

Київський національний торговельно-економічний університет

Кафедра комп'ютерних наук та інформаційних систем

ВИПУСКНА КВАЛІФІКАЦІЙНА РОБОТА

на тему:

«Модель SEO оптимізації багатосторінкового Web сайту.»

Студента 2 курсу, 3м групи,

спеціальності

122 «Комп'ютерні науки»

Лучука Артема

Тарасовича

підпис студента

Науковий керівник

Доктор технічних наук, професор

Краскевич Валерій

Євгенович

підпис керівника

Гарант освітньої програми

доктор фізико-математичних наук,

професор

Пурський Олег

Іванович

підпис керівника

Київ 2021

Київський національний торговельно-економічний університет

Факультет інформаційних технологій
Кафедра комп'ютерних наук та інформаційних систем
Спеціальність 122 «Комп'ютерні науки»

Зав. кафедри _____ **Затверджую**
Пурський О.І.
«20» грудня 2020р.

**Завдання
на випускну кваліфікаційну роботу студенту**

Лучуку Артему Тарасовичу
(*прізвище, ім'я, по батькові*)

1. Тема випускної кваліфікаційної роботи

«Модель SEO оптимізації багатосторінкового Web сайту.»

Затверджена наказом ректора від «05» листопада 2020 р. № 3311

2. Строк здачі студентом закінченої роботи 26 листопада 2021 року

3. Цільова установка та вихідні дані до роботи

Мета роботи: Обґрунтування та розробка моделі SEO оптимізації багатосторінкового сайту, з урахуванням вимог пошукових систем

Об'єкт дослідження: Процес створення та налаштування SEO моделі

Предмет дослідження: Засоби створення моделі SEO оптимізації

4. Перелік графічного матеріалу _____

5. Консультанти по роботі із зазначенням розділів, за якими здійснюється консультування:

Розділ	Консультант (прізвище, ініціали)	Підпис, дата	
		Завдання видав	Завдання прийняв
1	Краскевич В.Є.	15.12.2020 р.	15.12.2020 р.
2	Краскевич В.Є.	15.12.2020 р.	15.12.2020 р.
3	Краскевич В.Є.	15.12.2020 р.	15.12.2020 р.

6. Зміст випускної кваліфікаційної роботи (перелік питань за кожним розділом)

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ

ВСТУП

РОЗДІЛ 1. ФУНКЦІОНАЛЬНІ ВИМОГИ ДО САЙТІВ ТА ЕЛЕМЕНТИ КОМПЛЕКСНОЇ SEO ОПТИМІЗАЦІЇ СТОРІНОК САЙТУ

1.1. Засоби та технології інформаційного пошуку

1.2. Склад і структура автоматизованих інформаційно-пошукових систем

1.3. Семантичне ядро та власний контент сайту, як підґрунтя SEO

1.4. Технічна оптимізація сайту

РОЗДІЛ 2. ОРГАНІЗАЦІЯ РОЗРОБКИ МОДЕЛІ SEO ОПТИМІЗАЦІЇ

2.1. Алгоритми пошукових систем

2.2. Інформаційно-пошукові мови

2.3. Узагальнений алгоритм пошуку інформації в Інтернеті

2.4. Стандарти XML. Оптимізація пошуку

РОЗДІЛ 3. РЕАЛІЗАЦІЯ МОДЕЛІ SEO ОПТИМІЗАЦІЇ БАГАТОСТОРИНКОВОГО ВЕБ САЙТУ

3.1. Модель SEO оптимізації веб-сайту

3.2. Створення SEO ядра багатосторінкового сайту

3.3. Оптимізовані XML файли для покращеного парсингу сторінок пошуковими роботами

РЕЗУЛЬТАТИ ТА ВИСНОВКИ

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

ДОДАТОК А

ДОДАТОК Б

7. Календарний план виконання роботи

№ Пор.	Назва етапів випускної кваліфікаційної роботи	Строк виконання етапів роботи	
		За планом	фактично
1	2	3	4
1	<i>Вибір теми випускної кваліфікаційної роботи</i>	<i>01.11.20</i>	<i>01.11.2020</i>
2	<i>Розробка та затвердження завдання на випускну кваліфікаційну роботу</i>	<i>05.12.2020</i>	<i>05.12.2020</i>
3	<i>Вступ</i>	<i>01.06.2021</i>	
4	<i>РОЗДІЛ 1. Функціональні вимоги до сайтів та елементи комплексної SEO оптимізації сторінок сайту</i>	<i>25.06.2021</i>	
5	<i>РОЗДІЛ 2. Організація розробки моделі SEO оптимізації</i>	<i>02.09.2021</i>	
6	<i>РОЗДІЛ 3. Реалізація моделі SEO оптимізації багатосторінкового Веб сайту</i>	<i>21.10.2021</i>	
7	<i>Результати та висновки</i>	<i>02.11.2021</i>	
8	<i>Підготовка статті у збірник наукових статей магістрів</i>	<i>18.06.2021</i>	
9	<i>Здача випускної кваліфікаційної роботи на кафедрі науковому керівнику</i>	<i>05.11.2021</i>	
10	<i>Попередній захист випускної кваліфікаційної роботи</i>	<i>01.12.2021</i>	
11	<i>Виправлення зауважень, зовнішнє рецензування випускної кваліфікаційної роботи</i>	<i>03.12.2021</i>	
12	<i>Представлення готової зшитої випускної кваліфікаційної роботи на кафедрі</i>	<i>06.12.2021</i>	
13	<i>Публічний захист випускної кваліфікаційної роботи</i>	<i>За розкладом роботи ЕК</i>	

8. Дата видачі завдання «15» грудня 2020 р.

9. Керівник випускної кваліфікаційної роботи

Краскевич В.Є.

(прізвище, ініціали, підпис)

10. Гарант освітньої програми

Пурський О.І.

(прізвище, ініціали, підпис)

11. Завдання прийняв студент-дипломник

Лучук А.Т.

(прізвище, ініціали, підпис)

АНОТАЦІЯ

Мета дослідження полягає в обґрунтуванні та розробці SEO моделі багатосторінкового web сайту.

У процесі роботи досліджувалися мови програмування, алгоритми розкрутки, SEO, які використовуються для підвищення індексації сайту.

Результатом роботи є SEO модель багатосторінкового web сайту, опис процесу побудови моделі та готові xml-файли.

Обсяг роботи: 28 сторінки, 21 ілюстрація, 7 формул, 1 таблиця, 27 використаних джерел, 2 додаток.

ABSTRACT

The objective is to substantiate and develop an SEO model of a multi-page web site.

In the course of work programming languages, algorithms of promotion, SEO which are used for increase of indexing of a site were investigated.

The result is an SEO model of a multi-page web site, a description of the model building process and ready-made xml files.

Explanatory note: 28 pages, 21 illustrations, 7 formulas, 1 table, 27 used sources, 2 appendixes.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ	8
ВСТУП.....	9
РОЗДІЛ 1. ФУНКЦІОНАЛЬНІ ВИМОГИ ДО САЙТІВ ТА ЕЛЕМЕНТИ КОМПЛЕКСНОЇ SEO ОПТИМІЗАЦІЇ СТОРІНОК САЙТУ	11
1.1. Засоби та технології інформаційного пошуку	11
1.2.Склад і структура автоматизованих інформаційно-пошукових систем.....	13
1.3. Семантичне ядро та власний контент сайту, як підґрунтя SEO.....	17
1.4.Технічна оптимізація сайту.....	20
РОЗДІЛ 2. ОРГАНІЗАЦІЯ РОЗРОБКИ МОДЕЛІ SEO ОПТИМІЗАЦІЇ.....	22
2.1.Алгоритми пошукових систем.....	22
2.2. Інформаційно-пошукові мови.....	30
2.3. Узагальнений алгоритм пошуку інформації в інтернеті.....	32
2.4. Стандарти XML. Оптимізація пошуку.....	33
РОЗДІЛ 3.РЕАЛІЗАЦІЯ МОДЕЛІ SEO ОПТИМІЗАЦІЇ БАГАТОСТОРИНКОВОГО ВЕБ САЙТУ	36
3.1. Модель SEO оптимізації веб сайту	36
3.2. Створення SEO ядра багатосторінкового сайту.....	39
3.3. Оптимізовані XML файли для покращеного парсингу сторінок пошуковими роботами.....	41
РЕЗУЛЬТАТИ ТА ВИСНОВКИ	Ошибка! Закладка не определена.
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	51
ДОДАТОК А.....	53
ДОДАТОК Б.....	64

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ

SEO	—	Search Engine Optimization.
ІПМ	—	інформаційно-пошукова мова.
ПОЗ	—	пошуковий образ запиту.
ІПС	—	інформаційно-пошукова система.
ПОД	—	пошуковий образ документу.
ІР	—	пошукове розпорядження.
ТІЦ	—	тематичний індекс цитування.
ІП	—	інформаційний пошук.
URL	—	Uniform Resource Locator.
СУБД	—	система управління базами даних.
SQL	—	Structured Query Language.
XML	—	eXtensible Markup Language.

ВСТУП

Використання методів SEO просування – один з альтернативних методів позиціонування компанії та інформування цільової аудиторії про товар. Саме в Інтернеті користувачі шукають інформацію про товар та магазини, де цей товар можна придбати. Завдяки інформації, яка розміщена на сторінках сайту користувачі можуть отримати уявлення про якість товару чи послуг та переваги однієї компанії, порівняно з компаніями-конкурентами.

Створення моделі пошукової оптимізації дозволить збільшити попит на товар підприємства, сповіщати клієнтів про актуальні товари, ціни та оновлення за рахунок правильних ключових запитів. Всебічний аналіз потреб клієнтів дозволить вчасно реагувати на зміни в попиті та проводити вчасно зміну пошукових запитів компанії з метою виведення компанії на більш вигідні позиції, порівняно з компаніями-конкурентами.

Це потребує в свою чергу значних зусиль. Так, ефективність пошукової моделі залежить від різних факторів, основні з яких:

- контент сайту;
- доступність викладеної інформації;
- правильно збудована семантика сторінок сайту;
- коректне відображення на програмному забезпеченні різного року оновлення;
- обрана оптимізація (біла, сіра, чорна).

Актуальність дослідження полягає в тому, що створення моделі SEO оптимізації багатосторінкового сайту дозволить збільшити рівень цитування сайту в Інтернеті та збільшити наплив клієнтів з чистої видачі.

Під **об'єктом** дослідження розуміємо процес створення та налаштування SEO моделі.

Предметом дослідження постають засоби створення моделі SEO оптимізації.

Мета дослідження – полягає в обґрунтуванні та розробці моделі SEO оптимізації багатосторінкового сайту, з урахуванням вимог пошукових систем.

Для досягнення поставленої мети дослідження головними завданнями є:

- аналіз трендів серед топових Інтернет-ресурсів;
- обрання оптимального методу створення моделі пошукової оптимізації.
- розробка моделі пошукової оптимізації web-сайту;
- розробка семантичного ядра web-сайтів;
- реєстрація сайту в пошукових системах.

Методи наукових досліджень, що використовувались для написання роботи:

- емпіричні (порівняння, вимірювання);
- комплексні (моделювання);
- теоретичні (сходження від абстрактного до конкретного).

Практична цінність даної роботи полягає в створенні універсальної моделі SEO оптимізації багатосторінкового сайту для пришвидшення ранжування сайту в пошукових системах.

Інформаційною базою дослідження виступають наукові публікації вітчизняних та зарубіжних фахівців та електронні ресурси Інтернет.

Наукова новизна одержаних результатів полягає в розробці SEO моделі, що забезпечує мультиплатформенність відносно пошукових систем та швидкість парсингу пошуковими роботами.

Публікації. Результати дослідження опубліковано у збірнику наукових статей студентів, які здобувають освітній ступінь магістра за спеціалізацією «Комп'ютерні науки» КНТЕУ на тему: «Модель SEO оптимізації багатосторінкового Web-сайту», 2021 р.

РОЗДІЛ 1

ФУНКЦІОНАЛЬНІ ВИМОГИ ДО САЙТІВ ТА ЕЛЕМЕНТИ КОМПЛЕКСНОЇ SEO ОПТИМІЗАЦІЇ СТОРІНОК САЙТУ

1.1. Засоби та технології інформаційного пошуку

Робота пошукових роботів відбувається в такий спосіб: пошукові роботи— спеціальні програми цілодобово переглядають простір мережі й копіюють усі знайдені сторінки, переходять за гіперпосиланнями, знаходять нові документи, копіюють їх тощо. З таких сторінок складається спеціальна база даних, іменована покажчиком. Завдяки цим індексним файлам запит клієнта може бути оброблений практично миттєво. Кожна пошукова система проводить індексацію за своїми алгоритмами, які є комерційною таємницею.

