

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧОРНОМОРСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ПЕТРА МОГИЛИ

ЛАБАРТКАВА СЕРГІЙ ПАВЛОВИЧ

УДК 004.51

НОВИННО-РОЗВАЖАЛЬНИЙ ВЕБПОРТАЛ ДЕРЖАВНИХ ЗМІ

Галузь знань 12 «Інформаційні технології» за спеціальністю

122 «Комп'ютерні науки»

122 - БКР.А - 401.1710116

Автореферат

бакалаврської кваліфікаційної роботи на здобуття освітньої кваліфікації

«бакалавр з комп'ютерних наук»

Миколаїв – 2021

Бакалаврська кваліфікаційна робота є рукопис.

Робота виконана в Чорноморському національному університеті імені Петра Могили Міністерства освіти і науки України на кафедрі інтелектуальних інформаційних систем

Науковий керівник:

к.т.н., доцент (б.в.з.)

Г. В. Горбань

Рецензент:

професор, доктор технічних наук

М.Т. Фісун

Захист відбудеться «_25_» червня 2021 р. о 9⁰⁰ год. на засіданні екзаменаційної комісії у Чорноморському національному університеті імені Петра Могили за адресою: 54003, м. Миколаїв, вул. 68-ми Десантників, 10.

З бакалаврською кваліфікаційною роботою можна ознайомитися в бібліотеці Чорноморського національного університету імені Петра Могили за адресою: 54003, м. Миколаїв, вул. 68-ми Десантників, 10.

Автореферат представлений «____» червня 2021 р.

Секретар

екзаменаційної комісії,

викладач кафедри ІІС

М.О. Таранов

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Тривалий час новинні портали були поза увагою вітчизняних і закордонних наукових досліджень із медіакомунікацій, журналістикознавства, соціальних комунікацій, видавничої справи. Останніми роками зростає інтерес науковців до детального вивчення мережових медіа, провідних інформаційних та вебтехнологій, стратегічного планування роботи й розвитку інтернет-ресурсів, їх комунікативного потенціалу, що зумовлено масовим переходом аудиторії у всесвітню мережу, динамічним ростом довіри до інтернет-медіа, конвергентних медіа процесів тощо.

Метою бакалаврської кваліфікаційної роботи є дослідження задачі створення новинно-розважального вебпорталу державних ЗМІ, автоматичних засобів для заповнення контенту та аналітичної складової для вивчення аудиторії.

Практичне значення отриманих результатів. Під час виконання бакалаврської кваліфікаційної роботи було розроблено вебпортал для того, щоб ділитися новинами.

Структура кваліфікаційної роботи. Пояснювальна записка до бакалаврської кваліфікаційної складається із вступу, 4 розділів, висновків. Загальний обсяг роботи складає 62 сторінок, 32 рисунків, 2 таблиці та 30 посилань на літературні джерела.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У наш час новини у першу чергу ми дізнаємося з інтернету. Читаємо різноманітні джерела від різних авторів. Тому щоб не відставати від тенденцій треба намагатися охоплювати аудиторію онлайн. Для цього ЗМІ повинні створити всі умови для користувача, аби залишатися на перших сходинках серед конкурентів. Зараз інтернет схожий з іншими медіа-середовищами, такими як радіо, телебачення та преса. Люди можуть дізнаватися усі актуальні новини нашого часу завдяки вебсайту. Деякі з таких засобів масової інформації зареєстровані як звичайні ЗМІ, але це вже не так популярно, тому це швидше виняток. Інтернет видання керуються у своїх роботах принципами журналістики, так само як і друковані ЗМІ.

Багато друкованих журналів та газет мають свої представництва в інтернеті. Вони викладають там матеріали своїх випусків, а буває навіть випускають самостійне інтернет-видання. Ви можете знайти будь-яку інформацію — є новинні, літературні, науково-популярні, дитячі, жіночі сайти тощо. Тому за жанром інтернет видання будуть схожими.

Останні роки спостерігається падіння тиражів газет. Це відбувається саме через зростання кількості відвідувань інтернет-ЗМІ. Процес пов'язаний з тим, що інтернет стає все доступнішим і швидшим, а також свідчить про те, що люди дедалі більше звертаються за новинами до інтернет-ЗМІ.

Перший розділ.

