4', 'webm', 'ogg', 'wav'). Logic: Hàm này sử dụng thư viện yt-dlp (một phiên bản thay thế cho youtube-dl) để tải video từ YouTube. Đầu tiên, nó thiết lập các tùy chọn cho quá trình tải xuống dựa trên download-format được cung cấp. Sau đó, nó sử dụng YoutubeDL() từ yt-dlp để tạo một đối tượng tải xuống YouTube. Cuối cùng, nó sử dụng phương thức download() của đối tượng này để tải video từ video-url với các tùy chọn đã thiết lập trước đó. Output: Một tuple gồm message và errorType. Nếu quá trình tải xuống thành công, message sẽ chứa thông báo "Downloaded successfully!" và errorType sẽ là 0. Nếu có lỗi xảy ra trong quá trình tải xuống, message sẽ chứa thông báo lỗi tương ứng và errorType sẽ là 1.

2.4 Cách thức cài đặt ứng dụng

Làm theo hướng dẫn trên link: [here](#)

2.5 Nhiệm vụ, vai trò của từng thành viên trong nhóm

Tên	Công việc	Email
Nguyễn Thanh Thiên Tứ	Thực hiện đề án	nguyenthalthientu18nt@gmail.com

3 Tài Liệu Tham Khảo

Tài liệu

- [1] Yt-dlp “link: <https://github.com/yt-dlp/yt-dlp>”,
- [2] Youtube Data “link: <https://developers.google.com/youtube/v3/docs>”, *API Reference*.