

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Київський національний університет імені Тараса Шевченка
Інститут філології
Кафедра української мови та прикладної лінгвістики

**Розроблення веб-застосунку
для SEO-аналітики українськомовних текстів**

Кваліфікаційна робота магістра
за спеціальністю 035 «Філологія»,
спеціалізацією 035.10 «Прикладна
лінгвістика»,
галузі знань 03 «гуманітарні науки»
ОП "Прикладна лінгвістика
(редакторсько-перекладацька
та експертна діяльність)"
Діани БОНДАРЕНКО

науковий керівник:
к.філол.н.,доц. Оксана ЗУБАНЬ

рецензент: д.філол.н. Наталія ДАРЧУК

Київ-2021

Зміст

ВСТУП.....	3
РОЗДІЛ I. ПОШУКОВА ОПТИМІЗАЦІЯ САЙТУ: БАЗОВІ ТЕОРЕТИЧНІ ПОНЯТТЯ ТА ІНСТРУМЕНТИ SEO-АНАЛІТИКИ.....	6
1.1. Ключові слова - основа SEO-оптимізації.....	7
1.2. Особливості внутрішньої розмітки сайту.....	24
1.3. Українськомовні запити в SEO.....	32
Висновки до першого розділу.....	40
РОЗДІЛ II. МЕТОДИ І МОДЕЛІ АВТОМАТИЧНОГО ВИЛУЧЕННЯ КЛЮЧОВИХ СЛІВ.....	41
2.1 Статистичні моделі автоматичного вилучення ключових слів.....	42
2.2 Гібридні моделі автоматичного вилучення ключових слів.....	44
Висновки до другого розділу.....	48
РОЗДІЛ III. СТВОРЕННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ SEO-ЗАСТОСУНКУ.....	49
3.1 Визначення вимог для реалізації застосунку.....	50
3.2 Загальна характеристика технологій та програмних інструментів для розроблення застосунку.....	54
Висновки до третього розділу.....	57
ВИСНОВКИ.....	58
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	60
Додатки.....	67

ВСТУП

У наш час, все більшого значення набувають інформаційні технології, мобільний зв'язок, телебачення, провідну роль відіграють бездротова й провідна мережа Інтернет, яка є не тільки джерелом інформації, але й універсальним середовищем для спілкування, розваг та навчання. Кількість користувачів мережі Інтернет зростає з кожним днем. Інтернет надає можливість миттєвого пошуку потрібної інформації. Нині корисною і пізнавальною інформацією всесвітньої павутини користуються усі верстви населення: студенти, науковці, фахівці різних галузей. Через Інтернет можна зв'язатися з людиною майже вбудь-якій точці світу, відправляти і отримувати листи за хвилини, робити різні покупки. Інтернет у сучасному житті людини це не тільки розваги, отримання інформації та спілкування, а й реальне робоче місце. Уже нікого не дивують такі поняття як інтернет-підприємець, онлайн-бізнес, мережевий маркетинг, фриланс.

З Інтернетом пов'язана й поява SEO-індустрії, яка допомагає просуванню веб-сайтів у пошуковій видачі за допомогою комбінацій ключових слів та якісного контенту. Потреба в написанні SEO-оптимізованих текстів щорічно збільшується. Це зумовлено конкуренцією, яку створюють веб-сайти у боротьбі за нового покупця чи клієнта. Кожен із конкурентів хоче отримати право на першість, тому в наш час SEO-копірайтинг – це один із найбільш необхідних та затребуваних видів діяльності людини, який дійсно покращує позиції сайту, що в результаті призводить до появи нових покупців та клієнтів. Таким чином, SEO-оптимізовані тексти та SEO-копірайтинг можна і потрібно вважати головним помічником у мережі Інтернет, адже завдяки цьому виду діяльності, ми можемо бути впевнені, що користувач швидко знайде необхідну релевантну до запиту інформацію, послугу чи товар, а власник сайту – нового клієнта та прибуток.

Стрімкий розвиток новітніх технологій потребує автоматизації SEO-процесів, а також створення нових методик зв'язку між людьми та машинами. При майже необмеженому доступі до інформації через Інтернет необхідно

структурувати дані та дати можливість комп'ютерній системі робити найбільш точну відповідь на запити користувачів. На жаль, сьогодні бракує сервісів, які допомагають аналізувати українськомовні тексти для подальшої SEO-оптимізації. Тому розроблення веб-застосунку для SEO-аналітики українськомовних текстів є **актуальним завданням** сучасної української прикладної лінгвістики.

Об'єкт дослідження – SEO-структура інтернет-текстів.

Предмет дослідження – лінгвістичні та інформаційні характеристики SEO-оптимізації українськомовних текстів.

Мета магістерської роботи – розробити веб-застосунок для SEO-аналітики українськомовних текстів. Поставлена мета передбачає виконання таких завдань:

1. Опрацювати літературу за темою “SEO”;
2. Визначити базові теоретичні поняття та особливості автоматичного аналізу тексту у галузі SEO;
3. Ознайомитись із відомими програмними бібліотеками та технологіями, які будуть використані у побудові інформаційного лінгвістичного продукту;
4. Розробити програмне забезпечення системи автоматичного аналізу українськомовного тексту;
5. Створити веб-застосунок.

Матеріал дослідження – лінгвістичні модулі та програмні бібліотеки для мови Python, українськомовні тексти.

Практичне значення роботи полягає у тому, що розроблений веб-застосунок, представлений у вільному доступі за адресою <https://django-semantic.herokuapp.com>, послужить аналітичним сервісом як для досвідчених фахівців, так і для новачків у сфері SEO-копірайтингу.

Теоретичне значення дослідження полягає в тому, що досвід розроблення системи автоматичного аналізу українськомовного тексту можна

інтерпретувати як внесок у розбудову сучасної української комп'ютерної лінгвістики та SEO-копірайтингу.

Структура роботи: магістерська робота складається зі вступу, трьох розділів, висновків, списку використаних джерел і додатків.

РОЗДІЛ I

SEO, а точніше SEO (SearchEngineOptimization) - це оптимізація сайту під пошукові системи, або пошукова оптимізація. Це дії, спрямовані на підвищення позиції сайту в результатах видачі пошукових систем, що призводить до зростання відвідуваності ресурсу[36].

SEO забезпечує комплексну оцінку веб-сайту, його внутрішньої організації, змісту, структури, html-коду, внутрішніх та зовнішніх покликань тощо. Семантичне ядро повинне складати серцевину змісту веб-сайту, а його основною метою є залучення відвідувачів до ресурсу та підвищення його ефективності.

Внутрішня оптимізація пошукової системи - це результат, який спрямований на покращення самого веб-сайту та його рейтинг у результатах пошуку. Внутрішня оптимізація включає такі операції:

1. Формування семантичного ядра веб-сайту (список ключових слів, за якими планують просувати інтернет-ресурс);
2. Покращення внутрішньої структури (читабельні покликання);
3. виправлення технічних помилок (дублікати сторінок, розрив покликань, пришвидшення роботи веб-сайту);
4. Покращення дизайну та зручність використання сайту (так звана, юзабіліті);
5. Коректне відображення веб-сайту на різних пристроях та адаптація дизайну;
6. Оброблення контенту (унікальність змісту тексту, актуальність ключових слів, їх розподіл у тексті, оптимізація зображень);
7. Правильне внутрішнє повторне покликання (внутрішні покликання між сторінками сайту).

Зовнішня оптимізація - це операція, спрямована на отримання покликань на інтернет-ресурси. Їх можна отримати, розмістивши покликання

на сайт у різних базах даних, повідомленнях на форумах, у соціальних мережах тощо. Існують також біржі продажів покликань.

1.1. Ключові слова - основа SEO-оптимізації

Розглянемо інструменти аналізу продуктивності веб-сайтів, якими може користуватися будь-яка людина. За допомогою отриманих даних можна оптимально вибрати базу ключових слів на сайті, проаналізувати трафік та, за наявності, виявити проблему.

Спочатку потрібно перевірити, у якій позиції пошуку знаходиться веб-сайт. Якщо це нещодавно створений ресурс, знадобиться деякий час (до декількох місяців), щоб увійти йому в пошукову систему. Щоб пришвидшити цей процес, можна повідомити Google про його існування. Якщо сайт є в базі даних пошуку, він з'явиться в результатах пошуку. Зазвичай завдання просування веб-сайту полягає в тому, щоб веб-сайт увійшов у першу трійку запитів, або принаймні в першу десятку (іноді 11-та позиція вигідніша, ніж 10-а, оскільки ресурс знаходиться у верхній частині сторінки).

Видача у Google називається PR (PageRank), яка розраховується для всього веб-сайту, а не для однієї сторінки. Точний алгоритм, який використовується для обчислення PageRank, є загадкою, і він часто змінюється, оскільки деякі веб-спеціалісти намагаються "обдурити" його, щоб вказати на нерелевантні ключові слова, зловживати покликаннями чи іншими хитрощами, щоб отримати найвищий рейтинг у результатах пошуку.

Значення PR може становити від 1 до 10, при цьому добре розкручені сайти можуть мати коефіцієнт 4-5. Для одного проєкту майже неможливо досягти значення вище 6-ти, якщо, звісно, мова не йде про таких гігантів, як Microsoft, Adobe та Wikipedia. Ви можете знайти PR веб-сайту на сайті «Checkpagerank»[37] (вибравши більш детальний режим перегляду, можна

проаналізувати загальну кількість відвідувань веб-сайту, категорії відвідувачів та іншу корисну маркетингову інформацію).

Тому, перш ніж формувати семантичне ядро, SEO-спеціалісти оцінюють поточну ситуацію та планують способи просування веб-сайту.

Всі пошукові запити можна поділити на кілька категорій. (Див таблицю 1)

Таблиця 1. типи пошукових запитів користувачів

За змістом	За частотністю	За конкурентністю
<ul style="list-style-type: none"> ▪ інформаційні (як вибрати ноутбук?) ▪ навігаційні (сайт про роботу) ▪ транзакційні (купити комп'ютер, скачати музику) 	<ul style="list-style-type: none"> ▪ високочастотні (ноутбуки) ▪ середньочастотні (комп'ютери Asus) ▪ низькочастотні (планшет Samsung Galaxy Tab) 	<ul style="list-style-type: none"> ▪ висококонкурентні (запиту відповідає багато сайтів) ▪ середньоконкурентні (запиту відповідає середня кількість сайтів) ▪ низькоконкурентні (запиту відповідає мало сайтів)

Ключові слова (англ. Keywords) представляють собою слова (або їх набір), які описують тему сайту або текст, і використовуються пошуковими системами для забезпечення точної відповідності (релевантності) результатів пошуку. Термін «ключові» вжито не дарма. Він якнайкраще передає всю важливість слів. Іншими словами, саме за допомогою ключових слів ви повинні описати свій сайт, а пошукові машини Google, Yandex, Bing та ін. на підставі такого «опису» включають ваш сайт у пошукову видачу при обробці

запитів користувачів, які відповідають вашій тематиці. Для полегшення розуміння суті ключових слів, їх можна порівняти з хештегом, використовуваними в Інстаграм або Твіттері.

Уявіть, що ви робите послуги з виготовлення меблів для кухні під замовлення. Ваші виробничі потужності знаходяться в місті Києві, відповідно послуги обмежені певним регіоном. Ви створили дуже симпатичний і сучасний сайт і цілком очікувано сподіваєтеся, що він приведе до вас багато нових замовників. В ідеалі, до вас на сайт повинні заходити жителі Києва та області, які хочуть замовити меблі для кухні. Тисячі «чудо-майстрів», які цікавляться самостійним виготовленням меблів, відвідувачі із Закарпаття і безліч інших «нерелевантних» відвідувачів навряд чи принесуть вам користь. Але як же залучити на сайт саме свою цільову аудиторію? Складання ключових слів дозволяє чітко ідентифікувати ваш сайт в очах пошукових машин. Правильно підібрані «ключі», використані в потрібних місцях, є гарантією того, що пошуковик включить покликання на ваш сайт в усі релевантні запити.

SEO – це аббревіатура з англійської мови «SearchEngineOptimization», яка перекладається як «оптимізація пошукових систем». Авторитетне інтернет-джерело «Ahrefs» [41] пропонує наступні визначення терміну «SEO»:

- «SEO - це процес, який допомагає сканувати, індексувати і ранжувати контент глибоко в Google та інших пошукових системах.»
Баррі Шварц, генеральний директор компанії RustyBrick
- «SEO - це спосіб поліпшення “видимості” сайтів у пошукових системах. При правильному виконанні, з використанням передових практик, тестування і наполегливої праці, він зазвичай працює. Іноді це буває складно.»
Сайрус Шепард, засновник компанії ZyrpySEO.

- «SEO - це процес оптимізації інформації для збільшення її видимості в пошукових системах. Це не залежить від типу пошукової системи, типу контенту, або навіть сайту. Якщо хтось хоче знайти інформацію в інтернеті, а хтось інший хоче опинитися в результатах цього пошуку, то SEO - це допомога творцеві контенту зв'язатися з тим, хто шукає.»
Рея Драйсдейл

- «У найпростішій формі, SEO - це спосіб допомогти пошуковій системі знайти ваш сайт, зрозуміти його зміст і оцінити його популярність та релевантність. Цього можна досягти багатьма різними способами.» Карл Хенді

- «SEO - тобто пошукова оптимізація, це процес оптимізації сайту, щоб він ранжувався вище в пошукових системах. По суті, два головні чинники ранжування - це контент і покликання. Контент, який відповідає очікуванням користувачів і сприймається пошуковими системами як досить авторитетний.»

Росс Хадженс, засновник і генеральний директор компанії «SiegeMedia»

В умовах постійного розвитку ринкових відносин інформаційного суспільства та приросту числа компаній, на одне з перших місць виходить позиціонування організації в середовищі інтернет. Для того, щоб компанія лідирувала у видачі пошукових систем, компанія звертається за послугами до спеціально-навчених людей – SEO-спеціалістів.

Розглянемо SEO-аналітику та поглянемо як працюють ключові слова у пошуковій видачі веб-сторінок. А як же працює SEO-видача, якщо звичайний середньостатистичний користувач мережі інтернет бажає придбати, наприклад, чайник.

Користувач вводить бажаний запит у пошукову систему (для нашого експерименту використовуємо пошукову мережу Google–дивись додаток 1). Наприклад «купити чайник». У додатку 1 наведено приклад пошукової видачі за обраним запитом.

Іноді природні результати видачі за запитом (які мають сайти в результаті їх пошукової оптимізації) плутають з контекстною рекламою. Контекстна реклама буде поза межами нашого дослідження. У будь якому разі ми спочатку побачимо контекстну рекламу, а нижче - органічну видачу.

Давайте спробуємо змінити деталі нашого пошуку, уточнивши фірму чайника (дивитись додаток 2) та глянемо чи вплине це на пошук. Веб-сайти змінять свої позиції, у видачі з'явилися нові позиції. Уточнимо модель чайника (дивитись додаток 3), що призведе до повторної зміни позицій.

Розглянемо приклад з додатку 1, додатку 2 та додатку 3 – сайт «comfy.ua» та прослідкуємо за його позиціями у таблиці 2.

Таблиця 2. Рейтинг сайту

Запит	Позиція сайту comfy.ua
Купити чайник	8
Купити чайник Philips	3
Електрочайник PHILIPS HD9351/91	10

Як свідчать дані, наведені у таблиці 2, при різних пошукових запитах видача сайтів у пошуку значно змінюється. В одних запитах сайт знаходиться в ТОП-3, в інших – ледве у ТОП-10, хоча категорія товару одна і та ж. Звідси впливає простий висновок: завдання SEO просування – створити таку аналітику, яка допоможе підібрати найрелевантніший запит та підняти сайт у топ-видачі. Сукупність таких пошукових запитів ще називають семантичному ядром або семантикою.

Загалом, пошукові системи враховують певні параметри контенту, перш ніж розмістити сайт на певній позиції.

Насамперед враховується взагалі релевантність тематики. Якщо тематика підходить – аналіз на запам'ятованість. Адже дуже просто було б добитись лідерських позицій, добавивши на свій сайт суцільні ключові слова. Для цього існують складні алгоритми пошукових систем, які здатні визначати частоту ключових слів. Такий аналіз здатний розрізнити пошуковий спам, та зрозуміти, коли слово зустрічається занадто часто (занадто часто вважається близько 8-10% від усього контенту) або, навпаки, занадто рідко. Найкраще, якщо ключові слова чи фрази будуть зустрічатись у тексті розміром 1800 символів – 3-4 рази.

Аналізуючи цю інформацію, можна зробити порівняльну таблицю (дивитись таблицю 3).

Таблиця 3 Пошукові слова у тексті

Заспамленість	Занадто рідко	Оптимальна частота
8-10%	1-3%	4-7%

Отже, ми дійшли до висновку, що саме від успішної SEO-оптимізації, залежить ранжування сайту в пошуковій видачі. Тепер необхідно детальніше з'ясувати, що саме аналізує SEO-оптимізація та як спеціалісти визначають ключові слова.

Для формування списку ключових слів нині існує багато безкоштовних та платних сервісів. Найпопулярнішим в Україні сервісами для підбору ключовиків із контенту веб-сторінок є:

- Semrush [51]

- Keywords Everywhere [45]
- Keyword Tool [44]
- Keycollector https [44]
- Rushanalytics [50]
- Spywords [57]
- Serpstahttps [54]
- Just-magic [43]

Розглянемо, як працює один із цих веб-сайтів – Serpstat[51]. На обраному сайті спробуємо знайти основні ключові слова для сайту Київського національного університету імені Тарса Шевченка[65].

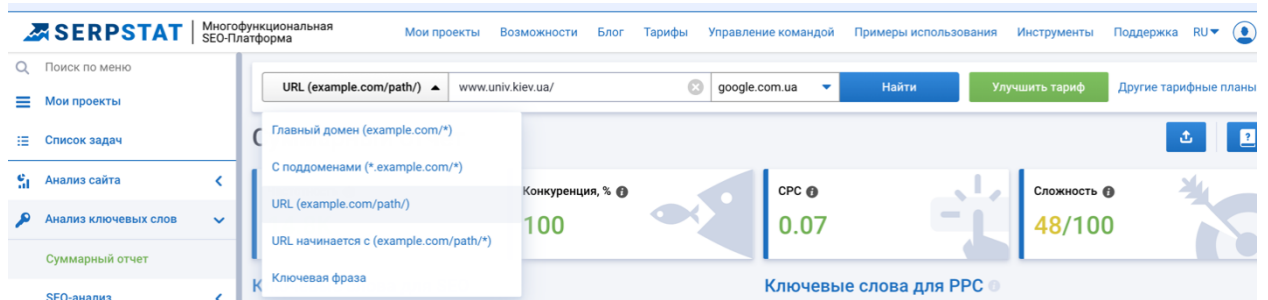


Рис 1. Аналіз сайту serpstata

Для цього необхідно ввести деякі параметри, зображенні на рисунку 1 (вказівка на те, що пошук ключових слів буде здійснюватися за покликанням) та вказати назву країни та пошукової системи, як показано на рисунку 2.