Основним засобом передачі інформації в часі і просторі є документ. Документ визначається як засіб закріплення на спеціальному матеріалі певної інформації про факти, події, явища об'єктивної дійсності і розумової діяльності людини [1]. Документи мають різну форму подання. В автоматизованих інформаційно-пошукових системах – це текстова інформація на природній мові.

Інформаційний запит — інформаційна потреба, яка сформульована на природній мові. Результат «перекладу» інформаційного запиту на інформаційно-пошукову мову (ІПМ) називають пошуковим образом запиту (ПОЗ). Синтаксис і семантика ІПМ визначається структурою та наповненням документів, та загальними задачами системи.

Інформаційний пошук розрізняють наступним чином:

- в залежності від мети – адресний пошук (формально-механічний) та семантичний (тематичний);
- в залежності від об'єкту пошуку – документний та фактографічний;
- в залежності від ступеню використання технічних засобів – ручний або автоматизований;

- в залежності від функціональної ролі – домінуючі/другорядні, центральні/периферичні, сталі/ситуативні потреби [3].

Інформаційно-пошукова система (ІПС) – це комплекс пов'язаних між собою окремих частин, який призначений для виявлення в будь-якій множині елементів інформації, яка відповідає заданому інформаційному запиту. Масив елементів інформації, в якому виконується інформаційний пошук, називається пошуковим масивом [4].

В загальному вигляді процедура інформаційного пошуку складається з чотирьох етапів:

- уточнення інформаційної потреби і формулювання запиту;
- визначення сукупності інформаційних масивів;
- вилучення інформації з інформаційних масивів;
- ознайомлення користувача з отриманою інформацією і оцінювання результатів пошуку.

Розташування Інтернет-ресурсів у верхніх рядках пошукових систем можна досягти різними способами: як легальними, так і не зовсім. З огляду на те існують різні види оптимізації веб-сайтів: біла, сіра та чорна.

Біла оптимізація – робота над ресурсом без застосування офіційно заборонених пошуковими системами методів “розкручування ресурсу” – без впливу на пошукові алгоритми.

Передбачає роботу над Інтернет-ресурсом, а саме над внутрішньою навігацією й вмістом та зовнішнім середовищем сайту, тобто просування за допомогою оглядів, прес-релізів, реєстрації в партнерських програмах тощо.

До сірої пошукової оптимізації належать засоби додання великої кількості ключових слів у текст сторінки [6].

До чорної оптимізації належать всі методи, які суперечать правилам пошукових систем і спричиняють блокування цільового проекту. Це зокрема: використання doorway (сторінок і ресурсів, створених спеціально для роботів пошукових систем, часто з великою кількістю ключових слів на сторінці),

клоакінг (користувачеві відображається одна сторінка, що легко читається, а пошуковому роботу – інша, оптимізована під будь-які запити), застосування прихованого тексту на сторінках сайту, використання “однопиксельних посилань”.

Найбільш загальний вигляд алгоритму пошуку, що проводиться незалежно від форми носіїв і рівня автоматизації, показано на рис. 1.1.



Рис. 1.1. Загальний вигляд алгоритму пошуку

1.2.Склад і структура автоматизованих інформаційно-пошукових систем

Інформаційно–пошукові системи (ІПС) — це різновид автоматизованих інформаційних систем, в яких завершальна обробка даних не передбачається. Ці системи призначені для пошуку текстів (документів, їх частин, фактографічних записів) в сховищах (базах даних) за формальними характеристиками. Тому в роботі ІПС можна виділити два основних етапи: перший — збір і зберігання інформації, другий — пошук і видача інформації користувачам.

ІПС відрізняються одна від одної за багатьма ознаками, але при вирішенні задач збору, зберігання і видачі інформації мають такі спільні процедури:

- аналіз документів і їх добір;
- створення пошукового образу документів (ПОД);
- запис документів і їх пошукових образів на прийнятні носії;
- зберігання документів і ПОД;
- аналіз запитів;
- видача документів користувачам.

ІПС можна класифікувати за такими ознаками:

- родом виконуваних операцій;
- режимом пошуку;
- типом інформаційно-пошукової мови (ІПМ);
- типом критерію відповідності;
- ступенем автоматизації [12].

Комплексні ІПС містять сукупність елементів документальних, фактографічних і логічних ІПС.

За режимом пошуку ІПС діляться на системи, що працюють в режимі вибіркового розподілу інформації, тобто за постійними інформаційними запитамі в масивах постійно поповнюваних документів, і системи ретроспективного пошуку за разовими змінними запитамі абонентів з пошуком документів з даної тематики в масивах.

За типом інформаційно-пошукової мови ІПС класифікуються на системи з природними (людськими) і інформаційними мовами. Інформаційно-пошукові мови — це синтетично створені мови для ідентифікації і пошуку документів за запитом [18].

За критерієм відповідності, тобто сукупністю правил визначення ступеня формальної відповідності між ПОД (ПОД — виражений у термінах мови ІПС основний зміст документа, що займає значно менше машинної пам'яті від повного змісту документа) і пошуковим розпорядженням ІП (ІП — виражений у термінах ІПС пошуковий запит), ІПС поділяються на системи з використанням логічних схем (І, ЧИ, НЕ); з використанням аналітичних функцій (статистичні і векторні

критерії); з використанням «вагових» функцій чи коефіцієнтів; з аналізом критерію відповідності на збіг чи часткове входження (пошукового розпорядження в ПОД).

За ступенем автоматизації ІПС класифікуються на системи з автоматичною класифікацією масивів (формування в ПОД породження класів за мітками, на основі асоціативних зв'язків між термінами в ПОД), системи з автоматичним пошуком для сортування і порівняння ПОД з пошуковим розпорядженням, системи з автоматичним індексуванням для автоматизації процесів згортання при реферуванні, анотуванні, виборі ключових слів і т. ін., системи з автоматичною видачею різновидів інформації (бібліографічних описів, ПОД, копій документів), системи з автоматичним управлінням (при наявності зворотного зв'язку і зміни режимів пошуку в ІПС).

Пошукові системи зазвичай мають три компоненти:

- агент (павук, кроулер або робот), який переміщується по мережі і збирає інформацію;
- база даних(індексатор) яка містить інформацію, що зібрано павуками;
- пошуковий механізм(пошукова машина), який користувачі використовують як інтерфейс для взаємодії з базою даних.

Засоби пошуку типу агентів, павуків, кроулерів і роботів використовуються для збору інформації про документи, які знаходяться в мережі Інтернет. Це спеціальні програми, які займаються пошуком сторінок в мережі, збирають гіпертекстові посилання з цих сторінок і автоматично індексують інформацію, яку вони знаходять для побудови бази даних. Кожний пошуковий механізм має власний набір правил, якими визначається збір документів.

- Спайдери (Spider, Crawler, Robot) — програма, що відвідує веб-сервери, зчитує й індексує цілком або частково їхній вміст і далі рухається через посилання, знайдені на сервері. Спайдер повертається через визначені періоди часу, наприклад, кожного місяця, і повторює індексацію сторінок.
- Агенти є найінтелектуальнішими з пошукових засобів.

Павуки здійснюють загальний пошук інформації в Інтернеті. Вони повідомляють про зміст знайденого документа, індексують його і добувають підсумкову інформацію. Вони також можуть переглядати заголовки, деякі посилання і відправляти проіндексовану інформацію до бази даних пошукового механізму [14].

- Кроулери переглядають заголовки і повертають користувачу тільки перше посилання.
- Роботи можуть бути запрограмовані таким чином, щоб переходити по різним посиланням різної глибини вкладеності, виконувати індексацію і перевіряти посилання в документі. Але, вони можуть застрягати в циклах, адже, проходячи за посиланнями, їм потрібні значні ресурси мережі. Існують методи, що забороняють роботам пошук по сайтах, власники яких не бажають, щоби вони були проіндексовані.

Різні пошукові системи використовують різні алгоритми ранжування, однак основними принципами визначення релевантності є наступні:

- кількість слів запиту, який є у текстовому вмісті документу;
 - теги, в яких ці слова розташовуються;
 - місцеположення шуканих слів у документі;
 - питома вага слів, відносно яких визначається релевантність сторінки, у загальній кількості слів документу;
 - час — тобто як довго сторінка знаходиться в базі пошукового сервера.
- Спочатку здається, що це недолугий принцип;
- індекс цитованості показує як багато посилань на дану веб-сторінку ведеться з інших сторінок, що зареєстровані у базі пошукової системи.

База даних виводить ранжований таким чином перелік документів з html-кодами і повертає його користувачу, який зробив запит. Різні пошукові механізми вибирають різні способи показу отриманого переліку — деякі відображають лише посилання, інші — виводять посилання з декількома першими реченнями документу чи заголовком документу разом з посиланням. Коли користувач

звертається до по-силання на один з документів, цей документ завантажується з сервера, на якому він знаходиться.

1.3. Семантичне ядро та власний контент сайту, як підґрунтя SEO

Семантичне ядро - це набір ключових фраз і запитів, які допомагають сайту знайти свою цільову аудиторію, а значить, потенційних клієнтів. Семантичне ядро визначає зовнішній попит на товари й послуги, представлені на сайті. Правильно складені ключові фрази і пошукові запити є основою просування сайту в Інтернеті. За допомогою цього інструменту досягається збільшення продажів і реалізація послуг.

Тому правильно спочатку скласти семантичне ядро, а тільки потім на його основі будувати сайт, а не навпаки. Від даних ядра безпосередньо залежить зміст ресурсу, вибираються сторінки та статті для просування, складається правильне опис товару або послуги. Втім, якщо сайт вже є, він непопулярний і не приносить прибутку, досвідчені фахівці можуть його оптимізувати, подавши потрібні зміни відповідно до семантичного ядра.

Складання семантичного ядра вимагає ретельного аналізу, досвіду і спеціальних програм, щоб знайти найтісніший взаємозв'язок між попитом і пропозицією, між змістом вашого сайту й тими запитами, які будуть надходити від користувачів.

Процес створення можна розділити на кілька етапів:

- аналіз тематики і текстів сайту. На його основі фахівець складає первинний список ключових запитів. Їх може бути від 50 до 100 і більше;
- підбір синонімів, слів з одним коренем, а також словосполучень зі схожим змістом. На основі первинного списку, за допомогою спеціальних програм і баз даних пошукових систем, фахівець розширює поле можливих запитів шляхом підбору споріднених слів і фраз. Потім він вручну має відсіяти те, що не стосується тематики сайту;
- список доповнюється термінами та словами, які використовуються, як в українській, так і в іноземних мовах;

- визначення можливих жаргонних висловів чи слів з помилками;
- аналіз частоти всіх запитів. Отримані результати для зручності обробки необхідно експортувати в таблицю MS Excel. Далі визначається пріоритет усіх складених запитів на основі баз даних запитів пошукових систем. Все рідкісне, невизначене і те, що не стосується вашого сайту видаляється зі списку;

Систематизація готового семантичного ядра. Отриманий список ключових запитів необхідно розподілити за темами. Це дозволить легко визначитися зі структурою сайту, його просуванням і змістом нових текстів. В іншому випадку, буде втрачено багато часу, коли кожного разу будуть перебиратися кілька-сот ключових слів. Наприклад, якщо складається ядро для Інтернет-магазину, запити розбиваються на загальні, запити за категоріями і запити за виробниками [19].

Крім того, запити, що входять у ядро, можна поділити за якістю. Це також дозволить заощадити час і сили, приділяти більше уваги важливим запитам, але й не забувати про другорядні.

Запити легше розділити на три категорії:

- високочастотні запити (трапляються найчастіше і складаються із загальних фраз або слів без конкретики, наприклад, «ремонт», «будівництво», «плита» і т. д.);
- середньочастотні запити (трапляються трохи рідше й містять уточнюючі слова, наприклад, «зняти квартиру в Києві», «просування сайтів у Львові» і т. д.);
- низькочастотні запити (трапляються набагато рідше, але вони найбільш цінні, тому що виходять від потенційних покупців. Зазвичай вони містять, як уточнюючі слова, так і назви марок і брендів. Наприклад, «орендувати квартиру в центрі Львова недорого», «купити айфон у Києві легально» і т.д.).

Контент сайту – це пряма калька з англійської, де слово content означає “зміст, вміст”. Це збірний термін для будь-якої інформації, яка міститься в

інформаційному ресурсі. Якщо мова йде про веб-ресурси, то чим якіснішим він буде, тим краще для просування сайту. Контент може бути постійним, а може регулярно поповнюватися[20].

Унікальним може називатися будь-який контент, який ще жодного разу не був опублікований в мережі. Неунікальний – це передрук, запозичення, що є порушенням авторських прав. Закон про інтелектуальну власність поширюється і в Інтернеті.

Неунікальний – часто навіть не індексується пошуковими машинами.

Унікальний контент також повинен володіти цінністю, а не бути набором незрозумілих слів. Це важливо не тільки з точки зору відвідувачів сайту, але і з точки зору пошукових систем, які розрізняють унікальний контент і навіть виділяють спам. Якщо до створення унікального контенту застосувати основи SEO-оптимізації і використовувати семантичне ядро, то цінність його з точки зору пошукових систем підвищиться.

