УДК 004.738.5+004.652

Висоцька В. А., Чирун Л. Б., Чирун Л. В.

АНАЛІЗ МЕТОДІВ ТА ЗАСОБІВ ОПРАЦЮВАННЯ ТЕКСТОВОЇ ІНФОРМАЦІЇ В СИСТЕМАХ УПРАВЛІННЯ КОНТЕНТОМ

У статті автори пропонують застосовувати метод контент-аналізу текстової інформації в системах електронної комерції для автоматизації електронного бізнесу і підтримки прийняття рішень модератором.

Ключові слова: *контент, контент-аналіз, електронна комерція, кількісний та якісний контент-аналіз, функціональні сервіси керування контентом.*

В статье авторы предлагают применять метод контент-анализа текстовой информации в системах электронной комерции для автоматизации електронного бизнеса и поддержки принятия решений модератором

Ключевые слова: *контент, контент-анализ, электронной комерции, количественный и качественный контент-анализ, функциональные сервисы управления контентом.*

In the given article authors proposed to apply the technique of content analysis of textual information in electronic commerce systems for automation of e-business and decision making by moderator.

Key words: *content, content analysis, electronic commerce, quantitative and qualitative content analysis, content management interoperability services.*

Вступ

Рівень складності завдань, що вирішуються за допомогою систем електронної комерції, постійно зростає [1-4]: від автоматизованого збирання, інтерпретації та реалізації інформації/контенту до управління, проектування, моделювання і прогнозування різноманітних бізнес-процесів. Але більшість складних завдань вирішуються не системами електронної комерції автоматично, а особами (модераторами), що супроводжують процес роботи відповідної системи. Виникла необхідність пошуку нетрадиційних підходів використання інформаційних технологій та математичних методів в процесах підготовлення й прийняття рішень в системах електронної комерції на основі отриманої інформації від форумів, коментарів користувачів продукції, електронного листування, результатів роботи пошукових систем та агентів. Більшість контенту, який отримує система електронної комерції від користувачів, опрацьовується модераторами. Вивчення інформаційних потреб для функціонування аналогічних систем свідчить, що в процесі прийняття рішень поряд із використанням фактографічних матеріалів, важливим стає залучення текстової інформації та автоматичне її опрацювання за короткий проміжок часу з мінімальними фінансовими затратами [4-8].

Актуальність теми зумовлена наступними чинниками: швидкі темпи зростання потреб в достовірній інформації для ведення бізнесу; необхідність формування множини оперативної інформації (контенту) для автоматизації процесів електронного бізнесу та автоматичної фільтрації небажаної інформації на сайтах; нерівномірність функціонування бізнес-процесів відповідно до регіонів (країн, областей тощо); зменшення часу опрацювання інформаційних ресурсів системами електронної комерції (СЕК); необхідність автоматизованого збирання, інтерпретації та реалізації контенту; необхідність в автоматизованому управлінні, проектуванні, моделюванні та прогнозуванні різноманітних бізнес-процесів; персоналізація у наданні послуг СЕК; інтегрованість систем електронної контент-комерції.

Мета дослідження – розроблення моделей та архітектурних принципів побудови систем керування контентом. Мета і задачі дослідження полягають в застосуванні аналітичного методу опрацювання інформаційних ресурсів системи електронної комерції у вигляді текстової інформації для можливості подальшого керування бізнес-процесами. Мета роботи визначає необхідність розв’язання наступних задач: здійснити аналіз функціонування і класифікувати системи управління контентом; розробити моделі управління контентом як складової частини систем електронної комерції; розробити архітектурні принципи побудови систем управління контентом. *Об’єкт дослідження* – моделі та процеси управління бізнес-процесами та методи моделювання і проектування архітектури систем управління контентом. *Предмет дослідження* – структура систем управління контентом і зв’язки, що зумовлюють взаємодію внутрішніх елементів системи.

Аналіз останніх досліджень та публікацій

Корпорації EMC, IBM і Microsoft, до яких приєдналися компанії Alfresco, Open Text, Oracle і SAP, розробляють специфікації Content Management Interoperability Services (CMIS) на інтерфейс Web-сервісів, що покликані забезпечити взаємодію між системами електронного документообігу [1-3]. *Цілі стандарту*: забезпечення роботи нових застосувань із наявними репозиторіями (сховищами архівних даних) і накопиченим в них контентом; забезпечення незалежного управління контентом різних репозиторіїв за допомогою Web-сервісів; забезпечення загальних Web-сервісів та інтерфейсів Web 2.0 для спрощення розроблення нових застосувань; створення платформи, яка є незалежною від мовної приналежності контенту; підтримка розроблення нового контенту, що складається із контенту декількох джерелі і який виглядає як єдине ціле.

Інструментарій управління контентом характеризується поглибленою спеціалізацією (рис. 1). Система документообігу забезпечує колективну роботу з документами, частина з яких може стати документами суворої звітності та потрапити в іншу категорію, і які є необхідними для надання зовнішнім організаціям (податковій інспекції, судовим органам тощо). Для зберігання і відстеження життєвого циклу документів суворої звітності потрібна система, відмінна від системи документообігу.

