УДК 004.738.5

В. В. Кучковський, В. А. Висоцька, С. З. Нитребич, Р. М. Оливко
Національний університет “Львівська політехніка”,
кафедра інформаційних систем та мереж

ЗАСТОСУВАННЯ МЕТОДІВ ІНТЕРНЕТ-МАРКЕТИНГУ ДЛЯ АНАЛІЗУ WEB-РЕСУРСІВ В МЕЖАХ РЕГІОНУ

© Кучковський В. В., Висоцька В. А., Нитребич С. З., Оливко Р. М., 2015

Запропоновано методи аналізу контенту для інтернет-газети. Модель описує процеси опрацювання інформаційних ресурсів у системах аналізу контенту та спрощує технологію автоматизації управління контентом. Проаналізовано основні проблеми синтаксичного та семантичного аналізу контенту та функціональних сервісів управління контентом. Швидкий темп росту глобальних мереж та електронної контент-комерції сприяють спрощенню процесу публікації та модерації статей, у результаті чого запропоновано інтелектуальну систему прийняття маркетингових рішень з поширення електронних видань. Розглянуто питання розроблення методів та програмних засобів опрацювання інформаційних ресурсів в інтернет-системах. Сформульовано новий підхід до застосування та впровадження бізнес-процесів для побудови таких систем. Розроблено методи опрацювання контенту та інформаційного ресурсу. Розроблено програмні засоби опрацювання контенту та інформаційного ресурсу.

Ключові слова: контент, інформаційний ресурс, контент-аналіз, бізнес-процес, система управління.

This article presents the content analysis techniques for online newspapers. The model describes the processing of information resources in content analysis and simplifies automation technology of content management. In this paper the basic problem of the syntactic and semantic analysis of content and functionality of content management services is analysed. The fast growth of global networking and online content-commerce simplified process of publishing and moderation of articles. As a result Marketing Intellectual System for Online Publishing was proposed. The article discusses the development of unified methods and software tools for processing information resources in the Internet systems. A new approach to application and implementation of business processes is formulated for the construction of these systems. The methods of content and information resource processing are developed. Software tools for content processing are developed.

Key words: content, information resource, content analysis, business-process, content management system, content lifecycle.

Вступ. Загальна постановка проблеми

Відомою технологією управління контентом є *інтернет-маркетинг* (англ. Internet marketing, online-маркетинг) з інтернет-інтеграцією, інформаційним менеджментом, Public Relation, службою роботи з покупцями та продажем у різних сферах [1, 3, 4]. *Інтернет-маркетинг* використовує всі аспекти та основні елементи традиційного маркетингу у поєднанні з новими методами дослідження (наприклад, вірусний маркетинг) та аналізу інформації за допомогою сучасних IT [1, 3, 4]. Ці методи з можливістю перебувати у відносно постійному контакті з користувачами є ефективними завдяки автоматичному відстеженню статистики, для аналізу якої використовують коефіцієнти окупності інвестицій (англ. Return on Investment, ROI), прибутковості електронного бізнесу (англ. Rate of Return, ROR), ефективності відвідин або конверсію сайту (англ. conversion rate). Основна мета інтернет-маркетингу – отримання максимального ефекту від потенційної аудиторії сайту з можливістю миттєвого одержання статистики продажів, відвідувань, попиту тощо [1, 3, 4].

Аналіз останніх досліджень та публікацій

Інтернет-маркетинг передбачає використання стратегій та напрямів традиційного маркетингу прямого відгуку та спеціальних напрямів дослідження, які застосовують до бізнесу Інтернет-простору (рис. 1, табл. 1). Інтернет-маркетинг – це не лише торгівля контентом, але й інформаційним простором, ПЗ, бізнес-моделями тощо [1, 3-4]. Компанії Google, Yahoo і MSN підняли на новий рівень і сегментували ринок інтернет-реклами, пропонуючи електронному бізнесу послуги з локальної реклами. Внаслідок автоматизації процесу дослідження аудиторії рентабельність інвестицій зростає, а витрати знижуються. Основні переваги інтернет-маркетингу: інтерактивність, можливість максимально точного таргетингу (табл. 2–4), можливість постклік-аналізу для максимального підвищення показників конверсії сайту і ROI/ROR інтернет-реклами [1, 3, 4].

Конверсія сайту – це відношення кількості відвідувачів сайту, що виконали на ньому цільові дії (приховані/прямі вказівки рекламодавців, продавців, творців комерційного контенту, тобто покупку, реєстрацію, підписку, відвідування певної сторінки сайту, перехід за рекламним посиланням), до загальної кількості відвідувачів сайту [1, 3, 4]. Успішна конверсія по-різному трактується продавцями (операція купівлі споживачем, який зацікавився продуктом, натиснувши на відповідний рекламний банер), рекламодавцями та постачальниками контенту (дії, очікувані від відвідувачів, наприклад, реєстрація відвідувачів на сайті, підписка на поштову розсилку, скачування ПЗ).

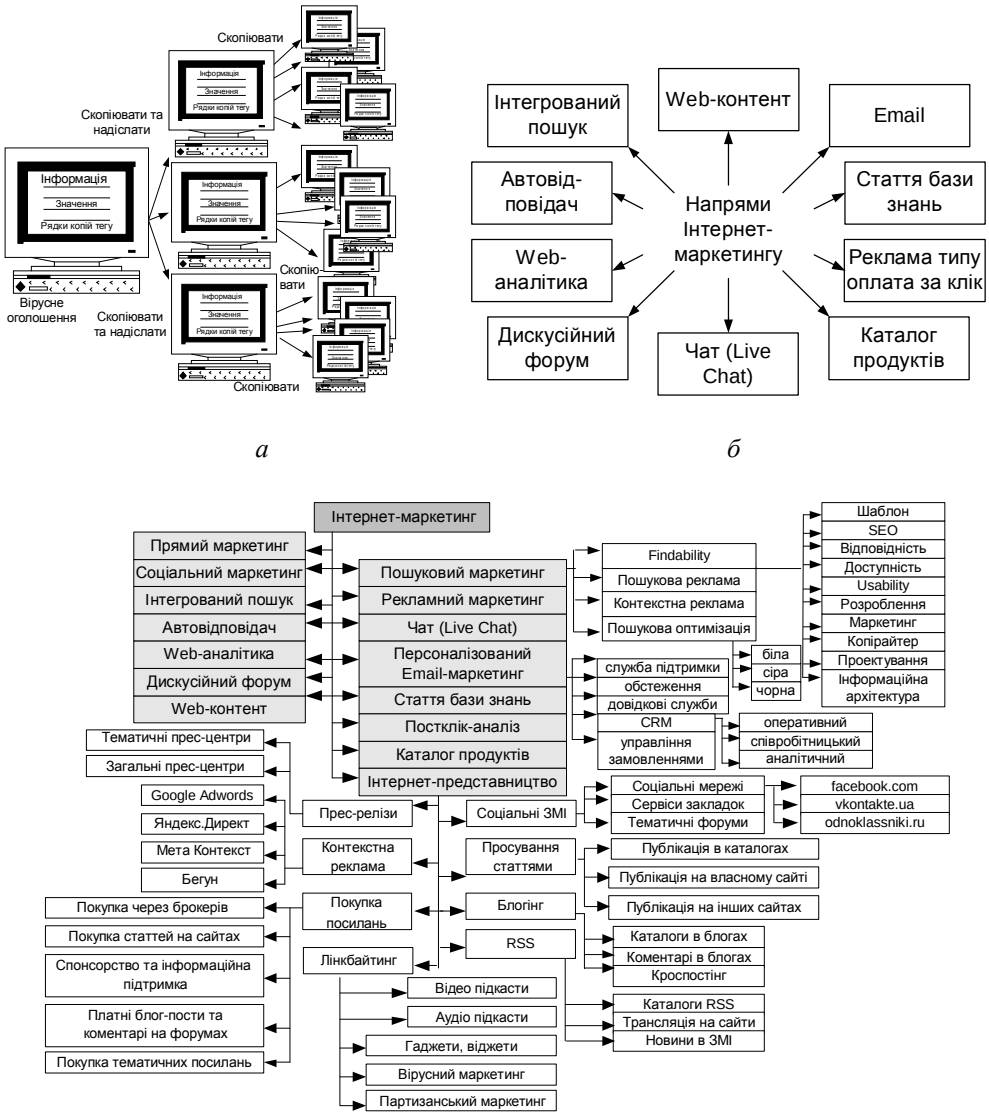


Рис. 1. Вірусний маркетинг (а); основні напрями інтернет-маркетингу (б); напрями інтернет-маркетингу (в)

Таблиця 1

Сфери впливу інтернет-маркетингу

Назва	Результат впливу Інтернет-маркетингу
Музична індустрія	У музичній індустрії багато споживачів почали купувати і завантажувати музику у форматі MP3 через Інтернет замість того, щоб купувати CD.
Банківська справа	Онлайн-банкінг є зручним для клієнта, оскільки позбавляє його необхідності відвідувати кожного разу банк або його філії.
Аукціони	Інтернет-аукціони завоювали популярність. Унікальні речі, які раніше можна було знайти на блошиних ринках, тепер продаються на онлайн-аукціонах, таких як eBay тощо. Розвиток аукціонів сильно вплинув на ціни на унікальні та антикварні речі.
Рекламний ринок	Вплив на рекламну індустрію був і залишається справді величезним. Впродовж декількох років обсяг онлайн-реклами нестримно зріс і сягає десятків мільярдів доларів на рік. Рекламодавці розпочинали активно змінювати свої пріоритети, і сьогодні інтернет-реклама вже займає велику ринкову нішу.
Ринок портативних пристроїв	Збільшення обсягів та географії продажу портативних електронних пристроїв (мобільні телефони, плеєри тощо) завдяки застосування сучасних маркетингових методів просування товару на Інтернет-ринок.
Е-комерція контенту	Збільшення обсягів та скорочення циклу виробництва/продажу нематеріальних товарів та надання відповідних послуг.
Інтернет-майданчики	Торговельні онлайн-майданчики вже давно перестали бути дошками оголошень, деякі з них перетворилися на великі корпорації, що надають низку маркетингових послуг. Зростають і ціни за участь на таких майданчиках (привілейоване членство), незважаючи на те, що кількість їх збільшується.

Таблиця 2

Основні елементи Інтернет-маркетингу

Назва	Англ.	Визначення
Товар	Product	Об’єкт, який продають за допомогою Інтернету через Web-сайт; має власну ціну та гідну якість, користується попитом, конкурує з аналогічними об’єктами інших Web-сайтів і традиційних магазинів.
Ціна	Price	Контрольований кількісний показник вартості товару з такими властивостями: на значно нижчий, ніж в звичайному магазині завдяки економії на витратах; регулярно порівнюваний із цінами на аналогічні конкуруючі товари.
Просу-вання	Promotion	Комплекс методів просування сайту в Інтернеті і товару, який містить величезний арсенал інструментів (пошукове просування, контекстна реклама, банерна реклама, e-mail маркетинг, партнерський (афіліативний, англ. affiliate) маркетинг, вірусний маркетинг, прихований маркетинг, інтерактивна реклама, робота з блогами тощо.
Місце продажів	Place	Точка продажів (Web-сайт), де велику роль відіграє графічний дизайн та використовуваність (юзабіліті, англ. Usability – дослівно “зручність користування”) сайту, якість опрацювання заявок із сайту, швидкість завантаження, робота з платіжними системами, умови доставляння, робота з клієнтами до, під час і після продажу.
Напрями марке-тингу	Marketing Trends	Процеси, спрямовані на створення/збільшення попиту та досягнення цілей електронного бізнесу за допомогою Інтернет-технологій через максимальне задоволення потреб споживачів товару/послуги.

Таблиця 3

Напрями інтернет-маркетингу

Назва	Характеристика
1	2
Медійна реклама (англ. Media advertising) або банерна реклама (англ. Banner advertising)	Ефективний спосіб популяризації (збільшення одсягу відвідувань) та потужний інструмент покращення іміджу власників Web-сайту. Статична/анімована графічна реклама (найчастіше Web-баннер розміром 468×60 пікселів), пов’язана через гіперпосилання із сайтом рекламодавця або сторінкою з додатковою інформацією. Разом із контекстною рекламою є одним із основних форматів реклами в Інтернет. Класичний Web-банер – графічний файл формату GIF, SWF або JPG із статичним/анімованим зображенням. Web-банери також створюють за технологіями Flash або Java, де на відміну від традиційних (растрова графіка) використовують векторну графіку для реалізації анімаційніх та звукових ефектів при невеликому розмірі, що підвищує ефективність банера як рекламного носія.

