

В. В. Литвин¹, В. А. Висоцька¹, В. В. Кучковський¹, С. Ю. Дуткевич¹, О. М. Наум²¹⁻²Національний університет “Львівська політехніка”, кафедра інформаційних систем та мереж,²Дрогобицький державний педагогічний університет імені Івана Франка

МЕТОД ІНТЕГРАЦІЇ ТА УПРАВЛІННЯ КОНТЕНТОМ МЕРЕЖІ ІНФОРМАЦІЙНИХ РЕСУРСІВ ТУРИЗМУ ЗГІДНО З ПОТРЕБАМИ КІНЦЕВОГО КОРИСТУВАЧА

© Литвин В. В., Висоцька В. А., Кучковський В. В., Дуткевич С. Ю., Наум О. М., 2018

У роботі описано метод інтегрованого опрацювання неоднорідних інформаційних ресурсів Web-систем для їх інтеграції та подальшого управління, який ґрунтується на моделі подання даних як узгодженого поєднання значень, правил їх зображення, правил інтерпретації та структури. Сформульовано новий підхід застосування та впровадження бізнес-процесів для побудови таких Web-систем. Розроблено методи та програмні засоби опрацювання контенту та інформаційного ресурсу як етап життєвого циклу контенту в Web-системах.

Ключові слова: інформаційний ресурс, контент-аналіз, контент-моніторинг, контентний пошук, інтеграція даних, бізнес-процес, система управління контентом, життєвий цикл контенту, Інтернет-туризм.

The paper of the method of integrated processing of heterogeneous information resources Web-systems for their integration and the subsequent management is described. A new approach to application and implementation of business processes for the construction of these Web-systems is formulated. The methods and software tools of content and information resource processing as the content life cycle stage in Web-systems are developed.

Key words: information resources, content analysis, content monitoring, content search, data integration, business-process, content management system, content lifecycle, Internet tourism.

Вступ. Загальна постановка проблеми

Сьогодні існує достатньо багато інформаційних ресурсів, які дають найрізноманітніший контент [1]. Проте значна частина з них є досить незручною у плані юзабіліті, інша – недостатньо інформаційно наповнена і не завжди вчасно оновлена [2]. Наприклад, коли користувачеві необхідно знайти розклад автобусів за певний проміжок часу і номери автобусів, то, враховуючи це, потрібно контент оновлювати кожного тижня і доповнити зручні фільтри пошуку для користувача, тоді достовірна інформація і зручний інтерфейс буде запорукою неодноразового повернення користувача на такий інформаційний ресурс [3]. Враховуючи те, що Інтернет надає можливість людям легко спілкуватись, зокрема в межах певної спільноти або регіону, наприклад, міста, та публікувати свої матеріали, тема побудови інформаційної системи (ІС), яка міститиме всю необхідну інформацію про певне місто, є досить актуальною [4–7]. Мета роботи полягає у створенні методу інтеграції та управління різним контентом з різних джерел в одній ІС, яка міститиме всю необхідну інформацію по певному місту згідно з потребами кінцевого користувача. Процес інтеграції та управління різнотипним контентом має базуватися на технологіях опрацювання інформаційних ресурсів в Інтернет-середовищі [8–16]. Об'єктом досліджень є процеси опрацювання контенту з різних джерел стосовно різних міст та створення інформаційних каталогів по містах. Предметом досліджень є мережа інформаційних ресурсів, яка містить та відображає необхідний користувачу контенту. Практична цінність роботи – створення єдиної інформаційної бази, яка дозволить знайти та надати потрібний контент про міста України. Така система дозволить користувачам спостерігати за погодою на певний час, шукати розклад громадського транспорту, переглядати афішу кінофільмів у кінотеатрах, новини, переглядати довідники наявних телефонів, каталог ресторацій, усе про туризм, фото міста, пам'ятники, адреси закладів тощо. В цьому випадку ІС виступає як каталог з інтегрованим контентом, оформленим у вигляді довідника кожного окремого інформаційного ресурсу.

Аналіз наукових публікацій

Мережа Web-сайтів “міські сайти” повинна охоплювати міста України, тобто сайти мають бути зроблені з переадресацією один на одного для міст, в яких населення понад 10 тис. осіб [1–7]. У сучасному світі все більшу роль починають відігравати інформаційні технології (ІТ) [8–12]. Їхні переваги: швидкість опрацювання запитів і видачі результатів, велика доступність, надійність яких неможливо переоцінити [13–19]. Застосування сучасних ІТ для полегшення орієнтування простому користувачу Інтернет та швидкому знаходженні у великих обсягах інформації необхідного контенту є першочерговою задачею для популяризації українських міст, їх сервісів та послуг [20–23]. Така ІС забезпечує підвищену надійність пошуку інформації про міста країни [24–28]. Також забезпечується централізоване збереження інформації про міста і підтримання її актуальності [29–33]. Існує багато ІС, які спеціалізуються на подачі статей, пов’язаних із містами [7]. Але їхня популярність напряму залежить від потужності трафіку на сайт. В Інтернеті є три види трафіку на сайт [8]. Перший – це з пошукових систем, другий – переходи з сайтів, третій – переходи із закладок і прямі заходи на сайт через URL стрічку браузера [9]. Треба орієнтуватись на перший вид трафіку для цього Інтернет-проекту. Для цього необхідно вчасно оновляти Web-сайт і оновлювати на ньому тільки актуальний контент, що є неможливим робити вручну модератором із-за швидких інформаційних змін в Інтернет-просторі [10]. Розвивати та рекламувати кожен сайт у мережі, запровадити переадресацію між сторінками різних сайтів мережі, побудови інфраструктури для розповсюдження інформації [11]. Для прикладу у табл. 1 наведено тільки сайти для міста Львова, їхні особливості, переваги та недоліки.

Таблиця 1

Порівняльні характеристики популярних інформаційних ресурсів “Мережі міських сайтів”

| Ресурс | Особливості | Недоліки | Переваги |
|------------------------|---|---|---|
| 032.ua | Новинний сайт міста Львова | Застарілість інформації, парсер з інших сайтів для забивання низькочастотних запитів у пошукових системах | Гарний трафік з пошукових систем за ключовими запитами, багато розділів з інформацією |
| loda.gov.ua | Новинний сайт Львівської обласної державної адміністрації, державний сайт | Прес служба ЛОДА веде сайт, Розвиває і публікує інформацію на ньому спеціально створений підрозділ, а саме: прес-відділ. | Актуальність новин є дуже високою |
| lviv.travel | Містить контент про місто для туристів | Здебільшого інформація тільки для туристів, редакція сайту на зарплаті | Авторський контент, вузькотематичний сайт, багатомовність |
| vorobus.com | Блог про міста очима жителя цього міста | Тільки авторська інформація, в основному новинний ресурс | Високі позиції за високочастотними запитами, високий ТІЦ |
| for.lviv.ua | Форум про Львів | Форум, не проста підтримка спільноти, відсутній довідник по місту, найм модераторів, постійна активність для розвитку форуму | Постійні відвідувачі форуму |
| posydenky.lvivport.com | Головний сайт форуму – це стрічка новин на тему міста Львова | Не проста підтримка спільноти, відсутній довідник по місту, найм модераторів, постійна активність для розвитку форуму | Постійні відвідувачі форуму |
| lvivport.com | Новини Львова | Збирач новин з інших ресурсів | Сайт є сателітом для форуму |
| city-adm.lviv.ua | Державний сайт з новинами від міської ради | Здебільшого інформація новинного типу, багато новинних розділів, прес служба веде сайт | Актуальність новин є дуже високою |
| dlab.com.ua | Інформація про заклади, підприємства міста | Не розвинена соціальна складова, є непотрібні застарілі функції, такі як друк, відправка по пошті | Високий трафік з пошукових систем, займає високочастотні запити, багатомовність, можливість додавати свої фірми |
| lviv-onilne.com | Мережа новинних міських сайтів | Виступає як реферальний сайт, не проста підтримка, постійне оновлення ресурсу потрібно, в групах у соціальних мережах дуже багато підписників | Дуже сильно розкручені групи в соціальних мережах |

Аналіз отриманих наукових результатів

Web-система мережі міських сайтів складається з підсайтів, об'єднаних у мережу. Вона повинна бути побудована в сучасному стилі з необхідним функціоналом, оптимізована під пошукові системи та сервіси, розподілена по серверах. Модель такої Web-системи подаємо як

$$W = \langle X, C, T, H, S, U, Q, Y, \alpha, \beta \rangle, \quad (1)$$

де $X = \{x_1, x_2, \dots, x_n\}$ – вхідні інтегровані дані з різних інформаційних ресурсів; $C = \{c_1, c_2, \dots, c_m\}$ – сформований контент (інформаційне наповнення, новини, статті, ресторація тощо); $T = \{t_1, t_2, \dots, t_l\}$ – часовий показник актуальності сформованого контенту для підтримки його релевантності; $H = \{h_1, h_2, \dots, h_p\}$ – список класифікованих URL інформаційних ресурсів, звідки інтегруються дані; $S = \{s_1, s_2, \dots, s_k\}$ – соціальний контент (результати ведення груп у соціальних мережах, репости записів тощо); $U = \{u_1, u_2, \dots, u_r\}$ – UGC або результати генерування контенту за допомогою користувачів (можливість користувачам додавати контент на сайт через промодерацію, а також коментарі без промодерації); $Q = \{q_1, q_2, \dots, q_w\}$ – запити користувачів IC необхідного контенту; $Y = \{y_1, y_2, \dots, y_f\}$ – релевантний контент як результат запитів користувачів IC; α – процес інтеграції необхідного контенту з різних джерел інформації; β – процес управління контентом для формування списків релевантного контенту згідно з запитом користувачів Web-системи.