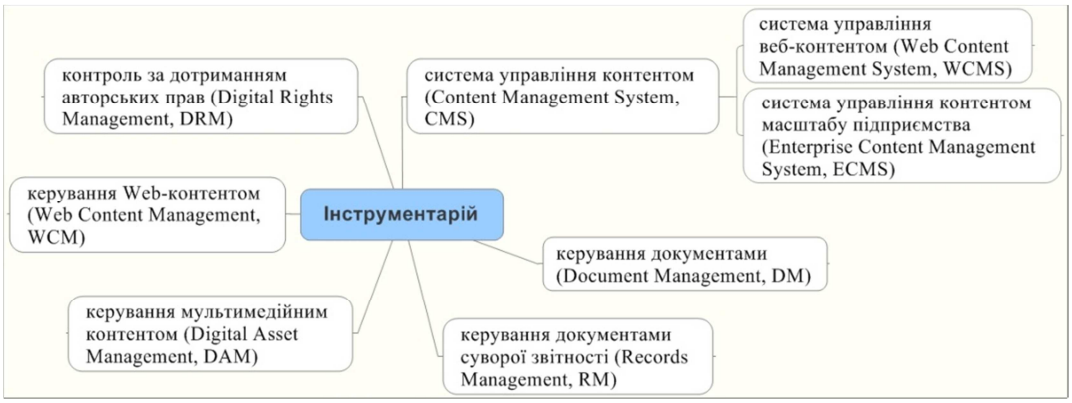


Рис. 1. Інструментарій управління контентом

Застосування Interoperable Content Application, що взаємодіє з контентом із різних репозиторіїв, може бути розміщеним на будь-якій платформі та здійснювати цю взаємодію за допомогою сервісного інтерфейсу та спеціальної надбудови (CMS), яка розробляється кожним учасником CMS для своїх репозиторіїв самостійно. Основні концепції CMIS 1.0, що лежать в основі репозиторію, визначені в [4].

CMIS не визначає налаштування параметрів безпеки – аутентифікація залишається за кожним репозиторієм ECM, за протоколами, що використовуються застосуваннями для взаємодії з конкретним репозиторієм за допомогою сервісів CMIS, і застосуваннями, що використовують ці протоколи для визначення того, як будуть визначені та верифіковані ідентифікатори користувачів. Всі сервіси CMIS допускають здатність конкретного репозиторію ідентифікувати користувачів за допомогою вказаних протоколів. Стандарт не дає жодного механізму для того, щоб вказати застосуванню, які операції користувач може виконувати над об’єктом в репозиторії за допомогою інтерфейсів CMIS чи іншим способом. За конкретним репозиторієм залишено право визначати на основі ідентифікації користувача чи буде його звернення задоволено, чи ні. CMIS – це єдиний сервіс, що дозволяє будь-якому застосуванню визначити, які операції, з числа визначених стандартом, можуть виконуватися в поточному (від імені конкретного користувача) контексті над конкретним об’єктом. Не передбачає стандарт єдиної системи безпеки, – спочатку необхідно побудувати систему, а потім запровадити систему безпеки із розмежуванням повноважень з доступу до конкретних даних. Спроба стандартизації в тому вигляді, яка існує в CMIS, не вирішує всіх проблем інтеграції.

Виділення проблем та формулювання мети

Вивчення динаміки потоку та побудова моделей керування потоком комерційного контенту є важливими і залишаються недослідженим [1-2]. При розгляді динаміки тематичних потоків контенту у рамках логістичної моделі виявлена обмеженість моделі, що відкриває шлях для подальших досліджень [2-3]. Під час дослідження механізмів породження текстової інформації виявлено, що від того, як побудоване ймовірнісно-лінгвістичне випробування та організовано вибірку з текстової інформації окремих його одиниць, залежить вибір тієї чи іншої моделі опису текстової інформації [3]. Ймовірнісне моделювання текстової інформації та його складових є вступним, підготовчим етапом опису функцій лінгвістичних одиниць в тексті. Вивчення функціонування мови та мовлення за допомогою ймовірнісного моделювання текстової інформації спирається на моделі теорії ймовірності, математичної статистики та комбінаторики. У лінгвістичних дослідженнях, особливо під час реалізації алгоритмів інформаційного пошуку, постійно виникають завдання, пов’язані з прогнозуванням появи в сегменті заданої довжини певної кількості словоформ чи словосполучень, що належать до певних класів. Ймовірнісне моделювання тексту. складів, словосполучень, граматичних класів дозволяє також визначати обсяг вибірки, необхідної для забезпечення із заданою ймовірністю появи хоча б одного разу відповідної лінгвістичної одиниці [6].

Одним із відомих методів аналізу текстової інформації є контент-аналіз – стандартна методика дослідження, предметом якої є аналіз змісту текстових масивів та продуктів комунікативної кореспонденції (наприклад, коментарі, форуми, електронне листування, статті тощо). Контент-аналіз – це кількісно-якісний аналіз текстової інформації та текстових масивів з метою подальшої змістовної інтерпретації отриманих кількісно-якісних закономірностей. Контент-аналіз застосовують при дослідженні джерел, інваріантних за структурою або змістом, але які існують як не систематизований, безладно організований текстовий матеріал [1; 2; 5]. Метод контент-аналізу полягає у формуванні з різного текстового матеріалу абстрактної моделі змісту тексту. Існує два типи контент-аналізу: кількісний і якісний.