Продовження табл. 3

1	2
Постклік-аналіз (англ. Post-click marketing)	Ефективний спосіб підвищення продажів і результатів маркетингу при зосередженні уваги на відвідувачах Web-сайту, які відповідають на діяльність інтернет-маркетингу (оплата за клік по рекламі, HTML електронної пошти, пошук з метою підвищення конверсії).
Точний таргетинг (англ. target – мета)	Рекламний механізм (табл. 4), який виділяє із всієї наявної аудиторії лише ту частину, яка задовольняє заданим критеріям (наприклад, цільову аудиторію), та показує рекламу саме їй.
Контекстна реклама (англ. Content-targeted advertising)	Принцип розміщення реклами, орієнтовану на зміст інтернет-сторінки або сайту, вручну або автоматично у вигляді банера чи текстового оголошення з тематикою. Одна з найсильніших сторін контекстної реклами – геотаргетинг (англ. Geotargeting), можливість обирати географію показу сторінок. Також застосовується рамкові обмеження по часу показу.
Пошуковий маркетинг (англ. Search engine marketing, SEM)	Комплекс заходів, спрямований на збільшення відвідуваності з пошукових систем. Раніше під “результатом пошуку” розуміли тільки посилання на відповідні сторінки (зазвичай з лівого боку сторінки), зараз до нього зараховують і рекламні оголошення в правій частині сторінки. Більшість технологій пошукового маркетингу не дозволяють досягти цілей реклами: не завжди безпосередньо прямо повідомляють про рекламований товар або послугу; не завжди переслідують мету продажу; за їх допомогою складно підвищити впізнаваність бренду; неможливо вивести новий товар на ринок. Результатів використання будь-якої технології пошукового маркетингу може бути два: залучення користувачів на сайт, причому для кожного окремого випадку характеристики цієї аудиторії відрізняються; поширення інформації про сайт або компанію в пошукових системах.
Оптимізація сайту в соціальних медіа (англ. Social media optimization, SMO)	Оптимізація під соціальні медіа, комплекс суто технічних заходів, спрямованих на перетворення контенту сайту таким чином, щоб його можна було максимально просто використовувати в мережевих спільнотах (форумах, блогах).
Соціальний маркетинг (англ. Social media marketing, SMM)	Просування або піар будь-чого у соціальних медіа (блоги, форуми, мережеві спільноти).
Прямий маркетинг (англ. Direct marketing) з використанням e-mail, RSS тощо	Вид маркетингової комунікації, в основі якого – пряма особиста комунікація з одержувачем повідомлення (споживачем B2C або клієнтом середовища B2B) з метою побудови взаємовідносин та отримання прибутку. Прямий маркетинг оснований на ставленні до клієнта як до індивідуума, передбачає зворотній зв'язок і не використовує для комунікацій інформаційних посередників. У західному маркетингу прямий маркетинг зараховують до заходів BTL (англ. below-the-line, використовують на противагу ATL – прямій рекламі), який об'єднує промоакції, Direct e-mail, виставки, POS тощо. Директ-маркетинг часто використовує пряму адресну розсилку повідомлень представникам цільової аудиторії за допомогою поштового зв’язку, e-mail, SMS тощо. Прямий маркетинг орієнтований не на цільові групи, а на окремих індивідуумів. Проведення заходів прямого маркетингу вимагає створення бази даних клієнтів – структурованої, постійно поновлюваної свіжою інформацією про фізичних або юридичних осіб і про їх споживчі переваги та потреби, з метою подальшого опрацювання цієї інформації та формування пропозиції продукції, яка відповідає їхнім запитам. Бази даних у прямому маркетингу тісно пов'язані з CRM-концепцією управління взаєминами із споживачами або клієнтами, що дозволяє консолідувати всю інформацію про клієнта, а також впорядкувати всі стадії взаємовідносин з клієнтами від маркетингу і продажів до обслуговування після продажу.
Вірусний маркетинг (англ. Viral marketing and viral advertising)	Методичні принципи, що базуються на заохоченні індивідів до передавання маркетингового повідомлення іншим особам, які утворюють потенціал для експоненційного зростання впливу цього повідомлення. Як і віруси, подібні техніки використовують будь-яку сприятливу можливість для збільшення числа переданих повідомлень.
Партизанський маркетинг (англ. Guerrilla marketing)	Концепція маркетингу, яка спрямована на пошук фірмою/підприємцем своєї маркетингової ніші, відмовою від відкритого конкурування зі своїми потужними конкурентами, концентрації зусиль на відокремлених ділянках “фронту” та використанні нетрадиційних, але ефективних способів реклами та просування своїх товарів/послуг. Партизанському маркетингу властиві гнучкість та мобільність, іноді його називають малобюджетним/маловитратним.
Прицільний маркетинг (англ. Precision marketing)	Метод росту успішності Інтернет-маркетингу через технологію збереження існуючих клієнтів, крос-продажу (англ. cross-selling) та допродажу (англ. up-selling). Можливість проведення точного або прецизійного аналізу (англ. Precision analysis) поряд з ROI/ROR електронного бізнесу та conversion rate, що дозволяє миттєво одержувати статистику продажів, попиту тощо. Підкреслює актуальність комерційного контенту, яка досягається безпосередньо через особисті переваги відвідувачів сайту, збирання та аналіз поведінкових і транзакційних даних користувачів.

Таблиця 4

Види таргетингу

Назва	Визначення
Підбір рекламних майданчиків	Найпопулярніший вид таргетингу. Здійснюється шляхом підбору рекламних майданчиків так, щоб їх відвідувачі відповідали цільовій аудиторії.
Тематичний таргетинг	Показ тематичної реклами на Web-сайтах, які відповідають цій тематиці.
Орієнтування за інтересами	Показ реклами (контекстна реклама) відповідно до інтересів відвідувачів рекламного майданчика.
Географічний таргетинг (геотаргетинг)	Показ реклами цільовій аудиторії, яка обмежена деяким географічним регіоном, обраним рекламодавцем.
Орієнтування за часом показу	Дозволяє обмежити показ реклами по часу протягом дня (ранок/вечір, будні/вихідні), тижня, року.
Соціально-демографічний таргетинг	Показ реклами, тематика якої спрямована на певний клас цільової аудиторії, тобто поділ реклами згідно класифікації цільової аудиторії за віком, статтю, доходом, посадою і т. д.
Обмеження кількості показів одному користувачеві	Дозволяє регулювати кількість показів рекламного носія одному унікальному користувачеві в процесі його взаємодії з рекламною площадкою. Найчастіше застосовується в банерній рекламі з оплатою за 1000 показів.
Поведінковий таргетинг	Найперспективніший сьогодні напрямок, суть якого зводиться до впровадження механізму збору інформації про дії користувача в Інтернет за допомогою cookie-файлів. Інформація збирається у так званих профілях і містить дані про переглянуті сайти, пошукові запити, покупки в інтернет-магазинах тощо. Одержавши такий профіль, рекламна служба може чітко уявити собі портрет об'єкта, дізнатися про його звички і пристрасті, стати власником контактних даних.
Геопове-дінковий таргетинг	Поняття вперше згадано компанією AlterGeo. Суть зводиться до того, що знаючи про переміщення об’єкта (точне місцерозташування аж до закладу, в якому знаходиться об’єкт, вміють визначати сучасні геосоціальні сервіси. Наприклад, AlterGeo.ru) і зупинках (“Чекіна” або “check-ins”) у деяких точках, можна чітко уявити звички і пристрасті об'єкта. Приклад: якщо об’єкт ходить до пивних, то й пиво, напевно, його цікавить.
Психологічний таргетинг	Також одне із самих перспективних напрямків. На підставі інформації про психологічні якості (психотип, соціотип) видається різна реклама. Теоретичне обгрунтування – соціоніка.
MSB (Matrix Shop Behaviours)	Новітня технологія, суть якої полягає в наступному: для кожного замовника, на основі специфіки спілкування його менеджерів з клієнтами створюється поведінкова матриця, яка містить не тільки стереотипні моделі мислення покупців у цій бізнес-моделі, а й вагові коефіцієнти їх значимості у відсотках. Така деталізація дозволяє оцінювати якість сайту (структура, юзабіліті, дизайн, вміст тексту) не на рівні добре/погано, а на рівні конкретних цифр.

Перебування у постійному контакті з користувачами є ефективним завдяки автоматичному відстеженню статистики, для аналізу якої використовують коефіцієнти окупності інвестицій (англ. Return on Investment, ROI) та прибутковості е-бізнесу (англ. Rate of Return, ROR) і конверсію інформаційного ресурсу (англ. conversion rate). Конверсія, або ефективність відвідування інформаційного ресурсу – це відношення кількості відвідувачів інформаційного ресурсу, що виконали на ньому цільові дії (приховані/прямі вказівки рекламодавців, продавців, авторів комерційного контенту, тобто покупку, реєстрацію, підписку, відвідування певної сторінки інформаційного ресурсу, перехід за рекламним посиланням), до загальної кількості відвідувачів інформаційного ресурсу [3]. Успішна конверсія по-різному трактується продавцями або авторами (операція купівлі споживачем, який зацікавився продуктом, натиснувши на відповідний рекламний банер), рекламодавцями або постачальниками контенту (дії, очікувані від відвідувачів, наприклад, реєстрація відвідувачів інформаційного ресурсу, підписка на поштову розсилку, скачування програмного забезпечення). Основні переваги інтернет-маркетингу: інтерактивність, можливість максимально точного таргетингу, можливість постклік-аналізу для максимального підвищення показників конверсії інформаційного ресурсу і ROI/ROR Інтернет-реклами [1, 3, 4]. Мета застосування технологій інтернет-маркетингу – отримання максимального ефекту від потенційної аудиторії інформаційного ресурсу з можливістю миттєвого одержання статистики продажів, відвідувань, попиту тощо (табл. 5) [1, 3, 4].

Таблиця 5

Основні переваги інтернет-маркетингу, розроблено за [1, 3, 4]

Назва	Визначення
Інтерактивність (Interaction)	Принцип організації системи, при якій мети досягають обміном контенту між елементами системи та навколишнім середовищем.
Пошуковий маркетинг (Search engine marketing)	Процес збільшення відвідуваності з пошукових систем, формування списків результатів пошуку та рекламних оголошень.
Точний таргетинг (Targeted advertising)	Рекламний механізм, що дає змогу виділити зі всієї аудиторії інформаційного ресурсу цільову, яка задовольняє задані критерії для показу їй реклами.
Постклік-аналіз (Post-click analysis)	Метод Post-click маркетингу, що максимально підвищує результативність і конверсію інформаційного ресурсу та ROI Інтернет-.

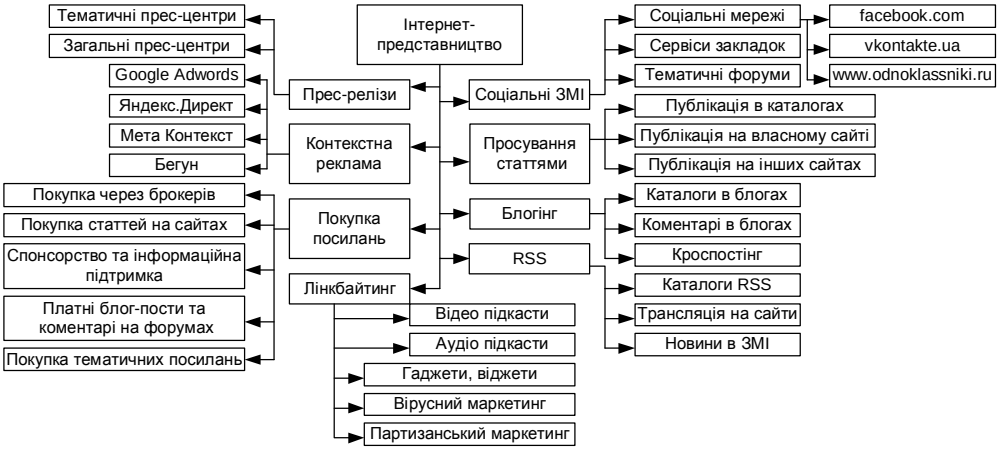


Рис. 2. Основні напрями інтернет-представництва

Пошуковий маркетинг (табл. 6) має такі характерні риси [1–5]: робота з конкретними запитами (ключові слова); зв'язок із пошуком (пошукові системи, пошук за інформаційним ресурсом); збільшення рівня знаходжуваності (англ. Findability) контенту інформаційного ресурсу; аналіз контексту (тематики контенту, інформаційного ресурсу тощо). Використовуваність (англ. Usability – зручність користування) – загальний коефіцієнт зручності об'єкта при використанні; концепція розроблення користувацьких інтерфейсів програмного забезпечення, орієнтована на максимальну психологічну/зорову зручність для користувача; коефіцієнт ефективності виконання дизайну меню і системи навігації по інформаційному ресурсі; зручність і простота використання застосування, дружелюбність і практичність інтерфейсу програмного забезпечення.