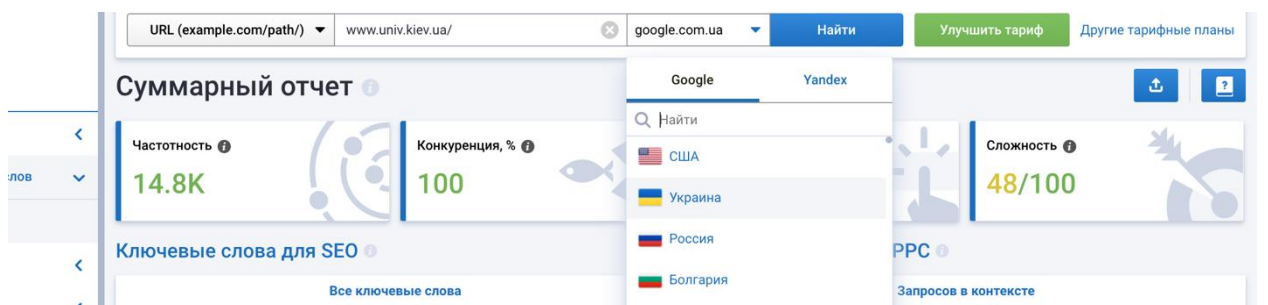


Рис 2. Вибір пошукової системи

У безкоштовній версії сервіс показав 10 ключових фраз (див рис 3). Перша фраза – назва навчального закладу, друга – аббревіатура, яку для пошуку, ймовірно, використовували студенти або працівники університету. Наступна фраза – запит російською мовою. Ключова фраза під номером 7 – пошук університетських вакансій.





















#	<input type="checkbox"/>	Ключевая фраза	Сложность ▼	Позиция ▼	Частотность ▼
1	<input type="checkbox"/>	університет ім шевченка   ▶	14	1	170
2	<input type="checkbox"/>	кду центр   ▶	N/A	1	1-5
3	<input type="checkbox"/>	сайт университета им т г шевче...   ▶	N/A	1	1-5
4	<input type="checkbox"/>	et univ   ▶	N/A	1	1-5
5	<input type="checkbox"/>	уніер  ▶	N/A	1	1-5
6	<input type="checkbox"/>	тарас шевченко    ▶	N/A	1	1-5
7	<input type="checkbox"/>	б ка ун та шевченко вакансии 20...   ▶	N/A	1	1-5
8	<input type="checkbox"/>	нпу ім шевченка   ▶	N/A	1	1-5
9	<input type="checkbox"/>	шевченка уні   ▶	N/A	1	1-5
10	<input type="checkbox"/>	інст шевченка   ▶	N/A	1	1-5

Рис 3. Ключові фрази

Сервіс запропонував ключові фрази, які відрізняються цільовою аудиторією та мовою пошуку, що свідчить про глибокий аналіз запитів.

Ми отримали 10 ключових фраз, проте, що саме seo-оптимізатори розуміють під поняттям «ключові фрази»? Адже з видачі видно, що це не слова, які узагальнюють зміст, а також не найбільш вживані слова серед контенту на веб-ресурсі.

Розробники цього сайту попереджають, що під ключовою фразою, вважається фраза, за якою домен ранжується в топ-100 Google.

Поруч із ключовою фразою відображаються іконки зі спеціальними елементами, які з'являються у видачі за цією фразою (наприклад, *карти, відео, зображення* ін.).

Також, на цьому сайті є параметр «частотність», де йдеться про частотність фрази в місяць згідно Google Ads (середня кількість запитів за ключовою фразою за останній рік).

Сервіс Serpstatзначає[59], що ресурс має такі інструменти як кластеризація ключових слів та текстова аналітика на основі штучного інтелекту. У розділі «Інструменти» Serpstat є цікавий розділ – текстова аналітика (рекомендації по оптимізації сторінки). Кластеризація - це розбивка фраз на групи, які називаються кластерами. У кожній групі повинні виявитися «схожі» фрази, а фрази з різних груп повинні якомога сильніше відрізнятися. Наприклад, такі фрази, як «кефір», «*iphone 6s*», «*iphone 7*», «йогурт», «*сир*» і «*macbook*» можна розбити на кластери: {*macbook*}, {*iphone 6s, iphone7*}, {*кефір, йогурт, сир*}.

Кластеризація потрібна для того, щоб:

- об'єднати фрази за змістом;
- провести аналіз семантичного ядра;
- створити групу фраз для розміщення на конкретній сторінці;
- розподілити фрази за сторінками й організувати повноцінну структуру сайту;
- знайти негрупові ключові фрази, які не належать ні до однієї з груп.

Основна проблема більшості інструментів кластеризації в тому, що вони формують всі групи на підставі центру кластера - високочастотної фрази, для якої підбираються фрази зі схожими результатами видачі.

У зв'язку з цим виникають труднощі:

1. Фрази виявляються в одній групі, але при цьому не мають семантичного зв'язку між собою.

2. Створюється безліч маленьких кластерів, які могли б бути об'єднані в більш великий кластер.

3. Потрібен час на визначення частотності. До того ж, ця процедура вимагає додаткових витрат та інструменти, які обчислюють точну частотність.

Навіщо веб-сайтам ключові слова та їх кластеризація? Ключові слова - це слова та фрази, за допомогою яких система визначає, за якими пошуковими запитами показувати веб-сайт чи рекламу. Типи відповідності ключових слів визначають, наскільки точно ключове слово має збігатися з пошуковим запитом користувача. Так, загальна відповідність дозволяє показувати веб-сайт або рекламу за більшою кількістю пошукових запитів, а точніше - зосередити зусилля на окремих запитах користувачів та просувати веб-сайт саме за конкретними запитами. Тобто, у сфері SEO - ключовими словами виступають зовсім не проаналізовані лінгвістами слова, а інтуїтивні запити користувача, тобто слова та фрази, які люди друкують у пошуковій системі. Саме їх прийнято називати як «пошукові запити» або «ключові слова SEO».

Дослідження ключових слів - це процес розуміння, аналіз «мови», яку використовують цільові відвідувачі сайту, коли шукають товар, послугу або контент. Отриманий результат потім включає в себе аналіз, порівняння та визначення пріоритетності можливостей використання ключових слів на веб-сайті.

Дослідження ключових слів - це єдиний спосіб з'ясувати, що люди друкують у пошукових системах. SEO-спеціалістам потрібно це знати, щоб не створювати контент про те, що ніхто не шукає. Багато власників веб-сайтів роблять цю помилку, і вона, швидше за все, є головною причиною, чому 90,63% сторінок не отримують трафіку з Google, як свідчить статистика на інтернет-ресурсі «Ahrefs»[58].

Інтернет-ресурс «Ahrefs»[58] надав дані про власне дослідження щодо аналізу трафіку веб-сторінок у видачі пошукової системи Google, згідно якому 90,63% містили контент, який не шукають користувачі (див рис. 4). Тобто, контент не містить необхідних ключових слів, за якими зазвичай і відбуваються запити.

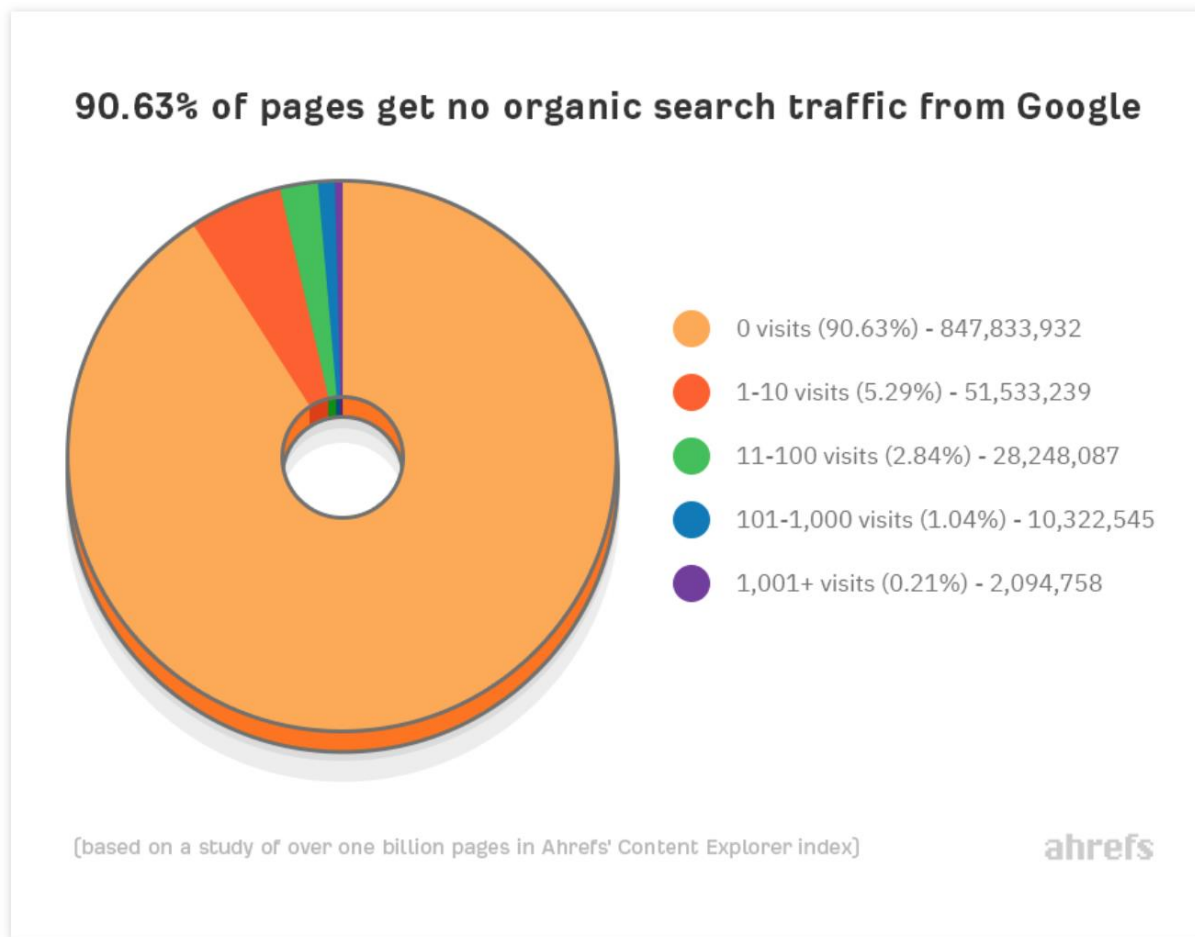


Рис.4 Видача пошукової системи Google

Серед усієї всесвітньої інтернет-павутини 90% веб сторінок ніколи ніхто не відвідував (див табл. 3), тому що їх не було видно в органічному пошуку google.

Таблиця 3. Статистика відвідуваності веб-ресурсів

0 відвідувань	90,6%
1-10 відвідувань	5,29%

11-100 відвідувань	2,84%
101 – 1000 відвідувань	1,04%
<1001 відвідувань	0,21%

Дослідження ключових слів також допомагає відповісти на такі питання:

1. Наскільки складно ранжуватиметься веб-сайт за цим ключовим словом?
2. Який контент потрібно створити, щоб ранжуватися за цим ключовим словом?
3. Яка ймовірність того, що люди, які вводять у пошук саме це ключове слова, перейдуть на потрібний нам веб-сайт?

Дослідження ключових слів починається з роздумів про те, як саме потенційні покупці могли б шукати ваш бізнес або веб-сайт. Після чого ви можете використовувати інструменти з дослідження ключових слів, щоб докладніше розкрити ці ідеї і знайти ще більше ключових слів.

Проаналізувавши декілька інтернет-ресурсів, в тому числі й ahrefs.com, можна помітити, що перший етап SEO-аналітики, це робота з ключами. Базові ключові слова є основою процесу дослідження. Кожен інструмент дослідження ключових слів просить базове ключове слово, яке потім використовується для генерації величезного списку ідей ключових слів. З базового ключового слова починаються розвиватись десятки інших словосполучень та фраз. Давайте розглянемо це на прикладі семантичного ядра львівського інтернет-магазину *FlugerPlast*[38]. В асортименті цього інтернет-магазину системи вентиляції, каналізації та водовідведення. Систематизуємо ці дані у таблицю 4.

Таблиця 4. Ключові запити інтернет-магазину «FlugerPlast»

Головний ключ	Другорядні ключі
Вентиляція	вентиляція львів
	вентиляції
	система вентиляції
	вентиляція приміщень
	види вентиляції
	кондиціонування
	вентилятори
	вентиляційні системи
	вентиляція повітря
	системи вентиляція
	системи вентиляції
	опалення вентиляція
	природна вентиляція
	розрахунок вентиляції
	вентиляція кондиціонування
	витяжна вентиляція
	види вентиляції
	вентиляція види вентиляції
	вентиляція виробничих приміщень
	вентиляція це
	дбн вентиляція кондиціонування
	загальнообмінна вентиляція
Опалення	опалення львів
	опалення будинку
	автономне опалення

	індивідуальне опалення
	системи опалення приватних будинків
	індивідуальне опалення ціна
	опалення приватного будинку
	індивідуальне опалення в квартирі
	системи опалення львів
	економне опалення
	система опалення будинку
Водопостачання	Водопостачання
	Водовідведення
	Водопостачання та водовідведення
	Система водопостачання
	Львів водопостачання

У таблиці 4 показано, що ключові запити користувачі можуть бути як одним словом, так і цілим реченням. Інколи користувач уточнює локацію товару, який його цікавить, що допомагає конкретизувати пошук: *«вентиляція львів»*, *«системи опалення львів»*.

Ключові слова по-різному вживають у текстах, багато чого залежить від цілі їх використання. SEO-експерти зазвичай ділять ключові слова на три типів, і вони використовуються по-різному в процесі пошуку. Але все залежить від тематики веб-ресурсу, оскільки серед популярних напрямів кількість запитів на високочастотні ключі може перевищувати 100 000, а в менш популярних - кількість запитів не перевищуватиме 1000. Наприклад, запити *«взуття»*, *«фото»*, *«фільм»* будуть однозначно популярнішими за запити *«синдром лікворної гіпертензії»* чи *«аллодоксофобія»*.

Ключі, які мають високу частотність, або «високочастотні ключі», це - ключові фрази з більш ніж із 3000 запитами на місяць. Працювати з такими ключами найтяжче, оскільки вони є найбільш конкурентними. Розглянемо ключ «взуття». Цей ключ можна вважати високочастотним.

Цей ключ є ядром таких ключів: «Чоловіче **взуття**», «Жіноче **взуття**», «Дитяче **взуття**».

Далі ці запити можна класифікувати, наприклад за сезонами: «Літнє взуття», «Осіннє взуття», «Весняне взуття», «Зимове взуття».

А ще можна додати: «Демісезонне взуття», «Пляжне взуття».

А також за конкретними функціями: «Танцювальне взуття», «Спортивне взуття», «Взуття для дому», «Бальне взуття», «Взуття для футболу», «Взуття для бігу».

Розглянемо запит «взуття для бігу», проаналізувавши видачу цього ключовика у пошуковій системі google (див рис. 5).

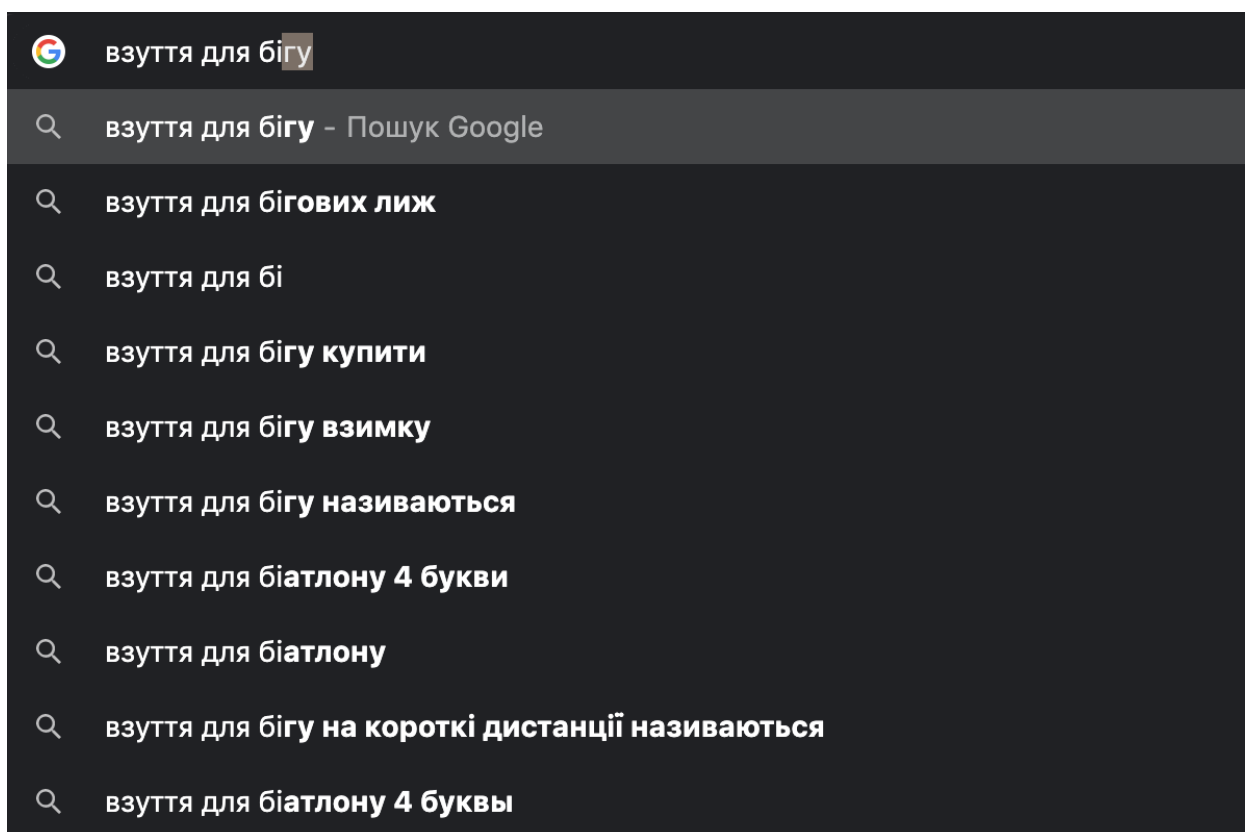


Рис. 5 Пошукова видача запиту «взуття для бігу»

Як свідчать дані у додатку 4, перші чотири позиції це інтернет-магазини спортивного взуття. Спробуємо задати менш конкретніший запит («спортивне взуття»), і побачимо, що видача сайтів змінилась (див додаток 5), не дивлячись на спільну спортивну тематику. Лідируючі сайти з попереднього запиту змінили свої позиції. При введенні основного ключа «взуття» (див додаток 6) – попередні веб-ресурси з першими позиціями пропали з видачі.