Формування списку релевантного контенту прямо пропорційно залежить від грамотно побудованих правил інтеграції даних з різних джерел $Y = \beta \circ \alpha$ (фільтрації, виділення спаму та дублювання, класифікації тощо), тобто $Y = \beta(Q, C, S, U, \alpha(X, H, T))$.

Процес інтеграції даних з різних джерел подаємо суперпозицією функцій

$$C = \alpha(X, H, T) = \alpha_7 \circ \alpha_6 \circ \alpha_5 \circ \alpha_4 \circ \alpha_3 \circ \alpha_2 \circ \alpha_1, \quad (2)$$

де α_1 – збір даних з різних джерел, наперед визначених модератором системи; α_2 – фільтрація зібраних даних (розбиття контенту на два масиви розпізнаного/нерозпізнаного, або ідентифікованого/не ідентифікованого, з першим працює далі система, з іншим – модератор); α_3 – виділення спаму та його ліквідація (перевірка зі словниками фільтрів за заблокованих лінгвістичних змінних); α_4 – виділення дублювання та його ліквідація (перевірка з наявним контентом в базі даних); α_5 – форматування даних згідно з вимогами, закладеними саме в конкретну Web-систему (найчастіше – в XML-формат); α_6 – загальна класифікація (наприклад, контент типу ресторації, проживання, локації, туризм, довідка та інше), при необхідності детальна класифікація (наприклад, для контенту типу ресторації – кав'ярні, кафе, бари, паби, ресторани, піцерії, бістро, клуби, коворкінг, тайм-клуби) та збереження відповідних даних; α_7 – відображення нових даних на інформаційному ресурсі Web-системи.

Процес управління контентом Web-системи подаємо суперпозицією

$$Y = \beta_6 \circ \beta_5 \circ \beta_4 \circ \beta_3 \circ \beta_2 \circ \beta_1, \quad (3)$$

де β_1 – функція опрацювання запиту користувача (ідентифікація, тобто спам/не спам, при не спамі збереження (для подальшого аналізу статистики звернень та пошуку), при спамі збереження запиту для подальшого аналізу модератором, виділення слів-маркерів (класифікаційних лінгвістичних змінних, за наявності нових слів збереження останніх для аналізу модератором) та класифікація, тобто до якої рубрики контенту належить (наприклад, пошук ресторації); β_2 – вибірка даних з бази даних Web-системи згідно з визначеними в запиті словами-маркерами та формування списку-відповіді релевантного контенту; β_3 – сортування релевантного контенту перед відображенням за певними критеріями (частота перегляду, час читання, для зареєстрованого користувача врахування

його вподобань тощо); β_4 – відображення відсортованого релевантного контенту; β_5 – збір та збереження статистичних даних щодо роботи користувача з відображеним релевантним користувачем (який контент переглянув, час читання цього контенту, час перегляду списку контенту, відмова перегляду); β_6 – перерахунок історії контенту, згідно з останніми статистичними даними маніпуляцій користувача Web-системи.

Концептуальна модель – це поєднання концепції розробника моделі і користувача. Це свого роду UML діаграми, логіки, зв'язки, обмеження тощо. Також це є абстрактне подання моделі для Web-системи, яка вирішує проблему і досягає її мети. На рис. 1а подано концептуальну модель Web-системи мережі міських сайтів. Мережа розділяється на сайти за певним містом.

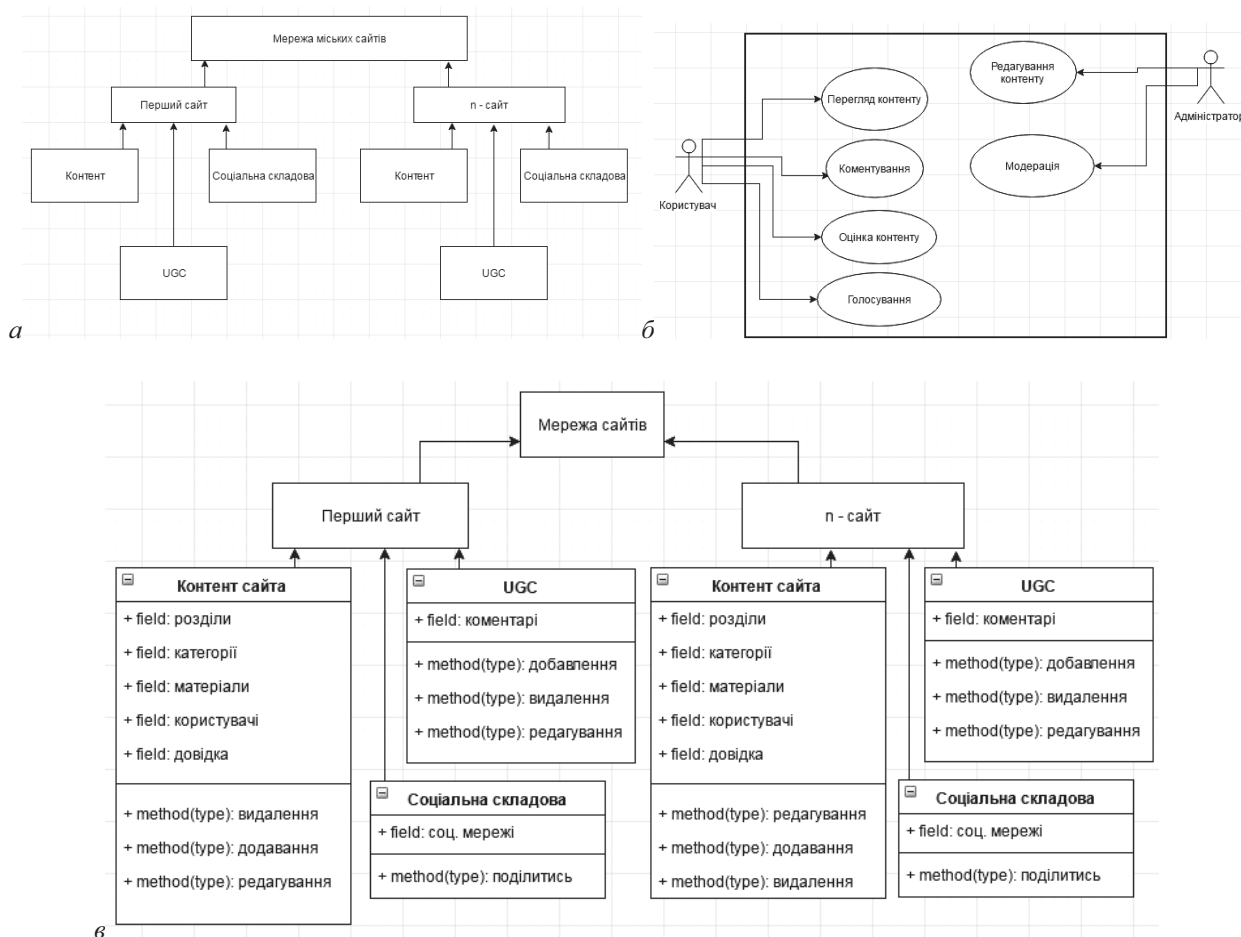


Рис. 1. Діаграма: а – концептуальної моделі; б – використання та класів ІС “Мережі міських сайтів”

На рис. 1б подано діаграму використання, де зображено всі актори до інформаційної системи і всі варіанти дій. Актор “Користувач” – це актор, який є відвідувачем сайту. Актор “Адміністратор” – це відвідувач, який має повні права на управління сайту. До таких прав належить редагування контенту і правка скриптів сайту. Користувач має такі варіанти використання як: перегляд контенту, коментування, оцінка контенту, голосування. Адміністратор має такі варіанти використання: редагування контенту, модерація (набір функцій для управління користувачами сайту, блокування порушників, правка і видалення недопустимих коментарів). На діаграмі (рис. 1в) зображено класи та атрибути ІС мережі міських сайтів.