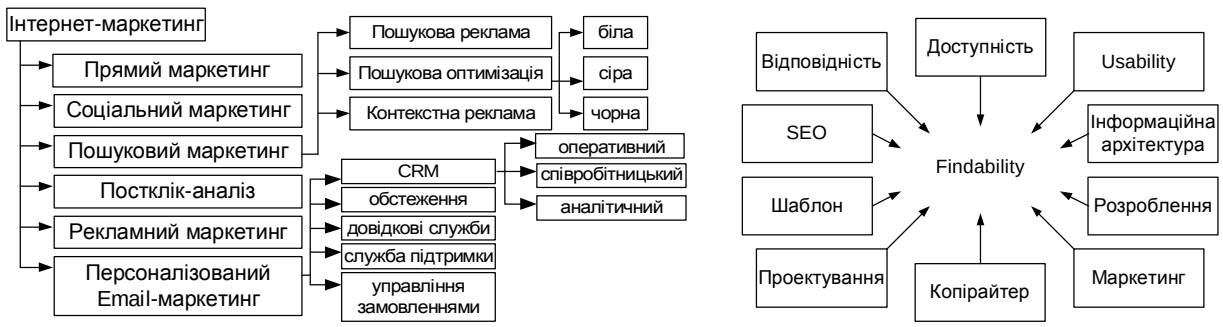
Таблиця 6

Технології пошукового маркетингу, розроблено за [1]

Технологія	Призначення технології пошукового маркетингу
Пошукова реклама	Поширення інформації в пошукових системах шляхом розміщення рекламних оголошень з ключовими словами.
Пошукова оптимізація (Search engine optimization)	Дії зі зміни стану інформаційного ресурсу (просування, розкручування) і елементів зовнішнього середовища для здобуття високих позицій в результатах пошуку.
Контекстна реклама	Розміщення оголошень на тематичних інформаційних ресурсах.

Технологія SEM (рис. 3, а) має характерні риси [1, 3, 4].

- Робота з конкретними запитами (ключові слова).
- Зв'язок із пошуком (пошукові системи, пошук по Web-сайту).
- Збільшення рівня знаходжуваності (англ. Findability) контенту Web-сайту (рис. 3, б).
- Аналіз контексту (тематики рекламного майданчика тощо).



а *б*

Рис. 3. Технології: пошукового маркетингу (а) та Findability (б)

Використовуваність (англ. Usability – зручність користування) – загальний коефіцієнт зручності об'єкта при використанні; концепція розроблення користувацьких інтерфейсів ПЗ, орієнтована на максимальну психологічну/зорову зручність для користувача; коефіцієнт ефективності виконання дизайну меню і система навігації сайту; зручність і простота використання застосування , дружелюбність і практичність інтерфейсу ПЗ. Пошуковий маркетинг не досягає цілей реклами через такі фактори [3]: не завжди прямо повідомляють про рекламований товар; не завжди переслідують мету продажу товару/послуги; за їх допомогою складно підвищити впізнавання бренду; неможливо вивести новий товар/послугу на ринок.

Результатом використання пошукового маркетингу є [3]:

1) залучення користувачів на інформаційний ресурс, де для кожного окремого випадку аудиторія відрізняється, тому привертають широку (збільшивши коефіцієнт відвідуваності інформаційного ресурсу) або зацікавлену аудиторію;

2) поширення контенту про інформаційний ресурс у пошукових системах.

Критерієм вдало обраної стратегії пошукового маркетингу є кількість відвідувачів інформаційного ресурсу та відповідність якості отриманої аудиторії необхідній. Простим критерієм перевірки популярності інформаційного ресурсу є динаміка кількості зовнішніх посилань на інформаційний ресурс і збільшення згадувань назви товару/послуги або торговельної марки компанії в Інтернеті. Граничний випадок пошукової і контекстної реклами полягає у розміщенні рекламних оголошень в результатах тематичного пошуку за інформаційним ресурсом. Розвиток Інтернету сприяв появі нових технологій пошукового маркетингу для соціальних мереж (англ. Social marketing optimization) та відео (англ. Video search marketing). Відокремлення пошукового маркетингу в окрему самостійну стратегію пов'язано з [1]: продовженням зростання інтернет-ринку; зростанням ринку контекстної/пошукової реклами; застосуванням пошукової оптимізації [1]; необхідністю оптимальної навігації/серфінгу в просторі контенту, що містить з текстовий/візуальний/анімований/звуковий контент і/або досвід користувачів системи; підтримкою складного процесу життєвого циклу контенту, який він проходить під час управління через різні етапи публікації. Процес проектування та створення СЕКК на основі результатів інтернет-маркетингу є ітеративним і проходить від аналізу, проектування, розроблення плану до створення прототипу і експериментальних випробувань, починаючи з формування специфікацій, верстки, створення шаблону, формування та контенту розміщення згідно із структурою інформаційного ресурсу (рис. 2). Концентрують рішення на бізнес-цілях і потребах кінцевих користувачів. На початкових етапах до визначення/розроблення функціональних вимог підключають користувачів за допомогою листів опитування, альтернатив проектування і прототипів різного ступеня готовності. Тобто збирають цінну інформацію, викликаючи у користувачів відчуття безпосередньої участі в процесі проектування та завойовуючи їх довіру.

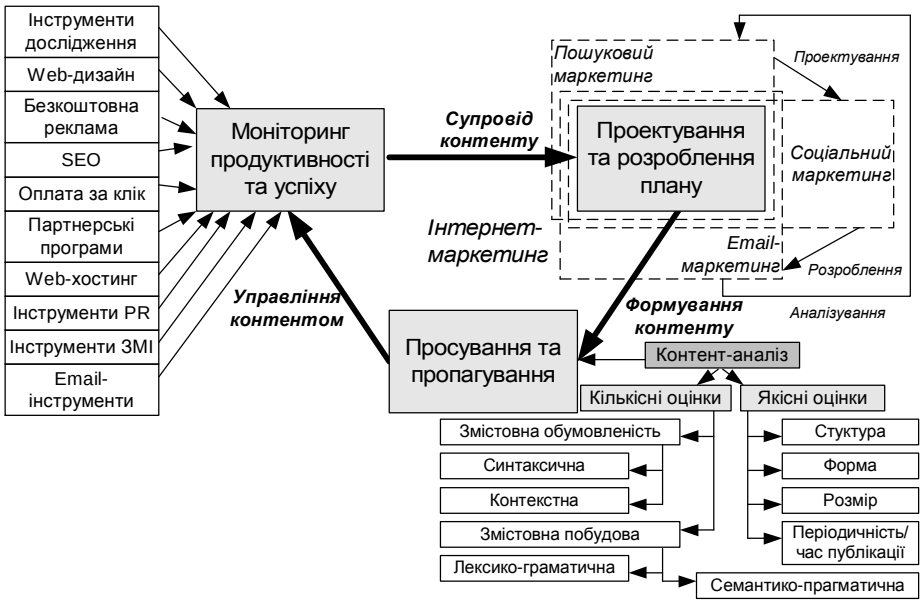


Рис. 2. Інтернет-маркетинг для систем електронної контент-комерції

Відомим методом аналізу текстової інформації є *контент-аналіз* – стандартна методика дослідження в області суспільних наук (рис. 2, табл. 7–9), предметом якої є аналіз змісту текстових масивів і комунікативної кореспонденції (коментарів, форумів, електронного листування, статей тощо). Поняття контент-аналізу не має однозначного визначення [4], тому системи, побудовані на основі різних підходів, є несумісними. Застосування контент-аналізу тексту в систем електронної контент-комерції дає ряд переваг для спрощення ведення бізнесу та вирішує низку проблем, що стоять перед учасниками бізнес-процесів, а саме:

1) фільтрація контенту користувача на інформаційному ресурсі;

2) можливість автоматичного створення *портрету* постійного користувача на основі аналізу його коментарів;

3) можливість автоматичного створення “портрету” цільової аудиторії на основі аналізу “портретів” постійних користувачів;

4) скорочення кількості модераторів інформаційного ресурсу в СЕКК;

5) скорочення часу для розміщення контенту користувача на інформаційному ресурсі із-за автоматичного його опрацювання, а не модераторами;

6) ліквідація мовного бар’єру завдяки автоматичному формуванню словників постійного користувача та автоматичного перекладу.

Таблиця 7

Неоднозначні визначення поняття контент-аналізу

Автор	Визначення
Д. Джері, ДЖ. Джері	методика об’єктивного якісного та систематичного дослідження змісту засобів комунікації;
Д. Мангейм, Р. Рич	систематичне кількісне опрацювання, оцінювання та інтерпретація форми і змісту інформаційного джерела;
В. Іванов	якісно-кількісний метод дослідження документів (характеризується об’єктивністю висновків і строгістю процедури) та квантифікаційного опрацювання тексту з подальшою інтерпретацією результатів; предмет дослідження – проблеми соціальної дійсності, які висловлюються і приховуються у документах, та внутрішні закономірності самого об`єкта дослідження;
Б. Краснов	складається із пошуку в тексті визначених змістовних понять (одиниць аналізу), виявлення частоти їх появи та співвідношення із змістом всього документа;
Е. Таршис	техніка дослідження для отримання результатів шляхом аналізу змісту тексту про стан і властивості соціальної дійсності.

Таблиця 8

Основні складові контент-аналітичного дослідження

Назва	Властивості основних складових контент-аналітичного дослідження
Спостереження	опрацювання масової сукупності текстів, використовуючи при цьому типові соціологічні процедури суцільного/вибіркового спостереження, з дотриманням вимог репрезентативності.
Структурування	припущення структурування, сегментації, розчленування текстів чи виділення із них змістовних інваріантів (повторення в усіх/ряді текстів) в досліджуваній масовій сукупності.
Формалізація	забезпечення однотипності сегментації і виділення інваріантів, застосування високого ступеня формалізації, суворих операційних правил і формальних алгоритмів в аналітичних процедурах.
Реферування	формалізований поділ цілісних текстів чи виділення окремих елементів їх для наступного збору із застосуванням аналітико-синтетичної процедури.
Аналіз	використання методів теорії ймовірності та математичної статистики для опрацювання текстів.

Таблиця 9

Типи методів контент-аналізу

Назва	Кількісний (змістовний)	Якісний (структурний)
Визначення	дослідження слів, тем та повідомлень, який зосереджується на змісті контенту.	дослідження, в якому досліджують не зміст контенту, а його форму та структуру.
Приклад	в якості першого кроку дослідник має створити словник, в якому кожне спостереження отримає визначення та буде віднесено до відповідної категорії.	визначення періоду часу або обсягу друкованого простору, який приділено темі в тому чи іншому джерелі, або скільки слів або стовпців приділено кожній темі відповідної категорії.
Особливість	перед проведенням аналізу обраних лінґвістичних одиниць передбачають їх зміст (створення словника) та визначають кожний можливий результат спостереження у відповідності із очікуванням дослідника.	розраховують питомі ваги P кожної теми і категорії ($P = R/T$, R – кількість одиниць даної категорії, T – загальна кількість одиниць) та проводиться порівняльний аналіз відповідних тем для подальшого прогнозування подій, процесів.

Більшість визначень контент-аналізу конструктивні, тобто процедурні (табл. 7). Через різні початкові підходи вони породжують різноманітні алгоритми, які інколи суперечать один одному [1]. Найбільші сумніви викликає ігнорування ролі контексту (табл. 8). Практичне значення методу дозволяє уникнути багатьох суперечностей. Об’єднання засобів і методів та їх природний відбір шляхом багатократного оцінювання отриманих результатів дають можливість виділення або підтвердження знань і фактичну силу/корисність інструментарію (табл. 9). Контент-аналіз – це кількісно-якісний аналіз масивів тексту для подальшої змістовної інтерпретації отриманих кількісно-якісних закономірностей. Метод полягає у формуванні з різноманіття тексту абстрактної моделі змісту контенту. Його застосовують для аналізу джерел, інваріантних за структурою/змістом у вигляді несистематизованого, безладно організованого тексту [1].

Результат контент-аналізу текстового комерційного контенту використовують для визначення тональності тексту, дублювання контенту, наявності спаму та виявленні нових подій для визначення тематичних сюжетів текстового контенту.

Контент-аналіз застосовують при дослідженні джерел, інваріантних за структурою/змістом та існуючих як не систематизований, безладно організований текстовий матеріал [6]. *Метод контент-аналізу* полягає у формуванні з різноманіття текстового матеріалу абстрактної моделі змісту тексту. Є два метода контент-аналізу: *кількісний* і *якісний*.