Ключ «взуття» об'єднував багато семантичних категорій, адже за цим запитом користувач може шукати будь-що, починаючи від визначення цього слова, закінчуючи біговим колекційним взуттям 1998 року випуску. Проте всі ці запити об'єднуюватиме одне ключове та високочастотне слово – «взуття», що в разі ускладнює роботу SEO-спеціаліста для вузьких галузей, пов'язаних із цим запитом. Адже яка ймовірність того, що користувач, уводячи високочастотний ключ – зайдет саме на ваш вузькопрофільний сайт? Фактично нульова, проте реальна.

Чим більше уточнень ми зазначаємо, шукаючи щось в інтернеті, тим більше знижується частота запиту. Так, середньочастотними вважаються запити, які вводили користувачі від одного до трьох тисяч разів на місяць. Саме такі ключі вважаються найлегшими у просуванні. Адже основний запит сформовано, уточнення добавлено. А саме уточнення і збільшує релевантну видачу, показуючи найоптимальніші запити.

Повернемося до прикладів, які ми аналізували вище. «Взуття» - високочастотний ключ. А от усі уточнення, які було перераховано у таблицях – середньочастотні ключі. Проте, шукаючи товар саме за таким запитом, користувач вийде на найбільш релевантний ресурс.

Розглянемо середньочастотний ключ *«літнє взуття»* детальніше. Ми можемо за допомогою цього ключа вивести користувача на націлений інтернет-магазин, проте це не гарантує що це буде цільовий клієнт. Адже і літнього взуття буває дуже багато видів: чоловіче, жіноче, дитяче, на підборах, закриті, відкриті. Тому, це найлегший спосіб привести трафік на сайт, проте це не гарантує цільового клієнта.

Низькочастотні ключова слова це найлегший спосіб залучити цільових клієнтів, адже у свій запит вони вкладають максимальну конкретику, як наприклад *«літнє жіноче взуття на підборах»*. Запит описано максимально точно, проте і тут є ряд мінусів, а саме – за такимивузькопрофільними ключами тяжко створювати контент, а трафік за такими ключами найменший.

Як вказує авторитетне інтернет-джерело «Outsourcingteam»[48], ключові слова можна поділити не лише за частотою, але ще й за додатковими параметрами, адже іноді, вказуючи свій запит, користувач бажає уточнити місце, тоді такий запит матиме назву –**геозапит**.

Геозапит допоможе користувачеві знаходити послугу чи товар у конкретному регіоні, місті, районі, наприклад, *«взуття Київ»*. Особливо такі ключі важливі для бізнесу, який націлено лише на локального клієнта. Наприклад, якщо йдеться мова про виготовлення об'ємних товарів (шафи-купе т. ін.) або про конкретну послугу (послуги манікюру, масаж, доставка їжі т. ін.).

Наступний вид запитів не пов'язаний із бізнесом чи наданням послуг, це – **інформаційні** ключові слова. Їх використовують для пошуку інформації, наприклад, *«біологія - це»*[48].

Загалом, якщо запит допомагає користувачеві виконати якусь дію, наприклад, *«замовити доставку», «купити книгу»*, тобто пов'язаний із комерцією, то такий запит називається **комерційний**[48].

Іноді користувача цікавить конкретний запити, як наприклад, «офіційний сайт київської міської ради» або «сайт borjomi », такі типи ключів називаються **навігаційні**, вони зорієнтовують на конкретний веб-ресурс[48].

Для пошуку фото та відео-контенту теж виділяють окремий вид ключів – **медійні**[48].

Виділяють ще один тип – **загальні** ключові слова, тобто ті, які не наділені уточнювальним змістом, наприклад, ключ «сайт», не уточнює про що саме йде мова. Можливо, користувача цікавить ціна за сайт, а можливо він шукає саме визначення.

Ключові слова використовуються для того, щоб користувач знайшов необхідний ресурс, а ресурс знайшов свого користувача. Для того, щоб така «угода» відбулась, необхідно правильно налаштувати сайт для використання його пошуковими машинами.

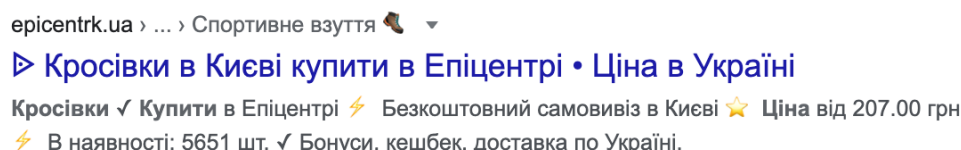
Потрібно пам'ятати, що цей процес не миттєвий, оскільки заявка на розміщення сайту повинна пройти ще певну модернізацію, потрапляючи у чергу на опрацювання. Для того, щоб пошукова система могла коректно реалізувати пошук, та внести ресурс до своєї бази даних, веб-сторінка повинна містити службову інформацію, а саме: опис ресурсу та ключові слова.

У підручнику «Основи веб-дизайну» авторів: Г. Пасічник, В. Пасічник, В. Стеценко зазначено: «для веб-сайту необхідно підібрати близько 20 ключових слів, які описали б тематику сайту, його функції, цільову аудиторію та призначення.» [5, 264]. Це можуть бути ключі «купівля одягу», «верхній одяг», «зимовий одяг», «куртка», «пальто», «спортивна куртка», «шапка» та ін.

1.2 Особливості внутрішньої розмітки сайту

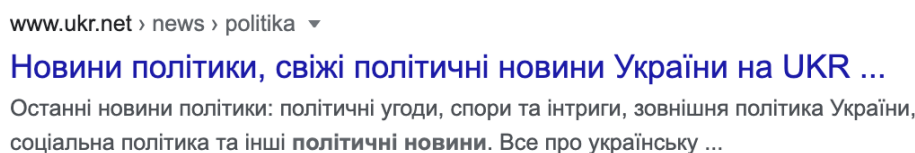
Опис сайту – це інформація, яка міститься одразу під назвою сайту. При формуванні сніпета використовується 3 варіанти:

1. За допомогою мета-тега description (як зображено на рис. 6)
2. ½ мета-тега description + релевантна фраза
3. Перші 10-20% статті (1-й абзац)(як зображено на рис. 7)



epicentrk.ua › ... › Спортивне взуття
▶ Кросівки в Києві купити в Епіцентрі • Ціна в Україні
Кросівки ✓ Купити в Епіцентрі ⚡ Безкоштовний самовивіз в Києві ★ Ціна від 207.00 грн
⚡ В наявності: 5651 шт. ✓ Бонуси, кешбек, доставка по Україні.

Рис. 6 Формування опису веб -сторінки за допомогою тега мета-тега description



www.ukr.net › news › політика
Новини політики, свіжі політичні новини України на UKR ...
Останні новини політики: політичні угоди, спори та інтриги, зовнішня політика України, соціальна політика та інші політичні новини. Все про українську ...

Рис. 7 Формування опису веб -сторінки за допомогою використання перших 10-20% статті (1-й абзац)

Якщо сайт створено декількома мовами, то й опис сайту теж необхідно зробити декількома мовами.

Зазвичай опис сайту не повинен перевищувати 250 символів із пробілами. Опис повинен виділити основну інформацію про сайт, проте не дублювати її з назви, як свідчить авторитетний інтернет-ресурс «Outsourcing.team»[47].

Релевантною вважається сторінка, яка має оптимальне співвідношення слів із введеного пошукового запиту до загальної кількості слів на сторінці. Чим більше відповідають результати пошуку поставленій меті – тим більш релевантною є сторінка. Тому, чим більше елементів на веб-сторінці містять

ключові слова, тим краще, проте не всі вони мають однакову вагу для пошукових систем.

Найважливішим вважають тег <title>, і саме в ньому повинна бути закладена ключова інформація про сторінку, саме цей тег має найбільшу вагу для пошукових роботів (як показано на рис.8, рис 9, рис 10).Контент, прописаний усередині тега, також буде відображатись не лише у програмному коді, його також бачитиме користувач у вигляді назви вкладки.

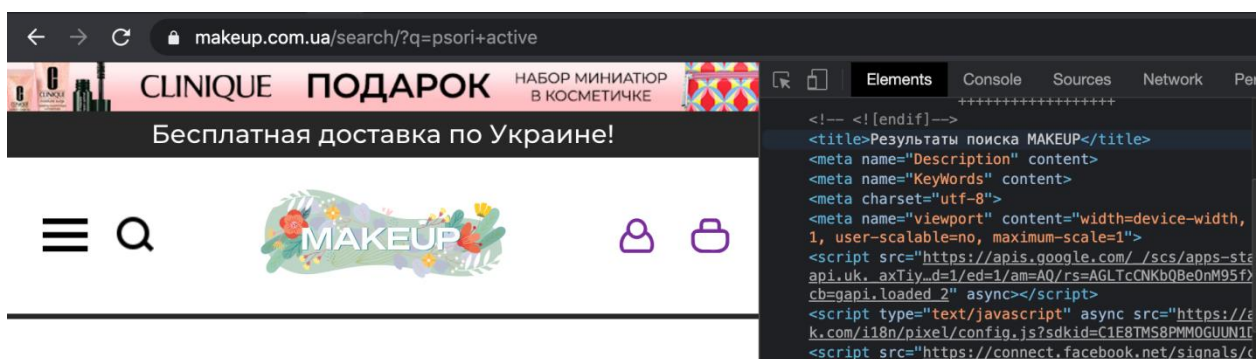


Рис.8 Аналіз інструментів розробника сайту makeup

```
<title>Результаты поиска MAKEUP</title>
```

Рис. 9 Інформація всередині тега

```
Результаты поиска MAKEUP  
makeup.com.ua
```

Рис. 10 Назва вкладки цієї веб-сторінки

Тег має обмежену кількість символів (зазвичай 60-65 символів), також сюди необхідно помістити найголовніший ключ цієї веб-сторінки. Аналізуючи веб-ресурси, можна звернути увагу, що зазвичай <title> має таку схему:

1. ключ на першому місці (див. рис. 11);
2. використання «вірусних слів» (таких як економна, дешево, знижки та ін.) (див. рис. 12);
3. використання назви компанії(див. рис. 13).

Вентиляція для приватного будинку - Клімат-Захід

5 груд. 2018 р. — Добре продумана і зроблена за всіма правилами вентиляція забезпечить ... Зараз більшість витяжок, навіть з дешевих моделей, ...

Рис. 11 Використання ключового слова на першому місці

Дешева вентиляція в приватному будинку. Облаштування ...

Дешева вентиляція в приватному будинку. Облаштування вентиляції в приватному будинку своїми руками: вибір схеми і складання проекту. Припливні ...

Рис. 12 Приклад використання використання «вірусних слів»

Приточная вентиляция дешёво — купить на Яндекс.Маркете

Приточная вентиляция дешёво — сравнить модели и купить в проверенном магазине. В наличии популярные новинки и лидеры продаж. Поиск по ...

Рис. 14 Приклад використання назви корманії при формуванні тега <title>

Далі пошукові роботи аналізуватимуть теги заголовків, тобто теги <h1>, <h2>, <h3>, <h4>, <h5>, <h6>. Саме тому смислові заголовки потрібно виділяти за допомогою перерахованих тегів, а не стилізувати звичайний текст під заголовки.

У цих тегах варто розміщувати найбільш релевантні ключі:

<h1> - головний ключ веб-сторінки ;

<h2> - другорядні ключові слова;

<h3> - використовується для назви списків, другорядних ключових слів.

Загалом, найоптимальнішою кількістю ключових слів вважається 7-8 ключовиків на 2500 символів з пробілами. Проаналізуємо ресурс «Galsad»[39]. У головному заголовку – головний ключ. Заголовок <h2> містить другорядний ключ. (Див. додаток 7)

Далі заголовок третього рівня (див. рис. 15), який є не лише другорядним ключовим виразом, але і списком.

Поради відносно загальних правил догляду за домашніми рослинами

Існують загальні правила ефективного догляду за рослинами, які утримуються вдома, в офісі або в інших приміщеннях. Дотримуючись таких простих порад, ви забезпечите для квітів, пальм, сукулентів, інших кімнатних рослин оптимальні умови здорового життя:

- воду для поливу попередньо відстоюють чи використовують талу воду;
- полив виконують теплою водою з температурою не менш, як 25C;
- добрива вносять згідно з рекомендаціями, наданими в інструкції до мінерального чи хімічного засобу;
- підгодовування не проводять, якщо рослину недавно придбали, переставили в інше місце, пересадили з заміною ґрунту. Потрібно почекати період «карантину» чи «адаптації», який складає приблизно 3-4 тижні;
- тонесеньке та ніжне листя обприскують, а товсте листя миють або протирають вологою серветкою з обох сторін;
- оптимальна температура для більшості кімнатних рослин складає 20-25C в теплий та 18-20C в холодний сезон;
- без потреби не переставляйте квіти постійно з місця на місце;

Ще важлива порада стосується врахування особливостей квітів, їх потреби в сонячному світлі. Так, багато сонця потребують рослини, що мають соковите листя та великі квіти – бегонії, лілії, амариліси, кактуси, герані, цикламени. Прямі сонячні промені не бажані для фікусів, аралії, сансевієрії, диффенбахії. Тінь гарно впливає на папороті, антуриуми, традесканції, калатеї, аспідстри, маранти, аспарагуси, азалії, англаонеми.

Якщо у вас виникли питання відносно того, які умови будуть найкращими для визначеного виду рослини, як підживити вазони, який догляд за вазонами потрібен в актуальному сезоні, то ви завжди можете отримати професійну консультацію від фахівців садового центру «GalSad». Телефонуйте або відправляйте свої питання через online форму, доступну на сайті.

Рис. 15 Приклад заголовка третього рівня

Для пошуку зображень зазвичай використовують опис, прописаний у параметрі `alt` та ``. Наприклад ``. Такий спосіб використовує більшість пошукових машин.

Взагалі, сам параметр спочатку був призначений для іншої цілі, а саме – відображення тексту, якщо картинка не завантажилась або була пошкоджена, проте згодом було реалізовано пошук графічних елементів саме за допомогою цього параметра. Альтернативний текст не повинен бути довгий, та також повинен містити головний чи другорядний ключ.

Мета-тег `keywords` - це фрагмент коду, який використовується для вказівки списку ключових слів, релевантних для конкретної сторінки. У кінці 90-х рр. він був важливим, тому що визначав список ключових фраз для просування кожної сторінки в пошуку. Оптимізатори використовували цю особливість на користь, намагалися вписати безліч ключових фраз. Часто зазначені в `meta keywords` ключі не відповідали змісту сторінки. У результаті

такого перенасичення ключовими фразами сенс тега повністю губився. Тег перестав ранжуватись, і від цього потерпала репутація самих пошукачів. Незабаром пошукові алгоритми стали застосовувати санкції до сайтів за надмірне використання фраз. Якщо алгоритм знаходив на сторінці велику кількість нетематичних ключових фраз у тезі keywords, сторінка падала в результатах видачі і потрапляла під санкції[55].

Ще в 2009 році в Google офіційно заявили, що пошукова система не буде враховувати keywords. У довіднику для вебмастера від Яндекс зазначено, що тег все ж може враховуватися при формуванні пошукової видачі[61].

Алгоритми пошукових систем регулярно ускладнюються. Цілком ймовірно, що в майбутньому зазначені в keywords ключові фрази не будуть ні на що впливати. Більшість пошукових оптимізаторів повністю відмовилися від використання тега, що вплинуло на пошукову аналітику.

Інтернет-ресурс «яндекс» пропонує такі правила заповнення keywords[63]:

1. Ключові фрази в тезі keywords потрібно прописувати через кому з пропуском.
2. Не допускати переспаму. Кількість запитів повинна бути невеликою, зміст - повністю відповідати тематиці сторінки.
3. Не застосовувати повтори слів. Часте повторення однакових фраз сприймається пошуковими системами негативно. Наприклад, правильно писати «купити ноутбук, ноутбук Київ, ціна» замість «купити ноутбук, купити ноутбук Київ, ноутбуки ціна».
4. Уписувати унікальні запити. Для кожної сторінки - набір із 2-5 фраз, які не повторюються для інших сторінок.
5. Не прописувати мета-тег keyword для технічних сторінок: форми зворотного зв'язку, контактів та ін.

Крім ручної вставки тега в код сайту, є актуальні спрощені способи впровадження keywords. Для сайтів на WordPress можна використовувати плагіни з вбудованим генератором тега ключових фраз. Або прописувати вручну запити на кожній сторінці в адмін-панелі, після чого вони автоматично будуть вписані в код сайту.

На веб-ресурсі «Joomla» [42] можна робити так само. На рисунку 16 зображено інтерфейс цього сайту та позначено, які поля варто заповнити SEO-спеціалісту.