Контент як засіб просування Таким чином, унікальний контент може розглядатися як один із найдієвіших засобів просування сайту. Унікальні фото і малюнки, унікальні картинки, унікальні тексти якісного виконання залучають на сайт реальних споживачів і проіндексуються пошуковими машинами. Сайт, який надає тільки унікальний контент, володіє певним трастом. Це поняття не має числового виразу, але воно визначає певну ступінь довіри з боку і користувачів, і пошукових машин. Цей фактор інколи має переваги навіть перед SEO-оптимізацією: сайт з високим трастом ранжується навіть краще, ніж грамотно оптимізований контент.

Контент є продуктом інтелектуальної діяльності і, відповідно, товаром. Власник сайту не зобов'язаний створювати контент, він може його купити на біржах копірайтингу, або замовити статті копірайтерам. Існує також поняття: користувальницький контент. Це та частина сайту, яка створюється відвідувачами сайту: коментарі, вірші, оповідання, рецензії, розміщені користувачами.

1.4. Технічна оптимізація сайту

Першопочатковою задачею перед SEOналаштуванням є технічна оптимізація сайту, що повинна складатися з наступних етапів:

- Встановити Google Analytics і Яндекс Метрика.

Установка лічильників статистики дає всю необхідну аналітику по сайту, допомагає відслідковувати трафік з пошукових систем, поведінкові фактори.

- Налаштувати цілі в Google Analytics і Яндекс Метриці.

Важливо мати розуміння про мету та цілі сайту: дзвінок, заявка, зворотний дзвінок, та вчасно вносити правки. Ефективність розкрутки сайту та саме показники ROI, не зробивши даний пункт, прораховувати не вийде.

- Додати сайт в панелі вебмастера Google і Яндекс.

Додавши сайт в Google Webmaster Tools і Яндекс Вебмайстер, демонструється повна картина по: індексації сторінок, помилки, дублях сторінок, битих посиланнях.

- Зв'язати Google Webmaster і Google Analytics.

Допомагає отримувати більш розширену інформацію про трафік і аналітиці сайту.

- Створити файл robots.txt і завантажити його в кореневий каталог сайту.

Robots.txt допомагає пошуковим роботам визначити розташування XML sitemap (XML-карти сайту). В даному файлі показується інформація, що потрібно індексувати, а що ні, заклавши доступ до певних сторінок або розділів сайту.

- Створити XML-карту сайту і помістити її в кореневий каталог.

Створена XML-карта поліпшить індексування всіх сторінок сайту пошуковими системами.

- Розмістити карту в вебмастер Google, Yandex.

Призначити основне дзеркало (з www, або без нього), налаштувати 301 редирект. Пошукові системи визначають url(урли) сайту з www і сайт як два різних. Якщо сайт давно існує в індексі пошукових систем, потрібно дивитися,

скільки разів сайт проіндексовано більше: з www або без, і вибрати більш популярний варіант.

- Оптимізувати url (урл) за принципом ЧПУ.

Прописати ключові запити. Слова, що частково дублюють H1, зробити максимально короткими. Між фразами потрібно робити поділ через дефіс (-). Не застосовувати препарат нижнє підкреслення (). Робити в структурі 1-2 підрівні, три вже багато в URL.

- Оптимізувати ЧПУ.

Це поліпшить індексацію сайту пошуковими системами, поліпшить поведінкові фактори, так як будуть зручніше користувачеві.

- Оптимізувати Canonical URL, дублі сторінок.

Зробивши потрібну адресу за допомогою атрибута rel = "canonical" для всіх версій сторінки сайту, визначити більш релевантну сторінку для пошукової системи, якщо можливе розміщення однієї і тієї ж інформації по декільком URL, що буде дубльованим контентом.

- Позбутися битих посилань.

Неробочі посилання, що ведуть на 404 сторінку видалити. Наявність таких погіршує ПЧ (Поведінкові чинники), роботи розцінюють сайт як неякісний і гірше його ранжують.

- Прискорити завантаження сторінки сайту.
- Прибрати посилання на зовнішні сайти.
- Перелінковка на сайті.
- Використовуємо хлібні крихти.

«Хлібні крихти», іншими словами навігаційна ланцюжок, важлива як для зручності переміщення клієнтом по сайту, так і для пошукових робіт.

РОЗДІЛ 2

ОРГАНІЗАЦІЯ РОЗРОБКИ МОДЕЛІ SEO ОПТИМІЗАЦІЇ

2.1. Алгоритми пошукових систем

Багато уваги в даний час відводиться також такій смисловий характеристиці, як пертинентність. Ця характеристика інформаційно-пошукових систем означає відповідність отриманих в результаті пошуку документів інформаційним потребам користувача, а не формальному відповідності документа запиту. Для обчислення показників якості пошуку прийнято розглядати таблицю, яку заповнюють за результатами пошуку в навчальній колекції документів. Цей підхід був запропонований в рамках створеної Американським Інститутом Стандартів (NIST) конференції з оцінювання систем текстового пошуку Text REtrieval Conference (TREC, <http://trec.nist.gov>). Для детального підрахунку показників використовують ряд формул, які вміщують в собі вхідні данні та змінні. Всі данні динамічні, так як для кожного процесу в різний проміжок часу надається різний результат. Таблиця результатів пошуку має наступний вигляд:

Таблиця 2.1

Результати пошуку сторінок

Документи	Видані	Не видані
Релевантні	a	c
Нерелевантні	b	d

Відповідно до таблиці розраховуються основні показники ефективності інформаційно-пошукових систем [27].

Коефіцієнт повноти (recall) – частина виданих релевантних документів серед всіх релевантних документів:

$$r=a/(a+c) \quad (2.1)$$

Коефіцієнт точності (precision) – частина виданих релевантних документів серед всіх виданих:

$$p=a/(a+b) \quad (2.2)$$

Коефіцієнт шуму – частина виданих нерелевантних документів серед всіх виданих:

$$l=b/(a+b) \quad (2.3)$$

Коефіцієнт осаду – частина виданих не релевантних документів до всіх нерелевантних:

$$q=b/(b+d) \quad (2.4)$$

Коефіцієнт специфічності – частина не виданих нерелевантних серед всіх нерелевантних:

$$k=d/(b+d) \quad (2.5)$$

Також з наведеної таблиці розраховується коефіцієнт акуратності та помилка:

$$acc=(a+d)/(a+b+c+d) \quad (2.6)$$

$$err=(b+c)/(a+b+c+d) \quad (2.7)$$

На практиці значення точності і повноти набагато зручніше розраховувати з використанням матриці неточностей (confusion matrix). У разі якщо кількість класів відносно невелике (не більше 100-150 класів), цей підхід дозволяє досить наочно представити результати роботи класифікатора.

Матриця неточностей – це матриця розміру N на N , де N – це кількість класів. Стовпці цієї матриці резервуються за експертними рішеннями, а рядки за рішеннями класифікатора. Коли ми класифікуємо документ з тестової вибірки ми збільшуємо число, що стоїть на перетині рядка класу, який повернув класифікатор і стовпчика класу до якого дійсно відноситься документ (рис. 2.1).

	0.91	0.96	0.94	0.75	1.00	0.83	0.85	0.97	1.00	0.86	1.00	0.79	1.00	0.75	1.00	1.00	0.96	0.90	0.81	0.89	0.94	0.98	0.86	0.89	0.94	0.92	0.96	
0.80	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		
0.95	1	94				3														1						1		
1.00	2		32																									
0.29	3			6		3	2		1						1	1				1		1	3				2	
1.00	4				2																							
0.50	5					5															1		2			1	1	
0.92	6	1					152			1								1	4	2	3					2		
0.97	7	1		1				256												1	2						2	
0.33	8								1											1							1	
0.97	9									69																		2
0.82	10					2					18											1	1					
0.87	11											34		4											1			
1.00	12												37															
0.57	13											9		12														
0.63	14														5													
0.50	15															2												
0.77	16					2	1										47		1	3	4			2		1		
0.87	17								1								1	69	1	2	5							
0.97	18					1	4			1										197	1							
0.78	19																	2	35	183	13			2		1		
0.97	20					10	3			1										4	702					6		
0.93	21																					56		2				
0.29	22																				1	1	1		6	2		
0.91	23																										1	
1.00	24																				1		3	6			115	
0.93	25																										16	
0.98	26																										1	196
																												78

Рис. 2.1. Матриця неточностей (26 класів, результуюча точність – 0.8, результуюча повнота – 0.91)

Як видно з прикладу, більшість документів класифікатор визначає вірно. Діагональні елементи матриці явно виражені. Проте в рамках деяких класів (3, 5, 8, 22) класифікатор показує низьку точність [8].

Маючи таку матрицю точність та повнота для кожного класу розраховується дуже просто. Точність дорівнює відношенню відповідного діагонального елемента матриці і суми всього рядка класу. Повнота – відношенню діагонального елемента матриці і суми всього стовпця класу. На практиці виявляється, що неважко побудувати систему, що володіє високою точністю пошуку при низькій повноті або високою повнотою при низькій точності.

З наведених параметрів оцінки якості пошуку найчастіше використовують два основних:

- повнота (recall)– частина знайдених релевантних документів у загальному числі релевантних документів масиву;
- точність (precision)– частина релевантного матеріалу у відповіді пошукової системи.

Для оцінки якості пошуку існує ще ряд критеріїв, наприклад, так званий 11-і точковий графік повноти/точності, що показує залежність точності від повноти при кроці повноти в 10%. Чим вище проходить даний графік, тим вища якість інформаційного пошуку, що може надати система.

При оцінці різних інформаційно-пошукових систем за допомогою 11-точкового графіка кращої вважається та система, в якій висока точність досягається при малій повноті, що свідчить про гарне ранжирування результатів пошуку. Крім того, кращою визнається та система, для якої площа під відповідною інтерполяційною кривою є найбільшою.

Інформаційні ресурси тільки веб-простору складають понад двадцять мільярдів документів, до яких можливий вільний доступ будь-якого користувача. Природно, для того, щоб знайти необхідну інформацію і цієї найбільшої розподіленої повнотекстової бази даних необхідно використовувати найпотужніші ІПС. Такі системи існують і конкурують один з одним. Мільйонам користувачів Інтернету відомі такі інформаційно-пошукові системи, як Google, Yahoo, AltaVista, AllTheWeb, MSN, Yandex, Rambler, які охоплюють мільярди веб-документів. В основу роботи всіх подібних систем покладені спеціальні алгоритми, які є модифікаціями основних підходів – моделей пошуку.

Модель інформаційного пошуку має три ключових аспекти.

1. Формат представлення документа. Під документом розуміється деякий об'єкт, що містить інформацію у зафіксованому вигляді. Документи можуть містити тексти на природній або формалізованій мові, зображення, звукову інформацію тощо.
2. Формат представлення запиту. Під запитом розуміється формалізований спосіб вираження інформаційних потреб користувача системи. Для цього використовується мова пошукових запитів, синтаксис якої змінюється в рамках різних систем.
3. Функція відповідності документа запиту. Ступінь відповідності запиту і знайденого документа (релевантність) – суб'єктивне поняття, оскільки результати пошуку, що задовольняють одного користувача, можуть не

задовольняти іншого.

В основу традиційних методів покладено три головні підходи, перший з яких базується на теорії множин (булева модель), другий – на векторній алгебрі (векторно-просторова модель), а третій – на теорії ймовірностей (імовірнісна модель). Ці підходи можуть застосовуватися на практиці і в канонічному вигляді, проте у них є спільний недолік, обумовлений припущенням, що зміст документа визначається безліччю слів і стійких словосполучень – термів (англ.– Terms), які входять в нього без урахування взаємозв'язків, і, більше того, вважаються незалежними. Таке припущення веде до втрати змістовних відтінків, проте воно дозволяє реалізувати пошук і групування документів за формальними ознаками.

Відомі такі основні недоліки традиційних моделей [9]:

- Булева модель – невисока ефективність пошуку, відсутність контекстних операторів, неможливість ранжирування результатів пошуку.
- Векторно-просторова модель пов'язана з розрахунком масивів високої розмірності і в канонічному вигляді малоприсадибна для обробки великих масивів даних.
- Імовірнісна модель характеризується низькою обчислювальною масштабованістю (тобто різким зниженням ефективності при зростанні обсягів даних), необхідністю постійного навчання системи.

Системи, побудована на «рафінованих» пошукових моделях, недостатньо оперативні і володіють слабо розвинутими пошуковими можливостями і засобами узагальнення даних. Крім представлених нижче, існують і інші моделі пошуку, наприклад, семантичні, в рамках яких робляться спроби організації смислового пошуку за рахунок аналізу граматики тексту, використання баз знань, тезаурусів, онтологій, які реалізують семантичні зв'язки між окремими словами та їх групами. Разом з тим, ефективність систем, що базуються на таких підходах поки, залишається невисокою [2].