Мережа сайтів складається з сайтів, кожен сайт складається з певних класів: контент, UGC, соціальна складова (табл. 2). На рис. 2а подано взаємодію користувачів з сайтом. Блоки розділені на користувача (User), систему управління сайтом (CMS), адміністратора (Admin) із взаємозв'язком між ними при додаванні коментаря та підтвердження його адміністрацією.

Опис атрибутів та методів класів ІС “Мережі міських сайтів”

| Клас | Складова | Назва | Опис |
|--------------------|----------|--|---|
| Контент сайту | Атрибут | Розділи | Це є ресторації, проживання, локації, туризм, галерея |
| | | Категорії | Розділ ділиться на певні категорії |
| | | Матеріали | Сторінки сайту в певній категорії |
| | | Користувачі | Зареєстровані користувачі сайту, їх профілі |
| | | Довідка | Інші матеріали, які не входять в чотири основні розділи |
| | Методи | Редагування | Можливість редагування розділів, категорій, матеріалів |
| | | Видалення | Можливість видалення розділів, категорій, матеріалів |
| Додавання | | Можливість додавання розділів, категорій, матеріалів | |
| UGC | Атрибут | Коментарі | Коментарі до матеріалів на сайті |
| | Методи | Редагування | Можливість користувача редагувати коментарі |
| | | Видалення | Можливість видаляти коментарі |
| | | Додавання | Можливість користувача писати коментарі |
| Соціальні складова | Атрибут | Соціальні мережі | Кнопки поділитись |
| | Метод | Поділитись | Соціальна функція для розповсюдження матеріалів |

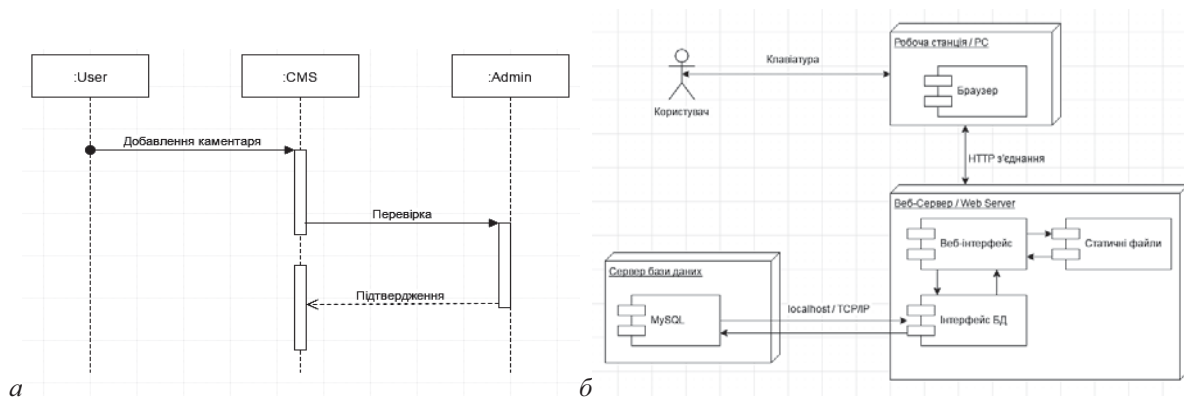


Рис. 2. Діаграма: а – послідовності та б – представлення системної частини ІС “Мережі міських сайтів”

Компоненти відповідають поданню робочих екземплярів одиниць коду. Компоненти, що не мають представлення під час роботи програми, на таких діаграмах не відображаються; натомість їх можна відобразити на діаграмах компонент. Діаграма розгортання відображає робочі екземпляри компонент, а діаграма компонент, натомість, відображає зв'язки між типами компонент. На діаграмі (рис. 2б) зображена серверна частина сайту, детальніше описано в табл. 3.

Таблиця 3

Діаграма розгортання інформаційної системи “Мережі міських сайтів”

| Модуль | Опис |
|-------------------|---|
| Браузер | Програма Web-переглядач сайтів |
| Web-інтерфейс | Дизайн та шаблон сайту |
| Статичні файли | Фото та відеофайли сайту, файли стилів та скрипти |
| Web-сервер | Програма для обробки програмних файлів сайту |
| Сервер бази даних | Програма для обробки SQL запитів |
| Інтерфейс БД | Тип бази даних та інтерфейс підключення |
| MySQL | База даних для зберігання інформації |

На діаграмі діяльності (рис. 3а) подано соціальні функції, зокрема такі, як коментарі та підписка на пошту. При успішній підписці, відвідувач буде занесений у базу даних розсилки і переадресований на шлюзову сторінку. При вдалому опублікуванні коментаря, він спершу модерується на дотримання правил, встановлених на сайті. Якщо модерація успішна і порушень нема, тоді коментар буде опублікований, в противному випадку – не буде опублікований, а користувач отримає блокування на 7 і більше днів. На діаграмі пакетів (рис. 3б) зображено структуру одного сайту у всій мережі, головний пакет “Мережа сайтів”, до нього входить пакет “Функціонал сайту”, “Ресторації”, “Локації”, “Проживання”, “Туризм” (табл. 4).

Таблиця 4

Вміст пакетів ІС “Мережі міських сайтів”

| Пакет | Складові | Опис |
|------------------|-------------------------------|---|
| Функціонал сайту | UGC | Контент, створений користувачами |
| | Соціальна складова | Можливість швидкого розповсюдження через соц. мережі |
| | Ресторації | Розділ сайту з описом закладів в місті |
| | Проживання | Розділ сайту зі списком готелів, motelів тощо |
| | Локації | Список закладів культури і мистецтва |
| | Туризм | Інформація для туристів міста |
| | Довідка | Інформаційний каталог закладів, організацій по місту |
| | Інше | Різне, що не увійшло в загальні категорії |
| Ресторації | Кав'ярні | Список кав'ярень в місті та їхні ціни |
| | Кафе/бари | Каталог кафе і барів у місті та їхні ціни |
| | Паби | Список пабів у місті |
| | Ресторани | Каталог інформації про ресторани міста, меню, столики |
| | Піцерії/бістро | Каталог піцерій у місті та їх доставка, години роботи |
| | Клуби | Список клубів та їхній час роботи, опис тощо |
| | Коворкінг/Тайм-клуби | Каталог клубів, їхній час роботи, ціна, опис |
| Проживання | Готелі | Каталог готелів, ціни, розташування |
| | Мотелі | Список motelів, ціни, розташування |
| | Хостели | Список хостелів, ціни, розташування |
| Локації | Музеї | Список музеїв міста, розташування, години роботи |
| | Театри | Список театрів міста, афіша, розташування |
| | Галереї | Розташування галерей у місті, години роботи |
| | Кінотеатри | Розташування в місті, години роботи, афіша |
| | Концертні зали | Години роботи, розташування, ціна |
| | Спортивні заклади | Години роботи, вид закладу, ціна |
| | Активний відпочинок | Вид відпочинку, опис його |
| | Визначні місця | Пам'ятники, скульптури тощо |
| Туризм | Туристичні фірми | Каталог фірм та їх розташування |
| | Консульства | Години роботи, розташування |
| | Центр туристичної інформації | Розташування, години роботи |
| Довідка | Карта міста | Карта міста та позначення на карті з розділів |
| | Розклад автобусів | Таблиця зі списком розкладу маршрутів |
| | Карта громадського транспорту | Карта руху громадського транспорту |
| | Установи та організації | Список державних адміністрацій та установ |
| | Медицина | Все, що пов'язано з медициною |
| | Послуги | Послуги змішаної тематики в місті |
| | Банки та фінанси | Організації, що займаються фінансами |
| | Шопінг | Все по роздрібній торгівлі міста |
| | Служби | Комунальні служби міста |
| | Освіта | Все, що пов'язано з навчальними закладами |
| | Торгівля | Список оптових продавців |

