

Basic Python. Lesson 11

Регулярні вирази. Робота з інтернет-запитами. Синтаксичний аналіз.



План

- Регулярні вирази
- Web-сторінка
- Робота з інтернет-запитами у Python
- Синтаксичний аналіз

Регулярні вирази

Документація: <https://www.regular-expressions.info/reference.html>

Вправи: https://regexone.com/lesson/wildcards_dot

Онлайн редактор: <https://regex101.com>

Модуль re в Python: <https://docs.python.org/2/library/re.html>

Web сторінка

URL: https://uk.wikipedia.org/wiki/Уніфікований_локатор_ресурсів

HTTP: <https://uk.wikipedia.org/wiki/HTTP>

DNS: https://uk.wikipedia.org/wiki/Доменна_система_імен

Сервер-клієнт

Модулі

urllib: <https://docs.python.org/3/library/urllib.html>

requests: <http://docs.python-requests.org/en/master/>

BeautifulSoup: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

Практичне завдання

Написати парсер для <https://docs.python.org/3/>

Результат роботи: коли користувач запускає програму, він бачить пронумерований список категорій(ми дістали його під час заняття). Після цього він вводить номер категорії і йому виводяться всі посилання основної частини сторінки(наприклад для <https://docs.python.org/3/reference/index.html> це всі посилання від 1 до 10 пунктів, а також посилання з абзацу “This reference...”). Виводити їх в форматі “<Назва>: <посилання>”

Questions?