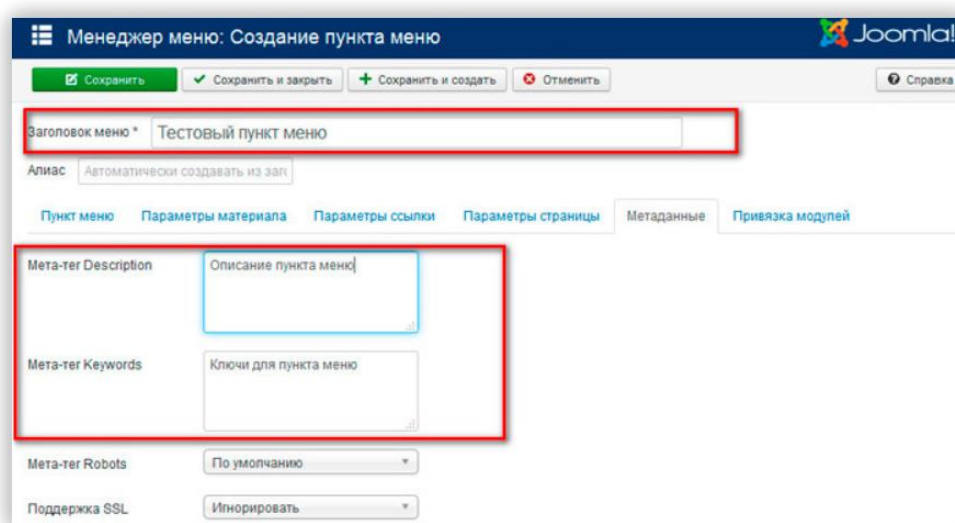


Рис. 16 Створення пункту меню

На сайтах OpenCart такої можливості немає. Тут додати meta keywords для Google і Яндекс можна такими способами:

- зайти в базу даних через хостинг-панель і додати скрипт, зображений на рис. 17.

```
ALTER TABLE `oc_information_description` ADD `meta_description` VARCHAR(
255 ) CHARACTER SET utf8 COLLATE utf8_bin NOT NULL ,
ADD `meta_keyword` VARCHAR( 255 ) CHARACTER SET utf8 COLLATE utf8_bin NOT
NULL
```

Рис. 17 запит для створення meta keywords

- внести зміни в файл admin/controller /catalog/information.php після 270 рядка (показано на рис. 18).

```

$this->data['entry_meta_keyword'] = $this->language-
>get('entry_meta_keyword');
$this->data['entry_meta_description'] = $this->language-
>get('entry_meta_description');|

```

Рис. 18 запит для створення meta keywords через admin/controller /catalog/information.php

- в файл admin/view/template/catalog/information_form.tpl додати ділянку коду, показану на рис. 19.

```

<tr>
<td><?php echo $entry_meta_description; ?></td>
<td><textarea name="information_description[<?php echo
$language['language_id']; ?>][meta_description]" cols="40" rows="5"><?php
echo isset($information_description[$language['language_id']]) ?
$information_description[$language['language_id']]['meta_description'] :
''; ?></textarea></td>
</tr>
<tr>
<td><?php echo $entry_meta_keyword; ?></td>
<td><textarea name="information_description[<?php echo
$language['language_id']; ?>][meta_keyword]" cols="40" rows="5"><?php
echo isset($information_description[$language['language_id']]) ?
$information_description[$language['language_id']]['meta_keyword'] : '';
?></textarea></td>
</tr>

```

Рис 19. запит для створення meta keywords через admin/view/template/catalog/information_form.tpl

- в файлі admin/model/catalog/information.php після 147 рядка додати код, який показано на рис. 20.

```

'meta_keyword' => $result['meta_keyword'],
'meta_description' => $result['meta_description'],

```

Рис. 20 запит для створення meta keywords через admin/model/catalog/information.php

Після виконання однієї з цих дій в адмін-панелі з'являються поля з keywords, які необхідно заповнити для кожної сторінки.

Отже, meta keywords можна використовувати на сторінках сайту, але це вже не обов'язково. Тег не впливає на ранжування. У довіднику для вебмастера від Яндексa зазначено, що пошукова система може враховувати зміст keywords в ранжуванні сторінок. В Google тег meta keywords вважають непотрібним, тому пошуковик його не аналізує.

1.3.Українськомовні запити в SEO

Оскільки набув чинності український закон "Про забезпечення функціонування української мови як національної"[7, 40], власники веб-сайтів почали активно створювати українські версії Інтернет-ресурсів, ТОМУ ЩО закон передбачає, що українська версія повинна бути основною версією і відкрита за замовчуванням, тоді як усі інші версії можуть бути факультативними.Думки експертів розділилися. Дехто вважає, що веб-сторінки потрібні негайно українізувати, інші – навпаки, схиляються до думки: "не чіпай, поки працює", адже переклад ресурсу українською мовою може значно знизити його трафік.Проте слід розглянути ситуацію з точки зору комерції. Розглянемо динаміку зростання запиту "купить" у порівняй українським запитом "купити".

Заданими сервісу “Serpstat”[66] (див. рис. 21) у середньому за період з 2017 по 2020 рік кількість запитів, що містять «купити» українською, складає майже 10-15%.

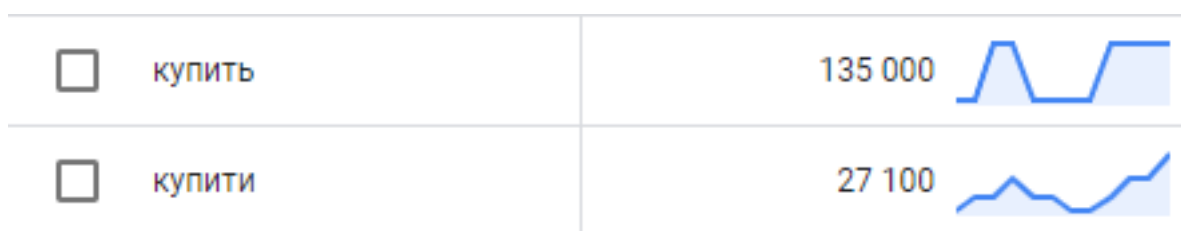


Рис. 21 Кількість запитів «купити» та «купить» у період з 2017 по 2020 рік

Нині цей запит став значно популярнішим, що означає збільшення популярності українськомовних запитів серед користувачів (див. рис. 22).

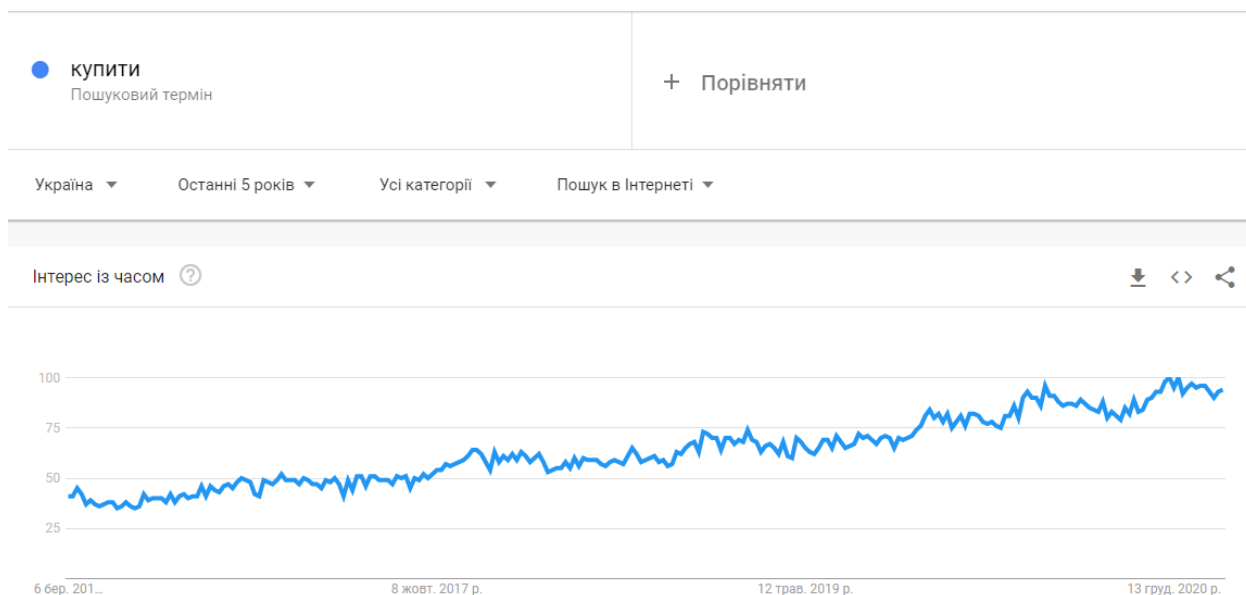


Рис. 22 Динаміка запиту «купити»

Якщо досліджувати популярність запиту за шкалою «інтересу» користувача, то початкове значення терміна «купити» – було 35 балів, а через три роки, цей запит набрав 100 балів.

Проаналізуємо динаміку аналогічного російськомовного запиту на тому ж сервісі “Serpstat”[66]. (див. рис. 23)

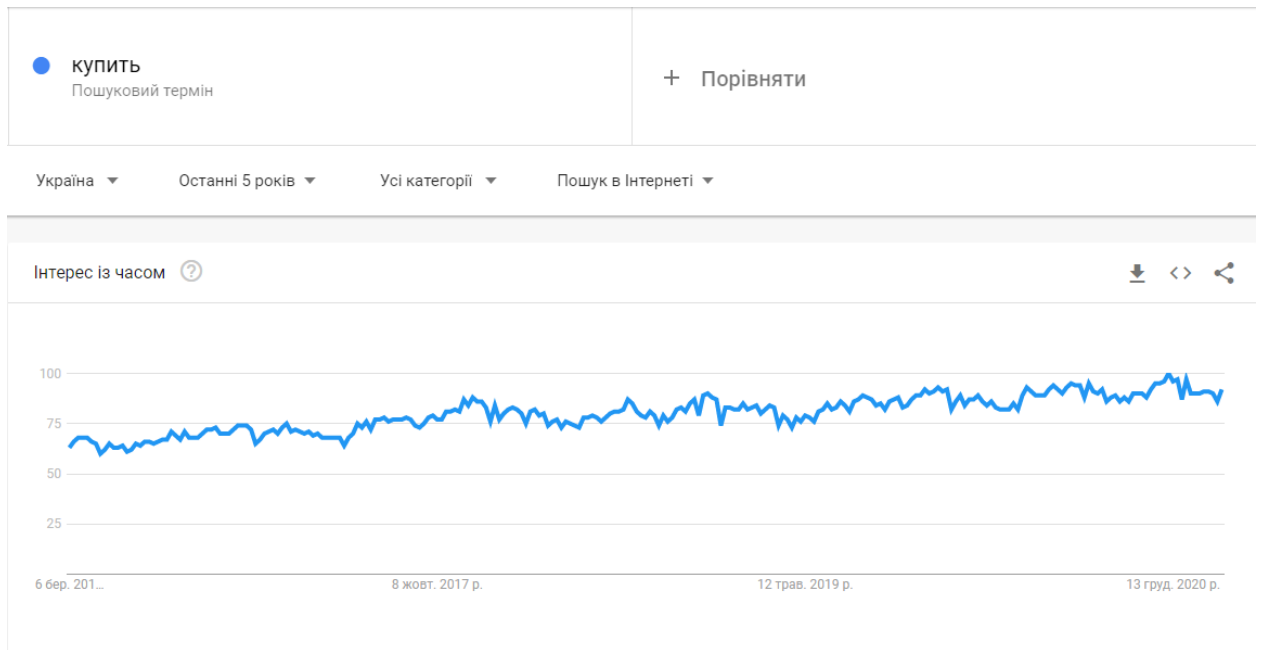


Рис. 23 Динаміка запиту «купить»

Динаміка цього запиту спочатку мала 62 бали, а згодом отримала 92 бали за такий ж період.

Україномовний запит набирає обертів повільніше, проте займає більш впевнені позиції, без різких стрибків та спадів, як це показано на рисунку 24.



Рис. 24 Динаміка запиту «купити» та «купить»

Аналіз регіональних особливостей (див. рис. 25) демонструє статистику використання вищенаведених запитів. «Купити» - позначено синім кольором, «купить» - червоні.



Рис. 25 Вживання термінів «купити» та «купить» в залежності від регіону

Зрозуміло, що основна маса запитів припадає на західні регіони, проте і східні регіони із загальної кількості додають 2-10 балів, що теж не мало.

Для аналізу ціни за клік, скористаємось сервісом Google Ads, який показує статистику, зображену на рисунку 26.

<input type="checkbox"/> Ключевые слова (по релевантности) ↓	Ср. число запросов в месяц	Уровень конкуренции	Процент полученных показов объявления	Ставка для показа сверху стр. (мин.)	Ставка для показа сверху стр. (макс.)
Указанные вами ключевые слова					
<input type="checkbox"/> купить	135 000	Средний	-	1,67 грн.	7,41 грн.
<input type="checkbox"/> купити	27 100	Средний	-	0,68 грн.	3,39 грн.

Рис. 26 Вартість кліку запитів «купити» та «купить»

Тобто, просувати українськомовні запити фактично вдвічі дешевше[60].

розглянемо конкретну ситуацію на прикладах. Припустимо, що ми маємо свою підприємницьку діяльність, наприклад, магазин жіночого одягу. У нас є сайт, який успішно функціонує вже декілька років, проте сайт російськомовний. Виникає запитання – що в перспективі може дати залучення розробників та SEO-фахівців для реалізації українськомовної версії категорії «женские платья»?

Результати зазначено на рисунку 27.

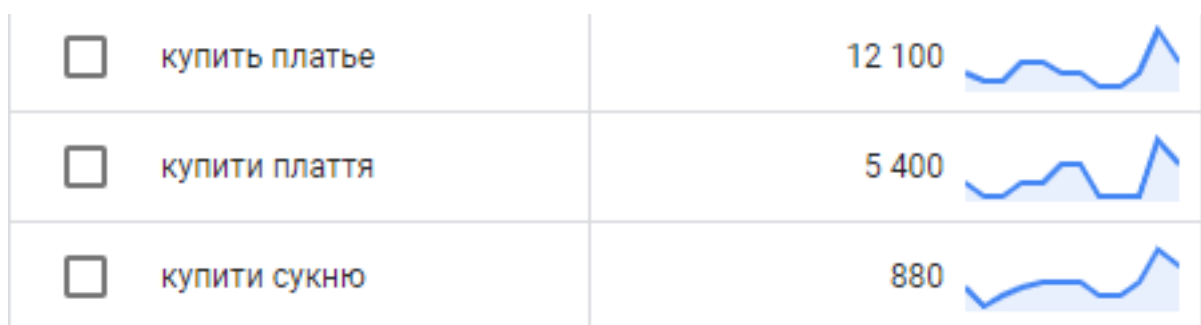


Рис. 27 Статистика запитів «купити плаття» та «купити сукню»

Якщо перевести ці дані у відсотки, то «купить платье» -100%, а українськомовні запити складатимуть лише 35%. Порівняємо ці дані за регіонами (див. рис. 28).

Порівняний розподіл за субрегіонами

● купити платъе ● купити плаття ● купити сукню

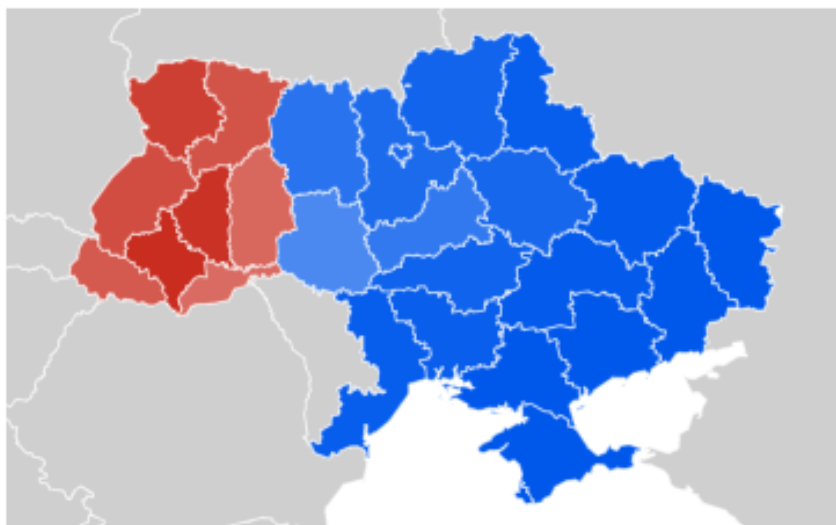


Рис. 28 Статистика запитів «купити плаття» та «купити сукню»

За рисунком 28 видно, що просуваючи українськомовні ключові запити, можна додатково залучити 7 західній областей. Більш детальні запити можна проаналізувати за допомогою сервісу «Serpstat» [54].

Кількість ключових запитів складає понад 16 тисяч, із яких 140 мають частотність більше ста (див додаток 8). Сюди ж, як синонім, запит «купити сукню» – понад три тисячі запитів, 15 з них – з частотністю більше ста.

Для порівняння, результати за запитом «купити платъе»: 50 тисяч запитів, з них 579 – частотність понад 100. При цьому, динаміка популярності, для порівняння, запитів «купити плаття» та «купити платъе» зростає приблизно в одному темпі (починаючи з 2010 року)(див. рис. 29).

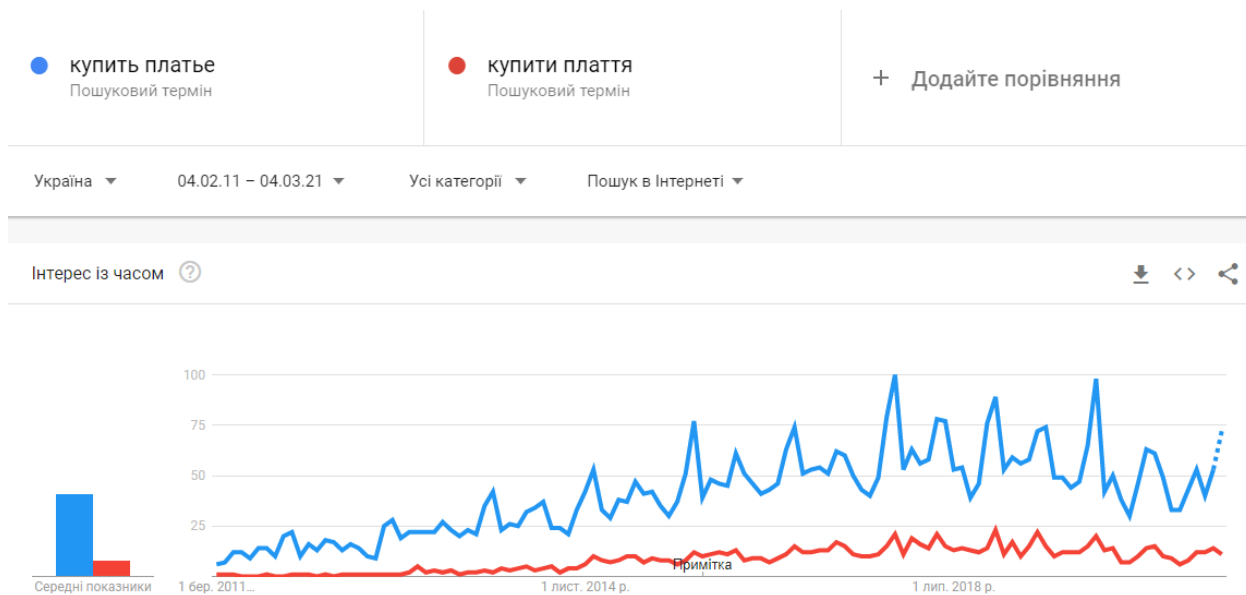


Рис. 29 Динаміка популярності запитів

Отже, аналізуючи дані з погляду комерції, розроблення української версії і використання українськомовних запитів (або тільки використання українськомовної семантики при просуванні сайту) - це можливості. І насамперед - можливості отримати джерело додаткового трафіку, плюс до того «механізму», який вже налагоджений і приносить дохід.