Булева модель базується на теорії множин і математичній логіці. Популярність цієї моделі пов'язана насамперед із простотою її реалізації, яка дозволяє індексувати і виконувати пошук у великих документальних масивах.

В рамках булевої моделі документи і запити представляються у вигляді безлічі термів – ключових слів і стійких словосполучень.

У булевій моделі запит користувача являє собою логічне вираження, у якому терми зв'язуються логічними операторами кон'юнкції (AND), диз'юнкції (OR) і заперечення (NOT). Відомо, що будь-який логічний вираз можна представити диз'юнкцією деяких висловів, з'єднаних між собою операцією кон'юнкції.

Якщо формула виконана в деякому документі, то вважається, що документ відповідає запиту. Існує декілька підходів до формування архітектури пошукових систем, відповідних булевій моделі, які знайшли своє втілення в реальних інформаційно-пошукових системах. Однією з реалізацій такої моделі була колись популярна система STAIRS корпорації IBM.

База даних цієї системи, що вже стала класичною, складається з наступних основних таблиць:

- текстової, що містить текстову частину всіх документів;
- покажчиків текстів, яка включає покажчики на місцезнаходження документів в текстовій таблиці;
- словникової, що містить всі унікальні слова, що зустрічаються в документах, тобто ті слова, за якими може здійснюватися пошук;
- інверсної, що містить списки номерів документів і координати окремих слів у документах.

Пошук по слову в базі даних системи такої архітектури здійснюється відповідно до алгоритму:

1. Відбувається звернення до словникової таблиці, за якою визначається, чи входить слово до складу словника бази даних, і якщо входить, то визначається посилання в інверсній таблиці на ланцюжок появ цього слова в документах.
2. Відбувається звернення до інверсної таблиці, за якою визначаються номери документів, що містять дане слово, і координати всіх входжень слова в

текстах бази даних.

3. За номером документа відбувається звернення до запису таблиці покажчиків текстів. Кожен запис цього файлу відповідає одному документу в базі даних.
4. За номером документа відбувається пряме звернення до фрагмента текстової таблиці – документу, після чого слідує висновок знайденого документа.

Наведений алгоритм охоплює випадок, коли запит складається з одного слова. Якщо ж у запит входить не одне слово, а деяка їх комбінація, то в результаті виконання пошуку по кожному з цих слів запиту формується масив записів, які відповідають входженню цього слова в базу даних. Після закінчення формування масивів результатів пошуку відбувається виявлення релевантних документів шляхом виконання теоретико-множинних операцій над записами цих масивів відповідно до правил булевої логіки.

Така модель іноді використовується у внутрішніх корпоративних системах пошуку, базах даних. Основним недоліком класичної булевої моделі є крайня жорсткість й непридатність до ранжування результатів пошуку за рівнем їх відповідності пошуковим запитам. Якщо слово, що вказане в запиті, присутнє у документі, то він вважається знайденим, в іншому випадку – не знайденим. Не буде знайдений документ, в якому знаходяться лише синоніми слова, у випадку, коли саме слово в документі не зустрічається [7].

Розширив класичну теорію множин Л.А. Заде і запропонував теорію нечітких множин, за ідеєю якої функція приналежності елемента множини може приймати довільні значення в інтервалі $[0,1]$, а не тільки 0 або 1.

За визначенням, нечіткою множиною на універсальній множині (будь-якої природи) є сукупність пар ступеней приналежності елемента нечіткій множині.

Ступінь приналежності – це число з діапазону $[0,1]$. Чим вище ступінь приналежності, тим більшою мірою елемент відповідає властивостям нечіткої множини. Функцією приналежності називається функція, яка дозволяє обчислити

ступінь приналежності довільного елемента універсальної множини до нечіткій множині.

Велика кількість відомих інформаційно-пошукових систем базуються на векторно-просторовій моделі опису даних (Vector Space Model), яка була запропонована Г. Солтоном в 1975 р і вперше застосованої в системі SMART.

Дана модель є класичною алгебраїчною. В рамках цієї моделі документ описується вектором в евклідовому просторі, в якому кожному терму, що використовується в документі, ставиться у відповідність його вагове значення, яке визначається на основі статистичної інформації про його появу як в окремому документі, так і в усьому документальному масиві. Опис запиту, що відповідає необхідній користувачеві тематиці, також являє собою вектор в тому ж евклідовому просторі термів. Для оцінки близькості запиту і документа використовується скалярний добуток відповідних векторів запиту і документа.

В рамках цієї моделі кожному терму в документі відповідає деяка невід'ємна вага. Кожному запиту, який являє собою також безліч термів, не поєднаних між собою ніякими логічними операторами, також відповідає вектор вагових значень.

Таким чином, кожен документ і запит можуть бути представлені у вигляді n -мірного вектора, де n загальна кількість термів в словнику моделі. Відповідно до розглянутої моделі, близькість документа до запиту, які як і в попередніх моделях розглядаються як інформаційні вектори, оцінюється як їх скалярний добуток. При цьому вагу окремих термів можна обчислювати різними способами. Один з можливих найпростіших підходів – використовувати нормалізовану частоту входження в документ як його вагу в документі.

Нормалізація необхідна для того, щоб врівноважити ваги документів з різною кількістю слів.

Векторно-просторова модель представлення даних забезпечує системам, побудованим на її основі, такі можливості, як:

- обробку запитів без обмежень їх довжини;
- простоту реалізації режиму пошуку подібних документів (кожен документ

може розглядатися як запит);

- збереження результатів пошуку з можливістю виконання уточнюючого пошуку.

Разом з тим у векторно-просторового моделі не передбачено використання логічних операцій у запитах, що істотно обмежує її придатність.

2.2. Інформаційно-пошукові мови

Інформаційно-пошукові мови є основними компонентами інформаційно-пошукових систем, за допомогою яких, зокрема, реалізуються інтерфейси між користувачами і системами.

На відміну від реляційних СУБД, у систем повнотекстового пошуку не існує стандартизованої мови запитів. У кожній системі цього типу існує свій спосіб завдання критеріїв пошуку.

Дуже часто мови запитів ІПС наближені до SQL, проте кожній з пошукових систем притаманний ряд індивідуальних особливостей, пов'язаних з такими моментами, як [10]:

- інтерпретація операцій, які задають порядок розташування слів у тексті (операцій контекстної близькості);
- обчислення рівня релевантності знайдених документів запитах для представлення результатів пошуку;
- застосування нестандартних для реляційних СУБД функцій, наприклад, таких як знаходження документів за принципом подібності змісту, побудова дайджестів з фрагментів документів, сніпетів (від англ. Snippet – фрагмент, уривок), що включаються пошуковими системами в списки знайдених документів і т.п.

У різних повнотекстових інформаційно-пошукових системах застосовуються різні архітектурні рішення, що охоплюють структури даних, алгоритми їх обробки, методи організації пошуку. Разом з тим, у сучасних інформаційно-пошукових систем багато спільних властивостей, наприклад, всі з

них забезпечують пошук хоча б по одному слову, більшість подібних систем реалізують граматичний пошук як результат застосування лінгвістичного аналізу. Більшість із сучасних систем здатні реалізовувати контекстний пошук фрази, укладеної в лапки (Google, Alltheweb, AltaVista, Яндекс тощо), пошук з використанням булевих операторів AND, OR і NOT, а також можливістю вказівки дужок для групування термів і операторів. Функції контекстної близькості свого часу отримали найбільший розвиток в системі Lycos, де були реалізовані за допомогою чотирьох операторів: ADJ, NEAR, FAR і BEFORE.

У найпопулярнішій в світі системі Google використовується досить лаконічний набір операторів, основні з яких – це кон'юнкція (використовується за замовчуванням, система видає документи, що містять всі слова запиту), диз'юнкція (OR) і заперечення (–).

Окремо розглядається можливість пошуку за параметрами документів, яка найчастіше дозволяє обмежувати діапазон пошуку значеннями URL, дат, заголовків. У більшій частині систем вийти на можливість пошуку за параметрами можна з режиму розширеного пошуку.

В Google, наприклад, забезпечується пошук по сайту («site:»), визначення посилань на сайт («admissionsite:»), пошук за цінами, наприклад «DVD player \$150..250», пошук по країнам, датам, доменам тощо. У багатьох системах забезпечується пошук не тільки за даними у форматі HTML, а й у форматах PDF, RTF, DOC (MsWord), PS [11].

Останнім часом набули поширення адаптивні інтерфейси уточнення запитів, найчастіше реалізовані шляхом застосування методів кластерного аналізу до результатів первинного пошуку. З'явилося таке поняття, як метод «папок пошуку» (Custom Search Folders), що об'єднує безліч підходів, загальна властивість яких – спроба згрупувати результати пошуку і представити групи найбільш пов'язаних документів (кластери) в зручному для користувачів вигляді.

Слова і словосполучення в так званих «інформаційних портретах», застосовуваних, наприклад, у корпоративних інформаційно-аналітичних системах

Галактика Zoom і Info Stream, також дозволяють адаптивне уточнення первинних запитів.

2.3. Узагальнений алгоритм пошуку інформації в Інтернеті

Інтерес до питання про пошук інформації в мережі Інтернет не слабшає протягом усього часу існування мережі. Пошук може вестися як користувачем-любителем, так і професіоналом. При проведенні пошуку інформації, що задовольняє інформаційним потребам користувача, необхідно знати, від чого залежить успішний пошук, і які проблеми виникають при роботі з інформацією.

Узагальнений алгоритм пошуку інформації в ІПС мережі Інтернет можна представити наступною послідовністю дій:

1. Формування завдання на пошук інформації.
2. Вивчення предметної області завдання на пошук.
3. Підготовка пошукового образу та пошукового припису.
4. Вибір інформаційно-пошукової системи Інтернет.
5. Введення URL-адреси.
6. Введення пошукового припису.
7. Аналіз пошукового результату або уточнення пошукового припису якщо результат відсутній.
8. Завершення пошуку, якщо результат задовольняє інформаційну потребу, або повтор пошуку починаючи з п. 2.

На основі запропонованого найпростішого алгоритму пошуку інформації розроблено модель, що дає можливість перейти до автоматизованого смислового пошуку інформації в розподілених ІПС мережі Інтернет.

Узагальнений алгоритм реалізації моделі пошуку в розподілених ІПС мережі Інтернет виглядає наступним чином:

1. «Зразок документа» (смілова задача), що представляє собою шаблон пошуку, вводиться експертом вручну.
2. З документа виділяється тема запиту, і визначаються пошукові приписи.
3. Розширюється тема запиту за рахунок синонімії та асоціативних запитів.

4. Формується пошуковий образ запиту на основі частотного словника з розбивкою його на окремі пошукові приписи.
5. Проводиться первинний пошук посилань на релевантні документи в існуючих пошукових Інтернет-машинах, загальний результат поміщається в сховищі даних.
6. Здійснюється закачування знайдених документів в сховище даних.
7. Для кожного документа в сховищі даних формується пошуковий образ документа (ПОД).
8. Проводиться ранжування документів відповідно до заданої теми (п.2). Чим точніший зміст в тілі документа, тим він має більш високий рейтинг.
9. Проводиться реферування знайдених документів і здійснюється передача рефератів для ознайомлення і аналізу експерту відповідно до рейтингу [13].

Запропонований підхід до організації смислового пошуку інформації в розподілених інформаційних системах мережі Інтернет дозволяє якісно поліпшити результати пошукових запитів до пошукових машин мережі Інтернет, дозволяє автоматизувати процес обробки релевантної інформації з ранжируванням смислової (пертенентної) інформації відповідно до заданої теми, що дає можливість експертам відійти від ручного послідовного перегляду знайдених ресурсів.

2.4. Стандарти XML. Оптимізація пошуку

Однією з найважливіших цілей створення платформи XML є привнесення в середу WEB метаданих, що описують властивості підтримуваних в ній інформаційних ресурсів. Насамперед, метадані дозволяють описувати структури XML-документів та їх смислового змісту (семантики). Завдяки цьому забезпечуються можливості автоматичної перевірки правильності структури XML-документів і зниження рівня інформаційного шуму при пошуку інформаційних ресурсів в WEB за допомогою різних пошукових машин.

Явний опис семантики XML-документів необхідно також для різноманітних просунутих WEB-додатків. Зокрема, стає можливим створення принципово нових

додатків високого рівня, заснованих на інтеграції інформаційних технологій і забезпечуючих інтеграцію неоднорідних інформаційних ресурсів. Цей напрямок активно розвивається в багатьох наукових центрах різних країн і пов'язан зі створенням інформаційних систем нового класу, що функціонують в середовищі WEB і які називають електронними бібліотеками.

У стандартах платформи XML передбачено кілька засобів опису та подання метаданих. Як уже вказувалося, для визначення логічної структури XML-документів спеціальні синтаксичні конструкції передбачені в мові XML.

