

Scraping YouTube Comments

Rodrigo Esteves de Lima-Lopes
State University of Campinas
rll307@unicamp.br

Contents

1	Introduction	1
1.1	What we need	1
2	Downloading the data	1
2.1	Preparing the command	1
2.2	Running it	2

1 Introduction

This script brings a fast solution to download YouTube comments. Here we are going to integrate R, Python and the base system in a very elegant script.

1.1 What we need

Some R packages are important for our exercise:

- `magrittr` - for the `%>%` operator
- `jsonlite` - for the manipulation of json files
- `stringr` - for string manipulation and command building

Outside R we also need:

- A working installation of Python
- A working installation of `youtube-comment-downloader`

2 Downloading the data

2.1 Preparing the command

We will:

1. Create a variable using the ID of a YouTube Video
2. Create a variable with the file name for our comments
3. Create a command to collect our data

```
Video.ID <- "Fv003Gysd7Q"  
#File name  
File.Name <- 'MourasComments'  
#creating a system query  
CMD <- str_glue("youtube-comment-downloader --youtubeid {Video.ID} --output {File.Name}.json")
```

2.2 Running it

Now we

1. Check how the command prints
2. Run the command
3. Import the JSONfile as a dataframe

```
cat(CMD) # Shows the command
system(CMD) # Actual download
# Importing as a dataframe
Comments <- stream_in(file("MourasComments.json"))
```