

МАТЕРІАЛИ ХХVII
МІЖНАРОДНОГО
МОЛОДІЖНОГО ФОРУМУ

МІНІСТЕРСТВО
ОСВІТИ І НАУКИ
УКРАЇНИ

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ
УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ

РАДІОЕЛЕКТРОНІКА
ТА МОЛОДЬ У ХХІ
СТОЛІТТІ



2023

ТОМ 4

ХАРКІВ

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ РАДІОЛЕКТРОНІКИ

МАТЕРІАЛИ 27-го МІЖНАРОДНОГО МОЛОДІЖНОГО ФОРУМУ
«РАДІОЕЛЕКТРОНІКА І МОЛОДЬ У ХХІ СТОЛІТТІ»

10 – 12 травня 2023 р.

Том 4

КОНФЕРЕНЦІЯ

**«ПЕРСПЕКТИВИ РОЗВИТКУ ІНФОКОМУНІКАЦІЙ ТА
ІНФОРМАЦІЙНО-ВИМІРЮВАЛЬНИХ ТЕХНОЛОГІЙ»**

Харків 2023

УДК 004:[621.317+621.391](06)

27-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у ХХІ столітті». Зб. Матеріалів форуму. Т.4. – Харків: ХНУРЕ. 2023. – 192 с.

В збірник включені матеріали 27-го Міжнародного молодіжного форуму «Радіоелектроніка і молодь у ХХІ столітті».

Видання підготовлено факультетом інфокомунікацій
Харківського національного університету радіоелектроніки

61166 Україна, Харків, просп. Науки, 14
тел./факс.: (057) 7021397

E-mail: mref21@nure.ua

Харківський національний університет
радіоелектроніки (ХНУРЕ), 2023

Програмний комітет конференції

Снігуров А.В. к.т.н., декан факультету ІК

Безрук В.М. д.т.н, зав. каф. ІМІ

Лемешко О.В. д.т.н., зав. каф. ІКІ

Захаров І.П. д.т.н., зав. каф. ІВТ

УДК 004:621.391]:004.6

АНАЛІЗ МОЖЛИВОСТЕЙ ВИКОРИСТАННЯ SELENIUM ДЛЯ ВЕБ-СКРАПІНГУ ЗАДЛЯ ОТРИМАННЯ ДАНИХ З ОНЛАЙН РЕСУРСІВ ТА АВТОМАТИЗАЦІЇ ВЗАЄМОДІЇ ПРОГРАМНИМИ ЗАСОБАМИ

Мишко М.М

Науковий керівник – доц. Костромицький А.І

Харківський національний університет радіоелектроніки, кафедра ІМІ
м. Харків, Україна

тел. +38(099) 031 29 33, email: maksym.myshko@nure.ua

Web scraping is the process of extracting information from websites by collecting data through automated methods. This technique has become increasingly popular in recent years as the internet has become more central to our lives, and businesses have turned to the web to collect data and insights. Python is one of the most popular programming languages for web scraping. However, some websites use complex user interfaces or require authentication, which can make web scraping difficult.

Selenium is a popular tool for automating web browsers, which makes it easier to perform web scraping tasks that require interaction with the web page. In the study, we will explore the use of Selenium for web scraping, including its benefits and drawbacks, and discuss specific use cases where Selenium can be helpful.

Веб-скрапінг - це процес вилучення інформації з веб-сайтів шляхом автоматичної збирання даних. Ця техніка стала все більш популярною в останні роки, оскільки Інтернет став центральним елементом нашого життя, і бізнеси звертаються до Інтернету, щоб збирати дані та інсайти. Python є однією з найпопулярніших мов програмування для веб-скрапінгу.

Однак деякі веб-сайти використовують складні інтерфейси користувачів або потребують автентифікації, що може зробити веб-скрапінг складним. Selenium є популярним інструментом для автоматизації веб-браузерів, що полегшує виконання завдань веб-скрапінгу, які потребують взаємодії з веб-сторінкою.

Метою роботи є аналіз можливостей використання Selenium для веб-скрапінгу, включаючи його переваги та недоліки, і обговорити конкретні випадки, де Selenium може бути корисним.

Випадки використання:

- Одним з найбільш поширених випадків веб-скрапінгу з використанням Python і Selenium є збір даних з веб-сайтів для дослідження або бізнес-цілей.

Наприклад, бізнес може використовувати веб-скрапінг для збору інформації про конкурентів, тенденції на ринку або поведінку споживачів.

Дослідники можуть використовувати веб-скрапінг для збору даних для аналізу, таких як дані соціальних мереж або наукові дані. [1]

- Інший випадок використання Selenium - отримання авторизаційного токена, який може бути використаний для API-запитів. Багато веб-сайтів використовують API-запити для надання доступу до своїх даних, але вимагають авторизаційного токена, що має бути включений до кожного запиту.

