Scraping U2 Lyrics with R

Gaston Sanchez

U2 Lyrics

The text of U2 lyrics are available at the AZLyrics website:

https://www.azlyrics.com/u/u2band.html

AZ A B C D E F G H I J K L M N O P Q R S T U V W X Y Z # [ ] Q Search

U2 Lyrics

↓ sort by album   ↓ sort by song

album: **"Boy"** (1980)

I Will Follow

Twilight

An Cat Dubh

Into The Heart

Out Of Control

Lyrics of "I Will Follow"

https://www.azlyrics.com/lyrics/u2band/iwillfollow.html

U2 Lyrics

Play "I Will Follow"
on Amazon Music Unlimited (ad)

"I Will Follow"

I will Follow...

I was on the outside when you said
You said you needed me
I was looking at myself
I was blind, I could not see

A boy tries hard to be a man
His mother takes him by his hand
If he stops to think he starts to cry
Oh why

R Scripts & Files

File structure

```
code/
    script1-scrape-album-year-and-song.R
    script2-download-song-html-files.R
    script3-scrape-u2-lyrics.R
data/
    u2band.html
    u2-songs-info.csv
    u2-lyrics.csv
html_files/
    Boy-1980-adaywithoutme.html
    ...
    Songs-Of-Experience-2017-yourethebestthingaboutme.html
```

File structure

code/
    script1-scrape-album-year-and-song.R
    script2-download-song-html-files.R
    script3-scrape-u2-lyrics.R
data/
    u2band.html
    u2-songs-info.csv
    u2-lyrics.csv
html_files/
    Boy-1980-adaywithoutme.html
    ...
    Songs-Of-Experience-2017-yourethebestthingaboutme.html

outputs

File structure

code/
    script1-scrape-album-year-and-song.R
    script2-download-song-html-files.R
    script3-scrape-u2-lyrics.R
data/
    u2band.html
    u2-songs-info.csv
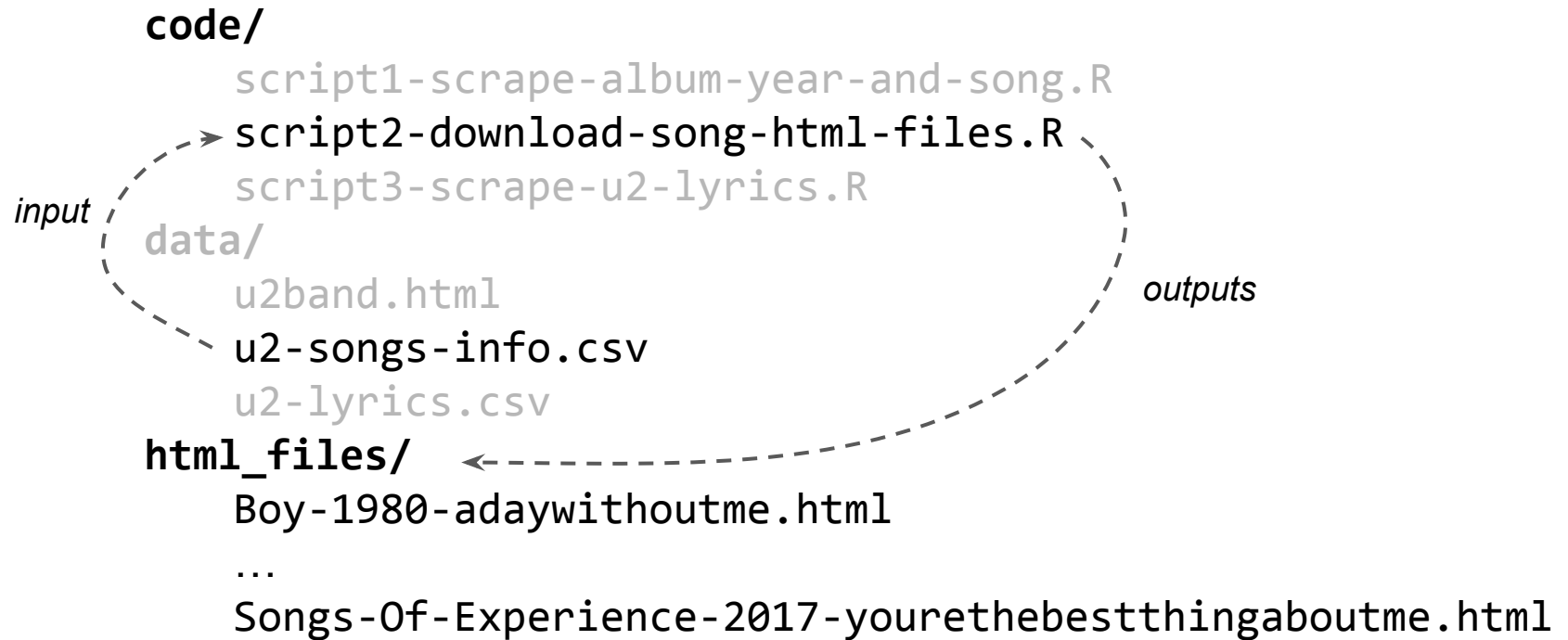   u2-lyrics.csv
html_files/
    Boy-1980-adaywithoutme.html
    ...
    Songs-Of-Experience-2017-yourethebestthingaboutme.html

input

outputs

File structure

code/
  script1-scrape-album-year-and-song.R
  script2-download-song-html-files.R
→ script3-scrape-u2-lyrics.R
data/
  u2band.html   *output*
  u2-songs-info.csv
  u2-lyrics.csv ←
html_files/
  Boy-1980-adaywithoutme.html
  ...
  Songs-Of-Experience-2017-yourethebestthingaboutme.html

inputs