

# #W8PA: Wikipedia Search Engine

---

1. Описание задания	<b>2</b>
2. Практические задачи	<b>2</b>
3. Задача со звездочкой: Inverted Index API	<b>2</b>
4. Полезные материалы	<b>3</b>

---



## 1. Описание задания

В рамках этого задания вам нужно написать и протестировать Web-сервис, который:

- Проксирует запросы к Wikipedia
- Парсит выдачу поиска по Wikipedia
- Форматирует выдачу для пользователя в виде удобной HTML и Json API
- Обрабатывает ошибки недоступности сервиса или неправильных обращений

## 2. Практические задачи

Задачи:

1. Реализуйте и протестируйте (TDD наше все) обращение к поиску Википедии с помощью библиотеки `requests`, предоставляя возможность пользователю вводить в URL слова для поиска в формате:
  - `.../<endpoint>?word=python&word=network&...`
2. Реализуйте и протестируйте парсинг выдачи Википедии с помощью библиотеки `BeautifulSoup (bs4)` и получение ответа через JSON API;
3. Зарегистрируйте обработчики ошибок 404, 503 (и другие в случае необходимости);
4. Реализуйте отображение результатов поиска в виде HTML-таблицы. Для хранения структур данных рекомендуется использоваться `namedtuple`, чтобы использовать конструкции вида `document.title`, а не `document[1]`;
5. Реализуйте юнит и интеграционные тесты на ваше приложение. Сделайте возможность запускать только юнит тесты, указав аргумент `--skip-integration` для `pytest` (см. `conftest.py`).

Стили и шаблон страницы 404 можно скачать по адресу:

- <https://github.com/big-data-team/python-course#study-artefacts>

Простые стили CSS для оформления таблицы можно положить в `<head>` HTML-документа, см. для примера:

- [https://github.com/big-data-team/python-course/table\\_style.head.html](https://github.com/big-data-team/python-course/table_style.head.html)

## 3. Задача со звездочкой: Inverted Index API

Возьмите решение ДЗ `python.inverted_index` и предоставьте возможность поиска по нему через Web-интерфейс.



## 4. Полезные материалы

Полезные материалы для расширения кругозора:

- [PEP 3333 -- Python Web Server Gateway Interface v1.0.1](#)

Всем удачи!