

Выполнение практического задания № 7

Урок 7. Selenium в Python

Предоставьте ваш Python-скрипт вместе с кратким отчетом (не более 1 страницы), который включает следующее:

URL сайта. Укажите URL сайта, который вы выбрали для анализа.

Описание. Предоставьте краткое описание информации, которую вы хотели извлечь из сайта.

Подход. Объясните подход, который вы использовали для навигации по сайту, определения соответствующих элементов и извлечения нужных данных.

Трудности. Опишите все проблемы и препятствия, с которыми вы столкнулись в ходе реализации проекта, и как вы их преодолели.

Результаты. Включите образец извлеченных данных в выбранном вами структурированном формате (например, CSV или JSON).

URL сайта.

<https://mail.ru/> (<https://e.mail.ru/inbox/>)

Описание.

Список всех писем:

Тема, Отправитель, Дата получения.

Подход.

Объясните подход, который вы использовали для навигации по сайту, определения соответствующих элементов и извлечения нужных данных.

Использовался пакет Selenium WebDriver с соответствующим драйвером браузера.

Использовался поиск и извлечение элементов страницы в сочетании с классом By.

Для навигации по сайту и имитации нажатия клавиш использовались классы Keys и ActionChains для одновременного нажатия кнопок.

Трудности.

Опишите все проблемы и препятствия, с которыми вы столкнулись в ходе реализации проекта, и как вы их преодолели.

Трудность состояла в том, что при прокрутке писем верхние письма терялись и возникали новые. Время загрузки сайта при малой скорости интернета. Поэтому требовалось время на загрузку страниц.

Одновременное нажатие клавиш

Решение: заходим в почту, выбираем самое верхнее письмо (первое) и заходим в него, считываем нужную информацию и сохраняем в базу писем, используем комбинацию клавиш для перехода к следующему письму.

Результаты.

Включите образец извлеченных данных в выбранном вами структурированном формате (например, CSV или JSON).

Например:

Новая Почта: с чего начать? Почта Mail.ru Вчера, 2:18

Выгодные предложения этого лета Utair 4 июня, 16:30

В Анталию со скидкой до 45% Utair 31 мая, 21:40

Очарование Востока Utair 30 мая, 18:33

Смотрите файл letters.json