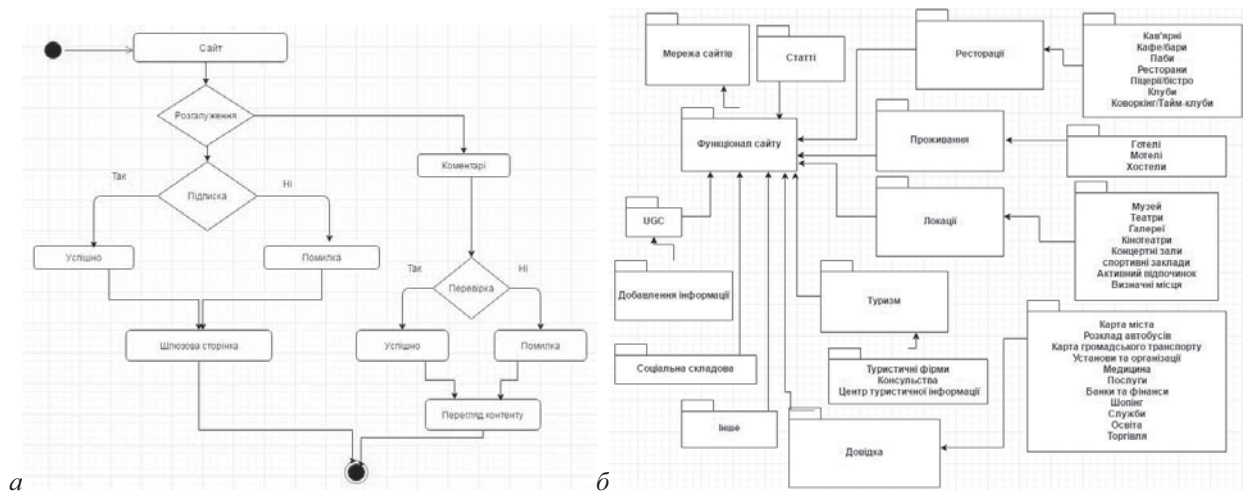


Рис. 3. Діаграма: а – активності для користувацьких функцій та б – пакетів ІС “Мережі міських сайтів”

Етапи створення інформаційної системи “Мережі міських сайтів”:

- реєстрація групи доменів, закупка серверів для проекту, налаштування серверів зі зв’язкою Apache, Nginx, обробник PHP, налаштування DNS під кожен домен;
- додавання dkim підписів на домен для поштових розсилок;
- створення децентралізованих баз даних на кожен сайт;
- налаштування технічної пошти admin@, abuse@, support@;
- створення технічних піддоменів для статичних файлів.

Етапи створення одного сайту для інформаційної системи:

- розробка концепції сайту, інформаційне проектування та розробка ескізу базового дизайну;
- створення працюючого шаблону сайту, включаючи повну розробку “візуала”, посилання, інтерактивні елементи налаштування зворотного зв’язку;
- програмування і підключення адміністраторського модуля, сайт повинен бути повністю інтерактивним і керованим адміністратором без залучення розробників до підтримки сайту;
- верстка сайту під макет шаблону та створення мобільної версії шаблону і верстка її;
- підготовка, редагування та налаштування вмісту сайту під пошукові системи;
- публікація сайту на сервері, тестування та генерація карт сайту;
- оптимізація сайту під Google, Yandex. Додавання сайту в пошукові системи;
- генерація.xml карти сайту та створення browserconfig.xml для Windows 8.

Основні вимоги до функціонування є:

- сайт розробляється під базову роздільну здатність екрана 1024x768 пікселів;
- коректне відображення браузером Google Chrome, Opera, Mozilla Firefox, Internet Explorer (по можливості), кросбраузерність та генерація мета-тегів для пошукових систем;
- авто генерація ключових слів для швидкого пошуку інформації;
- максимально простий інтерфейс сайтів та своєчасне оновлення інформації сайтів;
- швидка генерація сторінок сайту і доставка їх користувачу;
- авто оновлення інформації інформаційних модулів на сайті.

ІС мережі міських сайтів складається з міських сайтів. У мережу може входити необмежена кількість сайтів, але вони повинні бути з тематикою міського сайту. На кожному сайті повинна бути актуальна інформація. Методом вирішення проблеми з неактуальною інформацією є набір платних робітників на кожен сайт та введення функції правки неактуальних матеріалів. Щоб на сайтах не виникало проблем із пошуком інформації, буде можливість повнотекстового пошуку по матеріалах і можливість фільтрувати результати пошуку для покращення релевантності результатів. Додаткове введення автогенерування тегів з матеріалів сайту і формування хмари тегів. Модулі “Курси валют” та

“Погода” повинні підтягувати інформацію по сторінці через API інших ресурсів в базу даних і дані при завантаженні сторінки брати з бази даних. При високому навантаженні потрібно буде організовувати пошук через пошуковий сервер Sphinx і включення кешування Memcached. IC реалізована за допомогою таких технологій: PHP, Nginx, HTML 5, CSS 3, Apache, MySQL, Javascript. Сервера використовують операційну систему Debian 8. Для блокування певних легких DoS атак встановлено iptables. Переваги вибраних засобів для створення IC мережі міських сайтів приведено в табл. 5.

Таблиця 5

Переваги обраних засобів для створення IC “Мережі міських сайтів”

| Технологія | Переваги |
|-------------|--|
| Debian | Виходять нові версії, багато документації є по адмініструванню |
| ISP Manager | Оновляється, платна панелька, але дуже якісна, використана 4 версія (через навик користування нею), статистика користувачів, трафіка, графіки, файловий менеджер |
| Nginx | Топовий Web-сервер, який використовується як fronted на більшості сайтів світу, використовують дуже популярні сайти Яндекс, Wordpress.com і інші |
| Apache | Відкритий Web-сервер, популярний, легкий у користуванні як backend |
| PHP | Весь час оновляється, є ООП, зрозумілий синтаксис, дуже багато сайтів створених на даній мові програмування, один з них Wikipedia.org тощо |
| MySQL | Реляційна база даних, швидка обробка запитів |
| HTML 5 | Кросбраузерність, нові стандарти, легкий в освоєнні, мобільна верстка |
| CSS 3 | Кросбраузерність, нові стандарти |
| Javascript | Є готова бібліотека jquery, в якій дуже багато готових функцій |
| Sphinx | Швидка обробка запитів, повнотекстовий пошук, швидке сканування бази даних та індексування сканованих даних |
| Memcached | Є можливість керувати SQL запитами, є підтримка PHP |
| iptables | Написання правил для блокування порушників на протокольному рівні |
| proftpd | Відкритий FTP сервер простий в налаштуванні |
| PHP curl | Висока швидкодія, окрім HTTP запитів підтримує протоколи SMTP, IMAP, POP3, Telnet, FTP, LDAP, RTSP, RTMP |

Для зручного управління файлами через FTP-клієнт встановлено proftpd. Панель управління сервером ISP Manager 4 обрана через інтуїтивне управління і файловий менеджер, щоб не авторизовуватись по sftp через ftp-клієнт. Використовує система СУБД MySQL. Домени прикріплені до Яндекс.PDD, з них направлені А записи на сервери. На кожному сервері налаштована зв'язка Nginx, як фронтенд сервер, який приймає запити на.php файли і передає на порт Apache, який, своєю чергою, опрацьовує їх за допомогою оброблювача PHP, а статичні файли, фото, відео й інше віддає напряму в браузер користувачу. Для фронтенду кожного сайту використано PHP, для бекенду використано теж мову програмування PHP. За правильної побудови швидко працює і легко масштабується. Кросбраузерність реалізована за допомогою HTML 5 і CSS 3, верстка адаптивна, тому не потрібно заново створювати мобільну версію сайту на окремому піддомени. Використання JavaScript для певних модулів сайту та для полегшення користування і збільшення юзабіліті. При збільшенні обсягу інформації буде запроваджений пошук інформації через пошуковий сервер Sphinx і включене кешування через memcached. Мережа розділена на сайти, кожен сайт для кожного міста. Структура бази даних сайту для одного міста (рис. 4):

- content – головна таблиця, в якій міститься вся інформація про місто зі всіх розділів;
- category – список розділів, cat_sub використано для рекурсивного виводу підкатегорій;
- ratings – рейтинги матеріалу, де в атрибут rating_points додається +1 при голосуванні;
- comments – таблиця з коментарями на сайті;
- user – інформація про всіх зареєстрованих користувачів з їхніми правами в атрибуті user_group;
- pages – відокремлена таблиця від загального функціоналу для статичних сторінок, наприклад, таких як розклад автобусів, контакти тощо;
- polls – інформація про голосування та списки для голосувань у віджеті на сайті.

