YouTube

Contents

- Add video info
- YouTube loader from Google Cloud

How to load documents from YouTube transcripts.

```
from langchain.document_loaders import YoutubeLoader
```

```
# !pip install youtube-transcript-api
```

```
loader = YoutubeLoader.from_youtube_url("https://www.youtube.com/watch?v=QsYGlZkevEg", add_video_info=True)
```

```
loader.load()
```

Add video info

```
# ! pip install pytube
```

```
loader = YoutubeLoader.from_youtube_url("https://www.youtube.com/watch?v=QsYGlZkevEg", add_video_info=True)
```

```
loader.load()
```

to PDF ▶

Skip to main content

YouTube loader from Google Cloud

Prerequisites

1. Create a Google Cloud project or use an existing project
2. Enable the Youtube Api
3. Authorize credentials for desktop app
4. `pip install --upgrade google-api-python-client google-auth-httplib2 google-auth-oauthlib youtube-transcript-api`

🧑 Instructions for ingesting your Google Docs data

By default, the `GoogleDriveLoader` expects the `credentials.json` file to be `~/.credentials/credentials.json`, but this is configurable using the `credentials_file` keyword argument. Same thing with `token.json`. Note that `token.json` will be created automatically the first time you use the loader.

`GoogleApiYoutubeLoader` can load from a list of Google Docs document ids or a folder id. You can obtain your folder and document id from the URL: Note depending on your set up, the `service_account_path` needs to be set up. See here for more details.

```python
from langchain.document_loaders import GoogleApiClient, GoogleApiYoutubeLoader

# Init the GoogleApiClient
from pathlib import Path


google_api_client = GoogleApiClient(credentials_path=Path("your_path_creds.json"))


# Use a Channel
youtube_loader_channel =
GoogleApiYoutubeLoader(google_api_client=google_api_client,
channel_name="Reducible",captions_language="en")

# Use Youtube Ids

youtube_loader_ids = GoogleApiYoutubeLoader(google_api_client=google_api_client,
```

Skip to main content

```
# returns a list of Documents
youtube_loader_channel.load()
```