Кількісний контент-аналіз (змістовний) – метод дослідження слів, тем та повідомлень, який зосереджується на змісті повідомлення. Отже, перед проведенням аналізу обраних лінґвістичних одиниць, необхідно передбачувати їх зміст та визначати кожен можливий результат спостереження у відповідності із очікуванням дослідника. В якості першого кроку при проведені контент-аналізу цього типу дослідник має створити словник, в якому кожне спостереження отримає визначення та буде віднесено до відповідної категорії.

Якісний контент-аналіз (структурний) – метод дослідження, в якому дослідника цікавить не так зміст повідомлення, як форма та структура цього повідомлення.

Для автоматизованого опрацювання текстової інформації велике значення має не тільки частота появи тієї чи іншої категорії лінгвістичної одиниці, а чи взагалі присутня вона в досліджуваному тексті. Кількісний підрахунок дозволяє зробити об`єктивні висновки щодо спрямованості матеріалів за кількістю вживання одиниць аналізу (ключових цитат) в досліджуваних текстах. Якісний аналіз дозволяє зробити об`єктивні висновки, чи зустрічається певна лінґвістична одиниця в тексті та в якому контексті.

Аналіз отриманих результатів

Система електронної контент-комерції (СЕКК) – програмне забезпечення для автоматизованої підтримки процесів електронної контент-комерції, є різновидом СЕК. Інформаційний контент є важливим фактором ведення електронного бізнесу. Комерційний контент – це об'єкт купівлі/продажу між учасниками електронної комерції та є різновидом економічного контенту (рис. 3).

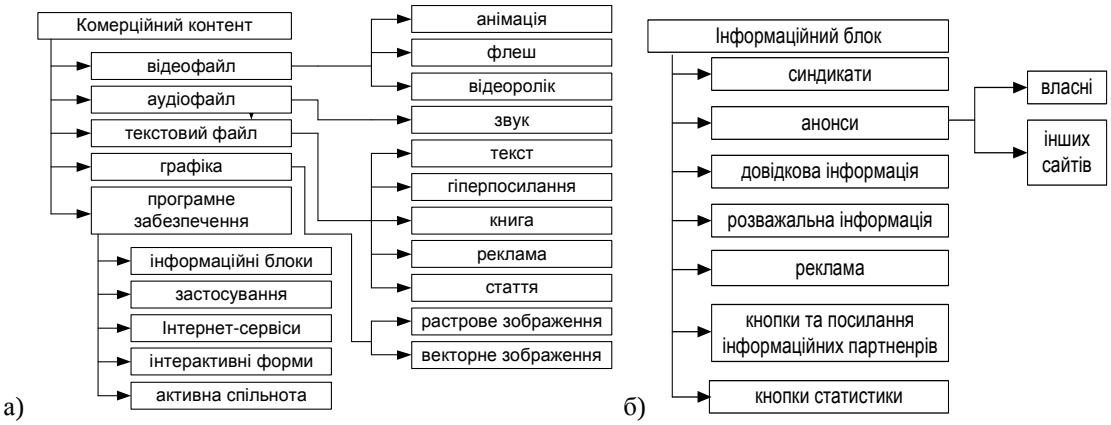


Рис. 3. Класифікація: а) комерційного контенту та б) інформаційних блоків

Система управління контентом (англ. Content management system, CMS) – це система, яка відповідає певному набору вимог (рис. 4а) та призначена для управління контентом. Такі системи використовуються для зберігання і публікації великої кількості документів, зображень, музики або відео та дозволяють управляти текстовим і графічним наповненням, надаючи користувачеві зручні інструменти зберігання і публікації інформації. На рис. 4б подана класифікація CMS.

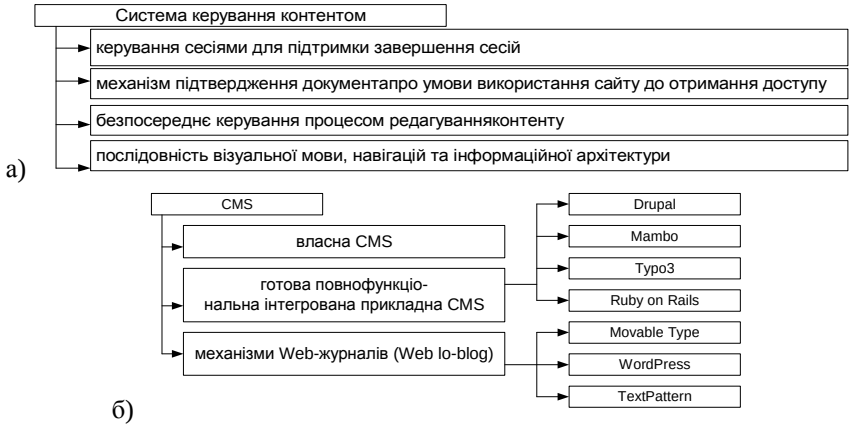


Рис. 4. а)Вимоги та б) класифікація систем управління контентом.

Системи управління Web-контентом розробляються для генерації та аналізу контенту всередині порталів (рис. 5), які мають ті ж проблеми (динамічне збирання, кешування контенту, безпека тощо), що й інші різновиди Web-застосувань. Цінність контенту визначає його привабливість для споживача. Інтеграція контенту робить портали привабливими, а інтеграція застосувань – корисними.