Під час дослідження механізмів породження текстової інформації в [6] виявлено, що вибір моделі опису тексту залежить від побудови ймовірнісно-лінґвістичного випробування та організації вибору з тексту окремих його одиниць. Ймовірнісне моделювання текстової інформації та його складових є вступним, підготовчим етапом до опису функцій лінґвістичних одиниць в тексті. Дослідження функціонування мови і мовлення за допомогою ймовірнісного моделювання текстової інформації ґрунтується на методах квантитативної лігвістики, теорії ймовірності, математичної статистики, теорії інформації та комбінаторики. У лінґвістичних дослідженнях, особливо під час реалізації алгоритмів контентного пошуку [6], постійно виникають завдання, пов’язані з

прогнозуванням появи в сеґменті заданої довжини певної кількості словоформ/словосполучень певних класів. Ймовірнісне моделювання тексту і складів, словосполучень, граматичних класів дає змогу визначати обсяг вибірки, необхідної для забезпечення із заданою ймовірністю появи хоча б одного разу відповідної лінґвістичної одиниці [6]. Кількісні оцінки змістовної інформації в тексті, словах та словосполученнях (рис. 3, а) отримують, спираючись на значення синтаксичної інформації та користуючись ідеєю контекстної обумовленості [6].

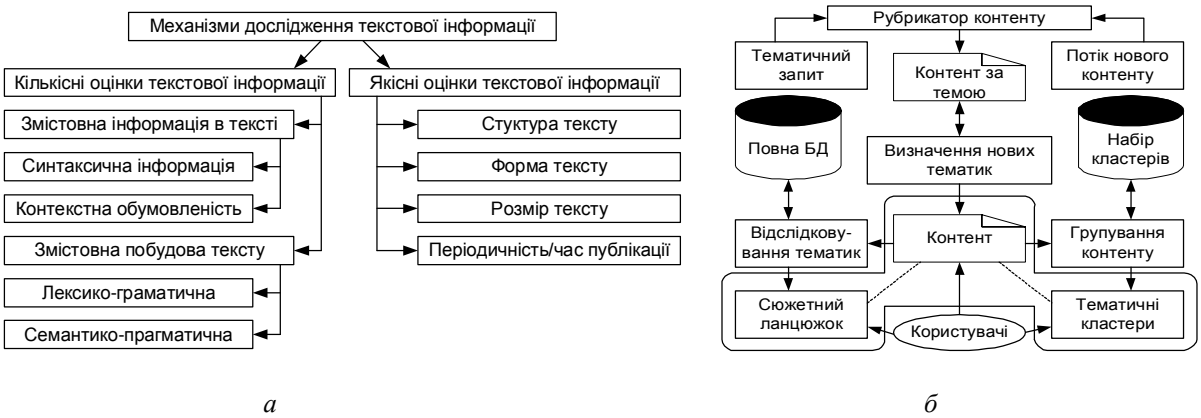


Рис. 3. Механізми дослідження (а) та формування контенту (б)

Під час експерименту із вгадування літер невідомого тексту в роботі [6] зауважено, що свої гіпотези про найімовірніші продовження тексту учасники експерименту будують, виходячи із двох типів комбінаторних обмежень: комбінаторики фіґур (літер та складів) та комбінаторики знаків (морфем, слів, словосполучень). Експеримент показує, що вже на 4-му або 5-му літерних кроках тексту комбінаторика літер та складів пригнічується обмеженнями, пов'язаними зі сполучуваністю морфем та слів. У міру розгортання тексту на комбінаторику слів нашаровуються обмеження зі сполученням словосполучень та речень з'являються обмеження, пов'язані з комбінаторикою параграфів, розділів, частин книги або статті. Отже, при вгадуванні літер, що розташовані на достатній відстані від початку тексту, учасник експерименту спирається не на статистичну комбінаторику літер та складів, а на змістовну (лексико-граматичну) побудову тексту. Якщо контент, вилучений із початкової ділянки тексту, є кількісною оцінкою дистрибуції (розподілу) та статистики літер, то синтаксична інформація, яка отримується з віддалених від початку тексту ділянок, слугує відображенням змістовної (семантико-прагматичної) інформації. Ці міркування дають змогу запропонувати метод контент-аналізу для кількісного оцінювання змістовної інформації в тексті та його сеґментах.

При виявленні нових подій з потоку контенту, що послідовно надходить на вхід СЕКК від засобів сканування або контентного роутера та відбирається за тематичним запитом (рис. 3, б), виявляють нові події, описані в контенті [2, 8]. Для них формують сюжетні ланцюжки подібного контенту. Контент, що відображає нові події, є основою кластерів взаємозалежного контенту (табл. 10) [8]. Кожен кластер може стати основою формування повноцінного сюжетного ланцюжка.

Визначення тональності контенту на основі аналізу текстів є складнішою за виявлення спаму. При виявленні спаму розглядають дві гіпотези (спам, не спам), а при визначенні тональності перевіряють емоційне забарвлення (позитивне, негативне, нейтральне) та їх комбінації. За методом Байєса для визначення спаму використовують базу оцінок – двох корпусів контенту, один із яких складений зі спаму, а інший – ні [8]. Для кожного контенту підраховують частоту використання кожного слова та вагову оцінку (від 0 до 1), тобто умовну імовірність, що контент з цим словом є спамом [2]. Значення ваг, близькі до ½, не враховують при інтегрованому розрахунку, тому слова з такими вагами ігнорують і видаляють. Простір гіпотез містить тональності $Tonality = H_{-1}$ (негативна), $Tonality = H_0$ (нейтральна) і $Tonality = H_1$ (позитивна). При гіпотезах H_1 із множини з позитивною тональністю обирають терми, характерні для цього контенту. Із них обирають терми

t із ймовірністю, розрахованою за формулою Байєса та більшою за ½. Рішення про тональність контенту приймають, враховуючи різницю значень вагових оцінок гіпотез H_1 і H_{-1} [2]. Відповідно до методу Paul Graham, якщо контент містить n термів із ваговими оцінками $w_1,...,w_n$, то оцінка умовної імовірності появи спаму основана на даних з оцінних корпусів та обчислюється як $Spm - \prod w_i / \left(\prod w_i + \prod (1 - w_i)\right)$ [2, 3]. Якщо S – подія, яка полягає в тому, що лист – спам, A – подія, яка полягає в тому, що лист містить слово t. Відповідно до формули Байєса $P(S \mid A) = \dfrac{P(A \mid S)P(S)}{P(A \mid S)P(S) + P\left(A \mid \overline{S}\right)P\left(\overline{S}\right)}$ [2].

Таблиця 10

Процеси виявлення нових подій

Автор	Етапи виявлення нових подій
G. Salton	1. Першому розглянутому контенту відповідає перший кластер. Кожен кластер зображається вектором термів (ключових слів), що входять у контент цього кластера. Нормований якимось чином вектор термів є центроїдом. Іноді центроїдом є контент, найближчий за деяким критерієм до вектора термів заданого кластера, що не змінює суті цього алгоритму. 2. Кожен наступний контент порівнюється з центроїдами наявних кластерів (для цього вводиться деяка міра близькості). 3. Якщо контент доволі близький до деякого кластера, то він приписується цьому кластерові, після чого перераховується відповідний центроїд. 4. Якщо контент не близький до наявних кластерів, то відбувається формування нового кластера, якому приписується такий контент. 5. Часовий діапазон розглянутого контенту прийнято називати “вікном спостереження”. Кластери, весь контент яких виходять за межі вікна спостереження, виносяться за межі розгляду. У результаті роботи алгоритму кожному новому кластерові, що виникає, відповідає нова подія, відображена в контенті цього кластера.
R. Papka	1. Формуються запити за відомими темами (при цьому використовуються техніки Text Mining – виявлення і вибору понять із контенту). 2. Новий контент, що надходить, порівнюється з наявними запитами. 3. Якщо контент не відповідає запитам, то він асоціюється з новою подією. 4. У систему включається новий запит, що відповідає такому контенту (опціонально).

Для виявлення подібного контенту c_i і дублікату c_j справедливе правило рефлективності, але не виконується умова транзитивності $c_i \mathbf{p} c_j, c_j \mathbf{p} c_k \not\Rightarrow c_i \mathbf{p} c_k$. Контент подібний до тексту з добірки, що його містить, але добірка не подібна до нього. Або контент подібний до інших двох, з яких він скомпільований, але оригінали істотно відрізняються. Для відношення дублювання виконуються симетричність і транзитивність, тобто $c_i \equiv c_j \Rightarrow c_j \equiv c_i$ та $c_i \equiv c_j, c_j \equiv c_k \Rightarrow c_i \equiv c_k$. Відношення із властивостями рефлексивності, симетричності та транзитивності є відношенням еквівалентності [8], тобто відношенням дублювання. Кожному контенту c_i за наведеним вище алгоритмом збігу термів у сигнатурах ставиться у відповідність вектор з елементами [2]:

$$a_{ij} = \begin{cases} 1, & c_i \equiv c_j, \\ 0, & \textit{інакше}, \end{cases} \text{, при } U_B = \begin{vmatrix} a_{11} \mathbf{K} a_{1n} \\ \ldots\ldots\ldots\ldots \\ a_{n1} \mathbf{K} a_{nn} \end{vmatrix}.$$

Під час аналізу критеріїв подоби використовують умови симетричності $\forall i, j : a_{ij} = a_{ji}$ та транзитивності $\forall i, j, k : a_{ij} = 1, a_{jk} = 1 \Rightarrow a_{ik} = 1$, змінюючи обсяг термів порівняння для знаходження відповідних коефіцієнтів [110, 111], тобто

$$\left(\left(\sum_i^N\sum_j^N\left|a_{ij}-a_{ji}\right|\right)\Big/\left(\sum_i^N\sum_j^N a_{ij}\right)\right)\to 0 \text{ і } \left(\left(\sum_i^N\sum_j^N\sum_k^N a_{ij}a_{jk}a_{ik}\right)\Big/\left(\sum_i^N\sum_j^N a_{ij}\right)\right)\to \max,$$

де N – кількість контенту. Коефіцієнт асиметричності асоціюється з визначенням дублікатів наближено, а рівень транзитивності – з повнотою.

При виявленні нових подій з потоку контенту, що послідовно надходить на вхід СЕКК від засобів сканування або контентного роутера та відбирається за тематичним запитом, виявляють нові події, описанні в контенті [2]. Для них формують сюжетні ланцюжки подібного контенту. Контент, що відображає нові події, є основою кластерів взаємозалежного контенту [2]. Кластер стає основою формування сюжетного ланцюжка при $u_{t_1}=sim\left(c_i,Dictionary\right)>a$, $u_{t_2}=\sum_{j=2}^{n}sim\left(c_i,Dictionary\right)>b$, де n – обсяг потоку контенту; c_1 – поточний контент; c_n – останній контент; c_i – i-й контент; $Dictionary$ – словник; $sim\left(c_i,c_j\right)$ – міра близькості контенту i до j; $sim\left(c_i,Dictionary\right)$ – міра близькості контенту i до словника, a і b – емпірично зумовлені параметри [2]. Якщо контент є множиною термів $c_i=\left\{w_{ik}\right\}=\left\{w:w\in c_i\right\}$, тоді $c_i+c_j=\left\{w:w\in c_i\mid w\in c_j\right\}$ – об'єднання термів із контенту c_i і c_j у вектор $E_i=\left\{e_{ik}\right\}$ розмірності N, який визначають як $e_{ik}=\begin{cases}1, & w_{ik}\in c_i,\\ 0, & w_{ik}\notin c_i,\end{cases}$ тобто $E=\begin{vmatrix}e_{11}\ \mathbf{K}\ e_{1n}\\ \dots\dots\dots\dots\\ e_{n1}\ \mathbf{K}\ e_{nn}\end{vmatrix}$. Міра близькості задається як $sim\left(c_i,c_j\right)=\left(\sum_{k=1}^{N}e_{ik}e_{jk}\right)\Big/N$. У [2] визначення міри близькості використовує апарат умовних ймовірностей (входження деякого терму w у контент c_i за умови, що він входить у контент c_j).

Для розроблення СЕКК використовують *систему управління контентом* (англ. Web Content management system, CMS) [1], якщо вона відповідає певному набору вимог (рис. 4).

