

Робота з internet-файлами

Файли для вивчення цієї теми і для виконання експериментів є в окремій папці "Робота з internet-файлами" навчальних матеріалів. Викладений там матеріал не є вичерпним, треба переглядати також авторську документацію системи програмування і різні дані, які можна знайти в інтернеті.

Посилання для вивчення про формат xml:




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

<https://support.microsoft.com/uk-ua/office/xml-для-новачків-a87d234d-4c2e-4409-9cbc-45e4eb857d44>

<https://support.microsoft.com/uk-ua/office/огляд-xml-у-excel-f11faa7e-63ac-4166-b3ac-c9e9752a7d80>

<https://microsoft.github.io/XmlNotepad/>

Поради щодо виконання завдань

1. Ця робота в значній мірі є експериментальною. Додані файли матеріалів показують один з можливих способів організації роботи з internet-файлами. Можуть бути інші варіанти, які можна спробувати самостійно.
2. Використання модулів Python, показане в прикладах, може бути інакшим, враховуючи модифікацію модулів в нових версіях бібліотеки. Можуть з'явитись нові модулі.
3. Виконуйте трасування сценаріїв прямої роботи в internet, щоб бачити місце та вірогідну причину можливих збоїв.
4. Будьте дуже уважні щодо надсилання запитів до серверів. Не можна надсилати однакові запити дуже часто, бо це може бути трактоване сервером, як спроба вивести з ладу   .

Задача 1. [Оберіть для себе тематику, яка вас цікавить]. В інтернеті знайти за своєю тематикою повну url-адресу сервера і веб-сторінки чи файла, розташованої на сервері. Для цього можна, наприклад, використати звичайний Internet Explorer, виконавши пошук в Internet до отримання кінцевої сторінки веб-сайту (ендпоінт). Internet Explorer показує повну url-адресу в рядку зверху вікна відображення сторінки. Цю адресу треба копіювати у свій сценарій (кириличні літери доведеться записати вручну, щоб зберегти кодування utf-8). Для веб-серверів загального доступу такі адреси зазвичай є постійними.

Задача 2. Складіть сценарій веб-клієнта, щоб отримати на свій комп'ютер знайдену веб-сторінку чи файл. Сценарій веб-клієнта бажано програмувати в режимі try-except, щоб на випадок відмови сайту побачити інформацію про вірогідну причину.

***Увага!** Не всі веб-сторінки дозволені для копіювання в обхід авторського сайту, спроба стороннього запиту на отримання веб-сторінки може бути заблокована. Такі речі треба перевіряти експериментально.*

а) може допомогти повторний запуск сценарію, але з паузою 10-15 секунд;

б) якщо не вдається, знайдіть в інтернеті іншу веб-сторінку.

Коли сторінка отримана, запусіть її на перегляд способами, розглянутими в темі "Деякі методи організації сценаріїв". Не забувайте, що для запуску перегляду в загальному випадку потрібно зазначити повний шлях до файла.

Текст сценарію обов'язково коментувати (#...). Так само коментувати кроки виконання (print()), друкуючи протокол. Подібні сценарії часто дають збій у виконанні, тому треба хоча б побачити трасування виконаної частини сценарію.

Задача 3. [Розшукайте окремим незалежним способом адресу деякого xml-файла за обраною тематикою]. Використовуючи одну з схем доступу до файлів на веб-серверах, складіть сценарій, виконайте, отримайте і збережіть на своєму комп'ютері такий xml-файл (файл даних). Запустіть його на перегляд, щоб переконатись про отримання і про зміст файла.

Задача 4. [Виконайте попередньо вправи на опрацювання xml-файлів, використавши наданий файл data.xml і приклад опрацювання цього xml-файла. Доведеться самостійно вивчити модуль xml.etree.ElementTree ☹].

Складіть самостійно деяку програму опрацювання свого xml-файла, отриманого в задачі 3. Продемонструйте в своїй програмі різні функції і способи опрацювання файла. Конкретний обсяг програми для цієї задачі не регламентуємо, це творча робота. Не забудьте описати задачі.

Надіслати у відповідь:

1) знайдені веб-адреси файлів і сторінок, використані в своєму завданні – окремий перелік; описати зміст таких файлів;

2) тексти сценаріїв чи програм задач 2, 3, 4; не забувайте про коментування і трасування; умови задач опрацювання xml-файла (задача 4) записати ОКРЕМИМ файлом так, щоб сторонні читачі розуміли зміст – він має бути відповідним до змісту xml-файлів з інтернету; не плутати зміст xml-файла і зміст умови задачі опрацювання xml-файла;

3) скріншоти підтвердженого виконання на своєму комп'ютері; скріншоти можна зібрати з підписами в один файл.

Файли надіслати окремими вкладеннями, НЕ АРХІВУВАТИ !
--