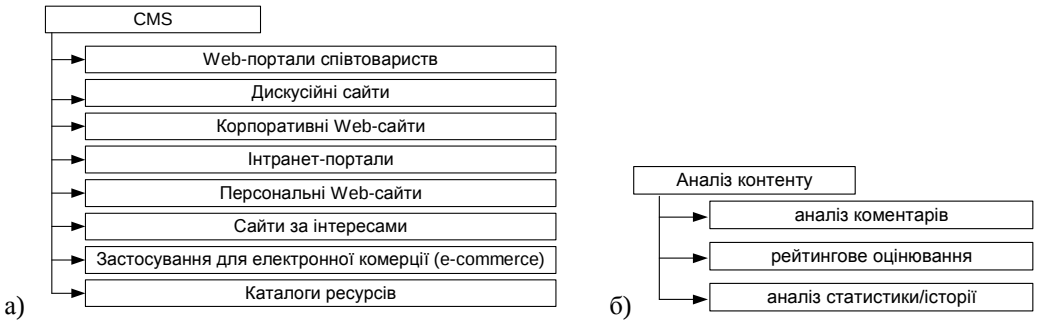


Рис. 5. а) Призначення CMS та б) методи аналізу контенту

Оскільки користувачі виконують на порталі щораз більшу кількість застосувань, то ці застосування (включаючи системи електронного документообігу), повинні виконуватися на самому порталі. Адміністративна панель системи управління Web-сайтом дозволяє змінювати або додавати нову інформацію для різних мовних версій Web-сайту. Зміни на сайті відображаються відразу ж після внесення і збереження. CMS Web-сайту представляють собою інструменти (табл. 1), що дозволяють моделювати розгалужені структури сайтів і керувати контентом.

Таблиця 1

Вимоги до ПЗ для CMS

CMS	Web-сервер	База даних	Мова
Ruby on Rails	Apache, FastCGI	MySQL, PostgreSQL, SQLite, Oracle, SQL Serer, DB2, Firebird	Ruby
Drupal	Apache IIS	MySQL, PostgreSQL	PHP
Mambo	Apache IIS	Apache IIS	PHP
Typo3	Apache IIS	Apache IIS	PHP
Movable Type	Apache IIS, Jetty, Tomcat	Apache IIS, Jetty, Tomcat	Perl
Word Press	Apache, mod_rewrite	Apache	PHP
Text Pattern	Apache	Apache	PHP

Базова система управління сайтом включає наступні можливості: швидке поновлення інформації на Web-сайті; пошук інформації на сайті; збирання даних про клієнтів та потенційних клієнтів; формування та редагування опитувань; аналіз відвідування Web-сайту.

Використання CMS сайту не вимагає встановлення спеціального програмного забезпечення на робочому місці. Для редагування й адміністрування використовується звичайний браузер (Internet Explorer або аналогічний). Інтуїтивний інтерфейс та простота роботи із CMS дозволяє полегшити управління сайтом і знизити подальші витрати на підтримку Web-ресурсу. Для роботи із системою необхідно мати лише початкові навички роботи в мережі Інтернет. Найпопулярнішою CMS є Drupal (табл. 2). Порівняємо Drupal із деякими іншими системами: Mambo, Typo3, Ruby on Rails, Movable Type, WordPress і TextPattern [1].

Таблиця 2

Порівняльні характеристики CMS

Вимоги	Ruby on Rails	Drupal	Mambo	Typo3	Movable Type	Word Press	Text Pattern
Простота встановлення	+/-	+/-	+	-	+	+	+
Крива навчання	+/-	+/-	+/-	-	+/-	-	-
Керування сесіями	+/-	+	+/-	+	-	-	-
Керування користувачами	+/-	+	+	+	+/-	+/-	+/-
Розширюваність	+	+	+/-	+	+/-	+/-	-
Маштабованість	+	+	+	+	+	-	-
Здатність використовувати теми	+/-	+	+/-	+/-	+/-	+/-	+/-
xHTML/CSS	+	+	+	-	+	+	+

Розглянемо управління контентом (рис. 6).

Рис. 6. Процес управління контентним потоком.

Добування контенту

Етапи добування контенту (рис. 7) реалізуються у вигляді комплексів контент-моніторингу, що виконують основну роботу зі збирання контенту з джерел (БД, СД, Інтернет тощо) наперед визначеними методами (рис. 8), і забезпечують створення документальних сховищ відповідно до інформаційних потреб споживачів.

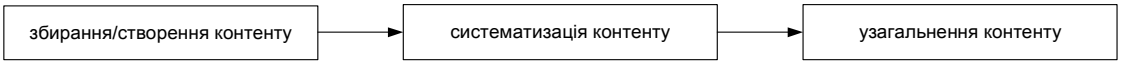


Рис. 7. Етапи добування контенту

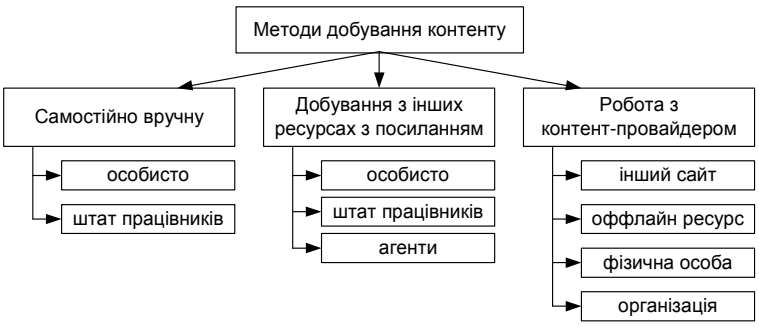


Рис. 8. Методи добування контенту

В результаті збирання і первинного опрацювання контент приводиться до єдиного формату, класифікується відповідно до визначеного рубрикатора та контенту приписується ряд дескрипторів, включаючи ключові слова. Добування контенту забезпечує постійне поповнення сховища оперативними повідомленнями, ефективний одночасний доступ до БД багатьма користувачами, зручні засоби пошуку необхідного контенту.