Також, використовуючи українськомовні ключові слова, просуваючи свою веб-сторінку українською мовою, сайт виявляє лояльність до українськомовних користувачів.

Сервіс Serpstat[54] надав свою статистику щодо власної бази даних українськомовних ключів. З кожним роком кількість ключових слів у базі даних стрімко збільшувалась.

- Січень 2013
запуск бази Google Україна
- Квітень 2017
У базі даних - 11 255 941

У пошуковій видачі - 9 055 132

- Листопад 2018

У базі даних - 21 556 324

У пошуковій видачі - 21 556 101

- Грудень 2018

У базі даних - 26 159 128

У пошуковій видачі - 26 159 128

- Січень 2019

У базі даних - 30 453 834

У пошуковій видачі - 27 009 462

- Лютий 2019

У базі даних - 42 345 576

У пошуковій видачі - 41 538 405

Висновки до першого розділу

За базове поняття в нашому дослідженні взято визначення «ключові слова, або, як їх іноді називають, «ключовики», це ті слова, які зустрічаються в тексті частіше, ніж інші. Вони відображають зміст контенту, описують загальну тему опублікованих матеріалів, наприклад, статті, сайту або блогу». Ключові слова є базовим елементом SEO-оптимізації і сприяють визначити пошуковим роботам, як класифікувати веб-ресурс.

Ключові слова не є семантичним ядром тексту у лінгвістичному розумінні. Для того щоб проаналізувати, чи відображає текст ключові слова, які ми можемо встановити за SEO-сервісами, спочатку потрібно знати які це слова та вирахувати їхню кількість у тексті. Тому частота вживання слів та ключові слова, які ми підбираємо між собою взаємопов'язані.

Для того щоб веб-сторінка займала лідируючі позиції, недостатньо оптимізувати лише контент, потрібно оптимізувати назву веб-ресурсу у видачі, тобто title, а також сформуванати опис (description). Якщо розробник не пропише опис сайту самостійно, він підтягнеться автоматично, сформувавшись з першого абзацу тексту. Такої спосіб не нашкодить веб-ресурсу, проте не додасть шансів підняти позицію сайту у пошуковій видачі. Тому, мета-тег <description>слід формувати самостійно використовуючи ключові слова та наприклад, вказівки локації. Або використовувати речення, з найбільшою смисловою вагою.

РОЗДІЛ II.

МЕТОДИ І МОДЕЛІ АВТОМАТИЧНОГО ВИЛУЧЕННЯ КЛЮЧОВИХ СЛІВ

Ключові слова - це одно- і багатокomпонентні лексичні групи, що відображають зміст документа [15, 140]. Автоматичне вилучення ключових слів є необхідним етапом оброблення тексту у таких важливих застосунках, як системи автоматичного інформаційного пошуку, анотування, реферування та ін.

Незважаючи на досить велику кількість досліджень таких науковців, як перерахувати із покликаннями, проблема автоматичного визначення ключових до цих пір не вирішена [23, 366]. Зокрема автоматичне вилучення багатокomпонентних ключових фраз, особливо, якщо робиться спроба автоматично отримати лексичні групи, наприклад, іменні групи.

Алгоритм верхнього рівня вилучення ключових слів універсальний і складається з таких етапів:

- а) формування списку «кандидатів» на ключові слова;
- б) фільтрація цього списку для отримання результативного списку ключових слів.

Досить часто до вилучення ключових слів із тексту видаляються стоп-слова

Стоп-слова - це слова, які не несуть ніякого смислового навантаження (артиклі, прийменники, сполучники, частки, займенники, вставне слово, вигуки і т. д.).

Відмінність методик формування списку ключових слів визначається процедурами оброблення тексту на кожному з етапів і кількістю необхідних для цього лінгвістичних моделей. Основні типи методів і моделей

автоматичного вилучення ключових слів можна розділити на власнестатистичні та гібридні. У межах цих підходів можна виділити методи, що вимагають наявності корпусу текстів однієї тематики, та методи, які не потребують такого корпусу текстів.

2.1. Статистичні моделі автоматичного вилучення ключових слів

Найбільш простий статистичний метод вилучення ключових слів передбачає побудову безлічі кандидатів ключових слів шляхом ранжування всіх словоформ або лексем документа за частотою вживання. Фільтрація полягає у відборі певної кількості найбільш частотних лексем. Цей метод є першим методом автоматичного вилучення ключових слів. Він розроблявся, щев роботах П. Луна [25, 309], Г. Піотровського [6] і широко використовується до сих пір. Поширеність методу відбору ключових слів тільки за частотою пояснюється його простотою.

При використанні частоти слова в документі як єдиного параметра для автоматичного вилучення ключових слів підрахунок загальної частоти словоформ з парадигми однієї лексеми найчастіше здійснюється наступним чином: загальна частота ключових слів підраховується шляхом порівняння словоформ, нормалізованих до однієї форми, як правило, до основи або леми. Автоматична ідентифікація словоформи є завданням автоматичного морфологічного аналізу і досить проблематична сама по собі.

У статистичних підходах до вилучення ключових слів використовуються прості евристичні алгоритми, найчастіше стемінг - ідентифікація словоформи методом відсікання від словоформи певної кількості букв отриманні квазіоснови. Такі алгоритми називають стемінг-алгоритмами, найбільш відомим з яких є стемінг-алгоритм Портера [30, 313]. Нормалізовані словоформи ранжуються за частотою і ті з них, чия частота вище заданого порогу, вважаються ключовими. Ключові слова, як правило,

видаються в усіченому вигляді квазіоснов. Статистичні методи вилучення багатокomпонентних ключових слів використовують обчислення n-грам [18, 152].

З одного боку, частота вживання слова безсумнівно характеризує важливість слова для документа, але, з іншого боку, ключові слова, як підкреслювали дослідники групи «Статистика мови» Г. Піотровського, і інші, не завжди є найбільш частотними [2, 195]. Часто саме унікальні терміни більш точно маркують тему документа.

Для обліку параметрів частотності й унікальності лексем тексту, для обчислення релевантності ключових слів документа широко використовується метод TF-IDF [19, 493] із застосуванням корпусу однакових за тематикою документів. Релевантність ключових слів у цьому випадку визначається як добуток двох величин: частоти слова в документі (TF = Term Frequency) і зворотної частоти слова в колекції документів (IDF = Inverse Document Frequency). Останнє означає - кількість документів у корпусі, де термін ужито принаймні один раз. Термін було введено британським вченим в області інформатики - Кареном Спарком Джонсоном [112].

Використання корпусу текстів для підвищення коректності вилучення ключових слів отримало досить широке застосування, однак відсутність таких корпусів для кожної конкретної предметної галузі робить застосування таких корпусних моделей і методів вельми проблематичним.

У прагненні більш точно відобразити зміст документа розробляються методики, що використовують комбінації значень різних параметрів лексем, наприклад, частоти лексеми в документі, розташування в певній частині тексту (наприклад, у заголовку або початку параграфа), статистики щодо повторів слів у документі / корпусі і їх дисперсії [26, 157] або відношення логарифмічних правдоподібностей [12, 61].

Перевагами власне статистичного підходу є універсальність алгоритмів визначення ключових слів і відсутність необхідності в трудомістких і часозатратних процедурах побудови лінгвістичних баз знань. Незважаючи на зазначені переваги статистичних методів вилучення ключових слів, ці методи часто не забезпечують задовільної якості результатів. При цьому сфера їх застосування обмежена мовами з бідною морфологією, такими як англійська, де частотність словоформ однієї лексеми велика. власне статистичні моделі визначення ключових слів, задовільно працюють, наприклад, на матеріалі англійської мови, та не придатні для природних мов з багатою морфологією, зокрема, для української мови, де кожна лексема характеризується великою кількістю словоформ із низькою частотністю в кожному конкретному тексті.

2.2. Гібридні моделі автоматичного вилучення ключових слів

Для підвищення коректності автоматичного вилучення ключових слів використовуються гібридні методики, у яких статистичні методи оброблення документів доповнюються однією або декількома лінгвістичними процедурами (морфологічним, синтаксичним і семантичним аналізами) і лінгвістичними базами знань різної глибини (словниками, онтологіями, граматиками, лінгвістичними правилами та ін.).

Гібридні методи вилучення ключових слів із документа, також як і статистичні, можуть вимагати або не вимагати корпусу текстів. Не вимагають корпусу гібридні методи вилучення ключових іменних груп описані, наприклад, в роботах: К. Баркера К. Корнакія Н. «Використання головних словосполучень для вилучення ключових фраз документа. Досягнення штучного інтелекту» [55], А. Хальса «Удосконалене автоматичне вилучення ключових слів з огляду на мовні знання»[16, 216], Б. Крулівіча «Аналіз інтересів користувачів для вилучення семантично значущих фраз»[46], С. Шереметьєвою «Ефективний патентний екстрактор ключових слів як ресурс перекладу»[28, 404], [55].

Метод Кена Баркера представлений в роботі «Використання головних словосполучень для вилучення ключових фраз документа. Досягнення штучного інтелекту» [55], включає пошук у тексті документа базових іменних груп (БІГ) із використанням морфо-синтаксичного аналізу на основі словників та обчислення релевантності БІГ. Ключовими вважаються іменні групи з показником релевантності вище заданого порогу. Іменна група (ІГ) - словосполучення, в якому іменник є вершиною, тобто головним словом, що визначає характеристику всієї складової.

Гібридний метод визначення ключових іменних фраз, розроблений С. Шереметьєвою [32, 25] для англійської мови, не вимагає наявності корпусу текстів, передбачає побудову безлічі кандидатів за допомогою обчислення всіх n-грам документа і фільтрацію отриманої множини за допомогою правил видалення n-грам, які не є іменними фразами, і обчислення релевантності «уцілілих» n-грам-іменних груп.

У гібридних методах визначення ключових слів на основі графів, основною процедурою є побудова зваженого графа, у вершинах якого стоять лексеми-кандидати в ключові слова, а дуги зважені відповідно до ступеня близькості кандидатів до вершин. Визначення ключових слів на основі графів представлено у роботах: К. Гіріша «Вилучення ключових слів з одного документа за допомогою мір центральності. Розпізнавання зразків та машинний інтелект.» [14, 503], М. Литвака «Виділення ключових слів на основі графів для узагальнення окремих документів» [22, 17], Ю. Мацуя «Ключовий світ: Вилучення ключових слів із малого світу документа. Наука про відкриття» [27, 271], Ю. Осава «Автоматичне індексування за графіком співіснування на основі метафори будівництва ADL '98» [29, 12], Р. Міхалча «TextRank: Наведення порядку в текстах.» [28, 404], Д. Усталова «Виділення термінів з російськомовних текстів при використанні графічних моделей.» [67]

Ключові слова відбираються в процесі обробки графаалгоритмами з теорії графів. Відмінність між цими методами полягає в способах відбору безлічі кандидатів-вершин і визначення близькості окремих кандидатів, які, поряд зі статистичними параметрами відбору (наприклад, близькістю розташування в тексті, що обчислюється за кількістю слів між двома термінами), засновані на морфологічному, синтаксичному, а іноді і семантичному аналізі, наприклад, за допомогою статей Вікіпедії (чому присвячені, наприклад, роботи М. Гриневої «Аналіз текстових документів для виведення тематично згрупованих ключових термінів» [4, 155], «Ефективне вилучення тематично згрупованих ключових термінів із тексту» [36].)

До числа гібридних методів вилучення ключових слів можна віднести методи на основі машинного навчання, де завдання визначення ключових слів розглядається як задача класифікації. Методи на основі машинного навчання для створення навчальної вибірки і побудови моделі-класифікатора, як правило, вимагають корпусу документів із розміченими ключовими словами. Розмічені ключові слова вважаються позитивним прикладом, інші слова - негативним прикладом. Далі вираховується релевантність кожного слова тренувального тексту шляхом зіставлення його вектора значень різних параметрів, наприклад, TF-IDF, довжини слова, частини мови, позиція слова в заголовку, позиція слова в першому абзаці, останньому абзаці, у списках літератури тощо фіксується відмінність значень векторів цих параметрів для ключових слів і неключових. Далі обчислюється ймовірність віднесення кожного слова до групи ключових і задається її поріг, тобто модель навчається. Витяг ключових слів із нового документа відбувається шляхом обчислення релевантності слів і їх ймовірності віднесення до ключових слів відповідно до побудованої моделі.

Серед методів на основі машинного навчання можна відзначити:

- баєсівські методи (розглянуто у наступних роботах:

Е. Френк, Пейнтер, Г. Віттен , С. Гутвін, Г. Невілл-Меннінг «Вилучення специфічних для домену ключових фраз» [13, 668], Х. Віттен «Практичне автоматичне вилучення ключових фраз» [20, 254], С. Вассерман, К. Фауст «Аналіз соціальних мереж: методи та застосування.» [35,403], Ю. Узун «Вилучення ключових слів за допомогою наївних Баєсів.» [60].);

- метод опорних векторів(розглянуто в наступних роботах: Р. Гербріх «Великі межі рангу границь для порядкової регресії» [15, 115], Х. Цзян «Рейтинговий підхід до вилучення ключових фраз» [17, 756], Чжан, Х. Сю, Дж. Тан, Дж. Лі«Вилучення ключових слів за допомогою машини підтримки Vector» [21, 85].);
- метод дерев рішень(П. Терні «Вивчення алгоритмів вилучення ключових фраз» [33, 303]);
- метод використання нейронних мереж (П. Лопес «Автоматичне вилучення ключових термінів із наукових статей у GROBID» [24, 248], К Саркар «Новий підхід до вилучення ключових фраз із використанням нейронних мереж.» [31, 16], Дж. Ван, Х. Пен, Дж.-С Ху. «Автоматичне вилучення ключових фраз із документа за допомогою нейронної мережі.» [34, 633]).

Аналіз існуючих методів автоматичного вилучення ключових слів показує, що для створення ефективних екстракторів ключових слів потрібно враховувати морфологічні типи природних мов (аналітичний, флективний, агглютинативний, ізолюючий), предметнусферу(тематичну підмову) і наявність необхідних лінгвістичних та програмних ресурсів.

Висновки до другого розділу

Зважаючи на існуючі методи, найбільш оптимальним буде вибір гібридного методу та використання семантичного аналізу для розрахунку частоти слів, а також використання методів, що використовують представлення тексту у вигляді графу.

В якості основного напрямку цієї роботи ми обрали дослідження одного з алгоритмів автоматичного вилучення слів і словосполучень з тексту, а саме TextRank, адже цей алгоритм широко відомий завдяки своїм високим результатам на матеріалі англійської мови, тому ми спробували застосувати його до українськомовних текстів і знайти оптимальний спосіб оцінки його ефективності для української мови.

Популярним підходом до виділення ключових слів є підхід на основі графів залежностей між словами. Мабуть, TextRank є найвідомішим методом. Основна ідея методу TextRank полягає в побудові графа зв'язків між словами з тексту. Такий граф містить в якості вершин всі слова, які зустрічаються в тексті. Цей набір слів можна обмежувати - наприклад, не включати слова з стоп-листа і / або включати тільки слова, що відносяться до певних частин мови, наприклад, тільки іменники і прикметники.

Деякі з методів припускають наявність словників або фонових корпусів, проте не всі. Так як більшість досліджень в цій області проводиться на матеріалі англійської мови, то для української мови відсутній великий вибір словників.

РОЗДІЛ III.

СТВОРЕННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ SEO-ЗАСТОСУНКУ

Сучасні методи SEO тісно пов'язані із задачами комп'ютерної лінгвістики, тому що автоматичне оброблення текстів спрощує процес формування списку ключовиків інтернет-тексту. Тому було поставлено завдання - розробити застосунок для seo, який спрямований на автоматичний аналіз українськомовних веб-ресурсів з метою формування ключовиків.

Створювана система повинна відповідати таким вимогам:

- швидкість роботи;
- можливість використовувати ПЗ на різних пристроях (кросплатформеність);
- наявність словників та алгоритмів, які дозволять працювати з українськомовними текстами
- функціонал для швидкої генерації тексту для SEO (наприклад «title» у HTML);
- функціонал для аналізу значущості окремих речень, семантичного та морфологічного аналізу тексту та/або його частин;
- можливість оновлення функціоналу та використання додаткових словників, за можливості – без необхідності оновлення додатка користувачами.

На основі описаних вище вимог було прийнято рішення створення ПЗ у вигляді односторінкового веб-застосунку. Перевагами такого підходу є кросплатформеність і можливість оновлення застосунка без втручання користувача.

Мовою програмування було обрано Python, тому що для цієї мови існує велика кількість готових модулів для аналізу тексту, автоматичних словників

для української мови, а також існує можливість досить простої реалізації веб-додатків із використанням програмного каркасу Django2 і веб-вервера gunicorn.

Попри наявність широкого вибору хостингів для веб-додатків, пріоритетними для цього проєкту є легкість розгортання, час доступності (“uptime”) та ціна процесорного часу, а також наявність підтримки Python/Django. Серед доступних хмарних платформ виділяється безкоштовний рівень PaaS-платформи Heroku: безкоштовне використання процесорного часу за умови активного використання, швидкий перезапуск вимкненого серверу за наявності запиту від користувача, підтримка інших використаних технологій за допомогою своїх модулів-з’єднувачів і наявність своєї реалізації системи «неперервної інтеграції» (англ. “continuous integration”, CI) для автоматичного тестування та розгортання нових версій додатка без втрати функціональності для поточних користувачів.