Тривалий час новинні портали були поза увагою вітчизняних і закордонних наукових досліджень із медіакомунікацій, журналістикознавства, соціальних комунікацій, видавничої справи. Останніми роками, як стверджує дослідник новинних порталів О. В. Ситник, зростає інтерес науковців до детального вивчення мережевих медіа, провідних інформаційних та вебтехнологій, стратегічного планування роботи й розвитку інтернет-ресурсів, їх комунікативного потенціалу, що зумовлено масовим переходом аудиторії у всесвітню мережу, динамічним ростом довіри до інтернет-медіа, конвергентних медіа процесів тощо. У зв'язку з відсутністю чіткої стандартизації технічних, технологічних і контентних складників

мережевих ресурсів мають місце їх різноманітні технологічні та змістові модифікації, гібридизація медій, які абсорбують ті чи інші ознаки, характерні для інших типів інтернет-ресурсів. Таким чином, виникають нові, які складно однозначно долучити до певного типу.

Як ми й говорили раніше, інтернет-портал це багатофункціональний сайт. Такі сайти можуть використовуватися під найрізноманітніші потреби, починаючи з інформування користувачів та продажу реклами на ресурсі, і закінчуючи налаштуванням комунікації між відділами компанії. Звичайно ж, в залежності від обраної тематики буде залежати прибуток, його ефективність, та кількість сил, які потрібно витратити на його розробку.

Розробка інтернет-порталу у 2021 році не відстає по перспективності від створення сайтів інтернет-магазинів і корпоративних ресурсів для офлайн компаній. Зважаючи на карантинні обмеження жителі міст та населених пунктів все частіше користуються інтернетом для того, щоб почерпнути цікаву для себе інформацію або знайти відповідь на питання, яке відвідало їх всього кілька хвилин тому. На допомогу таким користувачам мережі інтернет і приходять інтернет-портали.

Популярність розробки інтернет-порталів полягає в тому, що на відміну від звичайних комерційних сайтів, новинні вебпортали заробляють продаючи рекламу. З огляду на дану особливість ресурсу, принцип його створення дещо відрізняється від звичайних сайтів та інтернет-магазинів.

Створення новинного порталу вирішує кілька завдань:

- оперативність оновлення інформації. На відміну від друкованих та аудіовізуальних ЗМІ, розробка новинного порталу дозволяє відстежувати і публікувати новини в режимі реального часу. Таким чином, Вам доступно щохвилине висвітлення будь-яких останніх подій і зв'язок з аудиторією;
- скорочення витрат. Очевидною перевагою і однією з головних причин зробити новинний сайт є те, що не потрібно витрачатися на друк та витратні матеріали. У перспективі ціна створення новинного порталу і його підтримку Вам обійдеться дешевше, ніж запуск оффлайн проекту;

- широке охоплення аудиторії. Створити новинний сайт, значить покрити не тільки цільові запити окремих верств населення, але і окремого регіону, країни або навіть світу. В межах одного ресурсу можна сегментувати аудиторію і надати доступ до новин за інтересами. А інтеграція з соцмережами розширює сферу поширення інформації;
- аналіз громадської думки. Розробка сайту для ЗМІ надає можливість аналізу не тільки поведінки відвідувача на сайті, але і його думки щодо опублікованих матеріалів. Читач може легко дати свої коментарі з будь-якої теми. А розміщення опитувань допомагає збирати думки аудиторії безпосередньо.

Другий розділ.

Формалізуємо нашу задачу. Нехай в нас множина новин $U = \{u_1, u_2, \dots, u_n\}$ та множина користувачів $P = \{p_1, p_2, \dots, p_m\}$. Таким чином ми можемо створити матрицю відношень $R = (r_{ij})$ розмірності $n \times m$, де $i \in 1 \dots n$, $j \in 1 \dots m$.

Якщо відношення існує значення $r_{ij} = 1$, інакше $r_{ij} = 0$.

Кожна новина має множину векторів $P^* = \{p_{i1}, p_{i2}, \dots, p_{iN}\}$, де кожний вектор представляє собою користувачів, які переглянули новину, а p_i - окремий користувач.

Вхідні дані:

$U = \{u_1, u_2, \dots, u_n\}$ – множина новин;

$P = \{p_1, p_2, \dots, p_m\}$ – множина користувачів;

$P^* = \{p_{i1}, p_{i2}, \dots, p_{iN}\}$ – множина користувачів, які переглянули новину.

Треба отримати:

Для даної новини u_i знайти N-мірний вектор виду $\{p_{i1}, p_{i2}, \dots, p_{iN}\}$, де p_i – окремий користувач таким чином, щоб ці дані відповідали дійсності.

Можна розробити різні статистичні системи для відображення поточних даних та обрати що саме буде зібрано для аналітики користувачів. Можна обрати дані, що стосуються лише геопозиції або звідки перейшли до нашого джерела, а стосовно авторизованих користувачів ми можемо дізнатися навіть стать та вік людини.

