

WEB-SCRAPING В ЗАДАЧАХ ОРОБКИ ДАНИХ WEB-САЙТІВ

Донецький національний університет імені Василя Стуса, м. Вінниця

Останніми роками web-scraping (парсинг) став все більш популярнішим як цінний інструмент для вилучення та аналізу даних. Мається на увазі автоматизований процес вилучення даних із веб-сайтів, перетворення неструктурованого веб-вмісту в структуровані дані, які можна аналізувати та використовувати для різних цілей [1]. Зростання популярності парсингу можна пояснити кількома факторами:

1. Прийняття рішень на основі даних: у сучасному світі, підприємства та організації значною мірою покладаються на статистику на основі даних, щоб приймати обґрунтовані рішення. Web-scraping дозволяє швидко й ефективно збирати великі обсяги даних із різноманітних джерел. Під час парсингу веб-сайтів, підприємства можуть збирати інформацію про конкурентів, ринкові тенденції, відгуки клієнтів, деталі продукту тощо. Ці дані можна використовувати для отримання конкурентної переваги, визначення можливостей і оптимізації стратегій.

2. Дослідження ринку: web-scraping зробив революцію в галузі дослідження ринку. Компанії можуть аналізувати платформи електронної комерції, соціальні мережі, форуми та переглядати веб-сайти, щоб отримати цінну інформацію для споживачів. Ці дані допомагають компаніям зрозуміти вподобання клієнтів, відстежувати репутацію бренду, визначати нові тенденції та аналізувати настрої ринку. Використовуючи методи парсингу, дослідники ринку можуть заощадити час і ресурси, отримуючи актуальну інформацію[2].

3. Моніторинг конкурентів: web-scraping дозволяє підприємствам ефективно контролювати та відстежувати своїх конкурентів. Підприємства можуть збирати інформацію про стратегії ціноутворення, характеристики продукту, рекламні заходи та відгуки клієнтів. Ця інформація дозволяє компаніям приймати обґрунтовані рішення щодо власних пропозицій, цін і маркетингових кампаній. Це допомагає виявити прогалини на ринку та надає цінну інформацію, щоб залишатися попереду конкурентів.

4. Залучення потенційних клієнтів: web-scraping виявився цінним інструментом для залучення потенційних клієнтів. Збираючи інформацію з веб-сайтів, компанії можуть отримати контактні дані, адреси електронної пошти та іншу відповідну інформацію про потенційних клієнтів. Ці дані можна використовувати для створення цільових маркетингових кампаній та охоплення потенційних ринків.

5. Дослідження та наука: web-scraping відіграє важливу роль у дослідницькій та академічній сферах. Дослідники та науковці використовують парсинг для збору даних різноманітних досліджень, аналізу тенденцій у соціальних мережах, моніторингу громадськості та відстеження наукових публікацій. Це дозволяє їм обробляти великі обсяги інформації, яку в іншому випадку було б важко проаналізувати. Парсинг розширює можливості дослідників та сприяє прогресу в різних академічних дисциплінах [3].

6. Порівняння та моніторинг цін: web-scraping дає змогу споживачам порівнювати ціни на різних платформах електронної комерції та знаходити найкращі пропозиції. Інструменти аналізу цін і веб-сайти збирають дані від онлайн-магазинів і представляють їх у зручному форматі, що дозволяє користувачам приймати обґрунтовані рішення про покупку. Крім того, компанії можуть використовувати парсинг, щоб відстежувати зміни цін, визначати цінові тенденції та відповідно коригувати свої цінові стратегії.

Щоб підтвердити все вище сказане, наведемо декілька статистичних даних стосовно парсингу:

1. Згідно зі звітом Grand View Research, до 2027 року глобальний ринок web-scraping-у досягне 7,4 мільярда доларів США, що зросте на 22,4% з 2020 до 2027 року [4].

2. За даними SimilarWeb, провідної платформи веб-аналітики, web-scraping є однією з найпопулярніших технологій у сфері цифрового маркетингу [5].

3. В опитуванні, проведеному компанією Statista, 66% респондентів заявили, що парсинг є важливим для ініціатив бізнес-аналітики та аналітики їхніх організацій.

Підсумовуючи, популярність парсингу пояснюється його здатністю отримувати цінні дані як для компаній, дослідників, так і для споживачів. Оскільки цифровий простір продовжує розширюватися, web-scraping залишатиметься життєво важливим інструментом для збору даних.

Список використаних джерел

1. *How Web Scraping is Transforming the World with its Applications* [Електронний ресурс] – Режим доступу: <https://towardsdatascience.com/https-medium-com-hiren787-patel-web-scraping-applications-abf370d316f4>
2. *13 Reasons Why Web Scraping is Getting More Popular Since Past Decade* [Електронний ресурс] – Режим доступу: <https://www.promptcloud.com/blog/13-reasons-why-web-scraping-is-getting-popular/>
3. *11 reasons why you should use web scraping* [Електронний ресурс] – Режим доступу: <https://www.captaindata.co/blog/11-reasons-why-use-web-scraping>
4. *Global Web Scraper Software Market* [Електронний ресурс] – Режим доступу: <https://www.marketwatch.com/press-release/global-web-scraper-software-market-2023-2030-complete-edge-on-the-industry-with-our-statistical-report-on-top-countries-data-presented-tables-and-graphs-2023-04-05>
5. *What the Future of Web Scraping and Alternative Data Holds in 2023 and Beyond* [Електронний ресурс] – Режим доступу: <https://rayobyte.com/blog/future-of-web-scraping/>