Для безпосередньо морфологічного та семантичного аналізу текстів використовуються внутрішні інструменти роботи з текстом мови Python, а також сторонні модулі з відкритим кодом, які використовують NLTK-словники для української мови, або їх модифікації – модулі `summa`, `summarizer` та `rumorphy2`.

3.1 Визначення вимог для реалізації застосунку

Для створення шаблонного проєкту, готового до розгортання майже без модифікацій, було використано сервіс DjangoBuilder. На основі отриманого проєкту було розроблено односторінковий додаток, що виконує такі функції за HTTP-запитами:

- 1) У відповідь на GET-запит надсилає користувачеві основну сторінку з формою для введення тексту.
- 2) При відправці форми проводиться відправка POST-запиту із введеним текстом.

- 3) Отриманий у запиті текст обробляється на стороні сервера (за наявності ресурсів Django виконує оброблення в окремому потоці, що пришвидшує швидкість оброблення тексту декількома користувачами)
- 4) У відповідь на попередній запит сервер відправляє сторінку, на якій за допомогою DjangoTemplates виводяться результати оброблення – сумаризація, найважливіші речення та слова/словосполучення, а також повна статистика використання усіх слів після зведення їх до початкової форми.

Більша частина логіки реалізована як одне з «представлень» у файлі `views.py`, з перенаправленням запитів на правильну логіку у `urls.py`. Сервер також має стандартні внутрішні тести, які використовуються при розгортанні сервера на хмарну платформу.

За необхідністю постійного хостингу та необхідність постійно надавати послуги веб-спеціалістам створений додаток розгортається в автоматичному режимі на платформу Heroku за допомогою `buildpack/python` та `HerokuCI`. При появі нових змін у репозиторії на GitHub платформа автоматично завантажує нову версію, збирає та тестує її, і, за умови успішного проходження усіх тестів та за відсутності критичних помилок, публікує нову версію за старим покликанням.

Для досягнення можливості повністю автоматичного розгортання було зроблено такі зміни у додатку:

- Додано файл `requirements.txt` зі списком усіх модулів, які потрібно встановити перед розгортанням. З причини того що хмарні сервіси працюють в основному на контейнеризованих рішеннях (таких як `Docker`), без встановлення та/або перевірки версії доступних модулів не можна гарантувати працездатність ПЗ. Зазвичай контейнери, на яких

хмарні платформи запускають додатки, мають дуже обмежений набір заздалегіть встановлених модулів, тому крок перевірки їх наявності стає необхідним.

- Додано Procfile. Це спеціальний файл для платформи Heroku, що позначає процеси, які необхідно запуснути після виконання усіх кроків «збирання» проєкту. У такому випадку – це запуск веб-серверу gunicorn, що буде обробляти запити від користувачів.
- Додано django-heroku та його ініціалізаційні параметри. Це модуль від Heroku, який відповідає за інтеграцію додатка з платформою, можливості моніторингу внутрішніх змінних а також перенаправлення логування з внутрішніх логів сервера на системи логування, доступні на платформі.

Рymorphу2 - морфологічний аналізатор, розроблений мовою програмування Python. Виконує лематизацію і морфологічний аналіз слів, здатний здійснювати відмінювання за заданими граматичними характеристиками слів. Працює зі словником OpenCorpora, а для незнайомих слів будує гіпотези. Підтримує російську та українську мови[49]. Саме цей морфологічний аналізатор став основою для програмного забезпечення для українськомовних текстів.

Алгоритм TextRank - це алгоритм сортування тексту, який покращений порівняно з алгоритмом Google PageRank, який може витягувати текст із заданого тексту. Ключові слова чи група ключових слів використовується за допомогою методу автоматичного підсумовування.

Алгоритм TextRank перенесений з алгоритму PageRank, який є алгоритмом сортування за важливістю веб-сторінок: алгоритм PageRank обчислює важливість кожної сторінки на основі взаємодії покликань між сторінками в World Wide Web; Тому, коли мова йде про алгоритм TextRank, можна розглядати його на прикладі алгоритму PageRank.

Алгоритм PageRank з'явився у зв'язку з розробкою пошукових систем. Раніше пошукові системи зазвичай використовували метод каталогу класифікації, тобто ручну класифікацію веб-сторінок для сортування високоякісних веб-сайтів. Зі збільшенням кількості веб-сторінок метод ручної класифікації став нереалістичним, і люди почали намагатися використовувати метод пошуку тексту, тобто шляхом обчислення ступеня релевантності між ключовими словами, запитаними користувачем, і змістом веб-сторінки для отримання результатів пошуку. Цей метод долає обмеження кількості сторінок, але ефект цього методу не завжди дуже хороший, тому що деякі сайти навмисно «оперують» певними ключовими словами, щоб підвищити свій рейтинг у пошуку. Ця проблема була вирішена на сьомій Міжнародній конференції World Wide Web у квітні 1998 року - Ларрі Пейдж і Сергій Брін запропонували алгоритм PageRank. Алгоритм приблизно оцінював важливість веб-сторінок шляхом розрахунку кількості і якості покликань на веб-сторінки. Коли алгоритм був вперше протестований, його було застосовано для пошукової системи Google для ранжування веб-сторінок.

«Summa» - це бібліотека для мови Python, яка вилучає ключові слова на основі методу TextRank[59].

Визначення важливих словосполучень тексту відбувається за допомогою бібліотеки «Summarizer». Summarizer - це автоматичний алгоритм підсумовування [54]. Всі входження повинні супроводжуватися якоюсь формою модульного тестування.

3.1 Загальна характеристика технологій та програмних інструментів для розроблення застосунку

Набазі використання описаних інформаційних ресурсів було створено веб-застосунок, який слугує помічником для seo-оптимізаторів покликання на рисунок і аналіз функцій.

Для створення HTML-структури застосунку було використано блочну верстку (див. додаток 8, додаток 9). Для стилізації – фреймворк «Bootstrap».

Після введення того, як користувач ввів текст для аналізу, на сервері відбувається розгортання файлу(див. додаток 10) – скачування та запуск бібліотек, зазначених у файлі requirements.txt (див. додаток 11). Після запуску бібліотек відбувається їхнє тестування за допомогою файлів test_helpers.py та test_settings.py (див. додаток 12). Для розгорнення застосунку на хмарному сервісі, функції, розміщені у файлі settings.py запускають методи з бібліотеки “django-heroku. (див. додаток 13).

Інтерфейст застосунку мінімалістичний та інтуїтивно-зрозумілий для користувача. (див. рис. 30)

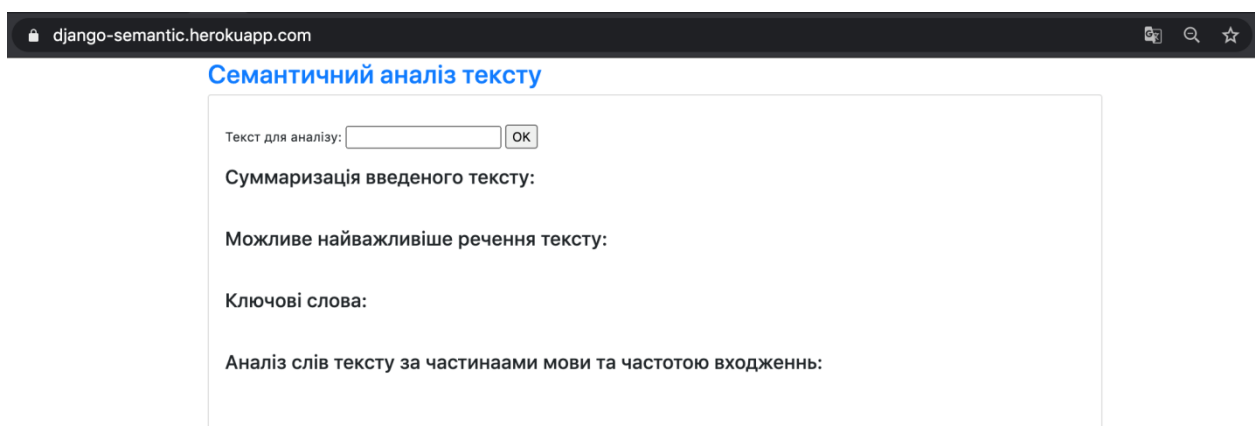


Рис. 30 Інтерфейс веб-додатка

Переглянути додаток можна за покликанням <https://django-semantic.herokuapp.com>, оскільки для зручності ресурс було опубліковано.

Текст з додатку 16 було введено у поле «текст для аналізу». Результат видачі зображено на рис. 31, рис. 32, рис. 33.

Семантичний аналіз тексту

Текст для аналізу:

Суммаризація введеного тексту:

["Python (найчастіше вживане прочитання — «Пайтон», запозичено назву[7] з британського шоу Монті Пайтон) — інтерпретована об'єктно-орієнтована мова програмування високого рівня зі строгою динамічною типізацією.", 'Інтерпретатор Python та стандартні бібліотеки доступні як у скомпільованій, так і у вихідній формі на всіх основних платформах.', "В мові програмування Python підтримується кілька парадигм програмування, зокрема: об'єктно-орієнтована, процедурна, функціональна та аспектно-орієнтована."]

Можливе найважливіше речення тексту:

В мові програмування Python підтримується кілька парадигм програмування, зокрема: об'єктно-орієнтована, процедурна, функціональна та аспектно-орієнтована.

Рис. 31 Аналіз застосунку. Реферування тексту

Ключові слова:

python
динамічною
програмування
високого
рівня
об
назву
з
британського
запозичено
привабливою

Рис. 32 Аналіз застосунку. Ключові слова

Аналіз слів тексту за частинами мови та частотою слововживань:

```
python None 4
динамічний ADJF 3
— None 2
мова NOUN 2
програмування NOUN 2
високий NOUN 2
рівня NOUN 2
(найчастіше ADVB 1
вживаний ADJF 1
прочитання NOUN 1
«пайтон», None 1
запозичити VERB 1
назву[7] None 1
британський ADJF 1
шоу NOUN 1
монті NOUN 1
пайтон) None 1
інтерпретований ADJF 1
об'єктно-орієнтований ADJF 1
строгий ADJF 1
типізацією.[8] None 1
розроблений ADJF 1
1990 None 1
рок NOUN 1
гідо NOUN 1
россумом. ADJF 1
структура NOUN 1
даний ADJF 1
семантика NOUN 1
зв'язування NOUN 1
робити VERB 1
привабливий ADJF 1
швидкий ADJF 1
розробка NOUN 1
програм, None 1
засіб NOUN 1
послдування NOUN 1
наявний ADJF 1
компонентів. ADJF 1
підтримувати VERB 1
модель. NOUN 1
```

Рис. 33 Аналіз застосунку. Частота вживання слів

Перейти до веб-застосунку можливо також за допомогою сканування QRкоду, який знаходиться у додатку 17. Для переходу на веб-ресурс через QR-код потрібно навести камеру на зображення, девайс автоматично здійснить перехід на веб-сайт. (Для деяких смартфонів потрібен спеціальний застосунок для считування QR-коду) .

Висновки до третього розділу

Оскільки SEO-оптимізація веб-сторінок це досить повільний та складний процес, було прийнято рішення створити застосунок, який допоможе SEO-спеціалістам вираховувати частоту слів у тексті для подальшого складання семантичного ядра та пошуку ключових слів. Також, застосунок допомагає формувати основні мета-теги для веб-ресурсу.

Переваги застосунка:

- мінімальний час на виконання поставленого завдання;
- адаптація під мобільний пристрій;
- інтуїтивно-зрозумілий інтерфейс;
- автоматичне реферування тексту, результат якого можна використати як сніпет до веб-ресурсу;
- визначення ключових слів відбуваються з відкиданням стоп-слів, таких як 'а', 'аби', 'абиде', 'абиким', 'абикого', 'абиколи', 'абикому', 'абикуди', 'абихто', 'абичий', 'абичийого', 'абичийому', 'абичим', 'абичию', 'абичия', 'абичие', 'абичиему', 'абичиею', 'абичиеї', 'абичиї', 'абичийї' та ін.;
- автоматичне визначення частини мови та частоти її входжень.

Додаток розроблено на мові Python через велику кількість існуючих бібліотек для автоматичного аналізу тексту. Основним морфологічним аналізатором було обрано «Rymorphy2», який після введення тексту виконує лематизацію та морфологічний аналіз слів. Серед усіх аналізаторів, доступних на мові Python, з українськими словниками працює лише «Rymorphy», використовуючи словник OpenCorpora(для незнайомих слів будує гіпотези).

ВИСНОВКИ

SEO-копірайтинг займає важливе місце в процесі створення і просування сайту. Застосування принципів SEO-копірайтингу на сьогодні є обов'язковим при написанні текстів для комерційних сайтів.

Наявність певних ключових слів з семантичного ядра на кожній сторінці сайту – обов'язкова умова грамотного пошукового просування. З якісно підготовленого тексту сторінки формуються влучні та доречні сніпети (фрагменти тексту із включенням слів запиту користувача, які знаходить пошукова машина на сторінці сайту), які виводяться в результатах пошуку.

Для полегшення просування сайтів у вільному доступі знаходиться чимало ресурсів, які допомагають відслідковувати пошукові запити користувачів у різних локаціях. Запити користувачів постійно змінюються, тому для правильного визначення ключових слів буде помилковим використання лише сервісів, які аналізують користувацькі запити.

Для того щоб правильно визначити ключові слова та сформувані в подальшому семантичне ядро, потрібно врахувати семантичний аналіз контенту, який розмішено на веб-ресурсі. Такий підхід дозволить здійснити орієнтацію на основний зміст контенту, що неодмінно збільшить трафік веб-ресурсу.

Проаналізувавши теоретичний матеріал та статистику українськомовних запитів, було створено застосунок для SEO-оптимізації веб-сторінок. Для зручності застосунок було опубліковано на хостинг для вільного доступу. Застосунок аналізує вхідний текст та одразу показує користувачеві результат:

- реферування вхідного тексту – невеличкий текст, який показує основний зміст введеного контенту. Його практичне використання: формування сніпета веб-сайту, опис ресурсу для контекстної реклами;
- найважливіше речення тексту, тобто речення, яке несе найбільше смислове навантаження. Практичне використання: формування сніпету, формування заголовків першого та другого рівня (<h1>, <h2>), створення мета-тегу<title>;
- ключові слова – слова, на які слід звернути особливу увагу SEO-спеціалісту для подальшої роботи з веб-ресурсом. Використовуючи список поданих ключових слів можна сформувати категорії веб-сторінки, створити сніпет, додавши до ключових слів «вірусні» слова (такі як «економно», «дешево», «знижки» та ін.)
- Аналіз слів тексту за частинами мови та частотою входжень – надасть інформацію про найвживаніші слова та вкаже, до яких частин мови вони належать. Практичне використання: найкраще, якщо ключові слова відносяться до різних частин мови, адже це допоможе автору досягти декілька цілей: наприклад розповісти про якусь подію або основний об'єкт тексту та, наприклад, назвати його основні ознаки або дати оцінку. Відповідно, з різними частинами мови SEO-спеціаліст зможе скомбінувати різні текстові композиції.

Вищеперераховані способи використовуються для якісного та ефективного просування веб-ресурсу, що і є кінцевою ціллю створеного додатка.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Агеев М. С., Добров Б. В., Лукашевич Н. В. Автоматическая рубрикация текстов: методы и проблемы // Учёные записки Казанского государственного университета. Серия Физико-математические науки. — 2008. — Т. 150, № 4. — С. 25–40
2. Алексеев П.М., Герман-Прозорова Л.П., Пиотровский Р.Г., Шепетова О.П. Основы статистической оптимизации преподавания иностранных языков. Статистика речи и автоматический анализ текста. Л., 1974. С. 195–234. [Alekseev P.M., German-Prozorova L.P., Piotrovskii R.G., Shepetova O.P. *Osnovy statisticheskoy optimizatsii prepodavaniya inostrannykh yazykov* (Basics of the Statistical Optimization of Foreign Languages Teaching). *Statistika rechi i avtomaticheskii analiz teksta* (Statistics of Speech and Automatic Analysis of the Text). Leningrad, 1974, pp. 195–234.]
3. Воронцов К. В., Потапенко А. А. Регуляризация вероятностных тематических моделей для повышения интерпретируемости и определения числа тем // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Бекасово, 4–8 июня 2014 г.). — Вып. 13 (20). — М: Изд-во РГГУ, 2014. — С. 676–687.
4. Гринева М., Гринев М. Анализ текстовых документов для извлечения тематически сгруппированных ключевых терминов. Труды ИСП РАН. 2009. Т. 16. С. 155–165.
5. О. Г. Пасічник, О. В. Пасічник, І. В. Стеценко. Основи веб-дизайну [Навч. посіб.]. — К.: Вид. група ВНУ. 2009. 336 с
6. Пиотровский Р.Г., Бектаев К.Б., Пиотровская А.А. Математическая лингвистика: учеб. пособие для пед. институтов. М.: Высшая школа, 1977. 383 с. [Piotrovskiy R.G., Bektaev K.B., Piotrovskaya A.A.