Статистична модель зазвичай визначається за допомогою математичних рівнянь, які відносяться до однієї/кількох випадкових величин і інших не випадкових змінних. Таким чином, «модель є формальним поданням теорії».

В цілому, статистичні моделі є частиною фундаменту статистичного висновування.

Значно збільшене відвідування вебсайтів вимагає систематичної служби підтримки для оновлення вмісту або реклами його в Інтернеті, що вимагає значного часу, витраченого кваліфікованим персоналом.

Ефективність таких дій проявляється зі значною затримкою часу. Прогнозування динаміки покращення відвідуваності після таких системних дій дозволило б раціонально використовувати службу підтримки людських ресурсів. У цьому розділі пропонується модель відвідуваності вебсайтів, побудована на основі системи звичайних диференціальних рівнянь Моно. Для його виявлення пропонуються методи, які залежать від засобів, що використовуються службою підтримки для забезпечення зростання відвідуваності. У той же час розрізняють короткочасне та довгострокове збільшення відвідуваності. Особливості запропонованих методів враховані в роботах та в ході подальшої презентації.

Спрощуючи аналіз аналізу розроблених вебсайтів, системне оновлення вмісту, яке проводиться епізодично. Побудова моделей ймовірності поширення з визначенням результатів та вхідних параметрів. Змінні стани системи вибирають 1x загальну доступність сторінки вебсайту WS, як фактор її представлення в рейтингах пошукових систем, а також наявність 2x результатів вебсайту, як засіб поширення цільової інформації.

В якості вхідних параметрів моделі ми вибираємо функції якості вмісту вебсайту. Значення цих функцій прирівнюється до відвідуваності цих тем. Динаміка відвідування тем залежить від інтенсивності їх поповнення. Для спрощення

моделювання таких відвідувань вебсайтів, для яких оновлення виконується епізодично, динаміка трафіку описується за допомогою набору дискретних станів.

У таких послідовностях ймовірність переходу від високого значення під час помітного поліпшення якості вмісту до вимірюваного та низького, тоді як якість вмісту може періодично зростати. Зі збільшенням якості вмісту спостерігається нелінійність у взаємозв'язку між функціями якості та загальними та ефективними оцінками можливостей. Пропонується моделювати цю нелінійність шляхом зниження якості контенту в попередні моменти часу, а також трансформації потужності значущих функцій якості контенту. Природно припустити, що інтенсивність дії підвищує якість вмісту, оцінюючи зменшення з часом. Найкращий такий ефект моделюється зовнішньою функцією з від'ємним показником.

Незважаючи на тиждень оцінки можливостей та факторів, що впливають на реальні вебсайти, їх динаміка визначається значними порушеннями. Це послаблює точність побудованих моделей та узагальнює прогнозування функцій якості вмісту. Тому ми перейдемо до згладжування певних показників. Тут використовується множинна колекція методів ковзного середнього з мінімальним, що, на відміну від інших методів, значно усуває аномальні викиди, а не лише їх приватну колекцію.

Демонстрація динаміки ймовірності. Для розрахунків використовуються триточковий шаблон ковзного середнього, що еквівалентно простому середньому для трьох судинних точок. Згаданий шаблон використовується тричі для одного набору даних. Аналіз згладженої ймовірності дозволяє розпочати її на інтервалі відновленої сталості та інтервалі основних натуралістів. З практичної точки зору, райони значного природокористування представляють найбільший інтерес. Моделювання оціночного бала лише для цих часових інтервалів може значно зменшити складність моделі.

Проблема полягає у виділенні такого інтервалу для кількох початкових значень. За допомогою цієї підтримки формальну артерію використовують для підвищення похідної ймовірності активного мінімального значення.

Величина мінімального значення підбирається для кожного вебсайту його дослідником індивідуально. Подана згладжена відвідуваність, яка за критерієм,

формально розбита на 4 інтервали приросту відвідуваності, початок яких відображено пунктирними лініями. Кінець природного інтервалу фіксується, коли ймовірність зменшується нижче початкового рівня у разі вичерпання через поточне оновлення вмісту або на початку етапу зростання після попереднього випуску через нове оновлення до кінця попереднього.

В даному розділі було проаналізовано існуючі методи заповнення контенту та шифрування даних для нашої бази даних, зроблений детальний аналіз кожного з методів відносно поставленої задачі. В ході аналізу було виявлено, що системи, що засновані на знаннях найкраще всього відповідають поставленій задачі. Спираючись на інформацію викладену в даному розділі найкраще усього вирішить завдання та проблеми система, що заснована на знаннях з гібридного шифрування. Сучасні бази даних мають дуже великі об'єми, що досягають сотень та тисяч гігабайтів та збільшуються з кожним днем. Для використання даних потрібні ефективні масштабовані алгоритми, що дозволяють вирішити задачу за прийнятний час, тому для формування бази знань продукційних правил треба використовувати швидкі та ефективні алгоритми. Проаналізувавши алгоритми було обрано алгоритм для заповнення бази знань та майбутня стратегія для створення контенту, що буде цікавим та актуальним для наших користувачів.

