

УДК 339.138

Мухамедова А. – ст. гр.. СНм-51

Тернопільський національний технічний університет імені Івана Пулюя

РОЗРОБКА ВЕБ-СЕРВІСУ “АНАЛІЗАТОР КОДУ HTML” ДЛЯ ВИЯВЛЕННЯ КРИТЕРІЇВ ОПТИМІЗАЦІЇ ВЕБ-РЕСУРСІВ

Науковий керівник: Козак Р.О.

Сучасні пошукові системи самостійно визначають найбільш вагомі слова для тієї чи іншої сторінки сайту. Однак повністю ігнорувати мета-тег keywords неможна. Так як пошукові системи, у разі неправильного утримання цього мета-тега, можуть не тільки знизити рівень видачі, але і накладати санкції.

В інтернеті можна знайти чимало програм аналізу текстів HTML сторінок, але більшість працюють в режимі онлайн і аналізують одиночні сторінки, що знаходяться на веб-серверах. Використання онлайн технології для аналізу всіх сторінок сайту неможливо, так як час аналізу в онлайн режимі займе дуже багато часу і ресурсів. При цьому, аналіз декількох сторінок, хоча і корисний, але не дає можливості отримати повний аналіз всього сайту в цілому.

Аналіз коду сайту (аналіз html, javascript) - етап дослідження коректності внутрішньої розмітки сайту, помилок в тегах і скриптах сайту, використання службових тегів і атрибутів.

Метою роботи є розробка веб-сервісу “Аналізатор коду HTML” для виявлення критеріїв оптимізації веб-ресурсів

Для досягнення поставленої мети розроблений веб-сервіс має розв’язати наступні задачі:

- перевірка на невірні (биті) посилання;
- підрахунок слів(тегів);
- оптимізація сторінок для пошукових серверів(рекомендації);
- переведення інформації з коду HTML у синтаксичне дерево.

Аналізатора вмісту HTML сторінок (або контент-аналізатор) дозволяє розбір html-тексту на складові так само, як це зробить пошукова система. Аналізує html-код сторінки, підраховує вагу і щільність ключових слів, створює звіт про правильність текстової оптимізації сайту. Використовується на етапі створення власного сайту, а також для аналізу сайтів конкурентів. Дозволяє аналізувати як локальні html-сторінки, так і он-лайн проекти. Підтримує особливості російської мови, тому може бути використаний для успішної роботи, як з англійськими, так і з російськими сайтами.