- Matematicheskayalingvistika. (MathematicalLinguistics). Moscow, Vysshayashkola, 1977, 383 p.]
7. Проект закону: Про забезпечення функціонування української мови як державної (Відомості Верховної Ради (ВВР), 2019, № 21, ст.81) URL: <https://zakon.rada.gov.ua/laws/show/2704-19#Text> (Дата звернення: 09.04.2021)
 8. Barker K. Cornacchia N. Using Noun Phrase Heads to Extract Document Keyphrases. *Advances in Artificial Intelligence*. 2000, vol. 1822, pp. 40–52.
 9. Bolelli L., Ertekin S., Giles C. L. Topic and trend detection in text collections using latent Dirichlet allocation // *ECIR*. — Vol. 5478 of *Lecture Notes in Computer Science*. — Springer, 2009. — pp. 776–780
 10. Chien J.-T., Chang Y.-L. Bayesian sparse topic model // *Journal of Signal Processing Systems*. — 2013. — Vol. 74. — pp. 375–389
 11. Dostal M. Automatic Keyphrase Extraction Based on NLP and Statistical Methods. *Proceedings of the Databases, Texts, Specifications and Objects. Pisek, Czech Republic, 2011*, pp. 140–145.
 12. Dunning T. Accurate Methods for the Statistics of Surprise and Coincidence. *Computational Linguistics – Special Issue on Using Large Corpora*. 1993, vol. 19, no. 1, pp. 61–74.
 13. Frank E., Paynter G.W., Witten I.H., Gutwin C, Nevill-Manning C.G.. Domain-Specific Keyphrase Extraction. *Proceeding of 16th International Joint Conference on Artificial Intelligence*. Stockholm, Sweden, 1999, pp. 668–673.
 14. Girish K.P. Keyword Extraction from a Single Document Using Centrality Measures. *Pattern Recognition and Machine Intelligence*. Springer Berlin Heidelberg, 2007, pp. 503–510.
 15. Herbrich R. Large Margin Rank Boundaries for Ordinal Regression. *Advances in Large Margin Classifiers*. MIT Press, 2000, pp. 115–132

16. Hulth A. Improved Automatic Keyword Extraction Given More Linguistic Knowledge. Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing. Sapporo, July, 2003, pp. 216–223.
17. Jiang X. A Ranking Approach to Keyphrase Extraction. Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval. Boston, MA, USA, 2009, pp. 756–757.
18. Jiao H. Chinese Keyword Extraction Based on N-Gram and Word Co-occurrence. Proceeding CISW '07 Proceedings of the 2007 International Conference on Computational Intelligence and Security Workshops. Harbin, 2007. pp. 152–155.
19. Jones K.S. A Statistical Interpretation of Term Specificity and its Application in Retrieval. Journal of Documentation. 2004, vol. 60, no. 5, pp. 493–502.
20. KEA: Practical Automatic Keyphrase Extraction. I.H. Witten, G. W. Paynter, G. W. Paynter, E. Frank, C. Gutwin, C. G. Nevill-Manning. DL '99 Proceedings of the Fourth ACM Conference on Digital Libraries. Berkeley, CA, USA, 1999, pp. 254–255.
21. Keyword Extraction Using Support Vector Machine. K. Zhang, H. Xu, J. Tang, J. Li. Advances in Web-Age Information Management. Springer Berlin Heidelberg, 2006, pp. 85–96.
22. Litvak M. Graph-based Keyword Extraction for Single-Document Summarization. Proceedings of the Workshop on Multi-source Multilingual Information Extraction and Summarization. Manchester, United Kingdom, 2008, pp. 17–24.
23. Liu Z., Huang W., Zheng Y., Sun M. Automatic keyphrase extraction via topic decomposition. Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing. Cambridge, Massachusetts, 2010, pp. 366–376.
24. Lopez P. HUMB: Automatic Key Term Extraction from Scientific Articles in GROBID. Proceedings of the 5th International Workshop on Semantic Evaluation. Uppsala, Sweden, 2010, pp. 248–251.

25. Luhn H.P. A Statistical Approach to Mechanized Encoding and Searching of Literary Information. *IBM Journal of Research and Development*. 1957, vol. 1, no. 4, pp. 309–317.
26. Matsuo Y. Keyword Extraction from a Single Document Using Word co-occurrence Statistical Information. *International Journal on Artificial Intelligence Tools*. 2004. V. 13, no. 1, pp. 157–169.
27. Matsuo Y. KeyWorld: Extracting Keywords from Documents Small World. *Discovery Science*. Springer Berlin Heidelberg, 2001, pp. 271–281. 31
28. Mihalcea R. TextRank: Bringing Order into Texts. *Proceedings of EMNLP 2004*. Barcelona, Spain, 2004, pp. 404–411.
29. Ohsawa Y. KeyGraph: Automatic Indexing by co-occurrence Graph Based on Building Construction Metaphor. *ADL '98 Proceedings of the Advances in Digital Libraries Conference*. Santa Barbara, CA, USA, 1998, pp. 12–18.
30. Porter M.F. An Algorithm for Suffix Stripping. *Readings in Information Retrieval*. Morgan Kaufmann Publishers Inc., 1997, pp. 313–316.
31. Sarkar K., Nasipuri M., Ghose S. A New Approach to Keyphrase Extraction Using Neural Networks. *International Journal of Computer Science Issues*. 2010, vol. 7, no. 2, pp. 16–25.
32. Sheremetyeva S. An efficient patent keyword extractor as translation resource. *MT Summit XII: Third Workshop on Patent Translation*. Ottawa, 2009. Pp. 25–32.
33. Turney P.D. Learning Algorithms for Keyphrase Extraction. *Information Retrieval*. 2000, vol. 2, no. 4, pp. 303–336.
34. Wang J., Peng H., Hu J.-S. Automatic Keyphrases Extraction from Document Using Neural Network. *Advances in Machine Learning and Cybernetics*. Springer Berlin Heidelberg, 2006, pp. 633–641.
35. Wasserman S., Faust K.. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press, 1995. 857 p.

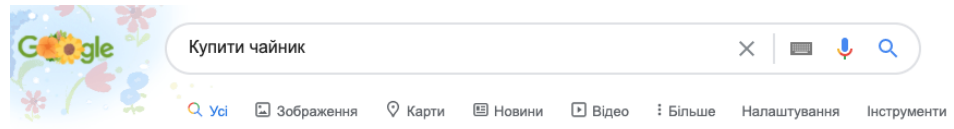
ЕЛЕКТРОНІ ДЖЕРЕЛА:

36. Ahrefs. Что такое SEO? URL: <https://ahrefs.com/blog/ru/chto-takoe-seo/> (Дата звернення: 04.04.2021)
37. Check PageRank. URL: <http://checkpagerank.net/index.php> (Дата звернення: 23.04.2021)
38. FlugerPlast. URL: flugerplast.com. (Дата звернення: 23.03.2021)
39. Gal sad. Блог. URL: <https://galsad.ua/blog/stati/doglad> (Дата звернення: 03.04.2021)
40. Google ads URL: <https://ads.google.com> (Дата звернення: 01.04.2021)
41. Grineva M. Effective Extraction of Thematically Grouped Key Terms From Text. Available at: <http://www.aaai.org/Papers/Symposia/Spring/2009/SS-09-08/SS09-08-010.pdf> (accessed: 30.04.2021).
42. Joomla. URL: <https://www.joomla.org/> (Дата звернення: 03.04.2021)
43. Just magic. URL: <https://just-magic.org> (Дата звернення: 15.04.2021)
44. Keycollector. URL: www.key-collector.ru (Дата звернення: 14.04.2021)
45. Keywordseverywhere. URL: keywordseverywhere.com (Дата звернення: 14.04.2021)
46. Krulwich B. Learning User Information Interests through Extraction of Semantically Significant Phrases. Available at: <http://www.aaai.org/Papers/Symposia/Spring/1996/SS-96-05/SS96-05-018.pdf> (accessed: 30.11.2014)
47. Outsourcing team. IT Blog URL: <https://outsourcing.team/uk/> (Дата звернення: 04.04.2021)
48. Outsourcing team. Що таке ключові слова? URL: <https://outsourcing.team/uk/blog/seo-prosuvannya/shho-take-klyuchovi-slova/> (Дата звернення: 04.04.2021)
49. Pymorphy. URL: <https://nlpub.ru/Pymorphy> (Дата звернення: 10.04.2021)
50. Rushanalytics. URL: <https://www.rush-analytics.ru> (Дата звернення: 15.04.2021)

- 51.Semrush. URL: semrush.com(Дата звернення: 14.04.2021)
- 52.SEO STUDIES & RESEARCH «90.63% of Content Gets No Traffic From Google. And How to Be in the Other 9.37% [New Research for 2020]» URL: <https://ahrefs.com/blog/search-traffic-study/> (Дата звернення: 04.04.2021)
- 53.Serpstat. URL: <https://uk.wikipedia.org/wiki/Serpstat> (Дата звернення: 04.04.2021)
- 54.Serpstat. URL: <https://serpstat.com/ru> (Дата звернення: 15.04.2021)
- 55.Serpstat. Нужно ли заполнять мета-тег keywords и как его использовать. URL: <https://serpstat.com/ru/blog/nuzhno-li-zapolnjat-meta-teg-keywords/> (Дата звернення: 07.04.2021)
- 56.Serpstat. Обновленная база GoogleUA: еще больше фраз для продвижения в Украине. URL: <https://serpstat.com/ru/blog/obnovlennaya-baza-google-ukraina-v-serpstat/> (Дата звернення: 01.04.2021)
- 57.Spywords. URL: <https://spywords.ru> (Дата звернення: 15.04.2021)
- 58.Summa. URL: <https://pypi.org/project/summa/> (Дата звернення: 10.04.2021)
- 59.Summarizer. URL: <https://pypi.org/project/summarizer/>(Дата звернення: 10.04.2021)
- 60.Uzun Y. Keyword Extraction Using Naive Bayes. Available at: http://www.cs.bilkent.edu.tr/~guvenir/courses/CS550/Workshop/Yasin_Uzun.pdf(accessed: 30.04.2021).
- 61.Webmaster. URL:<https://help.yandex.ru/webmaster/>(Дата звернення: 07.04.2021)
- 62.Wikipedia. KarenSpärckJonesURL: https://en.wikipedia.org/wiki/Karen_Sp%C3%A4rck_Jones(Дата звернення: 29.04.2021)
- 63.Yandex. Как правильно заполнять мета-теги URL: https://yandex.ru/q/question/computers/kak_pravilno_zapolniat_metategi_li_teg_bec772df/(Дата звернення: 07.04.2021)






64. ДистОсвіта. Поняття пошукової оптимізації та просування веб-сайтів
URL: <https://dystosvita.gnomio.com/mod/page/view.php?id=4389> (Дата звернення: 04.04.2021)
65. Київський національний університет імені Тараса Шевченка.
URL: www.univ.kiev.ua. (Дата звернення: 02.04.2021)
66. Урядовий кур'єр. Україномовні запити в SEO
URL: <https://ukurier.gov.ua/uk/articles/ukrayinomovni-zapiti-v-seo-ignoruvati-ne-mozhna-re/> (Дата звернення: 08.04.2021)
67. Усталов Д. Извлечение терминов из русскоязычных текстов при помощи графовых моделей. <http://koost.eveel.ru/science/CSEDays2012.pdf> (accessed: 30.04.2021).

Додаток 1



Приблизна кількість результатів: 2 140 000 (0,59 сек.)

Реклама · Купити чайник

 <p>Електрочай... 449,00 грн. ROZETKA</p>	 <p>Чайник-... 625,00 грн. thetea.ua</p>	 <p>Електрочай... 219,00 грн. ROZETKA</p>	 <p>Електрочай... 499,00 грн. ROZETKA</p>	 <p>Електрични... 11 599,00 г... Comfy</p>
---	--	--	---	--

Реклама · www.kofeeko.shop/чайники/на_плиту

Какой чайник купить - Чайники на плиту от 170 грн.

Продаем чайники от ведущих мировых брендов. Есть металлические, нержавеющие, со свистком. Разный дизайн, любая расцветка. Продукция сертифицирована.

Реклама · www.rozetka.com.ua/чайники/для_плиты 044 503 8080

Чайники для плиты - Интернет-магазин Rozetka

Товары для дома. Посуда для чая и кофе. Чайники для плиты. Заказывайте в...

epicentrk.ua > ... > Посуд для приготування

Чайники для плиты купити в Епіцентрі • Ціна в Україні

Чайники ✓ Купити в Епіцентрі ✨Ціна від 100 грн ⚡В наявності: 430 шт. 🚚Безкоштовний самовивіз ✓ Акції, знижки, доставка по Україні.

Скільки коштує Чайники в Епіцентрі?	▼
Які бренди купують в групі Чайники?	▼
Які недорогі товари в групі Чайники?	▼
▼ Показати більше	

www.foxtrot.com.ua > shop > elektrochayniki

Електрочайники - купити чайник електричний в Києві та ...

Купити електрочайник в Києві з будь-якого матеріалу дуже просто. Ви можете замовити пристрій будь-якої форми онлайн. В асортименті присутні ...

Які Електрочайники найдешевші?	▼
Які Електрочайники найбільш популярні в 2021 році?	▼
Які Електрочайники актуальні в 2021 році?	▼
Які Електрочайники належать до преміум сегменту?	▼


www.foxtrot.com.ua > shop > chainiki

Чайники – купити в інтернет-магазині < ФОКСТРОТ ...

В інтернет-магазині Фокстрот представлені чайники від перевірених виробників – RONDELL, KRAUFF, VINZER, LAMART, MAXMARK, GUSTO, KRAUFF, ...

Які Чайники найдешевші?	▼
-------------------------	---




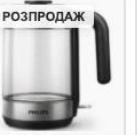

Додаток 2



Усі Зображення Покупки Відео Карти Більше Налаштування Інструменти

Приблизна кількість результатів: 896 000 (0,73 сек.)

Реклама · купить чайник philips

 Електрочай... 839,00 грн. ROZETKA	 Электрочай... 1 199,00 грн. ROZETKA	 Электрочай... 1 259,00 грн. ROZETKA	 РОЗПРОДАЖ Скляний... 1 258,98 гр... Philips UA	 Электрочай... 839,00 грн. ROZETKA
---	---	--	---	---

Реклама · www.philips.ua/
Мультиварка Philips - Доставимо безкоштовно - philips.ua
Купуйте мультиварку в офіційному інтернет-магазині Philips. Гарантія від виробника!
Доступні моделі · Функція "Мій рецепт" · Мультиварка-скороварка · +1 рік гарантії

Реклама · www.foxtrot.com.ua/ 0800 300 353
Купить Электрочайник - Philips. Официальная гарантия
Кешбэк на все товары. Множество распродаж и скидок на товары. Покупка в кредит и рассрочку

bt.rozetka.com.ua > ... > Электрочайники
Электрочайники Philips — ROZETKA — купить ...
ROZETKA © Электрические чайники Philips → Купить по выгодной цене ✓ Новинки от ...
Электрочайник PHILIPS Daily Collection HD4646/00 Белый.

eldorado.ua > Интернет-магазин > Электрочайники
Электрочайники Philips - купити електричний чайник Філіпс ...
Електрочайники Philips в інтернет-магазині Eldorado.ua ★Найнижчі ціни ✓ Знижки ✓ Акції
✓ Доставка ✓ Оплата частинами ✓ Замовляйте прямо ...

comfy.ua > ... > Дрібна техніка > Електричні чайники
Електричні чайники Philips купити за низькою ціною в Києві ...
Ціни на Електричні чайники Philips (актуально на березень 2021р.) Електричні чайники Philips, Ціна. Електричний чайник Philips Daily Collection HD4646/ ...

hotline.ua > elektrochajniki > Перекласти цю сторінку
Чайники Philips | сравнить цены и выгодно купить | Hotline
ЭЛЕКТРОЧАЙНИК ФИЛИПС от 520 грн! ✓СРАВНИТЬ цены и ВЫГОДНО купить с помощью Hotline. ✓ОБЗОРЫ, ВОПРОСЫ и ОТЗЫВЫ реальных ...

www.foxtrot.com.ua > ... > Электрочайники >
≡ Электрочайники PHILIPS – купити Електрочайники в ...
Купуй 【Електрочайники PHILIPS】 ЗАРАЗ та забирай уже сьогодні! >>>Інтернет-магазин ... HD9351/91, 1399 грн. Чайник PHILIPS HD9318/20, 799 грн ...

Додаток 3






Google

Електрочайник PHILIPS HD9351/91

Усі Зображення Покупки Відео Новини Більше Налаштування Інструменти

Приблизно кількість результатів: 37 200 (0,91 сек.)

Реклама - Електрочайник PHILIPS HD9351/91

 Чайник - Philips Daily... 1 399,02 грн. Philips UA	 Електрочайник PHILIPS... 1 399,00 грн. ROZETKA	 Електрочайник PHILIPS Viv... 1 699,00 грн. ROZETKA	 Електрический чайник Phili... 1 399,00 грн. Comfy	 Електрочайник PHILIPS Dai... 1 499,00 грн. ROZETKA
---	---	---	--	--

Реклама - www.comfy.ua/ - 0800 303 505
Електрический чайник Philips HD9351 купить по низкой цене
Халпй Кредит до 25 Месяцев или Покупай в Оплату Частями до 15 Месяцев! Широкий...

Реклама - www.foxtrot.com.ua/ - 0800 300 353
Philips Hd9351 91 - Всього 1159 грн - foxtrot.com.ua
Кешбэк на все товары. Заказывайте онлайн. Мы позаботимся о доставке.

bt.rozetka.com.ua > philips_hd9351_91
ROZETKA | Електрочайник PHILIPS HD9351/91. Цена ...
Електрочайник PHILIPS HD9351/91 купити на ROZETKA. Оперативная доставка →
Гарантия качества ☑ Лучшая цена \$
1 399,00 грн. · У наявності

hotline.ua > bt-elektrochajniki > Перекласти цю сторінку
цены на электрочайник Philips HD9351/91 ... - Hotline
Електрочайник Philips HD9351/91 ✓СРАВНИ предложения всех интернет-магазинов и ВЫБЕРИ самое выгодно! >HOTLINE знает, где ДЕШЕВЛЕ.
Индикатор уровня воды: есть Поворот на подставке на 360 град: есть
Объем, л: 1,7
★★★★☆ Оцінка: 4,2 · 3 відгуки · 1 119,00 грн. – 1 717,10 грн. · У наявності





epicentrk.ua > ... > Електрочайники > Philips >
Електрочайник Philips HD9351/91 Daily Collection ...
Електрочайник Philips HD9351/91 Daily Collection >>> Купити в Україні ✓ Епіцентр
Низька ціна та акції 🏪 Самовізіз та доставка по Україні 🚚 Відгуки ...
1 399,00 грн. · У наявності

www.philips.ua > ... > Чайник > Daily Collection Чайник >
Daily Collection Чайник HD9351/91 | Philips
Міцний чайник із безпечної, харчової нержавіючої сталі для тривалого та надійного щоденного використання. Компанія Philips втілює у цьому ...
★★★★☆ Оцінка: 4,6 · 24 відгуки · 1 399,02 грн. · У наявності

www.foxtrot.com.ua > ... > Електрочайники PHILIPS >
Електрочайник PHILIPS HD9351/91 - купити в інтернет ...
Електрочайник PHILIPS HD9351/91. купити за 1399 грн < ФОКСТРОТ > Інтернет-магазин

Электрочайник Philips
★★★★☆ 1 567 відгуків

Детали Відгуки Магазины

			
---	---	---	---

Реклама - Купити

1 399,02 грн. · Philips UA
Чайник - Philips Daily Collection - HD9351/91

1 399,00 грн. · ROZETKA
Електрочайник PHILIPS HD9351/91. Електрочайники

1 399,00 грн. · Foxyrot.ua
Електрочайник Philips Hd9351/91

1 399,00 грн. · ЕПІЦЕНТР
Електрочайник Philips HD9351/91 Daily Collection

1 399,00 грн. · МОЮ.UA
Електрочайник Philips HD9351/91

→ Порівняти ціни

Бренд: Philips
Матеріал: Нержавіюча сталь
Функція: 3 регулюванням температури
Об'єм: 1,7 літра

→ Докладніше

Відгуки користувачів

4,6

5	██████████
4	██████████
3	██████████
2	██████████
1	██████████

★★★★☆
1 567 відгуків

→ Більше відгуків

взуття для бігу



Реклама · <https://ua.factcool.com/> ▾ 094 711 1960

Взуття для бігу

Тільки у Factcool ви знайдете фірмовий одяг зі знижками до -70%.

