

Scrapy

1. 크롤링이란?

웹페이지를 그대로 가져와서 데이터를 추출해내는 행위

➔ 우리에게 어떻게 적용? : 음원사이트 음원차트나 음원사이트 자체에서 태그(상황, 기분, 날씨, 장소에 맞는 음악)별로 저장해둔 플레이리스트 정보를 추출해옴

2. Scrapy 란?

<https://github.com/scrapy/scrapy.git>

⇒ Scraping을 도와주기 위한 파이썬 기반 라이브러리

⇒ Scrapy를 이용하여 필요한 페이지로 접속하여 원하는 형태로 데이터를 가공하여 데이터를 저장할 수 있도록 도와줌.

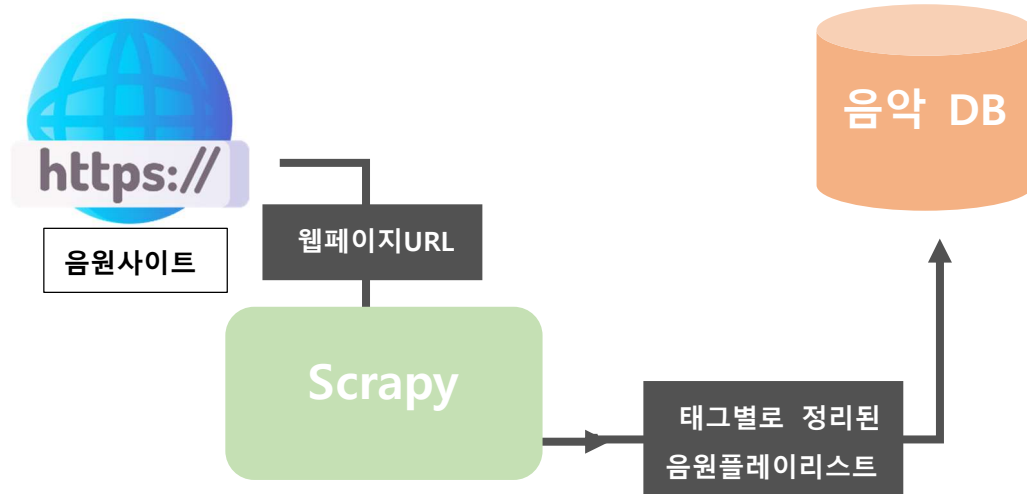
⇒ 포크 9.8k , 스타 45.1k로 GitHub에서 개발자들 사이에 인기있는 크롤러.

⇒ 크롤링을 할 URL을 입력받아 원하는 데이터를 추출함.

3. Scrapy을 이용해 웹 크롤링 예시

1) Scrapy을 이용해 네이버 뉴스 크롤링 -> CSV파일로 데이터 추출





태그별로 정리된 음원플레이리스트

인기 태그

#기분전환

#몽환적

#수능

#집중

#1990년대

#어쿠스틱

#포근한

#가을감성

#발라드

소개글




#카페 #재즈 태그로 유저들이 많이 찾는 곡입니다.

*업데이트 요일: 화요일

#탑백 #top100 #인기 #차트

수록곡 (100)

▶ 듣기 + 담기 ⬇ 다운 ⬇ FLAC ⚙ 선물

<input type="checkbox"/> 번호	곡정보	앨범	좋아요	듣기	담기	다운	유비
<input type="checkbox"/> 1	 Flower Johnny Stimson	Flower	♡ 94,941	▶	+	⬇	🗄
<input type="checkbox"/> 2	 City Of Stars (From "La La Land" So... Ryan Gosling, Emma Stone	라라 랜드 OST	♡ 130,089	▶	+	⬇	🗄
<input type="checkbox"/> 3	 Don't Know Why Norah Jones	Come Away With Me	♡ 54,805	▶	+	⬇	🗄