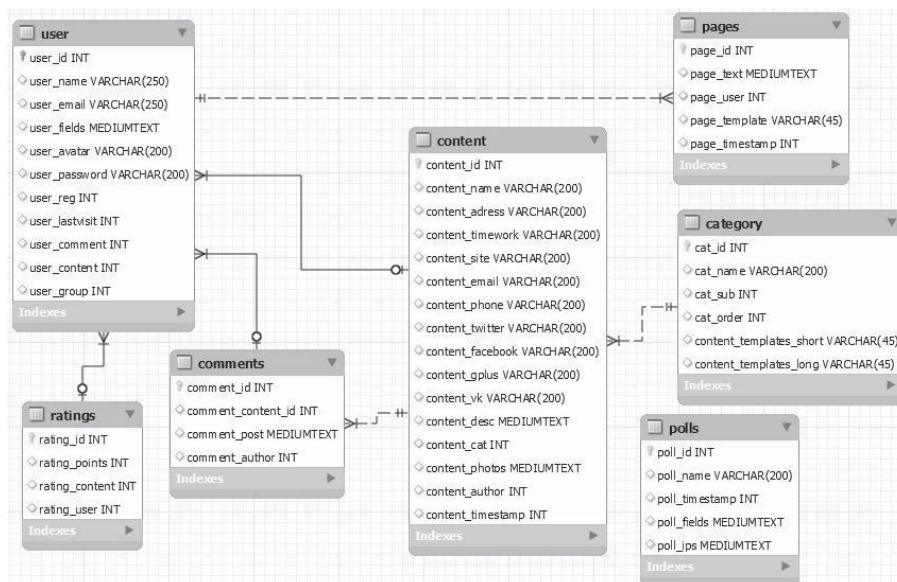


Рис. 4. Структура бази даних одного з сайтів

На рис. 5а подано прототип структури та верстки сайтів під всю мережу. Сайт ділиться на основні розділи: “Ресторації” (рис. 5, б), “Проживання” (рис. 5, в), “Локації” (рис. 5, г), “Туризм” (рис. 6, а), “Довідка” (рис. 6, б), “Галерея”, “Кав’ярні” (рис. 6в) та дочірні підрозділи. Оскільки на такі сайти будуть заходити здебільшого з мобільних телефонів, тому обов’язково потрібно передбачити мобільну версію сайту. При цьому, щоб не дублювати контент на піддомени, зазвичай – це **m.abcv.g.info** чи **mobile.abcv.g.info**, краще вже одразу використовувати мобільну верстку, беручи до уваги ще й те, що у 2016 році самі пошуковики навчились уже все визначати автоматично. Така мобільна верстка зображена на рис. 8, г.

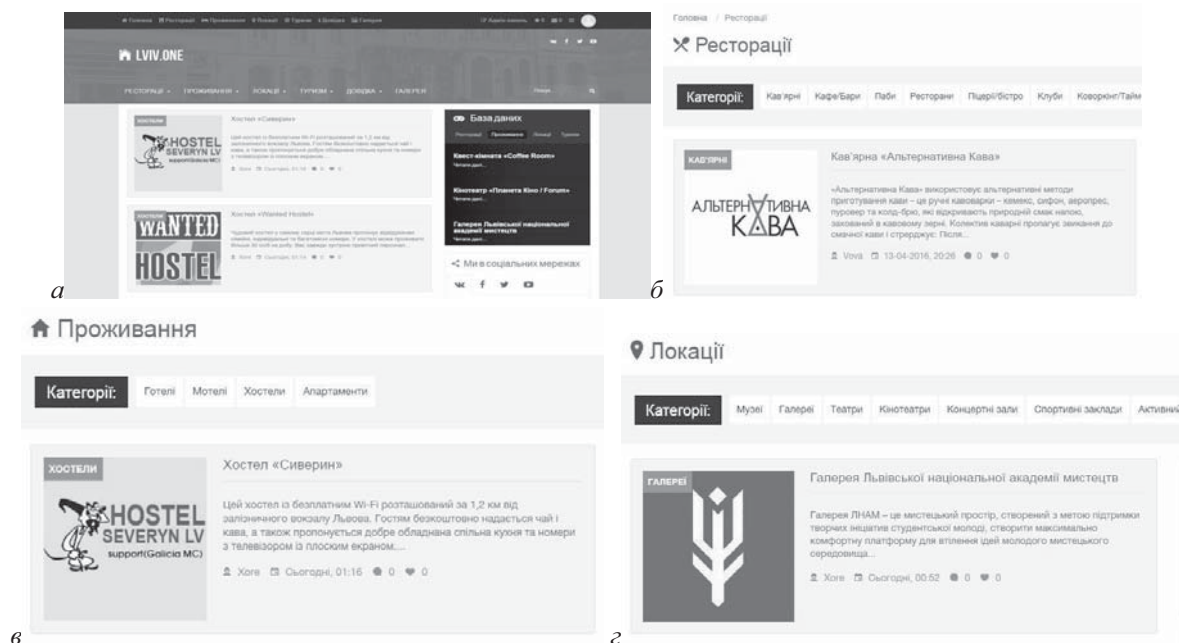


Рис. 5. Розділ сайту: а – головна сторінка; б – “Ресторації”; в – “Проживання” та г – “Локації”

Якщо є мобільна версія сайту, потреба в додатках до мобільного телефону відпадає. Оскільки пробитися в топ у магазинах Google Store і Apple Store дуже непросто за високої конкуренції і у короткий термін, ідея додатків була відкинута, а головний акцент зроблений саме на мобільній верстці. Вигляд меню та форма пошуку для навігації зображені на рис. 6, д. Вигляд повного

матеріалу на мобільній версії зображено на рис. 7, а. Календар, зроблений для зручнішого пошуку інформації на сайті і кращого юзабіліті по матеріалах, зображено на рис. 7, б. На сайті також присутній морфологічний пошук (рис. 7, в). Модуль опитування, для розбавлення функціоналу сайту, зображено на рис. 7, г, а також результати після голосування на рис. 7, д.

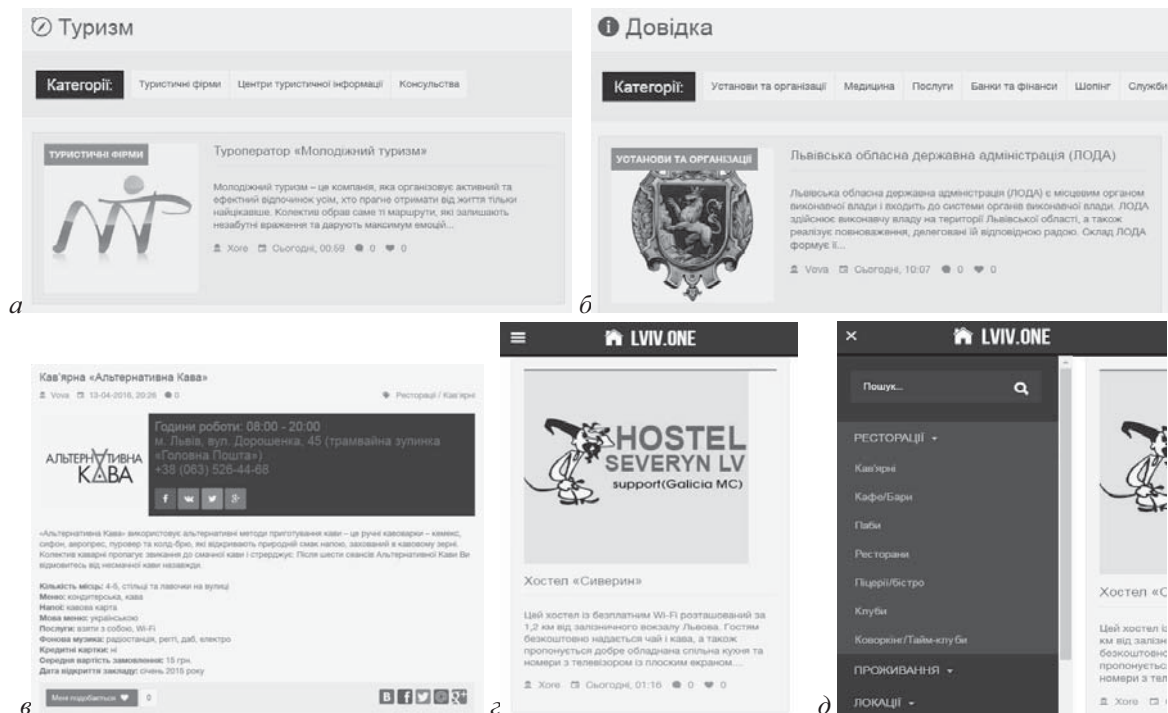


Рис. 6. Приклад вигляду розділу: а – “Туризм”; б – “Довідка”; в – “Кав’ярні” та мобільної верстки сайту; г – зі списком матеріалів; д – вигляд меню для навігації