Третій розділ.

Основою технологією була обрана система керування вмістом. Виходячи з цього, ми будемо легко вирішувати наші проблеми та майбутні оновлення та вдосконалення системи.

Серед переваг такого рішення можна відмітити:

- функціональність систем управління (основними характеристиками CMS є зручність використання, налаштування, простота відвідування, навігації; незначні зміни та редизайн без централізації вебсайту);
- універсальність системи управління (використовувати дані характеристик CMS для створення вебсайту будь-якого рівня складності, будь то візитна картка чи великий портал);

- адміністрування вебсайтів (дозволяють призначати різні групи користувачів різного рівня доступу, якісний захист системи від злочинів);
- інші характеристики CMS (наприклад, як дозволити перехід на іншу платформу без втрати інформації, збереження резервних копій, хороший відгук бази даних);
- своєчасна реакція команди розробників на запити клієнтів. Це важливо, оскільки важлива не лише інсталяція системи управління вмістом, а також додаткова інформація та технічна підтримка;
- вебсайт має панель управління, яка є лише частиною всіх програм, але достатньою для управління нею.

Ці переваги в комбінації надають найбільшу простоту та гнучкість при інтегруванні системи до будь-якого новинного вебпорталу.

Основною мовою розробки було обрано PHP7 через її орієнтованість до веб-розробки та тим, що більшість CMS написані саме нею. Серед CMS найбільше вирізняються Drupal та WordPress.

Задача вебпорталу актуальна в такій мірі, що майже всі великі корпорації мають свої електронні сторінки. Сфери для сайтів різнопланові: новини, ресторани, товари, фільми, музика. Розглянемо найбільш цікаві рекомендаційні технології, які існують на даний момент.

Як видно з порівняння, Drupal та WordPress мають приблизно однакові параметри. тому можна обирати будь-який. Коли мова заходить про надійність і Drupal, і WordPress стоять на рівні. Системи керування вмістом забезпечують надійну безпеку на декількох рівнях (включаючи паролі, аутентифікацію, SQL-ін'єкції, захист від підробки запитів, скрипти сайту, використання капчі і т. д.).

Системи керування вмістом WordPress було обрано за її неперевершені можливості, а саме:

- можливість перетворити свій вебсайт на мережу сайтів. Кожний з цих сайтів буде виконувати окремі визначені функції, що дуже важливо з точки зору UX/UI;
- ведення форуму, як чудовий спосіб будувати ком'юніті вашого бренду;
- ваш вебсайт легко може стати справжнім інтернет-магазином;
- може стати для вас інструментом для керування проектами, створюючи календарі та відкладені статті;
- ідеальне рішення для новачків, які хочуть запустити свої перші проекти.

Вагому роль також зіграло те, що за результатами тестів WordPress показала найкращу продуктивність серед аналогів. Вона має прості компоненти та використовує композицію замість наслідування.

Прийнято рішення програмну реалізацію розробляти базуючись саме на архітектурі WordPress.

Розділ з охорони праці.

У спеціальному розділі дипломної роботи «Охорона праці», було:

- розглянуто робоче місце інженера-програміста, яке є об'єктом дослідження з охорони праці;
- був проведений аналіз виробничих факторів, таких як температура повітря, освітлення, шум та інше на робочому місці;
- у ході роботи над розділом були розглянуті детальні питання щодо рівня шуму. Розрахунок у свою чергу показав, що загальний рівень шуму є помірним.

Було проаналізовано клімат внутрішнього середовища приміщень, який визначається діючими на організм людини поєднаннями температури, вологості і швидкості руху повітря.

Було проаналізовано техніку безпеки, бо більшість нещасних випадків спричиняє неправильна поведінка. Для того, щоб назавжди зменшити кількість нещасних випадків, усі співробітники мають проходити інструктаж і навчання поведінці, за якої дотримуються правила безпеки. Менеджери несуть відповідальність за забезпечення того, щоб співробітники в їхніх організаційних підрозділах мали належну підготовку в області безпеки і заходів запобігання нещасним випадкам.

