Скрапер книг из LitRes

Установка пакетов:

* pip3 install requests
* pip3 install BeautifulSoup4
* pip3 install transliterate

Подключение библиотек:

# --------------------подключение библ---------------

from bs4 import BeautifulSoup

import requests

from requests import get

import csv

from transliterate import translit

# ----------------------------------------------------

далее необходимо сформировать строку запроса

создаем переменную для основной части строки

url = 'https://www.litres.ru/pages/rmd\_search/?q='

параметры запроса получим от пользователя:

inputSearch = input('Введите запрос: ')

Создаем строку содержащую, строку запроса с параметрами

l = inputSearch.split()

req = '+'.join(l)

url = url + req

делаем гет запрос и сохраняем страничку

response = get(url)

html\_soup = BeautifulSoup(response.text, "html.parser")

находим все блоки, содержащие нужные данные

div\_data = html\_soup.find\_all('div', class\_='art-item search\_\_item item\_\_type\_art')

проверяем, на наличие элементов

if div\_data != []:

books.extend(div\_data)

вытаскиваем необходимые значения

while count < len(books):

info = books[int(count)]

# в транслите на англ, для избежания ошибок с стандартом кодирования символов

titleBook = translit(info.find('a',{"class":"art-item\_\_name\_\_href"}).text , language\_code='ru', reversed=True)

author = translit(info.find('a',{"class":"art-item\_\_author\_label rmd-author-href"}).text, language\_code='ru', reversed=True)

src = info.find('a',{"class":"art-item\_\_name\_\_href"})['href']

list.append([titleBook, author, 'https://www.litres.ru' + src])

count+=1

сохраняем в файл формата csv

for data in list:

with open ("data.csv","a") as file:

writer = csv.writer(file)

writer.writerow(data)