Також на сайті є соціальні функції такі, як авторизація, коментування, підписки на матеріали, оцінка матеріалів. Вигляд профілю зображено на рис. 8, а. Приклад вигляду галереї по місту зображено на рис. 8, б. В цю галерею завантажено тестові зображення. Приклад розгорнутої галереї з фото зображено на рис. 8в. Віджети-інформери, які працюють за допомогою Приват API (курси валют) та Yandex API (погода). Дані з API збираються і записуються в базу даних, а потім виводяться уже з бази даних по геолокації користувача і, отже, зменшується навантаження на сервер. Зображено два інформери на рис. 8, г. Також ще однією з додаткових соціальних функцій є кнопка “Мені подобається”, яка показує рейтинг матеріалу. Також треба брати до уваги те, що чим більше є такого різновиду функціоналу, то більша можливість того, що користувач повернеться на сайт. Соціальна кнопка зображена на рис. 8, д. Сьогодні одними з основних соціальних функцій контентних сайтів залишаються коментарі. Всі інші функції відійшли на другий і третій план. Форма додавання коментаря зображена на рис. 8, е та вигляд опублікованого коментаря – на рис. 8, ж. Коментарі можуть слугувати як відгуки про заклад, але, як іноді показує практика, така інформація не завжди буває достовірною, адже, як і з рейтингом, недобросовісні власники закладів можуть самі робити, так звані, “накрутки” за гроші.

На кожному сайті присутня форма додавання матеріалу для користувачів (рис. 9, а, б). Однією з додаткових функцій є додавання матеріалу в закладки (рис. 9, в). Необхідність цієї функції залежить від кількості аудиторії та її віку. Є функція пошуку подібних матеріалів, яка генерується за допомогою зібраних даних з історії переглядів користувача. На основі цих даних будується вивід подібних матеріалів (рис. 9, е). Для користувачів також зроблена статистика переглядів (рис. 9г). На рис. 9, е подані основні класи користувачів IC з характеристиками. У світі компанія Google активно просуває SSL сертифікати з 2014 року. Щоб йти в ногу з часом, потрібно виготовити SSL сертифікат. Є тільки два види сертифікатів – це китайської компанії Wosign, та відкритий SSL

Letsencrypt (рис. 10, а). RSS для Web-сайтів став як один з стандартів, тому він є обов'язковий для кожного сайту. Ще одним з таких стандартів, запровадженим Google, є використання домену для статичних файлів у вигляді назваusercontent.com. Зараз усі великі сайти монополії github.com, store.steampowered.com використовують цей стандарт. Ще одним стандартом для кожного сайту є зворотній зв'язок через різні ситуації. Наприклад, взяти інтерв'ю чи повідомити про баг на сайті, тощо (рис. 11). Також згенерована.xml карта сайту для пошукових систем.

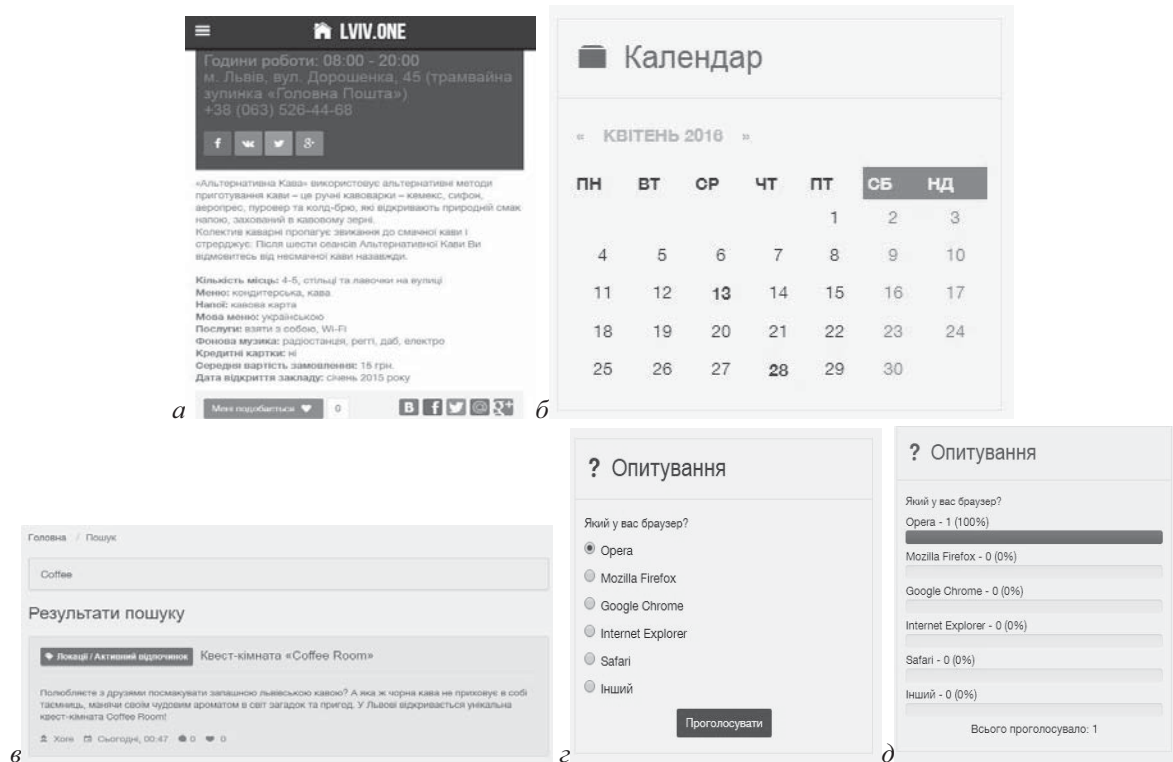


Рис. 7. Приклад: а – мобільної верстки сайту, вигляд повного матеріалу; б – календаря для кращої навігації; в – форми пошуку по сайті та вигляд результатів пошуку; г – опитування та д – вивід результатів опитування

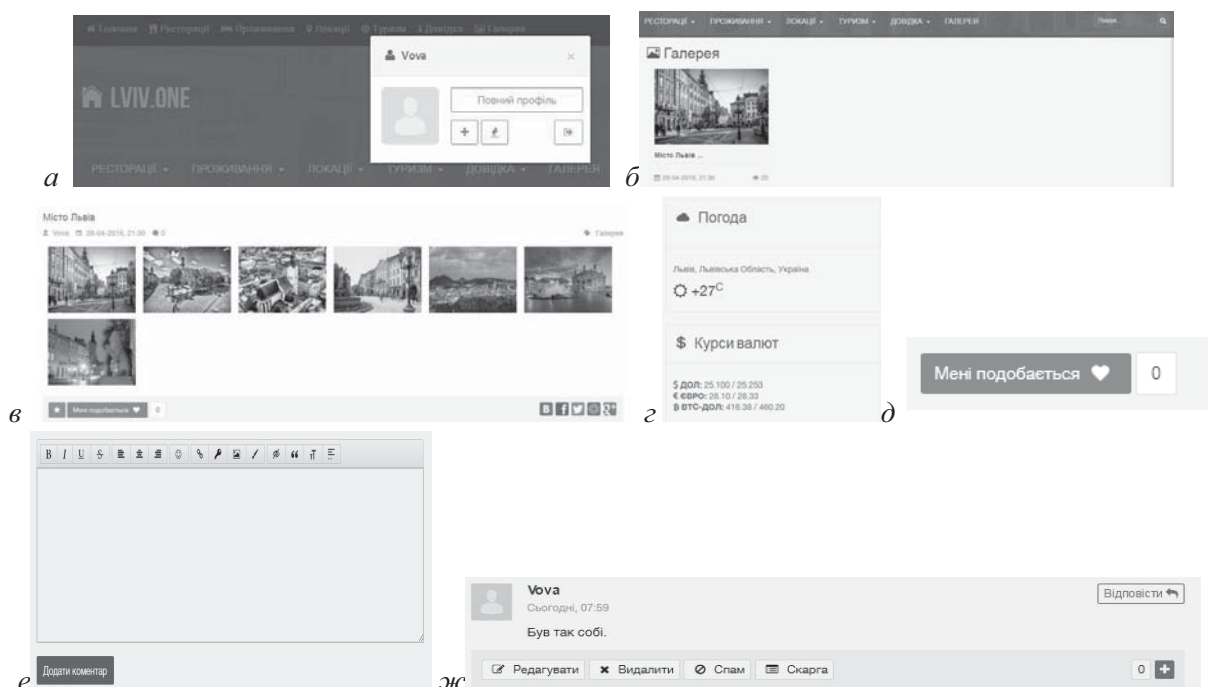


Рис. 8. Вигляд підзавантаженого профіля сайту; Інформер курсів валют і погоди, соціальна кнопка “Мені подобається”, форма додавання коментарів та вигляд опублікованого коментаря