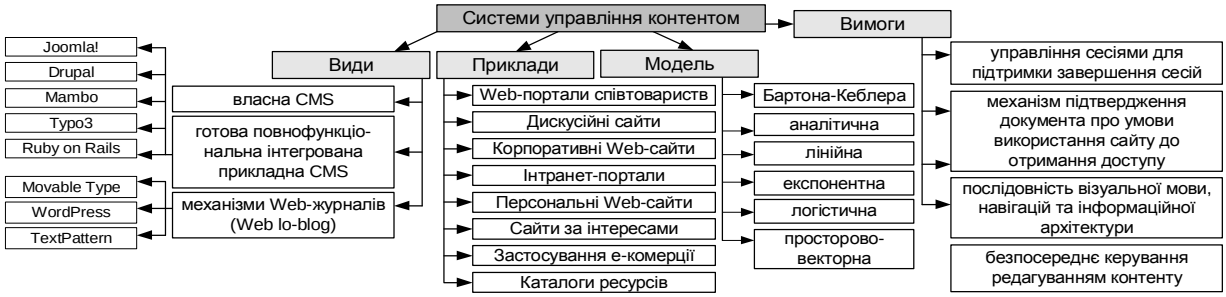


Рис. 4. Основні вимоги до систем управління контентом для побудови систем електронної контент-комерції

CMS не підтримують весь життєвий цикл контентного потоку та не вирішують основних проблем опрацювання інформаційних ресурсів – формування та супроводу контенту [4]. Основний недолік CMS полягає у відсутності зв'язку між вхідною інформацією, контентом та вихідною інформацією (рис. 5) [4–7]. Зазвичай такі системи використовують для зберігання і публікації великої кількості контенту (документів, зображень, музики, відео тощо). Подібні CMS дозволяють управляти текстовим і графічним наповненням, надаючи користувачеві зручні інструменти зберігання і публікації інформації.

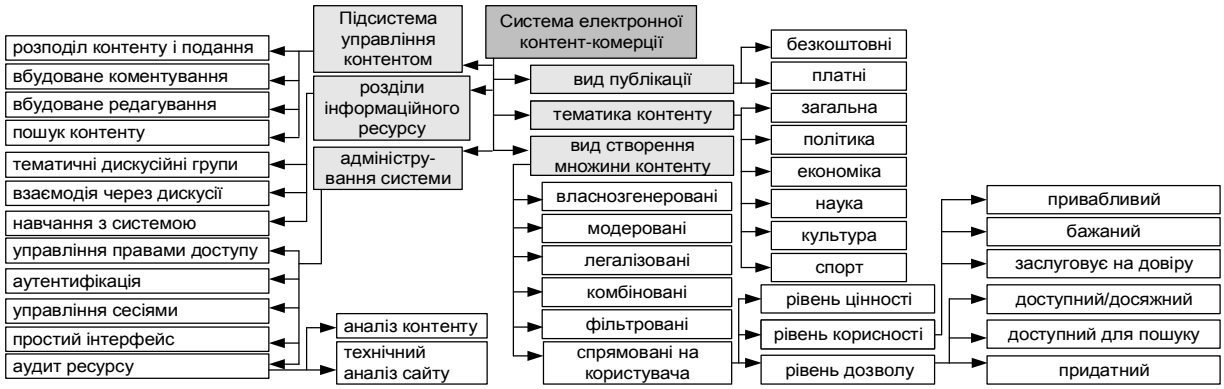


Рис. 5. Основні компоненти систем управління контентом для побудови СЕКК

CMS – це інструмент (табл. 11) для моделювання розгалужених структур інформаційних ресурсів в СЕКК і управління їх контентом [1] без наявності спеціальних технічних навичок як програмування або html-верстка. CMS розробляють для генерації контенту в застосуваннях з такими проблемами, як динамічне збирання, кешування контенту, його безпека тощо [2]. CMS забезпечує контроль над доступом до інформаційного ресурсу і внесенням змін та покликана максимально спростити процес управління інформаційним ресурсом при збереженні гнучкості налаштувань і контролю. Основні компоненти інформаційного ресурсу в CMS подано в табл. 12 [5].

Таблиця 11

Характеристика основних систем управління контентом для розроблення систем електронної контент-комерції, розроблено за [4]

№	Назва системи управління контентом	Вимоги до програмного забезпечення			Простота установки	Крива навчання	Управління сесіями	Управління користувачами	Розширюваність	Маштабованість	Використання тематик	xHTML/CSS
		Web-сервер	База даних	Мова								
1	Ruby on Rails	Apache,FastCGI	MySQL, PostgreSQL	Ruby	+/–	+/–	+/–	+/–	+	+	+/–	+
2	Drupal	Apache IIS	MySQL, PostgreSQL	PHP	+/–	+/–	+	+	+	+	+	+
3	Mambo	Apache IIS	Apache IIS	PHP	+	+/–	+/–	+	+/–	+	+/–	+
4	Typo3	Apache IIS	Apache IIS	PHP	–	–	+	+	+	+	+/–	
5	Movable Type	Apache IIS, Jetty	Apache IIS, Jetty	Perl	+	+/–	–	+/–	+/–	+	+/–	–
6	Word Press	Apache, mod_rewrite	Apache	PHP	+	–	–	+/–	+/–	–	+/–	+
7	Text Pattern	Apache	Apache	PHP	+	–	–	+/–	–	–	+/–	+
8	Joomla!	Apache	MySQL	PHP	+	+	+	+	+	+	+	+

Використання CMS не вимагає встановлення програмного забезпечення. Для редагування й адміністрування використовують браузер. Інтуїтивний інтерфейс і простота роботи із системою полегшують управління інформаційним ресурсом та знижує подальші витрати на його підтримку [1]. CMS включає такі можливості: швидке поновлення та пошук контенту в інформаційному ресурсі; збирання даних про клієнтів та потенційних клієнтів; формування і редагування опитувань; аналіз відвідування інформаційного ресурсу. В СЕКК присутні великі потоки та обсяги різного контенту. Значна частина цих потоків контенту складена з легко формалізованих і автоматизованих процедур та комерційного контенту. Але відсутній загальний підхід до процесу формалізації, проектування, розроблення та впровадження СЕКК. Не описано та не розкрито в таких системах зв’язки та залежності між вхідними даними, комерційним контентом, вихідними даними й процесами опрацювання інформаційних ресурсів (табл. 12).

Таблиця 12

Основні компоненти підсистеми опрацювання інформаційних ресурсів, розроблено за [4]

Назва	Характеристика Content management system
Пункти меню	Додавання, редагування, управління пунктами меню інформаційного ресурсу будь-якого рівня.
Статті	Додавання, редагування, планування та публікування статей (сторінок інформаційного ресурсу).
Новини	Додавання, редагування та публікування новин.
Фотогалерея	Можливість робити фотогалереї з підгалереями, автоматичне масштабування фото.
Дошка оголошень	Додавання оголошень з фотографіями, описом та контактними даними.
Налаштування	Зберігання всіх налаштувань інформаційного ресурсу та його системи управління.
Користувачі	Управління правами зареєстрованих користувачів.
Каталог фірм	Додавання, редагування, публікування фірм у підгрупах будь-якої вкладеності.
Опитування	Додавання/редагування опитувань, результати у вигляді графіків.

Цінність контенту визначає його привабливість для споживача. Інтеграція контенту робить інформаційний ресурс привабливим, а інтеграція застосувань – корисним [5]. Використання CMS не вимагає встановлення програмного забезпечення. Для редагування й адміністрування використовують браузер. Інтуїтивний інтерфейс і простота роботи із системою полегшує управління інформаційним ресурсом і знизити подальші витрати на його підтримку. [5]. CMS включає такі можливості: швидке поновлення та пошук контенту в інформаційному ресурсі; збирання даних про клієнтів та потенційних клієнтів; формування і редагування опитувань; аналіз відвідування інформаційного ресурсу.

Аналіз отриманих наукових результатів

Засоби управління контентом в електронній комерції

Існує декілька моделей життєвого циклу контенту (табл. 13, 14), автори яких пропонують та описують кілька етапів з набором властивостей, які підтримуються різними технологіями та процесами.

Таблиця 13

Визначення терміна “контент”

Призначення	Визначення терміна *“контент”*
Алгебра	загальний визначник серед коефіцієнтів, наприклад, значення серед множини даних (найбільший, найменший, середній час обчислення певної функції).
Теорія вимірювання	адитивна функція, яка визначає значення поля (змінної).
Web-контент	інформація, яка опублікована в Інтернеті (рис. 1.2, а).
Медіа-контент	інформація, опублікована користувачами, дизайнерами або адміністраторами Web-порталів, наприклад, аудіо/відео/графіка.
Кількісний визначник	об`єм, сфера, простір, розмір.
Прикладна лінгвістика	система визначень або семантика мови; тематика/питання, сформоване у вигляді книги або документа; частина мови, яка досліджується, наприклад, контент в реченні – іменник або дієслово; інформація, що передається в розмові та легко сприймається після прочитання і аналізу.
Форматований контент	закодований формат для відображення диних, наприклад, хеш-кодування пароля або приватний/публічний ключ для обміну даними в межах мережі.
Вільний контент	опублікований матеріал, юридично захищений авторськими правами за допомогою вільних ліцензій.
Відкритий контент	опублікований матеріал для вільного використання, на основі якого створюють матеріал для публікацій.
Математична лінгвістика	об’єкт, який містить контекстну інформацію (пояснення якоїсь області або її значення, пояснення про відмінність між об’єктми).
Контент-таблиця	структуровайй зміст документа, наприклад, зміст наукової книги або адміністративного документа.
Інформаційні технології	об’єкт інформаційного ресурсу наперед визначеної форми; текстове наповнення Web-сайту.

Таблиця 14

Класифікація життєвого циклу контенту

Автор	Етапи	Особливості моделі
McKeever Susan	збирання, доставка/публікація	Web Content Management (WCM) має чотири ступені ієрархії: контент, діяльність, вихід і аудиторія для ілюстрації широти WCM.
Bob Boiko	збирання, управління, публікація	Етап збирання – це збирання, створення, і редагування, управління, workflows, узгодження, управління версіями, архівування, синдикація, управління метаданими.
Gerry McGovern	створення, редагування, публікація	Знання, контент і картографічна інформація підходять для трьох ключових процесів: створення, редагування, публікація. Створення – це приблизне визначення знання для певної ідеї або набору ідей. Редагування – це процес модифікації професійного контенту з врахуванням цих ідей. Публікація – це подання контенту для відповідної людини у відповідний час. Для генерації переваг контенту необхідні процеси створення, редагування і публікації. Таким чином, переваги контенту описують формулою: $ContentBenefit = Create \times Edit \times Publish$.
JoAnn Hackos	створення, архівування, складання/ поєднання, публікація	Призначення: Content Management for Dynamic Web Delivery. В модулі добре висвітлена концепція метаданих. Hackos глибоко поринає в поняття метаданих і як вони повинні бути використані. Зосереджує увагу на легалізації даних, пошуку в БД і перетворення інформації в знання. Автор дає уявлення про відповідність нагромадження інформації твердій та стабільній інформаційної моделі. Hackos описує в моделі злиття ідеї юзабіліті, інформаційної архітектури, і управління контентом для зосередження на кінцевих користувачах.
Ann Rockley	створення, перегляд, управління, доставка	Етап створення складається із планування, проектування, авторизації та перевірки. Рекомендовано використання загальної стратегії контенту (англ. unified content strategy) – повторюваний метод ідентифікації всіх найважливіших вимог до контенту для створення послідовно структурованого контенту, для повторного використання, управління визначених джерел контнету, а також монтаж контенту залежно від вимог та потреб користувачів.
Russell Nakano	подання, порівняння, оновлення, злиття, публікація	Автор приділив особливу увагу процесу управління Web-контентом, зокрема управлінню версіями і одночасними змінами.
The State government of Victoria (Australia)	розроблення, затвердження якості, публікація, скасування публікації, архівування	Етап архівування складається із підетапів накопичення, архівування, утилізації. Крім того, ця модель відображає три основні аспекти потоку Web-контенту – це статус, процес та роль (контент автора, контент управління якістю, процес управління якістю, запису управління).
AIIM	захоплення/поглинення, управління, накопичення, доставка, зберігання	Основна увага в моделі приділена лише можливості організації збирання, управління, збереження, накопичення і постачання необхідної інформації необхідним людям в необхідний час.
CMP org-tion	планування, розроблення, управління, розгортання, зберігання, оцінювання	Кожнен етап поділений на підетапи, наприклад, етап планування складається із підетапів вирівняння, аналізу, моделювання, проектування. Етап розроблення – створення, збирання, колекціонування, класифікація, редагування.
Bob Doyle	організація, створення, накопичення, workflow, управління версіями, публікація, архівування	Модель базується на принципі – сім +/- два, тобто дозволяє внесення змін (збільшення/зменшення кількості етапів) в саму модель залежно від мети інтернет-маркетингу.
Woods Randy	легалізація, шаблон, створення, модифікація, управління версіями, обертання, моніторинг процесу та управління успіхом	В моделі увага акцентується на категорії вирішення питань, а не на самих етапах від початкового (створення контенту) до завершального (публікація контенту). Більшість питань з управління контентом належать до однієї із таких категорій: міграція легального контенту, шаблон міркувань, створення нового контенту, модифікація контенту та повторне використання, контрольоване управління версіями та відкат сайту, обертання контенту та кінець шляху, моніторинг процесу та управління успіхом (результатами).
Halverson	аудит, аналіз, стратегія, класифікація, структуризація, створення, перегляд, повторний перегляд, завершальний перегляд, затвердження, тестування, форматування, публікація, поновлення, архівування.	Ця модель акцентує увагу на стратегії контенту, надає високорівневий огляд переваг, ролі, діяльності та результатів, пов’язаних з утриманням стратегії. Призначення моделі: web-консалтинг, проектування і розроблення промисловості, описує процеси та методології, які можуть бути застосовані до всіх видів контенту: текст, зображення, відео та аудіо.