Ієрархічний принцип організації та функціонування добування контенту має певні переваги: економія ресурсів, необхідних для адміністрування (не треба сканувати й опрацьовувати дані безпосередньо з Інтернет); анонімність (при скануванні сайтів визначаються адреси робота-сканера); істотна економія Інтернет-трафіка співвідношення обсягів сканованого і вихідного контенту складає 50:1); не заперечується можливість самостійного сканування Інтернет.

Управління контентом

Аналіз та управління контентом сайту (рис. 9), їх моделювання сьогодні стає одним з найбільш інформативних методів кількісного вивчення динаміки окремих тематичних напрямів та проведення технічного аналізу сайту. Зміна величин управління контентом свідчить про швидкість розвитку як окремих тематичних напрямів, так і всього контентного простору. Стійкі статистичні зв'язки між окремим контентом свідчать про кореляції окремих тематик, про ефективність посилань на публікації попередників, більш ранні роботи, цитування, републікації тощо.

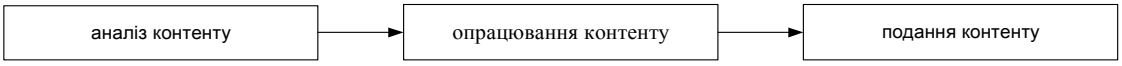


Рис. 9. Етапи управління контентом

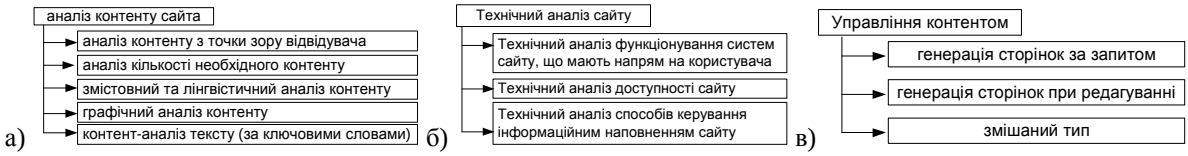


Рис. 10. Методи: а) аналізу контенту; б) технічного аналізу сайту та в) управління контентом

Механізми, що базуються на узагальнених методах кластерного аналізу, виявляють повідомлення в контентних потоках, що формують навколо себе нові тематичні напрями. Кластерний аналіз, теорія фракталів і автомодельних процесів при коректному застосуванні кількісно оцінюють ступінь зв'язку в тематичних контентних потоках.

Інтерпретація контенту

Актуальна проблема на фінансовому ринку з розповсюдження комерційного контенту – це відсутність загальних підходів для автоматичного аналізу комерційної діяльності відповідних компаній (маркетингових досліджень). Актуальність розроблення загальної архітектури системи управління комерційним контентом полягає у необхідності: отримувати оперативні та об'єктивні оцінки рівня конкуренції на сегменті фінансового ринку комерційного контенту; оцінювати рівень конкурентів та міру їх конкурентоспроможності на фінансовому ринку з розповсюдження комерційного контенту.

Отримані дані повинні враховуватися при створенні майбутнього Web-порталу або при оновленні вже існуючого. Модуль інтепретації контенту – це система покращення процесу підвищення рейтингу системи електронної контент-комерції з інтерактивним користувацьким інтерфейсом (рис. 11).

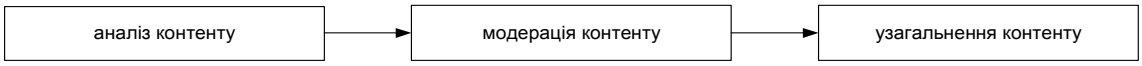


Рис. 11. Етапи інтерпретації контенту

Основні задачі, які покладаються на модуль інтерпретації контенту, є наступними: збирання маркетингової інформації з поширення комерційного контенту; підтримка діалогу з кінцевим споживачем комерційного контенту; формування каталогу комерційного контенту та послуг систем електронної контент-комерції; інформаційна підтримка кінцевого споживача контенту; формування віртуального кошика клієнта комерційного контенту; реєстрування кінцевого споживача комерційного контенту; опрацювання замовлення на комерційний контент.

Використання контент-аналізу текстової інформації в системах електронної комерції дозволяє визначити поширеність тієї чи іншої ознаки досліджуваної сукупності текстів. При цьому важливе не стільки абсолютне, скільки відносне значення ознаки, тобто характеристика її місця (частки) серед інших ознак. Основні етапи застосування контент-аналізу текстової інформації в системах електронної комерції:

1. Визначення сукупності джерел, що досліджуються, або повідомлень у відповідності до заданих критеріїв, яким відповідає кожне повідомлення: тип джерела (форум, електронна пошта, Інтернет-газета, чат, Інтернет-журнал); тип повідомлень (стаття, електронний лист, баннер, коментарій); сторони, що беруть участь в процесі комунікації (відправник, одержувач, реципієнт); співставлений розмір повідомлень (мінімальний обсяг або довжина); частота появи повідомлень; спосіб розповсюдження повідомлень; місце розповсюдження повідомлень; час появи повідомлень тощо.

