

Ира: аниме или манга



Дедлайн сдачи: 14.03.24 23:59

Формат сдачи:

1. тетрадка с расширением

.ipynb или ссылка на тетрадку в Google Colab с разрешением на комментирование;

2. полученный JSON-файл с отзывами и оценками, названный reviews.json.



Требования к коду:

1. Сохранить HTML-разметку со

всеми отзывами со всех страниц в файл reviews.html.

2. Оформить часть со скачиванием и сохранением в файл HTML-разметки как функцию

get_html(), которая принимает на вход ту часть ссылки, которая меняется при обращении к разным аниме (например, сравните ссылки на отзывы к "Monster": https://myanimelist.net/anime/19/Monster/reviews, — и к "Jujutsu Kaisen": https://myanimelist.net/anime/40748/Jujutsu Kaisen/reviews).

3. Оформить часть с выделением отзывов и оценок и сохранение их в JSONфайл под названием

reviews.json как функцию get_reviews_stars(), которая принимает на вход название файла, где находится вся HTML-разметка (в нашем случае reviews.html).

4. Следить за РЕР-8 и приличным названием переменных 🤥

У Вас есть сайт myanimelist.net с отзывами на аниме и мангу. Ваши задачи:

Ира: аниме или манга

- 1. Выбрать одно любое аниме (или мангу), у которого есть побольше отзывов на сайте (например, "Monster").
- 2. При помощи библиотеки requests скачать HTML-разметку как минимум первой страницы с отзывами.

NB! Особенно поощряться будет наличие всех отзывов. В этом Вам может помочь посмотреть на то, как, например, задается ссылка на вторую страницу:

https://myanimelist.net/anime/19/Monster/reviews?p=2

- 3. При помощи библиотеки beautifulsoup научиться автоматически доставать со страницы отзыв и оценку (Recommended / Not Recommended / Mixed Feelings) каждого пользователя.
- 4. Сохранить все отзывы и оценки как список словарей с ключами: 'id' (номер отзыва в списке), 'text_review' (полный текст отзыва) и 'rate_review' (рейтинг аниме или манги от пользователя, оставившего отзыв). Например:

```
[
    {'id': 1, 'text_review': 'Monster plays...', 'rate_review'
    {'id': 2, 'text_review': "Don't trust...", 'rate_review'
    ...
]
```

Полученный список сохраните как JSON-файл под названием reviews.json (не забудьте про параметры ensure_ascii и indent).

5. Сдайте в <u>гугл-форму</u> Ваш код целиком, а также полученный файл <u>reviews.json</u>.

Ира: аниме или манга