Були розроблені розрахунки шуму щоб попередити негативний вплив, шум не лише погіршує самопочуття людини і знижує продуктивність праці на 10-15%, але нерідко призводить до професійних захворювань. За результатами розрахунків рівень шуму знаходиться в межах норми та дозволяє виконувати свої прямі робочі обов'язки. Рівень шуму дорівнює 52.7 дБ, що є нормою для виробничих та офісних приміщень згідно з ДСН 3.3.6.037-99 Санітарні норми виробничого шуму, ультразвуку та інфразвуку.

ЗАГАЛЬНІ ВИСНОВКИ

У даній роботі були коротко розглянуто основні рішення, які використовуються в новинно-розважальних вебпорталах. Було розглянуто різні моделі відвідуваності, які використовують для контролю вмістом. Кожен підхід має свої переваги і недоліки. Наприклад, при використанні моделі відвідуваності із дискретною функцією якості контенту відвідуваності переходять із високих значень на момент помітного покращення якості контенту до помірних так низьких, причому якість контенту може періодично посилюватися та при посиленні якості контенту спостерігається нелінійність у співвідношеннях між функціями якості та загальними і результуючими відвідуваностями. Але для їх хорошої роботи потрібні відповідні вхідні дані про статистичні дані щодо користувачів. Чим більше є даних, тим більш точну систему керуванням вмістом можна розробити, використовуючи різні моделі.

В ході виконання роботи було проаналізовано основні системи керування вмістом. Було описано особливості, переваги та недоліки кожної з них та виконано порівняльний аналіз існуючих систем. Також було зроблено огляд допоміжних засобів та методів для збільшення ефективності застосування.

В результаті виконання роботи були досягнуті поставлені цілі та реалізованого вебпорталу було зроблено висновок, що система поводить себе коректно, новини – релевантні. Тому є сенс використовувати поточну реалізацію в контексті інтернет-новин, щоб просувати якомога більше релевантного та цікавого контенту користувачам. В подальшому можна зробити більш глибоке та детальне дослідження статистичних даних, ефективніший способів їх аналітиці, а саме – гібридних аналізів.

АНОТАЦІЯ

Лабарткава Сергій Павлович. Новинно-розважальний вебпортал державних ЗМІ. – На правах рукопису.

Бакалаврська кваліфікаційна робота на здобуття освітньої кваліфікації «бакалавр з комп'ютерних наук» в галузі знань 12 «Інформаційні технології» за спеціальністю 122 «Комп'ютерні науки».

Чорноморський національний університет імені Петра Могили, Миколаїв.

Об'єкт дослідження – новинно-розважальний вебпортал державних ЗМІ.

Предмет дослідження – математичні моделі генерації контенту та обробку даних користувачів.

Метою роботи є дослідження задачі створення новинно-розважального вебпорталу державних ЗМІ, автоматичних засобів для заповнення контенту та аналітичної складової для вивчення аудиторії. Результатом є існуючий сайт для новин.

Робота складається з фахового розділу і спеціальної частини з охорони праці. Пояснювальна записка складається зі вступу, трьох розділів, висновків.

У першому розділі наведено аналіз сучасного стану проблеми, огляд існуючих аналогів, виокремлення їх переваг та недоліків.

У другому розділі розглянуто існуючі методи заповнення контенту та шифрування даних. Обрано стратегію щодо майбутньої інтеграції контенту на вебпортал.

У третьому розділі описано процес реалізації новинно-розважального вебпорталу з використанням WordPress.

В результаті виконання роботи були досягнуті поставлені цілі та на основі реалізованого вебпорталу було зроблено висновок, що система поводить себе коректно.

Новини, аудиторія, контент, вебпортал, система керування вмістом, інформація.

ABSTRACT

Labartkava Sergey Pavlovich. News and entertainment web portal of state media. – On the rights of the manuscript.

Bachelor's qualification work for the educational qualification "Bachelor of Computer Science" in the field of knowledge 12 "Information Technology" in the specialty 122 "Computer Science".

Petro Mohyla Black Sea National University, Mykolaiv.

The object of the research – a news and entertainment web portal of state media.

The subject of research – mathematical models of content generation and user data processing.

The aim of the work is to study the problem of creating news and entertainment webportal of state media, automatic ways for filling content and analytical components for the study of the audience. The result is an existing news web-site.

The section 1 presents an analysis of the current state of the problem, an overview of existing analogues, highlighting their advantages and disadvantages.

In section 2 discusses existing methods for filling content and encrypting data. A strategy for the future integration of content on the web portal also has been developed.

The section 3 describes the process of creating the news and entertainment web portal using WordPress.

As a result of the work, the set goals were achieved and on the basis of the implemented web portal it was concluded that the system behaves correctly.

News, audience, content, webportal, content management system, information.