2. Формування обмеженої вибіркової сукупності повідомлень.

3. Виявлення лінґвістичних одиниць аналізу. Існують чіткі вимоги до вибору можливої лінґвістичної одиниці аналізу: достатньо велика для інтерпретації значення; достатньо мала, щоб не інтерпретувати багато значень; легко ідентифікується; кількість одиниць достатньо велика для проведення вибірки. У разі прийняття за одиницю аналізу теми, враховують такі правила: розмір теми не виходить за межі абзацу; нова тема виникає при зміні мети теми, призначення, категорії та особи, для якої створюється тема.

4. Виділення одиниць обчислення, які можуть збігатися зі змістовними одиницями або носити специфічний характер. У першому випадку процедура аналізу зводиться до підрахунку частоти появи виділеної змістовної одиниці, в іншому – дослідник на основі аналізованого матеріалу і цілей дослідження висуває одиниці обчислення, якими можуть бути: фізична протяжність текстів; площа тексту, заповнена змістовними одиницями; кількість рядків (абзаців, знаків, колонок тексту); розмір та вид файлу; кількість рисунків з певним змістом, сюжетом тощо.

В деяких випадках дослідники використовують й інші елементи обчислення. Принципове значення на цьому етапі контент-аналізу має строге визначення його операторів.

5. Безпосередньо процедура обчислення. В загальному вигляді схожа на стандартні прийоми класифікації за виділеними угрупованнями з формул математичної статистики та теорії ймовірності. Існують також спеціальні процедури підрахунку стосовно контент-аналізу.

6. Інтерпретація отриманих результатів відповідно до цілей і завдань конкретного дослідження. На цьому етапі виявляються і оцінюються такі характеристики текстового матеріалу, які дозволяють робити висновки про те, що хотів підкреслити або приховати його автор.

Застосування контент-аналізу текстової інформації в системах електронної комерції, на думку авторів, має ряд переваг для спрощення ведення бізнесу та вирішує низку проблем, що стоять перед учасниками бізнес-процесів, це: автоматизація процесу фільтрації текстової інформації, яку розміщує користувач на сайті системи електронної комерції; можливість автоматичного створення «портрету» постійного користувача на основі його коментарів; можливість автоматичного створення «портрету» цільової аудиторії на основі аналізу «портретів» постійних користувачів; скорочення кількості модераторів, які обслуговують систему електронної комерції; скорочення часу для розміщення текстової інформації постійного користувача на сайті за рахунок автоматичного опрацювання цієї інформації та відсутності проміжної ланки у вигляді модератора; ліквідація мовного бар'єру за рахунок автоматичного формування словників постійного користувача та використання систем автоматизованого перекладу.

Контент-аналіз розподілу текстової інформації користувачів дає можливість якісно оцінити контентний потік в системах електронної комерції для подальшого прийняття рішень відповідною особою. З метою залучення більшої кількості клієнтів на сторінки системи електронної контент-комерції додають опції аналізу контенту (рейтинги, відгуки і коментарі відкористувачів) про комерційний контент (стаття, книга тощо). Сайт із відгуками про комерційний контент привертає більшу кількість відвідувачів, але дана функціональність не включає модерації, що, в свою чергу, призводить до надмірної інформації на вітрині із представленим контентом. Кінцевий споживач відвідавши таку вітрину вимушений сегментувати

конструктивну інформацію, що вимагає зусиль і тим самим відштовхуває потенційного клієнта. Контент, який потрапляє на сайт, не фільтрується, тобто взаємодія людей із порталом є односторонньою.

СЕКК має наступні можливості покращення бізнесу:

1) підняття рейтингу комерційного контенту;

2) аналіз характеристик (коментарі, відгуки, побажання тощо), вплив на комерційний контент зі сторони споживача;

3) збирання, накопичення та опрацювання інформації про потреби кінцевого споживача комерційного контенту.