Представлені їх засобами метадані називаються визначенням типу документів (Document Type Definition, DTD). У DTD XML-документи даного типу описуються як ієрархічні структури, що складаються з елементів документів. Ці елементи можуть бути різних типів, описаних у DTD. Специфікація DTD може бути вбудована в XML-документ або зберігатися де-небудь в WEB. В останньому випадку в документі дається на нього посилання. Для більш витонченого опису структури XML-документів можуть використовуватися властивості стандарту XML Schema. У порівнянні з DTD, цей стандарт надає для опису XML-документів додаткові можливості, зокрема більш розвинену систему типів значень елементів і атрибутів елементів.

Семантика XML-документа може бути визначена явним чи не явним чином (за замовчуванням). Явна визначення може бути формалізовано в різному ступені. Один з формалізованих способів явного визначення семантики XML-документів забезпечується засобами складається з двох частин стандарту W3C – Resource Definition Framework (RDF). Таке визначення семантики XML-документів називається RDF-специфікацією.

Подальші роботи консорціуму з розвитку засобів представлення метаданих, що визначають семантику XML-документів, проводяться в рамках діяльності по створенню семантичного WEB.

Інформаційні ресурси в RDF - це ресурси WEB, ідентифіковані унікальним чином за допомогою їх URI. Вони можуть також являти собою колекції інших інформаційних ресурсів або літералів, які називають контейнерами.

Допускаються контейнери типу мультимножини, послідовності і альтернативи. Значення властивостей задаються літерально або можуть бути іншими ресурсами, які представляються, у свою чергу, їх властивостями. Таким чином, властивості можуть визначати і зв'язки між ресурсами.

Опис семантики властивостей в RDF називається RDF-схемою. По суті, RDF-схема повинна визначати онтологію предметної області. Онтології отримали в останні роки широке поширення у вирішенні проблем представлення знань та інженерії знань, семантичної інтеграції інформаційних ресурсів, інформаційного пошуку і т.д.

Під онтологією розуміється "специфікація концептуалізації предметної області". Така специфікація являє собою свого роду словник понять предметної області і сукупність явним чином виражених припущень щодо змісту цих понять. Рівень структурованості опису онтології може змінюватися в широкому діапазоні. У спрощених випадках онтологія представляється як ієрархія понять, пов'язаних відносинами деяких певних видів. Такі визначення онтологій використовуються в різних класифікаціях. Кілька великі можливості забезпечують схеми метаданих, наприклад, широко відоме дублінське ядро.

Розвинені визначення онтологій формалізуються засобами мов логіки першого порядку. Вони допускають можливості логічного висновку.

У специфікації стандарту RDF не регламентується спосіб завдання схеми для RDF-специфікації. Достатньо лише представити її як деякий ресурс WEB і використовувати URI цього ресурсу для посилання на неї в RDF-специфікації. У документації стандарту RDF розглядається, наприклад, варіант використання для цих цілей набору елементів метаданих Дублінського ядра.

Один з більш розвинених способів завдання схеми пропонується в проекті другої частини стандарту RDF, званої RDF Schema (RDFS). Цей спосіб заснований на об'єктній моделі, в якій використовуються концепції класів, властивостей і обмежень, асоційованих з класами і властивостями, підтримується ієрархічне відношення "клас-підклас".

РОЗДІЛ 3

РЕАЛІЗАЦІЯ МОДЕЛІ SEO ОПТИМІЗАЦІЇ БАГАТОСТОРІНКОВОГО ВЕБ САЙТУ

3.1. Модель SEO оптимізації веб-сайту

Провівши аналіз вимог пошукових систем та проаналізувавши методи оптимізації сайтів які займають лідируючі позиції в пошуку було створено наступну модель, що дозволяє пришвидшити ранжування сайту в пошукових системах та зменшити вірогідність блокування сайту через неякісний контент.

Індексування сайту:

- Створити акаунт в Google, Ukr.net.
- Додати сайт в інструменти Вебмайстра Google, Ukr.net
- Налаштувати файл Robots.txt.
- Налаштувати файл .htaccess.
- Створити HTML карту сайту.
- Створити XML карту сайту.

Аналіз відвідуваності:

- Встановити Яндекс Метрика і Google Analytics.
- Налаштувати цілі в Яндекс Метриці та Google Analytics.

Технічна оптимізація:

- Налаштувати ЧПУ.
- Налаштувати Canonical.
- Видалити биті посилання.
- Перевірити валідність коду.
- Не використовувати більше 100 посилань на сторінці.
- Видалити посилання, які ведуть самі на себе.

Ключові слова:

- Проаналізувати сайти конкурентів.

- Вибрати точні комерційні фрази.
- Скласти список інформаційних запитів.

Внутрішня оптимізація:

- Налаштувати перелінковку сторінок.
 - Скласти унікальні Title із закликом до дії для кожної сторінки зі змістом ключових слів.
 - Скласти унікальні Description, що містять ключові слова і описують контент станиці.
 - Використовувати H1 заголовки з використанням ключових слів.
 - Оптимізувати картинки, використавши зрозумілі назви та додавши опису в тег ALT.
 - Використати не більше 5 ключових запитів для комерційних сторінок.
 - Одна комерційна сторінка повинна відповідати одній групі запитів.
 - Використати не більше 7 ключових запитів для інформаційних сторінок.
 - Проаналізувати ТОП для розуміння якою має бути сторінка і текст на ній.
 - Використати в текстах слова синоніми.
 - Не використовувати прямі входження запитів в текст, якщо вони не живляються в мові.
 - Прибрати фрейми (текст з прокруткою в одній області) і прихований текст.
- Наповнення та контент:

- Зробити блог і періодично публікувати контент з посиланнями на внутрішні сторінки сайту.
- Охопити статтями всі інформаційні запити з ніші.
- Аналізувати ТОП для розуміння, який контент ранжируется краще.
- Щоб поліпшити індексацію і підвищити авторитетність, щодня публікувати 1 новину.
- Публікувати структурований контент.

Зовнішня оптимізація:

- Додати сайт в Яндекс Довідник і Google Place.

- Додати посилання в якісних тематичних каталогах.
- Розміщуючи посилання користуватися правилом: «Менше, але краще».
- Розміщувати посилання на сайтах однієї тематики та регіону.
- Купуючи посилання, віддавати перевагу посиланням, які можуть принести трафік.
- Не купувати багато дешевих посилань в короткий період.
- Не купувати посилання автоматично.
- Складати вручну тексти для посилань.

Авторитетність:

- Зробити на сайті Wiki сторінки і тезаурус.
- Всебічно залучати трафік на сайт.
- Встановити соціальні кнопки і розміщувати анонси статей в соцмережах.
- Домовляйтесь з авторитетними сайтами для публікації експертного контенту.

Монополія:

- Оптимізувати і просувати контент по 1-2 низькочастотних запитах.
- Монополізувати видачу своїм контентом.

Поведінкові фактори:

- Робити заголовок Title продаючим та цікавим.
- Слідкуватиме за текстом в сніплеті.
- Використовувати розширені сніппети.
- Використовувати посилання на сторінки з тезаурусом та вікі-сторінці.
- Надавати більше відповідей і подумати, що користувачеві знадобиться ще.
- Аналізувати трафік та ключові запити, зменшувати відмови дописуючи контент.
- Використовувати відео або презентації на сторінках.

3.2. Створення SEO ядра багатосторінкового сайту

Для того, щоб створити SEO ядро необхідно, вибравши потрібні ключові запити обрати ті, які будуть відповідати тематиці сторінок. Головною особливістю багатосторінкового сайту є те, що кожна сторінка просувається окремо. Але, кожан сторінка сайту та кожен запис на сайті доповнюють проект, та виступають як внутрішні посилання. Для того, щоб сторінка провести аналіз та підібрати правильні мета фрази необхідно створити наступні документа.

Першим документом є «Базові терміни.xlsx», що включає в себе наступні підрозділи (рис 3.1) (дод. Б):

	A	B	C	D	E	F	G	H	I	J	K	L
1	готовий набір лента											
2	готовий потолок											
3	готовий светодиодный комплект											
4	комплект светодиодной ленты											
5	комплекты подсветки											
6	лента комплект											
7	набір rgb											
8	набір підсвітки											
9	набір светодиодный											
10	светодиодные наборы											
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22												
23												
24												
25												
26												
27												
28												
29												
30												

Рис. 3.1. Базові терміни.xlsx

В даний документ необхідно вносити інформацію з сайту розбиту на блоки (підрозділи). До кожного підрозділу необхідно внести всі можливі ключові запити, що відповідають тематиці контенту на сторінках.

Наступним документом є «Семантичне ядро.xlsx». В даний документ потрібно занести наступні розділи (рис.3.2):

№	Фраза	Категорія	кількість фраз в групі	Частотність [WS]	Частотність " " [WS]
2	1 світлодіодна лента	01. Светодиодные ленты/светодиодная лента	194	257382	13988
3	1 світлодіодна лента купити	01. Светодиодные ленты/светодиодная лента	194	34299	3164
4	1 діодна лента	01. Светодиодные ленты/светодиодная лента	194	23629	948
5	1 led лента	01. Светодиодные ленты/светодиодная лента	194	11253	450
6	1 світлодіодна лента цена	01. Светодиодные ленты/светодиодная лента	194	4601	378
7	1 діодна лента купити	01. Светодиодные ленты/светодиодная лента	194	2171	229
8	1 купити світлодіодну ленту с пультом	01. Светодиодные ленты/светодиодная лента	194	1659	214
9	1 світлодіодні ленти для подсветки	01. Светодиодные ленты/светодиодная лента	194	3223	209
10	1 світлодіодна лента led	01. Светодиодные ленты/светодиодная лента	194	3916	123
11	1 комплект світлодіодної ленти	01. Светодиодные ленты/светодиодная лента	194	1592	121
12	1 где купити світлодіодную ленту	01. Светодиодные ленты/светодиодная лента	194	1082	99
13	1 led лента купити	01. Светодиодные ленты/светодиодная лента	194	923	90
14	1 заказать світлодіодную ленту	01. Светодиодные ленты/светодиодная лента	194	414	64
15	1 цена світлодіодної ленти за метр	01. Светодиодные ленты/светодиодная лента	194	695	57
16	1 стоимость світлодіодной ленти	01. Светодиодные ленты/светодиодная лента	194	714	51
17	1 світлодіодні ленти для дома	01. Светодиодные ленты/светодиодная лента	194	302	43
18	1 діодна лента цена	01. Светодиодные ленты/светодиодная лента	194	529	42
19	1 магазин світлодіодных лент	01. Светодиодные ленты/светодиодная лента	194	897	38
20	1 світлодіодна лента на потолок купити	01. Светодиодные ленты/светодиодная лента	194	154	35
21	1 света діодная лента	01. Светодиодные ленты/светодиодная лента	194	278	32
22	1 цветная світлодіодная лента купити	01. Светодиодные ленты/светодиодная лента	194	266	31
23	1 світлодіодная лента usb купити	01. Светодиодные ленты/светодиодная лента	194	197	29
24	1 набор світлодіодной ленти	01. Светодиодные ленты/светодиодная лента	194	376	27
25	1 круглая світлодіодная лента	01. Светодиодные ленты/светодиодная лента	194	204	26
26	1 світлодіодная лента интернет магазин	01. Светодиодные ленты/светодиодная лента	194	287	21
27	1 світлодіодная лента дешево	01. Светодиодные ленты/светодиодная лента	194	217	20
28	1 черная світлодіодная лента	01. Светодиодные ленты/светодиодная лента	194	407	20
29	1 купити комплект світлодіодной ленти	01. Светодиодные ленты/светодиодная лента	194	289	19
30	1 светодиодная лента цена за 1	01. Светодиодные ленты/светодиодная лента	194	320	19

Рис. 3.2. Семантичне ядро.xlsx. Розділ «Ключові фрази»

В розділі «Ключові фрази» розміщено всю інформацію про кожен з аналізованих запитів та їх відповідність блокам. Поруч аналізується загальна частотність згадувань цих запитів в ТОП сайтах с першої сторінок пошукових систем.

В розділі «Структура» розміщуються ті ж запити по категоріям, але з детальним аналізом кількості переходів по кожному з запитів та по кожній зі сторінок конкурентів. Це допоможе зрозуміти, яку фразу необхідно брати ключовою та ставити в H1, а які фрази варто прибрати з пошуку на сайті (рис. 3.3).