<https://www.decathlon.ua> > ... > Біг ▾

Бігові кросівки – купити взуття для бігу в інтернет-магазині ...

Замовляйте на Decathlon.ua бігові кросівки і **взуття для бігу** ✓ Гарантія від 2х років ▶▶
Безкоштовна доставка від 990 грн по всій Україні.

<https://extremstyle.ua> > catalog > vzuttya-dlya-bigu ▾

Спортивне взуття для бігу купити | Магазин спортивного ...

Купити краще бігове **взуття** в інтернет-магазині ✓EXTEMSTYLE✓ Акційні товари, Висока ...
Кросівки для бігу Asics (1011A767) GEL-KAYANO 27 2020.

<https://extremstyle.ua> > Головна > Статті ▾

Як вибрати кросівки для бігу та які краще для залу, для ...

Для залу і нетривалих тренувань (1-2 години) вибирайте кросівки бренду Asics. Які кросівки найкраще підібрати для бігу? Біг взимку на вулиці. Вибирайте ...

<https://www.sportmaster.ua> > ... > Біг > Чоловіче взуття ▾

Чоловіче взуття для бігу – підкори дистанцію! - Спортмастер

SPORTMASTER ▶▶ Купити → Чоловіче **взуття для бігу** 【НИЗЬКА ЦІНА】 ✓ Доставка по Україні! 📞 0 800 505 707 ✓ Гарантія • АКЦІЇ • ЗНИЖКИ ...

Додаток 5

взуття для спорту

<https://www.sportmaster.ua> > cholovche_vzuttya14

▶ **Взуття спортивне — купити з доставкою в інтернет ...**
SPORTMASTER ★ **[Взуття для чоловіків]** купити за доступною ціною ✓ Швидка доставка по Україні → Величезний вибір **спортивного взуття** і ...
Чоловічі черевки · Чоловічі напівчеревки · Чоловічі кеди · Чоловічі аквашузи

<https://www.sportmaster.ua> > gnoche_vzuttya13

▶ **Жіноче спортивне взуття — купити з доставкою в ...**
SPORTMASTER ★ **[Взуття для жінок]** купити за доступною ціною ✓ Швидка доставка по Україні → Величезний вибір **спортивного взуття** від відомих ...

<https://temposhop.com.ua> > Спортивне взуття > Блог

Спортивне взуття та види спорту
Якісний одяг, який не сковує рухів, плюс правильне спортивне взуття - це ключові моменти для занять будь-якими видами **спорту**. Далі в цій статті ми ...

<https://soccerpoint.com.ua> > sportyvne-vzuttia

Спортивне взуття, купити взуття для спорту з доставкою по ...
Спортивне взуття. Заняття **спортом** рекомендують для зміцнення і підтримки здоров'я. Але деколи вибір **неправильного спортивного** екіпування може ...

Додаток 6

взуття

<https://www.eobuv.com.ua> ▾
eobuv.com.ua • Жіноче, чоловіче і дитяче взуття • Сумки і ...
Перевір пропозиції магазину eobuv.com.ua 🌟 Понад 450 брендів ✓ 50 000 моделей взуття, сумок та аксесуарів ✓ Безкоштовна доставка і повернення.
[Жіноче взуття](#) · [Чоловіче взуття](#) · [Дитяче взуття](#) · [Шльопанці і босоніжки](#)

<https://intertop.ua> ▾
INTERTOP.UA: купити взуття в Україні, каталог взуття 2021 ...
Онлайн магазин взуття та одягу ❤️ INTERTOP ❤️ - придбайте фірмове взуття, одяг та аксесуари за демократичними цінами. Доставка по Україні. ☎ (044) ...
[Жіноче взуття](#) · [Чоловіче взуття](#) · [Дитяче взуття](#) · [Чоловікам](#)

<https://kasta.ua> ▾ market ▾ obuv ▾
• Модне ВЗУТТЯ 2021 ► купити на Kasta за найкращими ...
Часті запитання про взуття. ? Чи вигідно купувати взуття в Чорну П'ятницю-2020? Напис "sale" завжди привертає увагу споживачів, ...

Чи вигідно купувати взуття в Чорну П'ятницю-2020? ▾

Яке взуття купити на Чорну П'ятницю-2020? ▾

Де недорого придбати взуття у Чорну П'ятницю? ▾

▾ Показати більше

<https://g-outlet.com> ▾ vzuttya ▾
Взуття в магазині ► G-outlet ◄ | Купити стильне і недороге ...
Модне зручне взуття для жінок і чоловіків в G-outlet ❤️. Інтернет магазин недорогого і

Догляд за кімнатними рослинами від ГалСад

Головна - Статті

24 березня 2020 12:00 // Статті



Всі кімнатні рослини для активного і здорового росту, наявності рясного листя, гарних квітів або навіть плодів, обов'язково потребують правильного догляду. При цьому важливо пам'ятати, що догляд потрібно змінювати в залежності від сезону. Також слід враховувати вид рослини, індивідуальні особливості, важливі складники щоденного догляду, періоди життєвої активності та спокою.

Більшість квітів або інших різновидів кімнатних рослин розвивається, розмножується, росте, квітне у весняний та літній період. Та в осені, взимку настає час відпочинку, коли зупиняються життєві процеси і рослина набирається сил для наступного сезону. Тож, як доглядати за кімнатними рослинами, що необхідно робити та в якій послідовності? Радимо скористатися важливими рекомендаціями експертів.

Особливості сезонного догляду за кімнатними рослинами


































Весною, літом рослинам потрібні кілька факторів, що забезпечують активний зріст, життєздатність, здоров'я:

- тепло;
- багато світла;
- волога (регулярний полив);
- внесення добрив для живлення;

- сонце (у трафікарді);

Додаток 8

Ключові слова купити плаття ?

#	<input type="checkbox"/>	Ключова фраза	Частотність ↓
1	<input type="checkbox"/>	купити плаття  ▶	5.4K
2	<input type="checkbox"/>	плаття купити  ▶	3.6K
3	<input type="checkbox"/>	купити плаття не дорого   ▶	1.6K
4	<input type="checkbox"/>	купити плаття недорого   ▶	1.6K
5	<input type="checkbox"/>	купити плаття <i>великих</i> розмірів   ▶	880
6	<input type="checkbox"/>	купити плаття вечірнє   ▶	880
7	<input type="checkbox"/>	купити молодіжне плаття великих розмірів   ▶	720
8	<input type="checkbox"/>	купить вишите плаття  ▶	720
9	<input type="checkbox"/>	купити плаття в <i>хмельницькому</i>   ▶	720
10	<input type="checkbox"/>	купити вишите плаття   ▶	720
11	<input type="checkbox"/>	купити плаття з мереживом  ▶	720
12	<input type="checkbox"/>	купити вечірнє плаття   ▶	590
13	<input type="checkbox"/>	купити нарядні плаття <i>великих</i> розмірів недорого  ▶	590
14	<input type="checkbox"/>	купити плаття <i>великих</i> розмірів евеліна   ▶	590
15	<input type="checkbox"/>	вечірні плаття <i>великих</i> розмірів купити недорого  ▶	590
16	<input type="checkbox"/>	випускні плаття 2020 купити  ▶	480
17	<input type="checkbox"/>	випускні плаття 2020 купить   ▶	480
18	<input type="checkbox"/>	купити літнє плаття   ▶	480
19	<input type="checkbox"/>	купити плаття <i>великого</i> розміру   ▶	480
20	<input type="checkbox"/>	купити плаття жіноче   ▶	480

Додаток 9

```
<html>
<head>
  <title>Семантичний аналіз тексту</title>
  <link rel="stylesheet" href="https://stackpath.bootstrapcdn.com/bootstrap/4.3.1/css/bootstrap.min.css"
  integrity="sha384-ggOyR0iXCbMQv3Xipma34MD+dH/1fQ784/j6cY/iJTQU0hcWr7x9JvoRxt2MZw1T"
  crossorigin="anonymous">
</head>
<body>
  <div id="content" class="container">
    <h2><a href="{% url 'index' %}">Семантичний аналіз тексту</a></h2>
    <div class="card">
      <div class="card-body">
        <h5 class="card-title"></h5>
        <p class="card-text">{% block content2 %}{% endblock %}</p>
      </div>
    </div>
  </div>
</body>
</html>
```

Додаток 10

```
<div class="m-2"><h4></h4></div>
▼ <form action="/index/" method="post">
    <label for="enter_text">Текст для аналізу: </label>
    <input id="entered_text" type="text" name="entered_text" value="{{ enter_text }}">
    <input type="submit" value="OK">
</form>
▼ <div>
    <h4>Суммаризація введеного тексту:</h4><br>
    <p>{{summarized_text}}</p>
</div>
▼ <div>
    <h4>Можливе найважливіше речення тексту:</h4><br>
    <p>{{emphasized_text}}</p>
</div>
▼ <div>
    <h4>Ключові слова:</h4><br>
    <p style="white-space: pre-line;">{{keywords}}</p>
</div>
▼ <div>
    <h4>Аналіз слів тексту за частинами мови та частотою слововживань:</h4><br>
    <p style="white-space: pre-line;">{{analyzed_words_text}}</p>
</div>
```

Додаток 11

```
import os
import sys

if __name__ == '__main__':
    os.environ.setdefault('DJANGO_SETTINGS_MODULE', 'semantic-site.settings')
    try:
        from django.core.management import execute_from_command_line
    except ImportError as exc:
        raise ImportError(
            "Couldn't import Django. Are you sure it's installed and "
            "available on your PYTHONPATH environment variable? Did you "
            "forget to activate a virtual environment?"
        ) from exc
```

```
1  django-extensions
2  django-extensions
3  pymorphy2
4  pymorphy2-dicts-uk
5  summarizer
6  summa
7  gunicorn
8  django-heroku
9  whitenoise
10
11
```



```

'чей', 'чень', 'через', 'четвертий', 'чи', 'чий', 'чийого', 'чийогось', 'чийому', 'чийомусь',
'чийсь', 'чик', 'чим', 'чимось', 'чимсь', 'чир', 'численна', 'численне', 'численний', 'численным',
'численними', 'численних', 'численні', 'чию', 'чиюсь', 'чия', 'чиясь', 'чие', 'чиєму', 'чиємусь',
'чиесь', 'чиею', 'чиеюсь', 'чиеї', 'чиеїсь', 'чії', 'чіїть', 'чіїсь', 'чіїм', 'чіїми', 'чіїмись',
'чіїмось', 'чіїмсь', 'чіїсь', 'чіїх', 'чіїхось', 'чіїхсь', 'чіїхсь', 'чля', 'чого', 'чогось', 'чом', 'чому',
'чомусь', 'чон', 'чоп', 'чортзна', 'чос', 'чотири', 'чотирнадцятий', 'чотирнадцять', 'чу', 'чум',
'чур', 'чш', 'чїм', 'чїмось', 'чїмсь', 'чїт', 'ш', 'ша', 'шаг', 'шал', 'шам', 'шво', 'шед', 'шен',
'шиз', 'шир', 'шляхом', 'шостий', 'шістнадцятий', 'шістнадцять', 'шість', 'щ', 'ще', 'щем', 'щеп',
'щип', 'щир', 'що', 'щоб', 'щоби', 'щодо', 'щойно', 'щоправда', 'щось', 'щі', 'ь', 'ю', 'юз',
'юн', 'юнь', 'юс', 'ют', 'юхт', 'я', 'яв', 'яд', 'яз', 'язь', 'як', 'яка', 'якась', 'якби', 'яке',
'якесь', 'який', 'якийсь', 'яким', 'якими', 'якимись', 'якимось', 'якимсь', 'яких', 'якихось',
'якихсь', 'якого', 'якогось', 'якому', 'якомусь', 'якось', 'якою', 'якоюсь', 'якої', 'якоїсь',
'якраз', 'яку', 'якусь', 'якщо', 'які', 'якій', 'якійсь', 'якім', 'якімсь', 'якісь', 'ял', 'ям',
'ян', 'янь', 'яо', 'яп', 'яп', 'ярл', 'ясь', 'ять', 'є', 'єр', 'єси', 'ї', 'їбн', 'їд', 'їз', 'їз-за',
'їз-під', 'їззаду', 'їзм', 'їзсередини', 'їк', 'їкс', 'їкт', "їм'я", 'їмовірно', 'їнакша',
'їнакше', 'їнакший', 'їнакшим', 'їнакшими', 'їнакших', 'їнакшого', 'їнакшому', 'їнакшою',
'їнакшої', 'їнакшу', 'їнакші', 'їнакшій', 'їнакшім', 'їнколи', 'їноді', 'їнша', 'їнше', 'їнший',
'їншим', 'їншими', 'їнших', 'їншого', 'їншому', 'їншою', 'їншої', 'їншу', 'їнші', 'їншій',
'їншім', 'їнь', 'їч', 'їще', 'ї', 'їд', 'їй', 'їм', 'їх', 'їхнього', 'їхньому', 'їхньою',
'їхньої', 'їхню', 'їхня', 'їхне', 'їхні', 'їхній', 'їхнім', 'їхніх', 'її', 'її'
```

```

summarized_text = str(summarize("Python", text, 3))
print(texsum.summarize(text))
emphasized_text = texsum.summarize(text)
print(keysum.keywords(text))
s = keysum.keywords(text)
li = s.split()
keywords = '\n'.join(li) #str(keysum.keywords(text))
text = text.split()
endlist = []
for val in text:
    p = analyzer.parse(val)[0]
    endlist.append(p.normal_form)
m_dict = {i:endlist.count(i) for i in endlist}
for w in sorted(m_dict, key=m_dict.get, reverse=True):
    if analyzer.parse(w)[0].tag.POS != "CONJ" and w not in stopwords:
        analyzed_words_text += w + ' ' + str(analyzer.parse(w)[0].tag.POS) + ' ' + str(m_dict[w])
        + '\n'
}
context = {
    'username':username,
    'summarized_text':summarized_text,
    'emphasized_text':emphasized_text,
    'keywords':keywords,
    'analyzed_words_text':analyzed_words_text
}
return render(request, 'index.html', context)

```

Додаток 15

```
"""
Django settings for semantic-site project.
For more information on this file, see
https://docs.djangoproject.com/
"""

import os
import django_heroku

# Build paths inside the project like this: os.path.join(BASE_DIR, ...)
BASE_DIR = os.path.dirname(os.path.dirname(os.path.abspath(__file__)))
PROJECT_ROOT = os.path.dirname(os.path.abspath(__file__))
STATIC_URL = '/static/'
STATIC_ROOT = os.path.join(PROJECT_ROOT, 'static')

# Quick-start development settings - unsuitable for production
# See https://docs.djangoproject.com/en/2.1/howto/deployment/checklist/

# SECURITY WARNING: keep the secret key used in production secret!
SECRET_KEY = '^l)7d*%h&db4uft@dk%h-w&nup#pu%)a!d)c7jwgoixo5_hm0$'

# SECURITY WARNING: don't run with debug turned on in production!
DEBUG = True

ALLOWED_HOSTS = []

# Application definition
INSTALLED_APPS = [
    'django.contrib.admin',
    'django.contrib.auth',
    'django.contrib.contenttypes',
    'django.contrib.sessions',
    'django.contrib.messages',
    'django.contrib.staticfiles',
    'rest_framework',
    'semantic',
    'pymorphy2',
    'summarizer',
    'summa',
]
```

```

▼ MIDDLEWARE = [
    'django.middleware.security.SecurityMiddleware',
    'django.contrib.sessions.middleware.SessionMiddleware',
    'django.middleware.common.CommonMiddleware',
    'django.middleware.csrf.CsrfViewMiddleware',
    'django.contrib.auth.middleware.AuthenticationMiddleware',
    'django.contrib.messages.middleware.MessageMiddleware',
    'django.middleware.clickjacking.XFrameOptionsMiddleware',
]

ROOT_URLCONF = 'semantic-site.urls'

▼ TEMPLATES = [
    ▼ {
        'BACKEND': 'django.template.backends.django.DjangoTemplates',
        'DIRS': ['templates'],
        'APP_DIRS': True,
        ▼ 'OPTIONS': {
            ▼ 'context_processors': [
                'django.template.context_processors.debug',
                'django.template.context_processors.request',
                'django.contrib.auth.context_processors.auth',
                'django.contrib.messages.context_processors.messages',
            ],
        },
    },
]

WSGI_APPLICATION = 'semantic-site.wsgi.application'

# Database
# https://docs.djangoproject.com/en/2.1/ref/settings/#databases

▼ DATABASES = {
    ▼ 'default': {
        'ENGINE': 'django.db.backends.sqlite3',
        'NAME': os.path.join(BASE_DIR, 'db.sqlite3'),
    }
}

```

Додаток 16

Python (найчастіше вживане прочитання — «**Пайтон**», запозичено назву з британського шоу Монті Пайтон) — інтерпретована об'єктно-орієнтована мова програмування високого рівня зі строгою динамічною типізацією. Розроблена в 1990 році Гвідо ван Россумом. Структури даних високого рівня разом із динамічною семантикою та динамічним зв'язуванням роблять її привабливою для швидкої розробки програм, а також як засіб поєднання наявних компонентів. Python підтримує модулі та пакети модулів, що сприяє модульності та повторному використанню коду. Інтерпретатор Python та стандартні бібліотеки доступні як у скомпільованій, так і у вихідній формі на всіх основних платформах. В мові програмування Python підтримується кілька парадигм програмування, зокрема: об'єктно-орієнтована, процедурна, функціональна та аспектно-орієнтована.