За допомогою Selenium можна автоматизувати процес входу в систему та отримання авторизаційного токена, що дозволяє нам здійснювати API-запити.

Можливі проблеми та рішення:

- Однією з найбільших проблем веб-скрапінгу за допомогою Selenium є те, що він може бути повільнішим, ніж інші методи веб-скрапінгу. Оскільки Selenium потребує відкриття браузера та імітації взаємодії з веб-сторінкою, він може працювати повільніше, ніж використання запитів для отримання даних безпосередньо з сервера веб-сайту.

Однак, є декілька способів оптимізації продуктивності Selenium, таких як скорочення кількості взаємодій з веб-сторінкою або використання браузера у «headless» режимі для уникнення рендерингу веб-сторінки на екрані. [2]

- Іншою проблемою веб-скрапінгу за допомогою Selenium є те, що він може бути менш надійним, ніж інші методи веб-скрапінгу. Оскільки Selenium ґрунтується на взаємодії з веб-сторінкою, будь-які зміни у макеті веб-сайту або структурі HTML можуть зламати скрипт скрапінгу.

Для зменшення ризику виникнення цієї проблеми важливо використовувати стабільні селектори, які менш схильні до змін, і моніторити веб-сайт на зміни, які можуть вплинути на працездатність скрипта.

Список використаних джерел:

1. Martin P. What is Web Scraping and What is it Used For? [Електронний ресурс] / Perez Martin. – 2021. – Режим доступу до ресурсу: <https://www.parsehub.com/blog/what-is-web-scraping/>.

2. Nadkarni S. Selenium Headless Browser Testing [Електронний ресурс] / Shilpa Nadkarni. – 2021. – Режим доступу до ресурсу: <https://www.toolsqa.com/selenium-webdriver/selenium-headless-browser-testing/>.

АЛФАВІТНИЙ ПЕРЕЛІК

А

Акіменко А.С 25
Акіменко А.С. 21
Андрущенко О.В. 33, 35

Б

Белозьоров С. Ю. 86, 88
Білик О.С. 37
Божко О.В. 128
Бондаренко В.С. 17
Будянський В.С. 149

В

Вакуленко Д. В. 84
Войлов В.І. 64
Ворончихін О.А. 21
Ворончихін О.А. 25

Г

Гапонюк К.В. 90
Геворк`ян Л.А. 29
Гонтарь І. А. 106,108
Горяінова К.О 42

Д

Діденко Є.С. 94,96
Довгополий С.О. 174
Дригач К.В. 56
Дробяз М.О. 13

Є

Євсюкова О.О. 31
Євсюкова О.О. 112

З

Зражевець К.П. 74,76,78

К

Кабаченко В.О. 110
Канівець В.І. 133
Капуста Р.Д 42
Качан В.Є 54

Кобзєв.В.Д 139

Козін А.О. 155

Копиця А.А. 145

Котенко К.О. 19

Красніков В. О. 161

Красюкова В.В. 104

Кротінов А.П. 141

Кулічко-Павленко І.С. 186

Л

Ліннік М.В.163

Любарець І.О. 170

М

Магдаліна М.І. 120, 122, 124

Майба М.А. 92

Маньковський А.Г. 126

Маслакова 39

Меюс Ю.О.182

Мишко М.М 147

Муха Р.В. 23

Н

Назаров Б. А. 100, 102

Новіченко Є.О. 5, 131

Новіченко Є.О. 131

П

Пастушенко М.С. 44

Пашкова А.В. 66

Петраченко М.О 44

Петрачков М.О. 7

Поддельський В.М. 165

Показій.К.О 56

Поліщук В.Г. 68,70,72

Пономаренко І.О.184

Поповська Є.О. 116

Прийдак О.І. 118

Р

Радченко Р.В. 9

Резніченко Д.Ю. 98
Румянцева О.В 46, 48
Русанова Є.В. 180

С

Сізов Я.А. 15
Скиба Є.О. 82
Славгородський Я.В. 143
Соцька В.В. 153
Сошенко Д.Д. 176
Стахова А.П. 172
Степанов О.О. 135

Т

Твердохліб Л. 178

У

Усатий Д.О. 11

Усов 27

Ф

Фодченко А.В. 151
Фукс М.А. 50,52

Ш

Шалатов В.О. 137
Шедін Д.А. 80
Шлома О.К. 167
Шпількін А. Р. 114
Шрамко В.С. 157
Шульга М.Д. 58, 60, 62
Шумков І.М 33,35

Я

Ярова О. С 159