№	фраза	кількість	Частотність [WS]	Частотність " " [WS]	URL1	URL2	URL3	URL4	URL5	URL6	URL7	URL8	URL
4	Світлодіодні стрічки	91/1254	764670	38852									
6	світлодіодна стрічка	194	386746	20841	https://ww	https://ww	https://ww	https://ww	https://ww	https://ww	https://ww	https://ww	https://ww
7	світлодіодна стрічка		13988	https://clu	https://led	https://ww	https://ww	https://ww	https://ww	https://ru	https://ma	https://sv	https://sv
8	світлодіодна стрічка купити		34299	3164	https://len	https://ww	https://ww	https://ww	https://ww	https://led	https://ali	https://ali	https://ali
9	діодна стрічка		23629	948	https://len	https://ww	https://ww	https://ww	https://led	https://ww	https://ru	https://ww	https://ww
10	led стрічка		11253	450	https://ww	https://len	https://ru	https://ww	https://ww	https://ali	https://led	https://ww	https://ww
11	світлодіодна стрічка цена		4601	378	https://ww	https://ww	https://ww	https://ww	https://ww	https://ali	https://ww	https://led	https://led
12	діодний стрічка купити		2171	229	https://ww	https://len	https://ww	https://ww	https://ww	https://sv	https://ww	https://led	https://led
13	купити світлодіодну стрічку з пультом		1659	214	https://len	https://ww	https://ww	https://ww	https://ww	https://mc	https://ww	https://mc	https://mc
14	світлодіодні стрічки для підсвічування		3223	209	https://len	https://led	https://clu	https://rer	https://ww	https://hal	https://ww	https://ww	https://ww
15	світлодіодна стрічка led		3916	123	https://len	https://ww	https://ww	https://ww	https://ww	https://led	https://ali	https://led	https://led
16	комплект світлодіодним стрічки		1592	121	https://ww	https://len	https://ww	https://ww	https://ww	https://led	https://ww	https://mc	https://mc
17	де купити світлодіодну стрічку		1082	99	https://len	https://ww	https://ww	https://ww	https://ww	https://ww	https://led	https://ww	https://sh
18	led стрічка купити		923	90	https://len	https://ww	https://ww	https://ww	https://led	https://ww	https://led	https://ww	https://ww
19	замовити світлодіодну стрічку		414	64	https://len	https://ww	https://sv	https://ww	https://led	https://ali	https://ww	https://led	https://ww
20	цена світлодіодної стрічки за метр		695	57	https://ali	https://len	https://ww	https://ww	https://ww	https://led	https://ww	https://ww	https://ww
21	вартість світлодіодної стрічки		714	51	https://ww	https://len	https://ww	https://ww	https://ww	https://led	https://ww	https://led	https://led
22	світлодіодні стрічки для будинку		302	43	https://rer	https://len	https://clu	https://led	https://ww	https://ww	https://ww	https://ww	https://ww
23	діодний стрічка цена		529	42	https://len	https://ww	https://ww	https://ww	https://ww	https://ww	https://led	https://ww	https://sh
24	магазин світлодіодних стрічок		897	38	https://mc	https://led	https://ww	https://ww	https://led	https://sh	https://ww	https://ww	https://ww
25	світлодіодна стрічка на стелю купити		154	35	https://len	https://ww	https://led	https://ww	https://ww	https://ali	https://ww	https://ww	https://sh
26	світла діодний стрічка		278	32	https://ww	https://len	https://ww	https://ww	https://ww	https://led	https://ww	https://led	https://rer
27	кольорова світлодіодна стрічка купити		266	31	https://ww	https://ww	https://len	https://ali	https://ww	https://cle	https://led	https://ww	https://mc
28	світлодіодна стрічка usb купити		197	29	https://ww	https://ww	https://ali	https://ww	https://ww	https://mc	https://ww	https://yu	https://ww
29	набір світлодіодним стрічки		376	27	https://len	https://ww	https://ww	https://ww	https://ww	https://ww	https://cle	https://mc	https://mc

Рис. 3.3. Семантичне ядро.xlsx. Розділ «Структура»

Розділ «Топ тематики» допоможе знайти сайти конкуренти та сайти донори для подальшого просування (рис. 3.4).

№	Домен	Запитів в ТОП
1	https://aliexpress.ru	4107
2	https://leroymerlin.ru	2881
3	https://ledrus.org	2490
4	https://ledpremium.ru	2410
5	https://www.dns-shop.ru	2395
6	https://www.vseinstrumenti.ru	2221
7	https://www.avito.ru	1729
8	https://www.ozon.ru	1686
9	https://shopleds.ru	1620
10	https://moskva.tiu.ru	1454
11	https://www.youtube.com	1428
12	https://www.chipdip.ru	1258
13	https://zen.yandex.ru	1182
14	https://www.drive2.ru	1036
15	https://www.etm.ru	1021
16	https://svetodiod77.ru	945
17	https://market.yandex.ru	693
18	https://svetlovsem.ru	687
19	https://mo.tiu.ru	661
20	https://svetilnik.info	631
21	https://www.wildberries.ru	619
22	https://svetodiodinfo.ru	611
23	https://ledrock.ru	588
24	https://arlight.ru	584
25	https://ulight.ru	561
26	https://www.mdm-complect.ru	560
27	https://www.220-volt.ru	544
28	https://lampaexpert.ru	529
29	https://www.mir-svetodiodov.ru	486

Рис. 3.4. Семантичне ядро.xlsx. Розділ «ТОП тематики»

3.3. Оптимізовані XML файли для покращеного парсингу сторінок пошуковими роботами

За пошукову оптимізацію та за відображення сторінок та контенту на них в пошукових системах відповідають xml карти сайту. Створення грамотної структури xml документу та набору xml файлів гарантує сайту швидке ранжування при оновленнях та логічну подачу інформації для SEO просування.

Провівши аналіз сайтів донорів та сайтів, що займають ТОП позиції в пошуковій мережу Google розроблено наступний список xml документів:

- sitemap_index.xml (рис.3.5);
- post-sitemap.xml (рис.3.6);
- page-sitemap.xml (рис.3.7);
- forum-sitemap.xml (рис.3.8);
- topic-sitemap.xml (рис.3.9);
- reply-sitemap.xml (рис.3.10);
- album-sitemap.xml (рис.3.11);

- work-sitemap.xml (рис.3.12);
- people-sitemap.xml (рис.3.13);
- category-sitemap.xml (рис.3.14);
- post_tag-sitemap.xml (рис.3.15);
- topic-tag-sitemap.xml (рис.3.16);
- genre-sitemap.xml (рис.3.17);
- work_genre-sitemap.xml (рис.3.18);
- group-sitemap.xml (рис.3.19).

Зважаючи на те, зі збільшенням інформації на сайтах та збільшені сторінок в пошукових мережах просування по тематиці сторінок є недоцільним. Найоптимальнішим способом просування є просування по конкретним записам на контент, зображення, категорій, документів. Саме задля цього розроблено масив окремих xml документів, які включають в себе всі елементи сайту.

Результатом роботи є аналог класичним таблицям БД, з яких створюються багатосторінкові сайти. Це допоможе пришвидшити індексування окремих елементів сайту та збільшить цитування контенту сайту в Інтернеті.

Структура та опис кожного з документів (дод. А). Інформацію по кожному з xml файлів в зображеннях (3.5 – 3.20).

```

1 <?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="//test/main-sitemap.xml"?>
2 <sitemapindex xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
3   <sitemap>
4     <loc>https://test/post-sitemap1.xml</loc>
5     <lastmod>2021-07-10T17:38:30+02:00</lastmod>
6   </sitemap>
7   <sitemap>
8     <loc>https://test/post-sitemap2.xml</loc>
9     <lastmod>2021-07-09T21:18:27+02:00</lastmod>
10  </sitemap>
11  <sitemap>
12    <loc>https://test/post-sitemap3.xml</loc>
13    <lastmod>2021-07-07T18:34:21+02:00</lastmod>
14  </sitemap>
15  <sitemap>
16    <loc>https://test/page-sitemap.xml</loc>
17    <lastmod>2021-05-11T23:48:53+02:00</lastmod>
18  </sitemap>
19  <sitemap>
20    <loc>https://test/category-sitemap.xml</loc>
21    <lastmod>2021-07-10T17:38:30+02:00</lastmod>
22  </sitemap>
23  <sitemap>
24    <loc>https://test/post_tag-sitemap1.xml</loc>
25    <lastmod>2021-07-10T17:38:30+02:00</lastmod>
26  </sitemap>
27  <sitemap>
28    <loc>https://test/post_tag-sitemap2.xml</loc>
29    <lastmod>2021-07-09T21:18:27+02:00</lastmod>
30  </sitemap>
31  <sitemap>
32    <loc>https://test/post_tag-sitemap3.xml</loc>
33    <lastmod>2021-07-10T17:38:30+02:00</lastmod>
34  </sitemap>
35  <sitemap>
36    <loc>https://test/post_tag-sitemap4.xml</loc>

```

Рис. 3.5. Головна карта сайту (sitemap_index.xml)


```
category-sitemap.xml | forum-sitemap.xml | genre-sitemap.xml | group-sitemap.xml | page-sitemap.xml | people-sitemap.xml | post_tag-sitemap.xml | post-sitemap.xml | reply-sitemap.xml | sitemap_index.xml | topic-sitema...
1 <?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="//test/main-sitemap.xsl"?>
2 <urlset xmlns:xs="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitema...
3 <url>
4 <loc>https://test/maria-dyvnjuk-doveriaite-miru-esli-seichas-chto-to-ne-poluchaetsia-znachit-vpered-i-vas-zhdioi-luchshee</loc>
5 <lastmod>2021-07-10T17:38:30+02:00</lastmod>
6 </url>
7 <image:image>
8 <image:loc>http://test/uploads/2021/07/193737934_949766525815803_4726758052090830220_n.jpg</image:loc>
9 </image:image>
10 <image:image>
11 <image:loc>http://test/uploads/2021/07/img_8688.jpg</image:loc>
12 </image:image>
13 <image:image>
14 <image:loc>http://test/uploads/2021/07/190518417_943882369737552_326394246232341571_n.jpg</image:loc>
15 </image:image>
16 <image:image>
17 <image:loc>http://test/uploads/2021/07/177208163_925668898225566_2786148038127914064_n.jpg</image:loc>
18 </image:image>
19 <image:image>
20 <image:loc>http://test/uploads/2021/07/96091389_686156128843512_6970068808522268672_n.jpg</image:loc>
21 </image:image>
22 <image:image>
23 <image:loc>http://test/uploads/2021/07/189096214_943887949736994_6542559738056072327_n.jpg</image:loc>
24 </image:image>
25 <image:image>
26 <image:loc>http://test/uploads/2021/07/186525585_9397660080149781_7517184808687826304_n-1.jpg</image:loc>
27 </image:image>
28 <image:image>
29 <image:loc>http://test/uploads/2021/07/208400121_3959071450981824_8710431679230925546_n.jpg</image:loc>
30 </image:image>
31 </url>
32 <url>
33 <loc>https://test/mi-6a-vidremontuiut-za-spriianija-olega-voronka</loc>
34 <lastmod>2021-07-09T21:18:27+02:00</lastmod>
35 <image:image>
36 <image:loc>http://test/uploads/2021/07/img_1285-111_scaled.jpg</image:loc>
37 </image:image>
38 </url>
39 </urlset>
40
```

Рис. 3.6.Карта записів та текстів (post-sitemap.xml)

```
category-sitemap.xml | forum-sitemap.xml | genre-sitemap.xml | group-sitemap.xml | page-sitemap.xml | people-sitemap.xml | post_tag-sitemap.xml | post-sitemap.xml | reply-sitemap.xml | sitemap_index.xml | topic-sitema...
1 <?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="//test/main-sitemap.xsl"?>
2 <urlset xmlns:xs="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitema...
3 <url>
4 <loc>https://test/about-us</loc>
5 <lastmod>2021-05-11T23:48:53+02:00</lastmod>
6 </url>
7 <url>
8 <loc>https://test/privacy-policy</loc>
9 <lastmod>2020-12-20T14:07:57+03:00</lastmod>
10 </url>
11 </urlset>
12
```

Рис. 3.7.Карта сторінок сайту (page-sitemap.xml)

```
category-sitemap.xml | forum-sitemap.xml | genre-sitemap.xml | group-sitemap.xml | page-sitemap.xml | people-sitemap.xml | post_tag-sitemap.xml | post-sitemap.xml | reply-sitemap.xml | sitemap_index.xml | topic-sitemap.xml | topic-tag-sitemap.xml
1 <?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="/test/main-sitemap.xsl"?>
2 <urlset xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitem
3 <url>
4 <loc>https://test/data_1</loc>
5 <lastmod>2021-05-11T23:48:53+02:00</lastmod>
6 </url>
7 <url>
8 <loc>https://test/data_2</loc>
9 <lastmod>2020-12-20T14:07:57+03:00</lastmod>
10 </url>
11 </urlset>
12
```

eXtensible Markup Language file length: 716 lines: 12 Ln: 1 Col: 1 Pos: 1 Unix (LF) UTF-8 INS

