

СЕКЦІЯ 1

ДОКУМЕНТ В ІНФОРМАЦІЙНОМУ ПРОСТОРІ ХХІ СТ.

Вакуленко Юлія Олегівна

(наук. керівник – канд. екон. наук, доцент Щербіна О. С.)

Донецький національний університет імені Василя Стуса, м. Вінниця

КОНЦЕПЦІЯ ФОРМУВАННЯ КОМУНІКАЦІЇ З ВИКОРИСТАННЯМ ПРОГРАМ-ПАРСЕРІВ У ВІДДІЛІ КАДРІВ ЗАКЛАДУ ВИЩОЇ ОСВІТИ

На сьогодні основною та головною потребою людського життя є інформація, комунікація. Кожного дня люди дізнаються щось нове, спілкуються з іншими людьми, слухають музику тощо. Деколи особа, користуючись мережею Інтернет, витрачає багато часу для пошуку потрібної їй інформації, адже зазвичай для цього використовується не лише одна сторінка чи сайт, а великий масив, що може або надати користь, або навпаки.

Отож, виникає потреба в пришвидшенні пошуку інформації та отриманні хорошого результату, який за основу надає уже готові, конкретні дані, що потребує користувач мережі Інтернет.

Люди навчилися користуватися «парсерами» – програми, що проводять синтаксичний, лексичний, автоматизований аналіз, обробляють велику кількість інформаційних запитів, розбирають документи на частини, вилучаючи звідти необхідні корисні дані користувачу.

Дієслово «to parse» в дослівному перекладі означає «робити граматичний розбір або структурувати».

Мовою всіх, хто працює з даними на сайтах, це слово має своє трактування. Парсери – це програмні продукти, основною функцією яких є отримання необхідних даних, які відповідають заданим параметрам.

Основні завдання парсингу:

- **Пошук даних.** У програму-парсер завантажується вихідний HTML-код сторінки сайту. З кодом починає працювати скрипт, який розбиває весь текст на лексеми, виділяючи необхідну інформацію.

- **Витяг інформації.** Пошук даних відбувається завдяки певним наборам знаків, що описують мету пошуку. Цей набір також називається регулярними виразами. Вони дають змогу виділити з усього масиву тільки цікаві фрагменти.

- **Збереження даних.** Після отримання інформація зберігається у вигляді таблиць або вноситься в базу даних [1], [2].

Парсинг дає змогу здійснювати роботу з даними будь-якої тематики. Серед основних сфер застосування такої технології можна виділити: пошук і наповнення ресурсів текстовим і мультимедійним контентом; товари і ціни в інтернет-магазинах; дані з оголошень, розміщених на спеціальних ресурсах; пошук і збір контактних даних користувачів; в рамках соціальних мереж (наприклад, відгуки та коментарі); сайти, що спеціалізуються на публікації спортивних результатів [2].

Специфіка роботи програм-парсерів і загалом використання методу парсингу приводить до наступного питання: чи дозволено використовувати контент, розміщений у вільному доступі на інших сайтах, в своїх цілях?

Проте парсинг законний, якщо він стосується збору інформації, що знаходиться у відкритому доступі (все, що можна зібрати вручну).

Програми-парсери дають змогу прискорити процес і уникнути помилок через людський фактор. Інша справа, як власник зібраної бази розпорядиться інформацією.

Програма-парсер може не тільки обробляти великий масив різної інформації, а й брати конкретний запит. Наприклад, «парсити» ціну товару, відгуки, структуру сайту.

Нормальним бажанням будь-якого власника інтернет-ресурсу стане захист інформації, розміщеної на сайті. При наповненні сайту контентом, розробленим власними силами, його запозичення може бути вкрай неприємним. Існує кілька способів боротьби з небажаним парсингом: розмежування прав доступу; чорні та білі списки; тимчасова затримка між запитами; різні методи захисту від роботів тощо [3, 4].

Використання парсингу у відділі кадрів Донецького національного університету імені Василя Стуса є багатоплановим. Адже враховуючи значне накопичення об'ємів інформації в сукупності з прогресуючими темпами зростання її кількості та важливість новацій в області інформаційного пошуку, актуальними залишаються питання розробки, вдосконалення та впровадження пошукових алгоритмів [5, 6].

Google Spreadsheet – програма електронних таблиць, що входить до складу безкоштовного веб-програмного пакета програмного забезпечення, пропонованого компанією Google у межах служби Google Drive. Послуга також включає Документи Google і Google Слайди, текстовий процесор і програму презентації відповідно [7].

Задля роботи з великими масивами документів у Донецькому національному університеті імені Василя Стуса є фахівці, які можуть управляти ними правильно та забезпечувати процвітання закладу вищої освіти. Отже, доцільним використанням відділом кадрів Google Spreadsheet для швидкого пошуку інформації стосовно відпустки, звільнення, прийняття на роботу, відомостей про відрахованих студентів або випускників. Найдені дані формуються в електронні таблиці за конкретними ключовими словами.