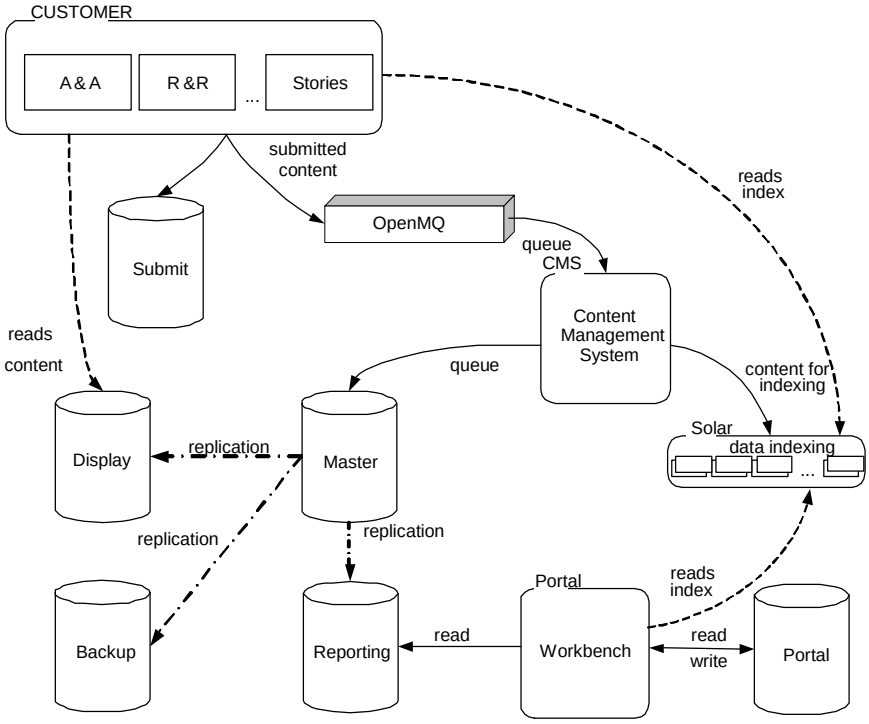


Рис. 12. Схема потоків даних в модулі інтерпретації контенту

Основні компоненти модуля інтерпретації контенту (рис. 12):

– Customer – потенційний/існуючий клієнт із власним сайтом, який купує аплікацію з метою підняття рейтингу власного товару та кращого розуміння потреб кінцевого користувача.

– A&A, R&R, Stories – типи контенту (питання/відповіді, рейтинги, відгуки, статті).

– OpenMQ – черга типу FIFO.

– CMS – модуль фільтрування контенту за наперед визначеними правилами, індивідуально для кожного клієнта.

– Workbench – індивідуальний кабінет клієнта з можливістю перегляду та аналізу статистики відносно своєї продукції.

– Solr – модуль індексування для швидкого пошуку контенту.

Бази даних модуля інтерпретації контенту (рис. 12).

– Master – головна БД, яка є основною ланкою системи, для розподілу навантаження реплікує дані на дочірні бази.

– Display – дочірня БД від Master. З нею працює сайт клієнта на зчитування того контенту, який має дозвіл на публікацію.

– Submit – попереджує дублювання інформації на першому етапі.

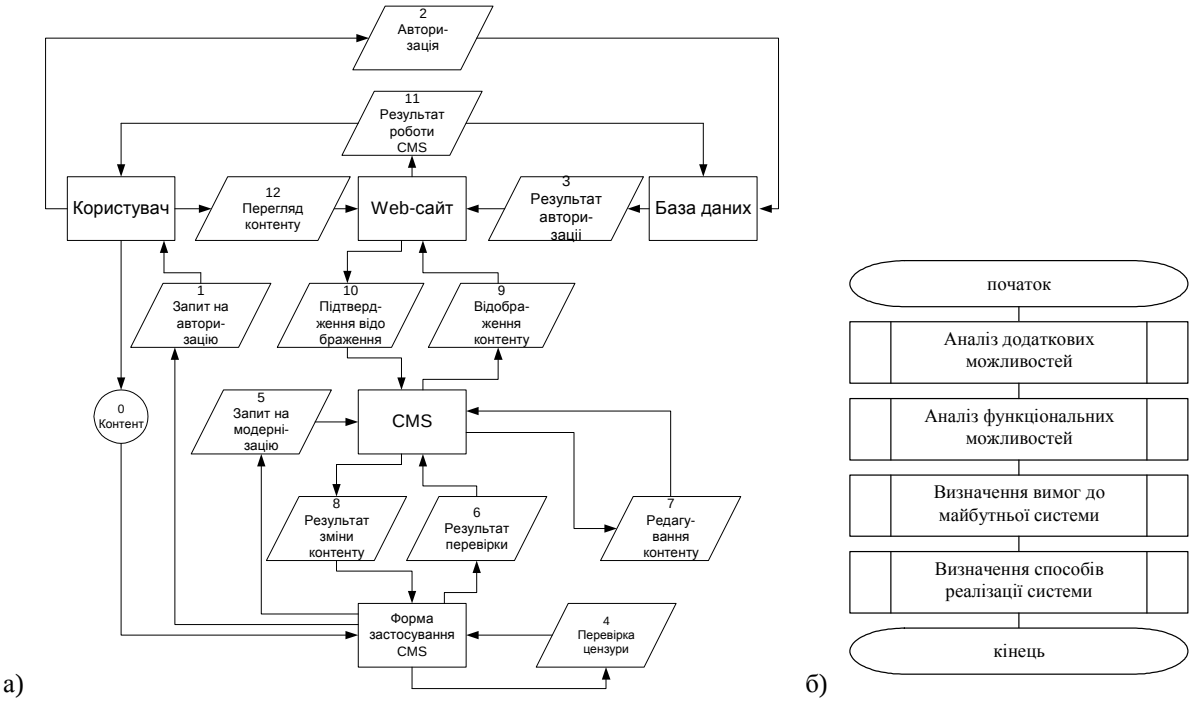
– Reporting – на її основі формуються статистичні звіти для клієнтів.

– Backup – додаткове СД.

– Portal – БД призначена для роботи з Workbench аплікацією.

Для підвищення попиту на комерційний контент необхідно викладати на сайті лише відфільтрований контент.

Запропонована модель інтерпретації контенту дозволяє: будувати моделі спілкування з кінцевим споживачем; керувати динамічним матеріалом, що генерується кінцевими споживачами; а також – кінцевим споживачам «будувати» бізнес клієнта. Аналізом, вибіркою, побудовою статистичних даних та модерацією контенту (рис. 13) займається CMS на етапі інтерпретації контенту за допомогою методу контент-аналізу текстової інформації, правил (табл. 3) та алгоритмів модерації.