Рис. 3.8.Карта записів форуму сайту (forum-sitemap.xml)

```
forum-sitemap.xml | genre-sitemap.xml | group-sitemap.xml | page-sitemap.xml | people-sitemap.xml | post_tag-sitemap.xml | post-sitemap.xml | reply-sitemap.xml | sitemap_index.xml | topic-sitemap.xml | topic-tag-sitemap.xml
1 <?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="/test/main-sitemap.xsl"?>
2 <urlset xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitem
3 <url>
4 <loc>https://test/data_1</loc>
5 <lastmod>2021-05-11T23:48:53+02:00</lastmod>
6 </url>
7 <url>
8 <loc>https://test/data_2</loc>
9 <lastmod>2020-12-20T14:07:57+03:00</lastmod>
10 </url>
11 </urlset>
12
```

eXtensible Markup Language file length: 716 lines: 12 Ln: 1 Col: 1 Pos: 1 Unix (LF) UTF-8 INS

Рис. 3.9.Картатематик сайту (topic-sitemap.xml)


```
forum-sitemap.xml | genre-sitemap.xml | group-sitemap.xml | page-sitemap.xml | people-sitemap.xml | post_tag-sitemap.xml | post-sitemap.xml | reply-sitemap.xml | sitemap_index.xml | topic-sitemap.xml | topic-tag-sitemap.xml
1 <?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="/test/main-sitemap.xsl"?>
2 <urlset xmlns:xsl="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitem
3 <url>
4 <loc>https://test/data_1</loc>
5 <lastmod>2021-05-11T23:48:53+02:00</lastmod>
6 </url>
7 <url>
8 <loc>https://test/data_2</loc>
9 <lastmod>2020-12-20T14:07:57+03:00</lastmod>
10 </url>
11 </urlset>
12
```

eXtensible Markup Language file length : 716 lines : 12 Ln: 1 Col: 1 Pos: 1 Unix (LF) UTF-8 INS

Рис. 3.10.Карта коментарів сайту (reply-sitemap.xml)

```
album-sitemap.xml | category-sitemap.xml | forum-sitemap.xml | genre-sitemap.xml | group-sitemap.xml | page-sitemap.xml | people-sitemap.xml | post_tag-sitemap.xml | post-sitemap.xml | reply-sitemap.xml | sitemap_index.xml
1 <?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="/test/main-sitemap.xsl"?>
2 <urlset xmlns:xsl="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitem
3 <url>
4 <loc>https://test/data_1</loc>
5 <lastmod>2021-05-11T23:48:53+02:00</lastmod>
6 </url>
7 <url>
8 <loc>https://test/data_2</loc>
9 <lastmod>2020-12-20T14:07:57+03:00</lastmod>
10 </url>
11 </urlset>
12
```

eXtensible Markup Language file length : 715 lines : 11 Ln: 11 Col: 10 Pos: 716 Unix (LF) UTF-8 INS

Рис. 3.11.Карта альбомів сайту (album-sitemap.xml)


```
1 <?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="//test/main-sitemap.xml"?>
2 <urlset xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitem
3 <url>
4 <loc>https://test/data_1</loc>
5 <lastmod>2021-05-11T23:48:53+02:00</lastmod>
6 </url>
7 <url>
8 <loc>https://test/data_2</loc>
9 <lastmod>2020-12-20T14:07:57+03:00</lastmod>
10 </url>
11 </urlset>
12
```

Extensible Markup Language file length: 716 lines: 12 Ln: 12 Col: 1 Pos: 717 Unix (LF) UTF-8 INS

Рис. 3.12.Карта документів сайту (work-sitemap.xml)

```
1 <?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="//test/main-sitemap.xml"?>
2 <urlset xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitem
3 <url>
4 <loc>https://test/data_1</loc>
5 <lastmod>2021-05-11T23:48:53+02:00</lastmod>
6 </url>
7 <url>
8 <loc>https://test/data_2</loc>
9 <lastmod>2020-12-20T14:07:57+03:00</lastmod>
10 </url>
11 </urlset>
12
```

Extensible Markup Language file length: 716 lines: 12 Ln: 1 Col: 1 Pos: 1 Unix (LF) UTF-8 INS

Рис. 3.13.Карта користувачів сайту (people-sitemap.xml)

```
album-sitemap.xml | category-sitemap.xml | forum-sitemap.xml | genre-sitemap.xml | group-sitemap.xml | page-sitemap.xml | people-sitemap.xml | post_tag-sitemap.xml | post-sitemap.xml | reply-sitemap.xml | sitemap.xml
1 <?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="//test/main-sitemap.xml"?>
2 <urlset xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitem
3 <url>
4 <loc>https://test/bez-rubriki</loc>
5 <lastmod>2021-05-26T14:55:10+02:00</lastmod>
6 </url>
7 <url>
8 <loc>https://test/news</loc>
9 <lastmod>2021-07-09T21:18:27+02:00</lastmod>
10 </url>
11 <url>
12 <loc>https://test/blog</loc>
13 <lastmod>2021-07-06T19:17:37+02:00</lastmod>
14 </url>
15 <url>
16 <loc>https://test/articles</loc>
17 <lastmod>2021-07-10T17:38:30+02:00</lastmod>
18 </url>
19 <url>
20 <loc>https://test/information</loc>
21 <lastmod>2021-06-04T17:41:46+02:00</lastmod>
22 </url>
23 </urlset>
24
```

Extensible Markup Language file length: 1 009 lines: 24 Ln: 1 Col: 1 Pos: 1 Unix (LF) UTF-8 INS

Рис. 3.14.Карта категорій сайту (category-sitemap.xml)

```
album-sitemap.xml | category-sitemap.xml | forum-sitemap.xml | genre-sitemap.xml | group-sitemap.xml | page-sitemap.xml | people-sitemap.xml | post_tag-sitemap.xml | post-sitemap.xml | reply-sitemap.xml | sitemap.xml
1 <?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="//test/main-sitemap.xml"?>
2 <urlset xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitem
3 <url>
4 <loc>https://test/tag/afghanistan</loc>
5 <lastmod>2021-02-15T19:27:01+02:00</lastmod>
6 </url>
7 <url>
8 <loc>https://test/tag/neschanski</loc>
9 <lastmod>2020-12-28T15:03:52+03:00</lastmod>
10 </url>
11 <url>
12 <loc>https://test/tag/kniga</loc>
13 <lastmod>2021-05-27T18:55:07+02:00</lastmod>
14 </url>
15 <url>
16 <loc>https://test/tag/ialinki</loc>
17 <lastmod>2020-12-27T19:56:20+03:00</lastmod>
18 </url>
19 <url>
20 <loc>https://test/tag/blagodinist</loc>
21 <lastmod>2020-12-27T19:56:20+03:00</lastmod>
22 </url>
23 <url>
24 <loc>https://test/tag/misiutin</loc>
25 <lastmod>2020-12-29T22:40:24+03:00</lastmod>
26 </url>
27 <url>
28 <loc>https://test/tag/gimnastika</loc>
29 <lastmod>2020-12-29T22:40:24+03:00</lastmod>
30 </url>
31 <url>
32 <loc>https://test/tag/olimpiiskii-chempion</loc>
33 <lastmod>2020-12-29T22:40:24+03:00</lastmod>
34 </url>
35 </urlset>
36
```

Extensible Markup Language file length: 21 198 lines: 795 Ln: 1 Col: 1 Pos: 1 Unix (LF) UTF-8 INS

Рис. 3.15.Карта відміток сайту (post_tag-sitemap.xml)


```
1 <?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="//test/main-sitemap.xsl"?>
2 <urlset xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitem
3 <url>
4 <loc>https://test/data_1/</loc>
5 <lastmod>2021-05-11T23:48:53+02:00</lastmod>
6 </url>
7 <url>
8 <loc>https://test/data_2/</loc>
9 <lastmod>2020-12-20T14:07:57+03:00</lastmod>
10 </url>
11 </urlset>
12
```

Extensible Markup Language file length: 716 lines: 12 Ln: 1 Col: 1 Pos: 1 Unix (LF) UTF-8 INS

Рис. 3.16.Карта глобальних тем сайту (topic-tag-sitemap.xml)

```
1 <?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="//test/main-sitemap.xsl"?>
2 <urlset xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitem
3 <url>
4 <loc>https://test/data_1/</loc>
5 <lastmod>2021-05-11T23:48:53+02:00</lastmod>
6 </url>
7 <url>
8 <loc>https://test/data_2/</loc>
9 <lastmod>2020-12-20T14:07:57+03:00</lastmod>
10 </url>
11 </urlset>
12
```

Extensible Markup Language file length: 716 lines: 12 Ln: 1 Col: 1 Pos: 1 Unix (LF) UTF-8 INS

Рис. 3.17.Карта рубрик альбомів сайту (genre-sitemap.xml)


```
<?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="/test/main-sitemap.xsl"?>
<urlset xmlns:xsl="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitem
<url>
  <loc>https://test/data_1</loc>
  <lastmod>2021-05-11T23:48:53+02:00</lastmod>
</url>
<url>
  <loc>https://test/data_2</loc>
  <lastmod>2020-12-20T14:07:57+03:00</lastmod>
</url>
</urlset>
```

eXtensible Markup Language file length: 716 lines: 12 Ln: 1 Col: 1 Pos: 1 Unix (LF) UTF-8 INS

Рис. 3.18.Мета-карта сайту (work_genre-sitemap.xml)

```
<?xml version="1.0" encoding="UTF-8"?><?xml-stylesheet type="text/xsl" href="/test/main-sitemap.xsl"?>
<urlset xmlns:xsl="http://www.w3.org/2001/XMLSchema-instance" xmlns:image="http://www.google.com/schemas/sitemap-image/1.1" xsi:schemaLocation="http://www.sitem
<url>
  <loc>https://test/data_1</loc>
  <lastmod>2021-05-11T23:48:53+02:00</lastmod>
</url>
<url>
  <loc>https://test/data_2</loc>
  <lastmod>2020-12-20T14:07:57+03:00</lastmod>
</url>
</urlset>
```

eXtensible Markup Language file length: 716 lines: 12 Ln: 1 Col: 1 Pos: 1 Unix (LF) UTF-8 INS

Рис. 3.19.Карта адміністраторів сайту (group-sitemap.xml)

РЕЗУЛЬТАТИ ТА ВИСНОВКИ

В роботі було проведено аналіз сучасних інформаційних технологій щодо SEO оптимізації багатосторінкового сайту. Створення моделі пошукової оптимізації дозволить економити ресурси та забезпечити зв'язки, які постійно є актуальними.

В першому розділі було проведено аналіз сучасних методів створення моделі пошукової оптимізації web-сайтів та визначено найактуальніші серед них. Провівши аналіз технологій було визначено метод, на основі якого створювалася SEO модель.

В другому розділі були досліджені основні моделі пошуку, що дало змогу порівняти структурний рівень кожної з моделей, їх пристосованість до оптимізації під даний тип інформації. Як висновок, було обрано Булеву модель, що дозволяю оптимізувати веб-сайти різної складності, з використанням мінімуму ресурсів.

Аналіз сучасних інформаційно-пошукових мов, дав змогу обрати оптимальну для використання мову, що не перевантажуватиме сервер, та дасть змогу зручного оптимізування та редагування інформації, яка міститься в БД. Стандарти XML дають змогу оптимізувати сайт без ризиків потрапляння в «чорний список» пошукової системи та втрати позицій в пошуку.

В третьому розділі здійснювалась розробка SEO моделі пошукової оптимізації web-сайту. Було розроблено алгоритм SEO моделі, що відповідає поставленій задачі та xml файли для покращення ранжування сайту.

Створена модель допомагає створити єдину структуру, що буде оптимізована під більшість пошукових систем. Перевагою використання такої моделі є збільшення продуктивності сайту, та можливість уникнути погіршення ранжування сайту шляхом погіршення контенту.

За результатами розробки отримано SEO модель багатосторінкового web сайту, яка буде стабільно працювати з будь-якою з існуючих пошукових систем.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Максимов Н. В. Інформаційні ресурси і пошукові системи / Н. В. Максимов, О. Л. Голіцина, Г. В. Тихомиров, П. Б. Храмцов, М: Київ, 2015. - 400 с.
2. Ашманов І.С., Просування сайту в пошукових системах / І.С. Ашманов, А.А. Іванов. -М .: Вільямс, 2017. - 304 с.
3. Байков В.Д. Інтернет: пошук інформації та просування сайтів /В.Д. Байков. - Дніпро, 2016. - 288 с.
4. Євдокимов Н.В. Розкрутка веб-сайту: практичне керівництво / Н.В. Євдокимов, І.В. Лебе-Дінської. - М .: Вільямс, 2015. - 288 с.
5. Зуєв М. Просування сайтів у пошукових системах / М. Б. Зуєв, П. А. Маурис. - К .: Біном, Лабораторія знань, 2017. - 304 с.
6. Крістіан Д. Пошукова оптимізація на ASP.NET / Д.Крістіан - М .: Вільямс, 2016. - 400 с.
7. Крохина О. І. Перша книга SEO-копірайтера. Як написати текст для пошукових машин і користувачів / О. І. Крохина, М.Н. Полосина. - М .: "Инфра-Інженерія", 2016. - 216 с.
8. Севостьянов І.О. Пошукова оптимізація. Практичне керівництво по просуванню сайту в Інтернеті / І.О. Севостьянов. - Вінниця, 2010. - 240 с.
9. Байков В. Д. Інтернет. Пошук інформації. Просування сайтів. - Одеса, 2017. 288 с.
10. Колісниченко Д. Н. Пошукові системи і просування сайтів в Інтернеті. М .: Діалектика, 2015. - 272 с.
11. Євдокимов М. В. Розкрутка Web-сайтів. Ефективна Інтернет-комерція. М .: Вільямс, 2017. - 160 с.
12. Євдокимов М., Лебединський І. Розкрутка веб-сайту. Практичний посібник. М .: Вільямс, 2019. - 288 с.
13. Євдокимов Н.В. Основи тематичної оптимізації. Ефективна інтернет-комерція і просування сайтів в інтернет. М: Вільямс, 2012. - 160 с.