Рис. 9. Форма а – реєстрації; б – входу; в – вигляд профілю користувача; г – форма додавання матеріалу; д – приклад матеріалу в закладках; е – подібні матеріали; ж – статистика сайту та з – список рівнів доступу та групи користувачів; і – персональні повідомлення

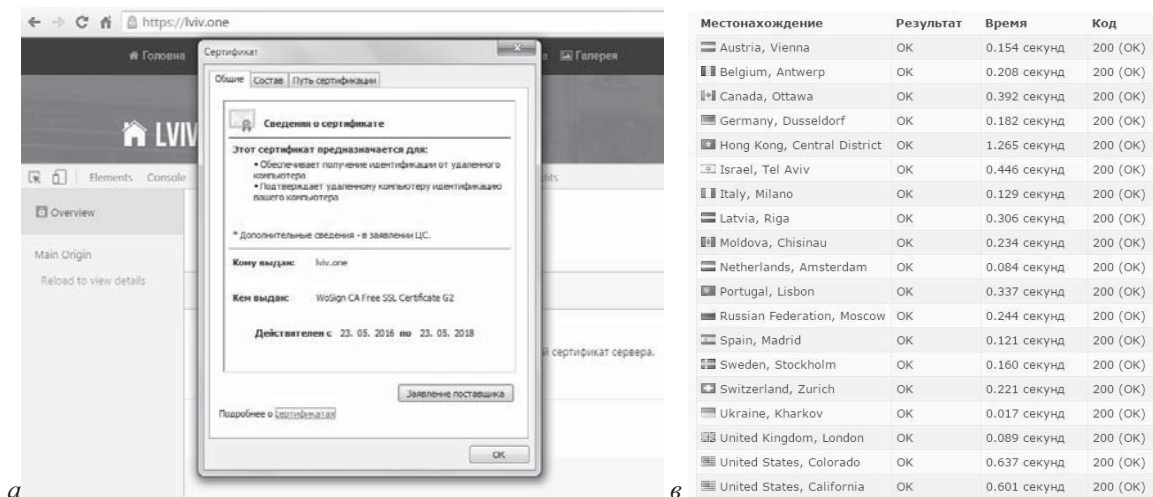


Рис. 10. SSL сертифікат домена

Наступним кроком у тестуванні є проведена перевірка Web-порталу на валідність, скориставшись безкоштовним сервісом onlinewebcheck.com. У результаті перевірки було виявлено 0 помилок, але 22 повідомлень уваги. Результати перевірки можна переглянути на рис. 13, а. Водночас у конкурента наявні 4 помилки, що можна побачити на рис. 13, б. Також перевірено швидкість віддачі сайту через сервіс check-host.net. Результати перевірки подані на рис. 10, в.

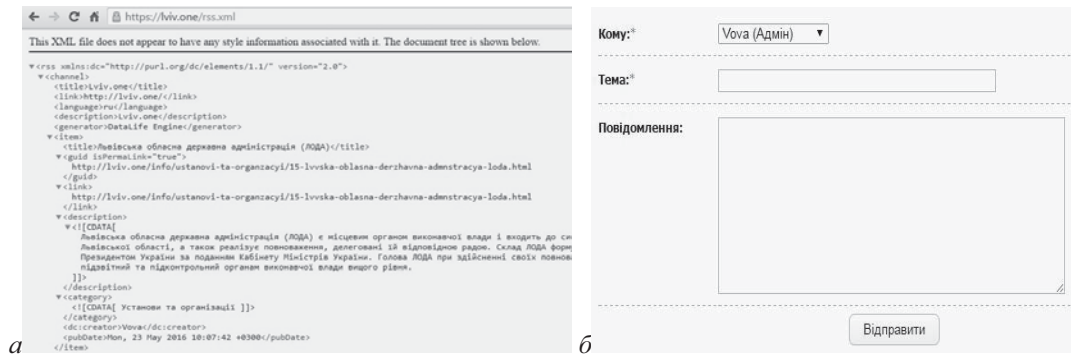


Рис. 11. Реалізація а – RSS стрічки сайту та б – зворотного зв'язку

Для аналізу функціональності розробленого інформаційного ресурсу використано сервіс навантажувального тестування сайтів loadimpact.com. Результат: при 25 одночасних підключеннях сервер працює у штатному режимі (рис. 12).

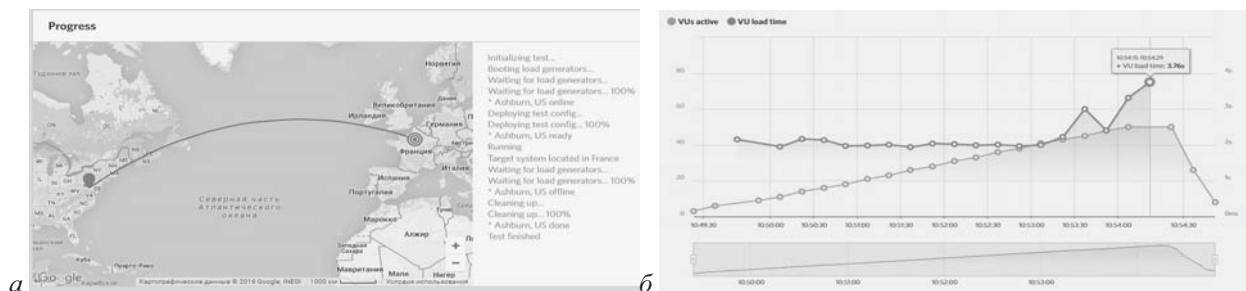


Рис. 12. Результати навантажувального тестування



Рис. 13. Результати перевірки Web-порталу на валідність та на швидкість віддачі

Основною перевагою такої IC є її масштабованість. Її можна легко масштабувати для будь-якого сайту і під будь-яку тематику. Все зроблено для максимальної швидкодії і роботи з великими масивами даних. Одними з кращих можливостей IC, порівняно з аналогами, є:

- швидкість генерації сторінки та присутність SSL сертифіката та шифрування TLS;
- якісніший контент, оскільки редагують редактори, а не є лише парсером;
- відсутність новин, усі новини перетекли у групи в соціальних мережах;

• відсутні такі неактивні розділи, як дошка оголошень, каталог із вакансіями, нерухомість тощо (оскільки дані розділи – це здебільшого парсер з сайтів, які є монополіями olx.ua, work.ua, lun.ua тощо (рис. 14), не буде, так званої, “сміттєвої” інформації для Інтернет-аферистів);

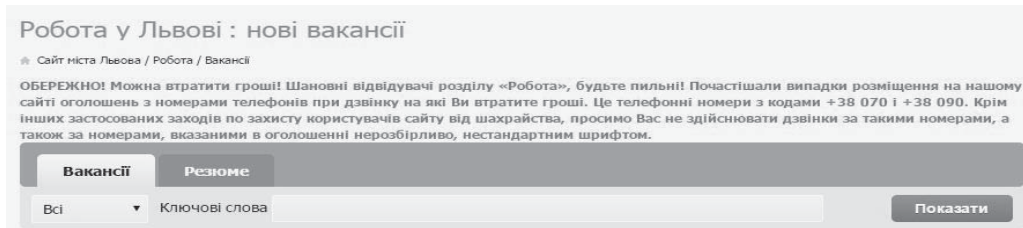


Рис. 14. Приклад неактуальності даних розділів на основі сайту конкурента

• мобільна верстка сайту без дубляжу контенту на піддоміні;
 • RSS стрічка для кожного розділу;
 • акцентування на соціальні мережі та їхніх відвідувачів, а також на мобільних користувачів.