Наведена програма має такі характеристики:

1. Інтерфейс програми-парсеру Google Spreadsheets (простий і дуже зручний інтерфейс) [8], [9].

2. Імпорт та експорт даних та файлів (імпорт файлів в трьох форматах – odf, xls і csv).

3. Робота Google Spreadsheets (всі основні можливості роботи з електронними таблицями) [10].

Підсумовуючи вищесказане, можна зазначити, що програми-парсери істотно прискорюють процес роботи, особливо якщо є ключові слова. Налаштувавши роботу, можна оперативніше підібрати необхідні для просування запити, тому що програма-парсер за короткий термін «обходить» тисячі сторінок, фільтрує представлені дані, відбираючи серед них потрібні, після чого формує отриманий результат для подальшої обробки, що дає можливість швидко копіювати інформацію з інших сайтів для розміщення на власному ресурсі тощо.

Google Spreadsheets спрямована на використання новітніх інформаційно-комунікаційних технологій та програм, що допомагають вдосконалити управління, покращують взаємодію працівників та сприяють соціальним інноваціям.

Список використаних джерел

1. Парсинг. Різновиди парсингу. Етапи парсингу. URL: <https://www.seonews.ru/glossary/parsing>

2. Що таке парсинг. Обмеження при програмі-парсері. Законність парсерів. Як захиститися від парсерів. URL: <https://blog.calltouch.ru/chto-takoe-parsing>

3. Парсинг. Парсери. Використання парсерів. Алгоритми роботи парсерів. Основи роботи парсерів. Етичні і технічні складнощі роботи з парсерами. URL: <https://ipipe.ru/info/parsing>

4. Говорущенко Т. О., Павлова О. О., Боднар М. А. Сучасні проблеми семантичного аналізу специфікацій вимог до програмного забезпечення. *Вчені записки ТНУ імені В. І. Вернадського*. Серія: Технічні науки. URL: <http://tech.vernadskyjournals.in.ua/journals/2019>
5. Парсинг. URL: <http://parsing.valemak.com>
6. Методи парсингу сайтів. URL: <http://ponka.vnukov.ru/content/metody-parsinga-saytov>
7. Grune D., C. Jacobs. *Parsing Techniques – A Practical Guide* / D. Grune, Chichester: Originally published by Ellis Horwood, 1990. 320 p.
8. Aho A. V. *The theory of parsing, translation, and compiling*. USA: Prentice-Hall, 1972. 121 p.
9. AngularJS. URL: <https://angularjs.org>
10. Popular Technology. URL: <http://www.populartechology.net>



Заблоцька Анна Олександрівна
(*наук. керівник – д-р філол. наук., професор Лукаш Г. П.*)
Донецький національний університет імені Василя Стуса, м. Вінниця

ДОКУМЕНТ В СУЧАСНОМУ ІНФОРМАЦІЙНОМУ СУСПІЛЬСТВІ

Уже не вимагає додаткових доказів те, що практично будь-який документ, який має інформаційну значимість, проходить в суспільстві процес функціональної еволюції, що підпорядковується чітко вираженій закономірності. Він створюється, переважно, для подальшого практичного використання, за винятком випадків, коли він виконує виключно автокомунікативну функцію. Потім він оцінюється з погляду необхідності архівного зберігання і вже як об'єкт такого виду зберігання може стати об'єктом джерелознавчого вивчення. Маючи цю рису, документ співвідносять з цілком об'єктивними, науково обґрунтованими критеріями [1].

Поширення різних видів документів і технологій передавання документованої інформації, характерних для сучасного суспільства, робить актуальним завдання вивчення тих функцій, які можуть виконувати і виконують документи в сучасному інформаційному просторі. При цьому очевидно, що місцевий розвиток письмової грамотності, збільшення числа кількості державних установ і громадських організацій, створення глобальних комунікаційних систем на рубежі ХХ–ХХІ ст. привели до того, що практика документування різних явищ дійсності стала реалізуватися практично необмеженим колом суб'єктів суспільства. Тож, порівняно з попередніми історичними періодами, ця практика вийшла далеко за рамки системи діловодства.

Минули часи, коли ми повністю покладалися на локальне сховище для завантаження своїх важливих документів. Сьогодні технологічна галузь зробила величезний стрибок, і зараз люди переходять від локального сховища до віддаленого серверного сховища, яке відоме як хмара. Це означає, що повна бібліотека, документи чи файли, до яких колись здійснювався доступ із місцевих медіа, тепер знайшли житло на серверах і можуть бути доступними із хмари. У подібний спосіб у житловому товаристві існує безліч файлів та документів, які зберігаються в кімнаті чи фізичному просторі. Використання хмари для зберігання цих файлів може надати великі переваги, такі як доступ до будь-якого місця та