

TUGAS BESAR_ BEAUTIFULSOUP MRIS



Oleh:

Joko Susilo

1914321012

UNIVERSITAS BHAYANGKARA SURABAYA

TEKNIK INFORMATIKA

2021

Ambil data dari <https://www.wavsource.com/people/people.html> Dengan BeautifulSoup

Link Github

<https://github.com/jokosby123/Tugas-Besar-Bs4-1914321012-Malam-Ubhara-Surabaya-.git>

Langkah – Langkah

A. Mengekstraksi halaman untuk mendapatkan link download

1. Mengimport Package yang di butuhkan seperti gambar di bawah

```
import requests as req
from bs4 import BeautifulSoup as bs
```

2. Memasuki halaman dan mengambil data yang di inginkan

```
header={"Host": "www.wavsource.com",
        "User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/90.0.4430.212 Safari/537.36"}
halaman=req.get("https://www.wavsource.com/people/famous.htm", headers=header)
data_halaman=bs(halaman.text, "html.parser")
data_c1=data_halaman.find_all("td", class_="c1")
```

3. Mengambil data dan Menampilkan Data yang terlah di ambil

```
for ic in data_c1:
    data_script=ic.find_all("script")
    if(len(data_script)>0):
        ambil_data=str(data_script[0])
        ambil_data=ambil_data.replace("<script type=\"text/javascript\">", "")
        ambil_data=ambil_data.replace("</script>", "")
        ambil_data=ambil_data.replace("s2p(", "")
        ambil_data=ambil_data.replace(")", "")
        ambil_data=ambil_data.replace("'", "")
        ambil_data=ambil_data.split(",")
        print("-"*10)
        print("Judul          : {}".format(ambil_data[4]))
        print("Tahun ditambahkan: {}".format(ambil_data[2].replace("added ", ""), ambil_data[3]))
        print("Link Lagu       : {}/{} {}".format('https://www.wavsource.com/snds_2020-10-01_3728627494378403',
                                          ambil_data[0], ambil_data[1]))
        print("-"*10)
```

4. Hasil atau output setelah source code di jalankan

```
Judul          :Patton no fun

Tahun ditambahkan:Dec 13 2001
link lagu      :https://www.wavsource.com/snds_2020-10-01_3728627494378403/
people/famous/patton_no_fun.wav
-----
Judul          :Swaggart sinned

Tahun ditambahkan:July 31 2001
link lagu      :https://www.wavsource.com/snds_2020-10-01_3728627494378403/
people/famous/swaggart_sinned.wav
-----
Judul          :Tutu Mandela

Tahun ditambahkan:Dec 13 2001
link lagu      :https://www.wavsource.com/snds_2020-10-01_3728627494378403/
people/famous/tutu_mandela.wav
-----
```

B. Download audio melalui link yang kita dapatkan dari output source code A

1. Mengimport Package yang di butuhkan seperti gambar di bawah

```
import youtube_dl
```

2. Mengambil data hanya audio

```
ydl_opts = {
    'format': 'bestaudio/best'
}
```

3. Memasukan semua link yang telah kita dapatkan




























```
with youtube_dl.YoutubeDL(ydl_opts) as ydl:
    ydl.download([
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/eastwood_lawyers.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/edison.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/aramis_nation.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/harvey_gonads.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/hawking01.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/hawking02.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/houdini_goodnight.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/ireland_ouch.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/jackson_allegations.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/king_injustice.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/king_nonviolence.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/king_war_no_more.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/leary.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/Lewis_prime-time_ratings.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/ll_cool_j_va_know.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/macarthur_return.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/malcolm_x_by_any_means1.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/manson_believe_me.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/manson_devil.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/mother_teresa.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/nimoy_spock.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/patton_no_fun.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/swaggart_sinned.wav',
        'https://www.wavsource.com/snds_2020-10-01_3728627494378403/people/famous/tutu_mandela.wav'])
```

4. Setelah source code di atas di jalankan maka akan download data wav yang ada di halaman tersebut

```
[download] 100% of 237.54KiB in 00:02
[generic] manson_believe_me: Requesting header
[download] Destination: manson_believe_me-manson_believe_me.wav
[download] 100% of 77.04KiB in 00:01
[generic] manson_devil: Requesting header
[download] Destination: manson_devil-manson_devil.wav
[download] 100% of 52.72KiB in 00:01
[generic] mother_teresa: Requesting header
[download] Destination: mother_teresa-mother_teresa.wav
[download] 100% of 148.40KiB in 00:02
[generic] nimoy_spock: Requesting header
[download] Destination: nimoy_spock-nimoy_spock.wav
[download] 100% of 171.50KiB in 00:02
[generic] patton_no_fun: Requesting header
[download] Destination: patton_no_fun-patton_no_fun.wav
[download] 100% of 220.34KiB in 00:02
[generic] swaggart_sinned: Requesting header
[download] Destination: swaggart_sinned-swaggart_sinned.wav
[download] 100% of 51.92KiB in 00:01
[generic] tutu_mandela: Requesting header
[download] Destination: tutu_mandela-tutu_mandela.wav
[download] 100% of 151.34KiB in 00:01
```

5. Data tersimpan di folder yang dilam terdapat file.py yang telah dijalankan di atas

s PC > Windows-SSD (C:) > My Data > Tugas > MRIS > UAS_BS

Name	#	Title	Contributing artists	Album
 download_wav				
 eastwood_lawyers-...				
 edison				
 ekstraksi_bs				
 graham_any_nation...				
 harvey_gonads-harv...		gonads.wav	Paul Harvey	
 hawking01-hawking...				
 hawking02-hawking...				
 houdini_goodnight-...				
 ireland_ouch-irelan...				
 jackson_allegations...				
 king_injustice-king_i...				
 king_nonviolence-ki...				
 king_war_no_more-...				
 leary-leary				
 lewis_prime-time_ra...				
 ll_cool_j_ya_know-ll...				
 macarthur_return-m...				
 malcolm_x_by_any_...				
 manson_believe_me...				
 manson_devil-mans...				
 mother_teresa-mot...				
 nimoy_spock-nimoy...				
 patton_no_fun-patt...				
 Speech_to_Text				
 swaggart_sinned-sw...				
 tutu_mandela-tutu_...				