У деяких моделях життєвого циклу контенту (табл. 15) передбачено концепції управління проектами, управління інформацією, інформаційну архітектуру, стратегії контенту, управління Web-сайтом, семантичний друк. Різні автори пропонують різні етапи життєвого циклу контенту [4]. Основні етапи (створення контенту, розроблення, перегляд, поширення та архівування) присутні майже у всіх запропонованих моделях [5]. Життєвий цикл процесів, дій, статусу контенту і ролі управління контенту відрізняються в моделях залежно від організаційних стратегій, потреб, вимог і можливостей [6].

Таблиця 15

Класифікація життєвого циклу контенту

№	Автор	Опрацювання інформаційних ресурсів		
		Формування	Управління	Реалізація
1	McKeever S.	+/–	–	+/–
2	Bob Boiko	+/–	+/–	+/–
3	McGovern G.	+/–	–	+/–
4	JoAnn Hackos	+/–	–	+/–
5	Ann Rockley	+/–	+/–	+/–
6	Russell Nakano	+/–	–	+/–
7	The State Victoria	+/–	–	+/–
8	AIIM	+/–	+/–	+/–
9	CMP organization	+/–	+/–	–
10	Bob Doyle	+/–	+/–	+/–
11	Woods Randy	+/–	+	+
12	Halverson	+	+/–	+/–

Моделі життєвого циклу Web-контенту № 1-7 не вирішують проблеми формування і реалізації контенту та вирішують далеко не всі проблеми управління конвентом (наприклад, подання множини контенту кінцевому користувачу згідно з його запитом, історії або інформаційного портфеля; автоматичне формування дайджестів та інформаційних портретів; автоматичне виявлення тематичних сюжетів; побудова таблиць взаємо'язку понять; розрахунок рейтингів понять; збирання інформації з різних джерел та її форматування; виявлення ключових слів та понять контенту; автоматична рубрикація контенту; виявлення дублювання змісту контенту; вибіркове поширення контенту). Моделі № 8–10 успішно справляються з проблемами управління контентом та деякими проблемами формування контенту, але не приділяють уваги проблемам реалізації контенту. Моделі № 11–12 вирішують проблеми формування контенту та управління контентом і частково вирішують проблеми реалізації контенту. Всі ці моделі не підтримують Web 2.0 (табл. 16).

Відомим засобом управління контентом є *система управління контентом* (англ. Content management system, CMS) – це програмне забезпечення для організації Web-сайтів або інших інформаційних ресурсів в Інтернеті чи окремих комп'ютерних мережах [1]. CMS має відповідати певному набору вимог (рис. 6, а). Зазвичай такі системи використовують для зберігання і публікації великої кількості контенту (документів, зображень, музики, відео тощо). Такі CMS дають змогу управляти текстовим і графічним наповненням, надаючи користувачеві зручні інструменти зберігання і публікації інформації. На рис. 6, б подано класифікацію CMS.

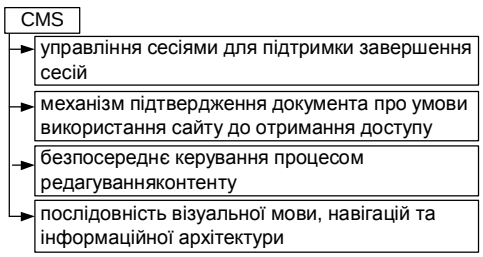
CMS не підтримують весь життєвий цикл контентного потоку та не вирішують основних проблем опрацювання інформаційних ресурсів – формування та реалізації контенту. Основний недолік CMS полягає у відсутності зв'язку між вхідною інформацією, контентом та вихідною інформацією. CMS часто використовують для побудови СЕК та СЕКК (табл. 17–20), наприклад, інтернет-газет (рис. 7, а) [1]. *Системи управління Web-контентом* (англ. Web Content management system, WCMS) розробляють для генерації контенту всередині порталів, які мають ті самі проблеми (динамічне збирання, кешування контенту, його безпека тощо), що й інші різновиди Web-застосувань [1]. Цінність контенту визначає його привабливість для споживача [1]. Інтеграція

контенту робить портали привабливими, а інтеграція застосувань – корисними, оскільки користувачі все частіше тяжіють до виконання на порталі щораз більшої кількості застосувань, включаючи ECM. Адміністративна панель WCMS дозволяє змінювати/додавати нову інформацію для різних мовних версій Web-сайту. Зміни на сайті відображаються відразу після внесення і збереження.

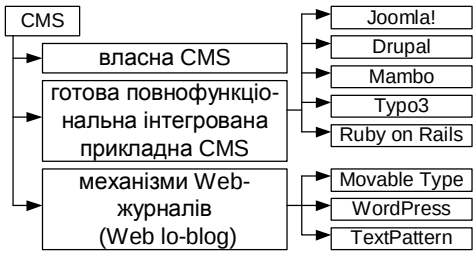
Таблиця 16

Порівняння параметрів Web 1.0 та Web 2.0

Назва	Web 1.0	Web 2.0	Нові властивості
Актор	Розробник-користувач або автор контенту і читач.	Користувач як співрозробник, читач як співавтор або товариство.	Право на участь; скасування сторонньої регламентуючої сторони (модерації).
ПЗ	ПЗ створювалось для ПК; ПЗ – товар; закриті вихідні коди, API; ліцензійний продаж; прив'язка ПЗ до обладнання; напрямленість на винахід; запланований реліз; для перегляду контенту використовують браузер.	ПЗ створюють для Web; ПЗ – сервіс, застосування; відкриті вихідні коди, API, open-source software; ПЗ може бути безкоштовним; софт поверх обладнання; пошук застосування вже винайденому; вічна бета версія; альтернативні засоби сприйняття.	Web як платформа; зняття та розмивання бар'єрів та обмежень (вільний доступ, універсальність, спрощення).
Контент	Поповнення БД: плата постачальнику контенту або наймання модераторів; таксономічна організація даних; ієрархія рубрик; засоби збереження даних – каталог, бібліотека, сховище; односторонні посилання; форма подання – персональні сторінки; статичний сайт; сторінка сайту має адресу; джерело – розум автора контенту; меню навігації сайту для роботи з його даними; копірайт; для сприйняття контенту відвідують сайт, переходячи по посиланню чи закладці.	Поповнення БД – те, що має один, відразу стає доступне кожному; дані організовують фолксономічно; засоби використання даних – API-інтерфейси; автоматичні двосторонні посилання; форма подання – блоги; динамічний сайт; адресу має мікроелемент контенту; джерело – колективний розум; інтерфейс для роботи з даними по всій мережі; "вільна" ліцензія GNU FDL; для сприйняття контенту не потребується відвідування сайту – можливість читати RSS –стрічки.	Мережа як єдиний колективний розум, атомізація контенту, агрегація, синдикація.
Події	Замовлення та виготовлення ПЗ; публікація контенту авторами і сприйняття його читачами; звертання до третьої особи – посередника для задіяня його ресурсів; великі, нечисленні угоди.	Співпраця через відділ технічної підтримки ПЗ; взаємодія, додавання властивостей, цінності, створення спільного контенту кожним учасником; самообслуговування, засноване на партнерській архітектурі сервісу, який лиш посередник між користувачами, що використовують власні ресурси; дрібні численні транзакції.	Співпраця; самодіяльність; масові одиничні взаємовідносини.
Цінність/ вартість	Цінність в ПЗ – власник ПЗ заробляє гроші; Інтернет цінний як джерело інформації.	Цінність в БД – власник БД та сервісів для роботи з ними заробляє гроші; Інтернет цінний як інструмент комунікацій.	Робота з БД; сервіс, а не продукт; економія часу та уваги.



а



б

Рис. 6. Вимоги (а) та класифікація (б) CMS

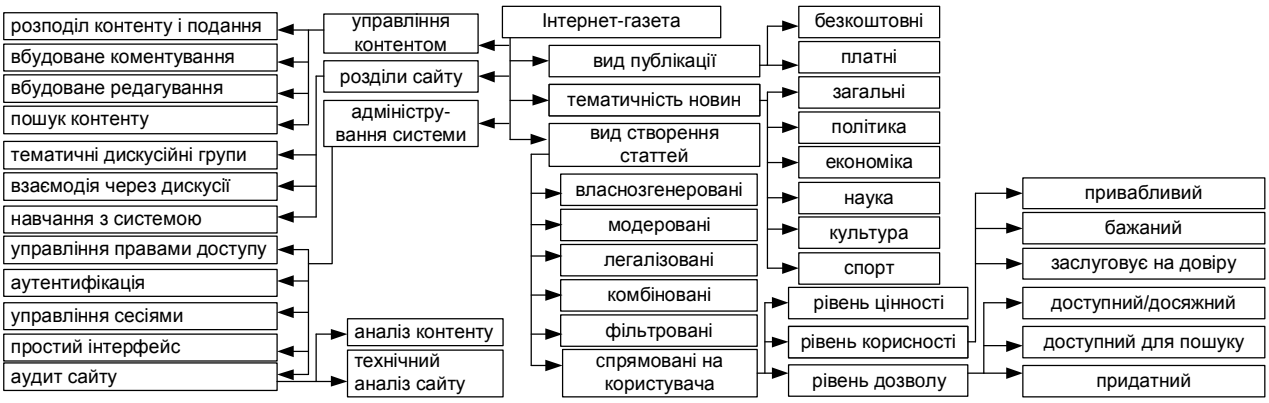


Рис. 7. Класифікація інтернет-газет

Таблиця 17

Основні відмінності СЕК і СЕКК

Характеристика	СЕК	СЕКК
Вид товару	матеріальний	нематеріальний (контент).
Обсяг товару	зменшується	стала.
Наявність складу	присутній	відсутній.
БД товару	опис товару	контент та його опис.
Просування товару	за ключовими словами в описі товару	за ключовими словами в описі контенту та змістом самого контенту.
Пошук товару	за ключовими словами в описі товару	за ключовими словами в описі контенту та змістом самого контенту.
Виявлення дублювання товару	вручну	автоматично (програмно), застосування відомих методів виявлення та ліквідації дублювання інформації.
Визначення старіння товару	вручну або за терміном	автоматично (програмно), застосування відомих методів дослідження інформації на старіння.
Визначення актуальності товару	вручну або за кількістю замовлень за останній період порівняно з іншими замовленнями або періодами	автоматично (програмно), застосування відомих методів дослідження інформації на актуальність.
Аналіз аудиторії	за замовленнями, відвідуванням та коментарями; обмежена кількість атрибутів товару для аналізу; коментарі не висвітлюють думку аудиторії	за замовленнями та змістом замовленого контенту, відвідуванням сторінок та коментарями.
Автоматичне формування дайджестів	неможливо; короткий опис формують вручну	можливо, наприклад, автоматичне формування описів (дайджестів, тобто короткого змісту) статей в інтернет-газеті.
Автоматичне формування товару	неможливо	можливо, наприклад, автоматичне формування нової статті в інтернет-газеті із набору новин, які надійшли з достовірного джерела.
Автоматичне форматування товару	неможливо	можливо, наприклад, автоматичне форматування статей в інтернет-газеті згідно з класом аудиторії (вік, стать, професія тощо).
Досвід користувача	суттєво не впливає на збільшення обсягу продажів товару	суттєво впливає на рівень підписки товарів.

Використання CMS не вимагає встановлення спеціального ПЗ на робочому місці. Для редагування й адміністрування використовують звичайний браузер (Internet Explorer або Opera). Інтуїтивний інтерфейс і простота роботи із системою полегшують управління сайтом і знижують подальші витрати на підтримку Web-ресурсу. Для роботи із системою необхідно мати лише початкові навички роботи в Інтернеті. Найпопулярнішими CMS є Drupal і Joomla!.