Висновки та перспективи наукових розвідок

У роботі розроблено загальну структуру ІС мережі міських сайтів, а також побудовані UML діаграми і описана предметна область. Розроблено методи та засоби вирішення проблеми побудови ІС мережі міських сайтів. Створено єдину інформаційну базу, яка дає змогу знайти та надати потрібну інформацію про певні міста України. Така система дозволяє користувачам спостерігати за погодою на певний час, шукати розклад громадського транспорту, переглядати афішу кінофільмів у кінотеатрах, новини, переглядати довідники наявних телефонів, каталог ресторацій, усе про туризм, фото міста, пам'ятники, адреси закладів тощо. В цьому випадку ІС виступає саме як інтерактивний каталог з динамічною оперативною інформацією, оформленою як довідник кожного сайту. В подальшому на основі зібраної статистики від постійних користувачів необхідно розробити алгоритми Machine Learning та правила навчання нашої ІС для формування пропозицій постійному відвідувачу згідно його вподобань маршрутів екскурсій та місць відпочинку.

1. Клименко Р. Web-мастеринг на 100 % / Р. Клименко. – СПб: Питер, 2013. – 512 с.
2. Ленгсторф Дж. Искусство общения в Интернет. Краткое руководство / Д. Ленгсторф. – М.: Вильямс, 2011. – 334 с.
3. Основні правила побудови семантично відкритих інформаційних систем / Жежнич П., Кравець Р., Пасічник В., Пелецишин А. // Вісник Нац. ун-ту “Львівська політехніка”. – Львів, 1999. – № 383. – С. 84–95.
4. Пелецишин А. Позиціонування сайтів у глобальному інформаційному середовищі / А. Пелецишин. – Львів: Нац. ун-т “Львівська політехніка”, 2007. – 260 с.
5. Methods based on ontologies for information resources processing / [V. Lytvyn, V. Vysotska, L. Chyrun, D. Dosyn]. – LAP Lambert Academic Publishing. Saarbrücken, Germany. – 2016. – 324 p.
6. Берко А. Системи електронної контент-комерції: монографія / А. Берко, В. Висоцька, В. Пасічник. – Львів: Нац. ун-т “Львівська політехніка”, 2009. – 612 с.
7. Математична лінгвістика / [В. Пасічник, Ю. Щербина, В. Висоцька, Т. Шестакевич] // Серія “Комп’ютинг”. – Львів: Новий світ-2000, 2012. – 359 с.
8. Висоцька В. Методи і засоби опрацювання інформаційних ресурсів у системах електронної контент-комерції: автореферат к.т.н.: 05.13.06 – інформаційні технології / В. Висоцька; Нац.ун-т “Львівська політехніка”. – Львів, 2014. – 27 с.
9. Vysotska V. Web Content Processing Method for Electronic Business Systems / V. Vysotska, L. Chyrun // International Journal of Computers & Technology. – Vol. 12, No 2. – December 2013. – P. 3211–3220.
10. Висоцька В. Моделювання етапів життєвого циклу комерційного web-контенту / В. Висоцька, Л. Чирун, Л. Чирун // Вісник Нац. ун-ту “Львівська політехніка”. – Львів, 2011. – № 715. – С. 69–87.
11. Висоцька В. Особливості проектування та впровадження систем електронної комерції / В. Висоцька // Вісник Нац. ун-ту “Львівська політехніка”. – Львів 2008. – № 631. – С. 55–84.
12. Vysotska V. Analysis and evaluation of risks in electronic commerce / V. Vysotska, I. Rishnyak,

L. Chyrun // *CAD Systems in Microelectronics, CADSM '07.* – 2007. – P. 332–333. 13. Vysotska V. Comprehensive method of commercial content support in the electronic business systems / V. Vysotska, L. Chyrun, L. Chyrun // *Вісник НУ “Львівська політехніка”.* – Львів, 2013. – № 777. – С. 21–30. 14. Berko A. Content formation method in the electronic content commerce systems / A. Berko, V. Vysotska, L. Chyrun // *Computer Science & Engineering (CSE-2013).* – 2013. – P.174–179. 15. Vysotska V. Uniform method of operative content management in the electronic content commerce systems / V. Vysotska, L. Chyrun, L. Chyrun // *Computer Science & Engineering (CSE-2013).* – 2013. – P.180–185. 16. Берко А. Метод формування контенту в системах електронної контент-комерції / А. Берко, В. Висоцька, Л. Чирун // *Вісник Нац. ун-ту “Львівська політехніка”.* – Львів, 2013. – № 751. – С. 108–118. 17. Берко А. Моделювання процесів опрацювання інформаційних ресурсів в системах електронної контент-комерції / А. Берко, В. Висоцька, Л. Чирун // *Вісник Нац. ун-ту “Львівська політехніка”.* – Львів, 2013. – № 771. – С.126–135. 18. Висоцька В. Уніфікований метод формування інформаційного продукту в системах електронної контент-комерції / В. Висоцька, Л. Чирун // *Наукові праці Чорноморського держ. Унів. ім. П. Могили.* – 2013. – №. 217(229). – С. 91–101. 19. Висоцька В. Уніфіковані методи опрацювання інформаційних ресурсів в системах електронної контент-комерції / В. Висоцька, Л. Чирун, Л. Чирун // *Наукові праці Чорноморського держ. унів. ім. П. Могили.* – 2013. – № 201(213). – С. 13–24. 20. Висоцька В. Уніфікований метод оперативного управління контентом в системах електронної контент-комерції / В. Висоцька, Л. Чирун, Л. Чирун // *Вісник Нац. ун-ту “Львівська політехніка”.* – Львів, 2013. – № 751. – С. 118–128. 21. Висоцька В. Особливості проектування та аналіз узагальненої архітектури систем електронної контент-комерції / В. Висоцька, Л. Чирун, Л. Чирун // *Вісник Нац. ун-ту “Львівська політехніка”.* – Львів, 2013. – № 770. – С. 83–101. 22. Берко А. Структура засобів опрацювання інформаційних ресурсів в системах електронної контент-комерції / А. Берко, В. Висоцька, Л. Чирун // *Вісник НУ “Львівська політехніка”.* – Львів, 2013. – № 770. – С. 12–21. 23. Щербина Ю. Утворення українських дієприкметників за допомогою породжувальних граматик / Ю. Щербина, Ю. Нікольський, В. Висоцька, Т. Шестакевич // *Вісник Нац. ун-ту “Львівська політехніка”.* – Львів, 2011. – № 715. – С. 354–369. 24. Висоцька В. Метод вибору оптимального алгоритму криптографічного захисту інформації / В. Висоцька, О. Гарасим // *Вісник Нац. ун-ту “Львівська політехніка”.* – Львів 2010. – № 673. – С. 220–233. 25. Щербина Ю. Науковий напрям та навчальна дисципліна “Математична лінгвістика” / Ю. Щербина, Т. Шестакевич, В. Висоцька // *Вісник Нац. ун-ту “Львівська політехніка”.* – Львів, 2010. – № 673. – С. 384–393. 26. Чирун Л. Застосування контент-аналізу текстової інформації в системах електронної комерції / В. Висоцька, Л. Чирун // *Вісник Нац. ун-ту “Львівська політехніка”.* – Львів, 2010. – № 689. – С. 332–347. 27. Висоцька В. Схеми моделювання систем керування контентом / В. Висоцька // *Вісник Нац. ун-ту “Львівська політехніка”.* – Львів, 2010. – № 689. – С. 90–108. 28. Висоцька, В. Умовна ентропія та ентропія поєднання контенту в системах електронної комерції / В. Висоцька // *Вісник Нац. ун-ту “Львівська політехніка”.* – Львів, 2008. – № 626. – С. 116–125. 29. Берко А. Аналіз і класифікація моделей систем електронної комерції / А. Берко, В. Висоцька // *Вісник Нац. ун-ту “Львівська політехніка”.* – Львів, 2007. – № 591. – С. 103–112. 30. Берко А. Методи та засоби оцінювання ризиків безпеки в інформаційних системах електронної комерції / А. Берко, В. Висоцька, І. Рішняк // *Вісник Нац. ун-ту “Львівська політехніка”.* – 2007. – № 591. – С. 81–87. 31. Висоцька В. Система опрацювання структури електронного підручника / В. Висоцька // *Вісник Нац. ун-ту “Львівська політехніка”.* – Львів 2003. – № 489. – С.49–63. 32. Vysotska V. Analytical methods for commercial web content processing of information resource in electronic business systems / V. Vysotska, L. Chyrun // *MEST Journal.* – Vol. 2, No. 2. – P. 285–300. 33. Vysotska V. Features of the content-analysis method for text categorization of commercial content in processing online newspaper articles / V. Vysotska, L. Chyrun // *Applied Computer Science. ACS journal.* – Vol. 11, Is. 1. – Poland, 2015. – P. 5–19.