а) б)

Рис. 13. а) Схема модерації контенту та б) процес аналізу сайту

Табиця 3

Правила роботи модуля інтерпретації контенту із самонавчанням

Правило (фільтр)	Характеристика	Режим роботи
нецензурної лексики	Перевірка за словником та заборона у разі TRUE	Напівавтоматичний, у разі FALSE остаточно перевіряє модератор, словник періодично поповнюється модераторами
за кількістю символів	Підрахунок кількості символів, порівняння із лімітом та заборона у разі FALSE	Автоматичний, ліміт визначає адміністратор або власник сайту
за наявністю URL-посилань	Визначення наявності URL-посилань, перевірка за списком та заборона у разі TRUE	Напівавтоматичний, у разі FALSE остаточно перевіряє модератор, список періодично поповнюється модераторами
за «blacklist» списком	Перевірка за списком та заборона у разі TRUE	Напівавтоматичний, у разі FALSE остаточно перевіряє модератор, список періодично поповнюється автоматично у разі перевищення ліміту
за рейтингом	Підрахунок рейтингу за визначений період, порівняння із лімітом та заборона у разі FALSE	Автоматичний, період і ліміт визначає адміністратор або власник сайту
за наявністю посилань на конкурентів	Перевірка за списком та заборона у разі TRUE	Напівавтоматичний, у разі FALSE остаточно перевіряє модератор, список періодично поповнюється модераторами
за IP адресами	Перевірка за списком та заборона у разі TRUE	Автоматичний, список періодично поповнюється автоматично у разі перевищення ліміту, ліміт визначає адміністратор або власник сайту
за ID користувачів	Перевірка за списком та заборона у разі TRUE	Автоматичний, список періодично поповнюється автоматично у разі перевищення ліміту, ліміт визначає адміністратор або власник сайту

Висновки

Незважаючи на різноманіття класів СЕКК та широкі можливості реалізації моделей, виділено основні закономірності переходу від побудови концептуальної моделі об’єкта моделювання до проведення експерименту з моделлю системи, для ефективного вирішення користувачем практичних завдань моделювання раціонально оформити у вигляді методики розроблення й процесів реалізації моделей. Найбільш істотним фактором, який варто враховувати при формалізації й алгоритмізації моделей, є використання інструменту дослідження апаратно-програмних засобів ІТ. При визначенні етапів моделювання СЕКК є необхідним залучення колективу розробників різних спеціальностей (аналітиків, алгоритмістів, програмістів). Ступінь розвитку Web-простору буде визначатися технологіями роботи з величезним обсягом інформації, що нагромаджується в Інтернет. Web наступного покоління буде характеризуватися переходом від мережі документів до мережі даних, що при необхідності агрегуються в семантично пов’язані за допомогою Web-сервісів документи. Перспективи використання інформаційного простору, очевидно, будуть залежати від створення і розвитку ефективної інфраструктури, у рамках якої будуть працювати програмні продукти з боку Web-серверів і користувачів.

ЛІТЕРАТУРА

1. CMIS – Content Management Interoperability Services [Електронний ресурс]. – Режим доступу : http://craigrandall.net/archives/2008 09/cmis/. – Назва з титул. екрану.
2. Content Management Interoperability Services (CMIS). Extending Office Document Management [Електронний ресурс]. – Режим доступу : http://www-01.ibm.com/software/data/content-management/cm-interoperablity-services.html. – Назва з титул. екрану.
3. Content Management Interoperability Services. Version 0.5 [Електронний ресурс] // Part I – Introduction, General Concepts, Data Model, and Services. EMC – Corporation, IBM Corporation, Microsoft Corporation. All Rights Reserved. 8/28/2008. CMIS Part I – Domain Model v0.5.pdf. – P. 76.
4. Алексеев А. Н. Контент-анализ в социологии и точки соприкосновения с другими отраслями знания / А. Н. Алексеев // Проблемы контент-анализа в социологии : М-лы Сибирского социологического семинара. – Новосибирск, 1970. – С. 11–12.
5. Іванов В. Ф. Контент-аналіз: Методологія і методика дослідження ЗМК : навч. посібник / В. Ф. Іванов ; [наук. ред. А. З. Москаленко]. – К., 1994. – 112 с.
6. Пиотровский Р. Г. Математическая лингвистика : учебное пособие / Р. Г. Пиотровский, К. Б. Бектаев, А. А. Пиотровская. – М. : Высшая школа, 1977. – 384 с.
7. Сорока М. Б. Використання методу контент-аналізу при створенні автоматизованих інформаційних систем / М. Б. Сорока, Н. В. Танатар // Бібліотека. Наука. Культура. Інформація: Наукові праці НБУВ. – 1998. – Вип. 1. – С. 318–322.
8. Федорчук А. Г. Контент-мониторинг информационных потоков [Електронний ресурс] // Б-ки нац. акад. наук: пробл. функционирования, тенденции развития. – Электрон. дан. (1 файл). – К., 2005. – Вып. 3. – Режим доступа : http://www.nbuv.gov.ua/articles/2005/05fagmip.html. – Назва з екрана.

Рецензенти: Фісун М. Т., д.т.н., професор;
Кондратенко Ю. П., д.т.н., професор

© Висоцька В. А., Чирун Л. Б., Чирун Л. В., 2010

Стаття надійшла до редколегії 08.12.2010 р.