При розгляді динаміки тематичних потоків контенту виявлено обмеженість моделей (табл. 20, 21), що відкриває шлях для подальших досліджень [2]. Моделі управління контентом призначенні для визначення процесів старіння/актуальності контентного потоку. Вони не вирішують проблеми

формування, реалізації контенту і вирішують не всі проблеми управління контентом: подання множини контенту кінцевому користувачу згідно з його запитом, історії або інформаційного портфеля; автоматичне формування дайджестів; автоматичне – інформаційних портретів; автоматичне виявлення тематичних сюжетів; автоматична побудова таблиць взаємозв'язку понять; розрахунок рейтингів понять; автоматичне збирання інформації з різних джерел; автоматичне форматування інформації; автоматичне виявлення ключових слів; автоматична рубрикація контенту; виявлення дублювання та вибіркове поширення контенту; залучення потенційних користувачів.

Таблиця 18

Класифікація інтернет-газет за типом створених статей

Статті	Характеристика
Власно згенеровані	Власні журналісти описують події, беруть інтерв'ю в учасників подій.
Модеровані	Журналісти/модератори самостійно шукають матеріал в Інтернеті, аналізують та систематизують отриманий матеріал, пишуть на основі цього матеріалу статті.
Легалізовані	Матеріал для майбутніх статей шукає модуль системи за посиланнями або підпискою, модератори аналізують та систематизують отриманий матеріал, пишуть на основі цього матеріалу статті.
Комбіновані	Комбінація перших трьох типів.
Фільтровані	Матеріал для майбутніх статей шукає модуль системи за посиланнями або підпискою та фільтрує згідно з закладеними в модулі словниками, модератори аналізують та систематизують отриманий матеріал, пишуть на основі цього матеріалу статті.
Спрямовані на користувача	Фільтровані, але при формуванні статті враховується рівень досвіду користувача, тобто формуються насамперед статті для користувачів з більшим рівнем досвіду.

Таблиця 19

Формальні моделі управління контентом

Назва	Формальна модель управління контентом
1	2
Бартона-Кеблера	$C =< A, B, T >$. Інтернет-простір поділяють на стабільну і динамічну складові з різними характеристиками з погляду інтеграції управління потоком контенту. Стабільна складова Інтернету містить "довгостроковий" контент, а динамічна складова – постійно поновлювані ресурси. Деяка частина цієї складової згодом вливається в стабільну, інша частина зникає або потрапляє у сегмент схованого Web-простору, не доступного користувачам за допомогою публічних інформаційно-пошукових систем. Найдинамічнішим є сегмент контенту у вигляді новин, який має найвищий рівень оновлення – у ньому генеруються і поширюються великі обсяги даних. Тому він є найкращим об'єктом досліджень. Процеси старіння контенту, втрата його актуальності описуються рівнянням з двох компонентів $m(t) = 1 - ae^{-T} - be^{-2T}$, де $m(t)$ – частина корисної інформації в загальному потоці через час T, перший від'ємник відповідає стабільним ресурсам, а другий – динамічним (новини). Контентна динаміка в мережі зумовлена багатьма факторами, які не піддаються точному аналізові. У межах моделювання розумним припущенням є загальний характер тимчасової залежності кількості тематичних публікацій, що визначається простими закономірностями. Для одержання адекватнішої залежності варто звернутися до складніших моделей.
Класична просторово-векторна	$C =< TF, IDF >$. Модель описують рівнянням $C = TF * IDF$, де TF – локальна частота терму (Term Frequency, рівень значущості терму в межах контенту), а IDF – величина, зворотна частоті появи повідомлень у всьому потоці контенту, що містить цей терм (Inverse Document Frequency, рівень унікальності терму у всьому потоці контенту). Добуток цих величин – це критерій визначення значущості терму (ваги). Контент (повідомлення, новини) старіє, втрачаючи свою актуальність з інтенсивністю, що визначається деяким емпіричним законом. Для ілюстрації припустимо, що це експонентний закон. Один із запропонованих підходів до такої частини узагальнення, як ранжування повідомлень, полягає у використанні параметричних множників, що залежать від часу. Наприклад, визначимо вагу повідомлення як добуток елементів типу $TF * IDF * e^{-at}$, де a – деяка константа, t – інтервал часу, що минув з моменту появи повідомлення в потоці контенту (значення a – це коефіцієнт напіврозпаду актуальності контенту, тобто, якщо передбачається застосування експонентної моделі, це $e^{-at} = 1/2$, де t – період часу, що визначають експертним шляхом, протягом якого повідомлення через старіння втрачає свою актуальність наполовину). Облік старіння контенту (втрати частини актуальності) має велике значення при аналітичних дослідженнях, створенні інформаційних продуктів (інформаційні портрети), основних сюжетів подій, ранжуванні результатів роботи інформаційно-пошукових систем. Наближене оцінення швидкості старіння контенту має практичну цінність у межах визначення множини найактуальнішого контенту.

Продовження табл. 19

1	2
Лінійна	$C=<Y,T,T_0,V>$. Динаміка управління тематичним контентом (актуальність/старіння) є лінійною []. Кількість повідомлень у момент часу t описується формулою $y(t)=y(t_0)\pm v(t-t_0)$, де $y(t)$ – кількість повідомлень на час t, v – середняя швидкість збільшення/зменшення інтенсивності потоку тематичного контенту в часі (наприклад, у результаті старіння). Змістовна складова контенту кількісно оцінюється як флуктуація (відхилення від норми) потоку контенту – зміна стандартного відхилення $s(t_i)=\sqrt{\frac{1}{i}\sum_{k=0}^{i}\{y(t_k)-(y(t_0)\pm v(t_i-t_0))\}^2}$. При зміні $s(t)$ процес зміни тематичного контенту є процесом з незалежними збільшеннями (не враховують зв'язки з попереднім контентом). У випадку поводження стандартного відхилення за часом як $s(t)\infty t^m$ що більше значення m (ступінь зв'язку між випадковими подіями у межах $\left[\frac{1}{2};1\right]$), то вища кореляція між поточним і попереднім контентом.
Експонентна	$C=<N,T_0,T,l>$. Процес росту актуальності/старіння контенту описують залежністю $N(t)=N(t_0)e^{l(t-t_0)}$, де l – середня відносна зміна інтенсивності потоку контенту. Відносна зміна інтенсивності у визначений момент часу $l(t_i)=(N(t_i)-N(t_{i-1}))/N(t_{i-1})$. Зміна флуктуацій величини $l(t_i)$ щодо середнього значення $s(t_i)=\sqrt{\frac{1}{i}\sum_{k=0}^{i}\{l(t_k)-l\}^2}$. При зміні $s(t)$ як корінь квадратний від часу процес є з незалежними збільшеннями (кореляція між окремим контентом несуттєва). При наявності значної кількості залежного контенту справедливо $s(t)\infty t^m$ в межах $\left(\frac{1}{2};1\right]$, що говорить про наявність довгострокової пам'яті системи. Такі системи породжують клас автомодельних процесів з врахуванням кореляції між контентом у різні моменти часу.
Логістична	$C=<N,T,M>$. Організації-генератори контенту в більшості працюють у стаціонарному режимі, що характеризується максимальною ємністю простору контенту N (розмірність параметрів та їхній вимір не враховують). Тобто кожна організація-генератор формує потік контенту у середньому постійний по кількості знаків і повідомлень. Змінюються в часі лише обсяги повідомлень тої або іншої теми. Ріст кількості контенту однієї теми супроводжується зменшенням обсягу контенту іншої теми, тому $\forall\Delta T$ маємо $\int_0^T\sum_{i=1}^{M}n_i(t)dt=NT$, де $n_i(t)$ – кількість контенту за одиницю часу, а M – загальна кількість усіх можливих тем. Частина $n_i(t)$ завжди дорівнює 0.
Аналітична	$C=<W,D>$. Вагу контенту визначають як $W_D=\frac{\sum_{w\in D}w}{\lvert D\rvert}$, де W_D – вага контенту, w – вага ключового слова з контенту (є монотонно зростаючою функцією від n), $\lvert D\rvert$ – кількість ключових слів у контенті ($1\le\lvert D\rvert\le 12$).Середня вага унікального ключового слова дорівнює $w(n)=n/v(n)=n^{1-b}/K$, де n – загальна кількість слів з потоку, K – кількість унікальних слів.

Таблиця 20

Складові CMS

Модель	Перевага	Недолік
Бартона–Кеблера	Описує процес старіння контенту, втрати його актуальності. Рівняння моделі має точний розв'язок у вигляді дуже простої і зручної функції – експоненти. Визначення швидкості розвитку окремих тематичних напрямів або всього контентного простору.	Є доволі сумнівною з погляду інтерпретації результатів. Монотонно зростає, не описує процесів, що за своєю природою повинні мати локальні екстремуми.
Просторово-векторна	Визначення значущого терму в усьому потоці контенту. Можливість визначення найактуальнішого контенту із множини наявних.	Обов'язкове ранжування контенту, використання параметричних множників, що залежать від часу
Лінійна	Визначення інтенсивності потоку тематичного контенту в часі (наприклад, у результаті старіння)	Застосовують при лінійній динаміці управління тематичного контенту.
Експонентна	Описує процес старіння контенту, втрати його актуальності.	Кореляція між окремим контентом несуттєва.
Логістична	Поєднює відносну простоту формулювання задачі з можливістю варіювати розв'язок за допомогою набору параметрів, що можуть мати більш-менш прозорий фізичний зміст.	Вивчення динаміки лише окремого тематичного потоку. Розмірність параметрів та їхній вимір не враховують.
Аналітична	Описує процес старіння контенту, втрати його актуальності.	Обов'язкова наявність словника ключових слів.

На рис. 8 подано контекстну діаграму типового регіонального інформаційного ресурсу, розроблена у CMS Joomla!. Застосуємо методи інтернет-маркетингу для аналізу Web-ресурсів у межах регіону. На рис. 9 подано IDF0 типового регіонального інформаційного ресурсу.

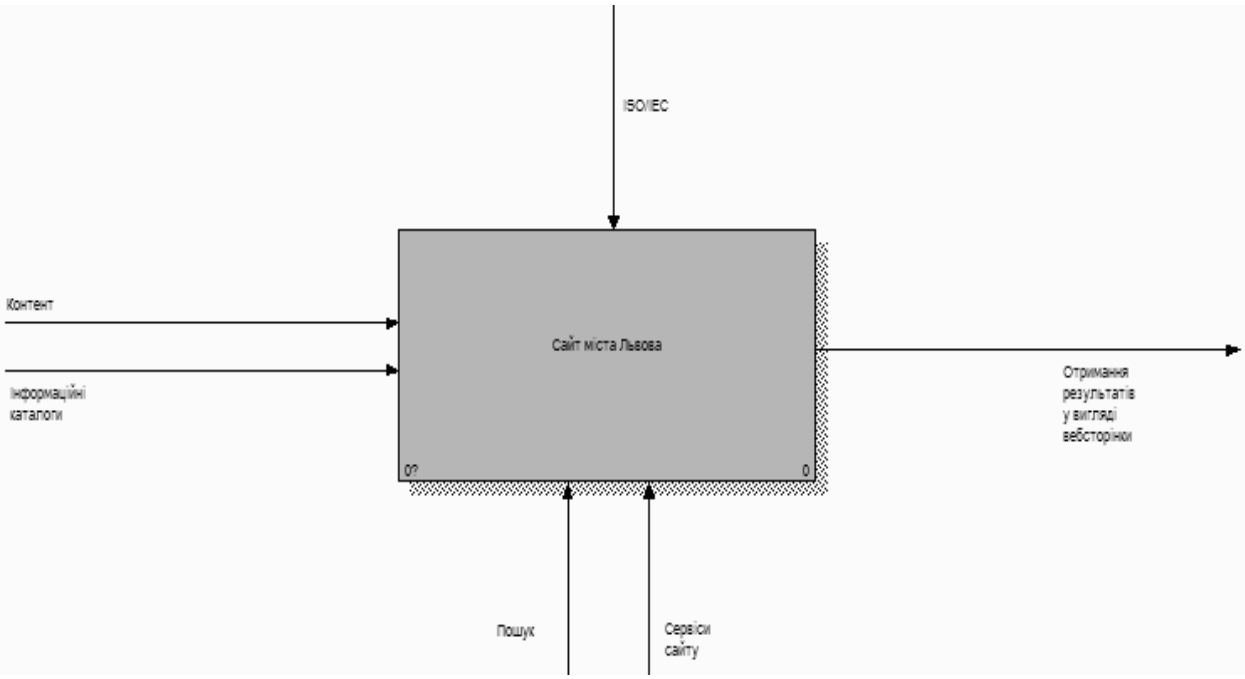


Рис. 8. Контекстна діаграма типового регіонального інформаційного ресурсу

Найефективніші методи залучення потенційних клієнтів

На інформаційних ресурсах деяких рекламодавців для користувачів наведено форми, через які вони мають право надіслати свої контактні дані й отримати додаткову інформацію про товари/послуги рекламодавця. Особи, які надали відомості за допомогою таких форм на інформаційному ресурсі рекламодавця, – це його потенційні клієнти. Аналіз цих даних та ефективне їх використання сприяє формуванню профілю потенційної аудиторії. Для цього необхідно скористатися такими методами.