14. Ашманов І.С., Іванов А.А. Оптимізація і просування сайтів в пошукових системах. Київ, 2018. - 400 с.
15. Яковлев А. А. Розкрутка і просування сайтів: основи, секрети, трюки. Харків, 2017. - 336 С.Ч
16. Яковлев А., Ткачов В. Розкрутка сайтів. Основи, секрети, трюки. Харків, 2017. 352 с.
17. Mladenic D. Text learning and related intelligent agents // IEEE Intelligent Systems. 14 (4): 44-54, 2020 року.
18. Халигов А. А. Методи підвищення швидкості ранжирування сайтів в пошуковій системі Google // реферат випускної роботи магістра Факультет обчислювальної техніки та інформатики ДонНТУ. 2019.
19. Електронний ресурс: Sergey Brin and Larry Page. The Anatomy of a Search Engine <http://www-db.stanford.edu/pub/papers/google.pdf>.
20. Сегаловіч І. В. Як працюють пошукові системи // Світ Internet. - 2018. - № 10.
21. Searching the Web. / A. Arasu, [and others] // ACM Trans. On Internet Technology. - 2016. - № 1 (1)
22. Арсірій Е.А., Ігнатенко О.А., А.А. Леус Розробка семантичного ядра сайту з динамічним контентом на основі асоціативних правил // 2018 Том 2, №1
23. Як скласти семантичне ядро сайту [Електронний-ресурс]. - Режим доступу: great-world.ua
24. Ефективна методика пошуку інформації в мережі Інтернет / А. Попов // CitForum. - Режим доступу: http://citforum.ru/pp/search_03.shtml.
25. Стандарти XML і електронні бібліотеки // Когаловскій М.Р. Аналітичний огляд 2020 року.
26. Mining the Link Structure of the Web
www.almaden.ibm.com/chakrabarti99mining.pdf.
27. Baeza-Yates R. and e. BerthierRibeiro-Neto. Modern Information Retrieval. Addison Wesley Longman Publishing Company, 2020.

Коди сторінок xml

Таблиця А.1

Головна карта сайту (sitemap_index.xml)

```

<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xsl"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>
    <lastmod>2021-05-11T23:48:53+02:00</lastmod>
  </url>
  <url>
    <loc>https://test/data_2</loc>
    <lastmod>2020-12-20T14:07:57+03:00</lastmod>
  </url>
</urlset>

```

Таблиця А.2

Карта записів та текстів (post-sitemap.xml)

```

<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xsl"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

```

```

xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>
    <lastmod>2021-05-11T23:48:53+02:00</lastmod>
  </url>
  <url>
    <loc>https://test/data_2</loc>
    <lastmod>2020-12-20T14:07:57+03:00</lastmod>
  </url>
</urlset>

```

Таблиця А.3

Карта сторінок сайту (page-sitemap.xml)

```

<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xsl"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>

```



```
<lastmod>2021-05-11T23:48:53+02:00</lastmod>
</url>
<url>
<loc>https://test/data_2</loc>
<lastmod>2020-12-20T14:07:57+03:00</lastmod>
</url>
</urlset>
```

Таблиця А.4

Карта записів форуму сайту (forum-sitemap.xml)

```
<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xsl"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>
    <lastmod>2021-05-11T23:48:53+02:00</lastmod>
  </url>
  <url>
    <loc>https://test/data_2</loc>
    <lastmod>2020-12-20T14:07:57+03:00</lastmod>
  </url>
</urlset>
```

Карта тематик сайту (topic-sitemap.xml)

```

<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xml"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>
    <lastmod>2021-05-11T23:48:53+02:00</lastmod>
  </url>
  <url>
    <loc>https://test/data_2</loc>
    <lastmod>2020-12-20T14:07:57+03:00</lastmod>
  </url>
</urlset>

```

Карта коментарів сайту (reply-sitemap.xml)

```

<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xml"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd

```

```
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>
    <lastmod>2021-05-11T23:48:53+02:00</lastmod>
  </url>
  <url>
    <loc>https://test/data_2</loc>
    <lastmod>2020-12-20T14:07:57+03:00</lastmod>
  </url>
</urlset>
```

Таблиця А.7

Карта альбомів сайту (album-sitemap.xml)

```
<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xsl"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>
    <lastmod>2021-05-11T23:48:53+02:00</lastmod>
  </url>
  <url>
```



```
<loc>https://test/data_2</loc>
<lastmod>2020-12-20T14:07:57+03:00</lastmod>
</url>
</urlset>
```

Таблиця А.8

Карта документів сайту (work-sitemap.xml)

```
<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xsl"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>
    <lastmod>2021-05-11T23:48:53+02:00</lastmod>
  </url>
  <url>
    <loc>https://test/data_2</loc>
    <lastmod>2020-12-20T14:07:57+03:00</lastmod>
  </url>
</urlset>
```

Таблиця А.9

Карта користувачів сайту (people-sitemap.xml)

```
<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
```

```

href="//test/main-sitemap.xml"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>
    <lastmod>2021-05-11T23:48:53+02:00</lastmod>
  </url>
  <url>
    <loc>https://test/data_2</loc>
    <lastmod>2020-12-20T14:07:57+03:00</lastmod>
  </url>
</urlset>

```

Таблиця А.10

Карта категорій сайту (category-sitemap.xml)

```

<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xml"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">

```

```

<url>
  <loc>https://test/data_1</loc>
  <lastmod>2021-05-11T23:48:53+02:00</lastmod>
</url>
<url>
  <loc>https://test/data_2</loc>
  <lastmod>2020-12-20T14:07:57+03:00</lastmod>
</url>
</urlset>

```

Таблиця А.11

Карта відміток сайту (post_tag-sitemap.xml)

```

<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xsl"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>
    <lastmod>2021-05-11T23:48:53+02:00</lastmod>
  </url>
  <url>
    <loc>https://test/data_2</loc>
    <lastmod>2020-12-20T14:07:57+03:00</lastmod>
  </url>

```



```
</urlset>
```

Таблиця А.12

Карта глобальних тем сайту (topic-tag-sitemap.xml)

```
<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xml"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>
    <lastmod>2021-05-11T23:48:53+02:00</lastmod>
  </url>
  <url>
    <loc>https://test/data_2</loc>
    <lastmod>2020-12-20T14:07:57+03:00</lastmod>
  </url>
</urlset>
```

Таблиця А.13

Карта рубрик альбомів сайту (genre-sitemap.xml)

```
<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xml"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
```

```

xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>
    <lastmod>2021-05-11T23:48:53+02:00</lastmod>
  </url>
  <url>
    <loc>https://test/data_2</loc>
    <lastmod>2020-12-20T14:07:57+03:00</lastmod>
  </url>
</urlset>

```

Таблица А.14

Мета-карта сайта (work_genre-sitemap.xml)

```

<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xsl"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>
    <lastmod>2021-05-11T23:48:53+02:00</lastmod>

```

```
</url>
<url>
  <loc>https://test/data_2</loc>
  <lastmod>2020-12-20T14:07:57+03:00</lastmod>
</url>
</urlset>
```

Таблиця 3.15

Карта адміністраторів сайту (group-sitemap.xml)

```
<?xmlversion="1.0" encoding="UTF-8"?><?xml-stylesheettype="text/xsl"
href="//test/main-sitemap.xml"?>
<urlsetxmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1"
xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
http://www.google.com/schemas/sitemap-image/1.1
http://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd"
xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://test/data_1</loc>
    <lastmod>2021-05-11T23:48:53+02:00</lastmod>
  </url>
  <url>
    <loc>https://test/data_2</loc>
    <lastmod>2020-12-20T14:07:57+03:00</lastmod>
  </url>
</urlset>
```


SEO проєкт

```

1 SQLite format 3.1
2 CREATE TABLE [tableKeyCollector_BGAuctionData] (
3   [ID] integer PRIMARY KEY AUTOINCREMENT NOT NULL,
4   [Key_ID] integer NOT NULL,
5   [AdText] varchar(128) NULL,
6   [AdUrl] varchar(128) NULL,
7   [AdCPC] double NULL
8 ), WITHOUT ROWID
9 CREATE TABLE [tableKeyCollector_YDAuctionDataKeyCollector_YDAuctionData] (
10  [ID] integer PRIMARY KEY AUTOINCREMENT NOT NULL,
11  [Key_ID] integer NOT NULL,
12  [AdTitle] varchar(128) NULL,
13  [AdUrl] varchar(128) NULL,
14  [AdSiteUrl] varchar(128) NULL,
15  [AdText] varchar(128) NULL
16 ), WITHOUT ROWID
17 CREATE TABLE [tableKeyCollector_GASeasonDataKeyCollector_GASeasonData] (
18  [ID] integer PRIMARY KEY AUTOINCREMENT NOT NULL,
19  [Key_ID] integer NOT NULL,
20  [FreqDate] datetime NULL,
21  [FreqValue] bigint NULL
22 ), WITHOUT ROWID
23 CREATE TABLE [tableKeyCollector_YWSeasonDataKeyCollector_YWSeasonData] (
24  [ID] integer PRIMARY KEY AUTOINCREMENT NOT NULL,
25  [Key_ID] integer NOT NULL,
26  [FreqDate] datetime NULL,
27  [FreqValue] bigint NULL
28 ), WITHOUT ROWID
29 CREATE TABLE [KeyCollector_KEIData] (
30  [ID] integer PRIMARY KEY AUTOINCREMENT NOT NULL,
31  [Key_ID] integer NOT NULL,
32  [PageURL] varchar(128) NULL,
33  [PageTitle] varchar(128) NULL,
34  [Snippet] varchar(128) NULL

```

Рис. Б.1. SQLiteformat 3.1 (google.com)

```

40 CREATE TABLE [tableKeyCollector_TaskQueueKeysKeyCollector_TaskQueueKeys] (
41  [ID] integer PRIMARY KEY AUTOINCREMENT NOT NULL,
42  [Key_ID] integer NOT NULL,
43  [Task_ID] integer NOT NULL,
44  [RetryCount] integer DEFAULT 0 NOT NULL,
45  [Order] integer NOT NULL,
46  [InUse] boolean DEFAULT 0 NOT NULL
47 ), WITHOUT ROWID
48 CREATE TABLE [tableKeyCollector_UserTabsKeyCollector_UserTabs] (
49  [ID] integer PRIMARY KEY AUTOINCREMENT NOT NULL,
50  [TabHeader] varchar(64) NOT NULL,
51  [Comment] varchar(256) NULL,
52  [DataGridCustomization] text NULL,
53  [TabOwner_ID] integer NOT NULL,
54  [Hidden] boolean NULL,
55  [OrderNumber] integer NULL,
56  [BgColor] varchar(10) NULL,
57  [FilterTemplateData] text NULL,
58  [TooltipText] varchar(512) NULL,
59  [OwnerId] integer NULL,
60  [IsExpanded] boolean DEFAULT 0 NOT NULL,
61  [IsRemoved] boolean DEFAULT 0 NOT NULL
62 ), WITHOUT ROWID
63 CREATE TABLE [tableKeyCollector_PreferencesKeyCollector_Preferences] (
64  [ID] integer PRIMARY KEY AUTOINCREMENT NOT NULL,
65  [ParamName] text NOT NULL,
66  [ParamValue] text NULL
67 ), WITHOUT ROWID
68 CREATE TABLE [tableKeyCollector_KeysForPackParsingKeyCollector_KeysForPackParsing] (
69  [ID] integer PRIMARY KEY AUTOINCREMENT NOT NULL,
70  [KeyText] varchar(256) NOT NULL,
71  [PackOwner_ID] integer NOT NULL,
72  [TargetTab_ID] integer NULL,
73  [AdditionalInfo] varchar(32) NULL,
74  [Order] integer NULL

```

Рис. Б.2. SQLiteformat 3 (yandex.ua)