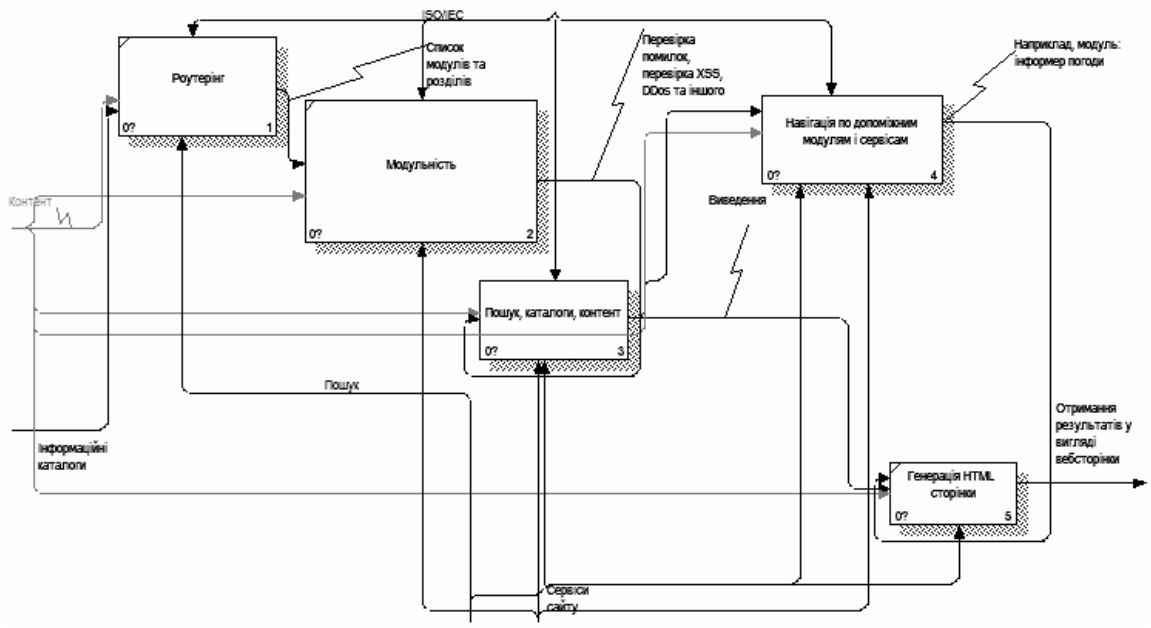


Рис. 9. IDF0 типового регіонального інформаційного ресурсу

1. Оптимізація кампаній AdWords. Після натискання оголошення відвідувачі розраховують виконати певні дії на цільовій сторінці. Важливо, щоб у відвідувача перед натисканням оголошення склалося правильне уявлення про те, на що можна розраховувати. Для цього оптимізують кампанію AdWords, установлюють ставки для необхідних ключових слів, за необхідності використовують мінус-слова (аналіз звітів пошукових термінів), використовують лаконічний, але змістовний і привабливий текст оголошення, а також налаштовують відстеження конверсій. Відстеження показника відмов і коефіцієнта конверсії цільової сторінки, а також їх порівняння для різних варіантів тексту оголошення, яке залучає трафік на сторінку, допоможе визначити, яке оголошення було ефективнішим.

2. Оптимізація цільової сторінки. Коли відвідувач потрапить на цільову сторінку, його очікування мають бути виправдані або перевершені. Цільові сторінки мають завантажуватися швидко, переходи між ними мають бути зручні і відповідати правилам Google щодо цільових сторінок. Для розширення клієнтської бази за допомогою форм на цільових сторінках необхідно зробити їх максимально простими та зрозумілими. Відвідувачі цільової сторінки можуть не захотіти залишати свої дані, якщо в них складеться враження, що сайт не заслуговує на довіру. Якщо скоротити обсяг інформації, яку має ввести відвідувач, і запитувати лише ті дані, що необхідні для подальшого зв'язку з ним, це підвищить вірогідність заповнення форми. Якщо відвідувач починає заповнювати форму, а потім залишає сайт, не надіславши її (наприклад, тому що форма є задовгою), програма Web-аналітики реєструє це як відмову.

3. Залучення зацікавлених клієнтів. Не кожен потенційний клієнт, який заповнив форму, стане реальним клієнтом. Причин може бути кілька: недостатньо зрозуміла цільова сторінка; у формі не передбачено підтвердження наданої інформації; форму автоматично заповнила спеціальна програма, яка сканує сайти й поширює спам тощо. Google не може контролювати поведінку користувачів на сайті. Окрім оптимізації кампанії AdWords і цільової сторінки, є ще кілька способів попередити реєстрацію фіктивних клієнтів на сайті.

1. *Додайте перевірку словом.*

2. *Активація функції автоматичного позначення в AdWords.* До URL-адреси цільової сторінки додано параметр GCLID для визначення перегляду сторінки як результат натискання оголошення.

3. *Реєстрація заявок*. Кожного разу, коли відвідувач надсилає форму, можна реєструвати його IP-адресу, параметр GCLID (якщо доступно), URL-адресу напрямку переходу й агента користувача. На основі цих даних можна визначити, чи надходить трафік від натискань оголошень і чи є випадки підозрілої активності. При підозрілій активості серед відвідувачів, які заходять на сайт через AdWords можна переслати компанії Google параметри GCLID чи IP-адресу цього відвідувача. Команда Google із підтримки якості рекламного трафіку перевірить ваш обліковий запис.

Перегляд даних Google Analytics у звітах AdWords

Переглянувши статистику Google Analytics про взаємодію із сайтом разом із відомостями про ефективність AdWords, можна дізнатися, що роблять користувачі після того, як натискають оголошення й переходять на цільову сторінку. У цій статистиці подано наведені нижче дані.

- **Показник відмов**. Якщо відвідувач сайту переглянув лише одну сторінку чи ініціював лише одну подію, Analytics зараховує це як *відмову*. Показник відмов – це відсоток перерваних сеансів.
- **Середня тривалість сеансу** – середній час, протягом якого користувач залишався на сайті.
- **Сторінки/сеанс** – середня кількість сторінок, переглянутих за один сеанс.
- **% нових сеансів** – приблизний відсоток сеансів користувачів, які відвідали сайт уперше.

Ця інформація показує, наскільки ефективний контент, розміщений на інформаційному ресурсі (кампанії та групи оголошень), і допомагає приймати рішення щодо оптимізації бюджету, ставок, цільових сторінок і тексту оголошень. Наприклад, порівнявши показник відмов і CTR для групи контенту-оголошень (рис. 10), можна отримати уявлення про те, наскільки сайт відповідає очікуванням користувачів, залучених оголошенням.

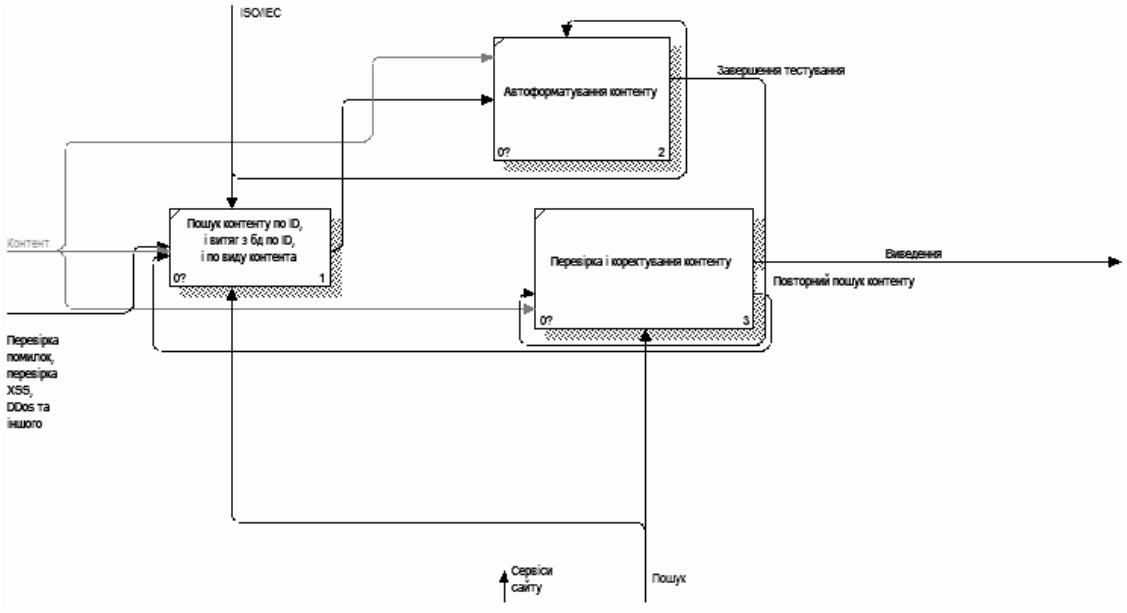


Рис. 10. IDF0 пошук контенту-оголошень

Наприклад, у табл. 21, що група оголошень *тема* 1 має не лише вищий CTR (8 %) порівняно з аналогічним показником для групи оголошень *тема* 2 (6 %), а й вищий показник відмов (60 %).

Таблиця 21

Приклад аналізу груп контенту-оголошень на інформаційному ресурсі

Тематика групи оголошень	Покази	CTR	Кліки оголошення	Показник відмов	Користувачі, які залишаються на сайті
Тема 1	1000	8 %	80	60 %	32
Тема 2	1000	6 %	60	30 %	42

Це означає, що більше половини користувачів, які відвідали сайт після натискання оголошень із групи *тема* 1, не залишилися на ньому, щоб детальніше переглянути пропозиції чи оформити покупку. Незважаючи на те, що група оголошень *тема* 2 отримує менше кліків, рентабельність інвестицій у них вища, оскільки залучені цією рекламою користувачі частіше затримуються на типовому регіональному інформаційному ресурсі (рис. 11).

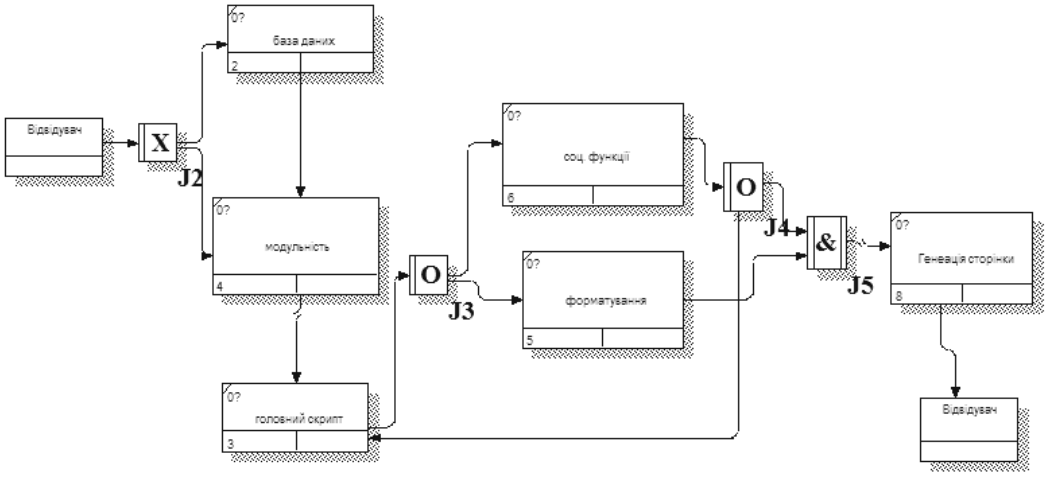


Рис. 11. IDF3 типового регіонального інформаційного ресурсу

Аналіз показників відмов відвідування інформаційного ресурсу

Показник відмов – це відсоток відвідувачів, які переглядають лише одну сторінку під час відвідування вашого сайту. Існує ряд причин, чому може бути високий показник відмов (рис. 12).

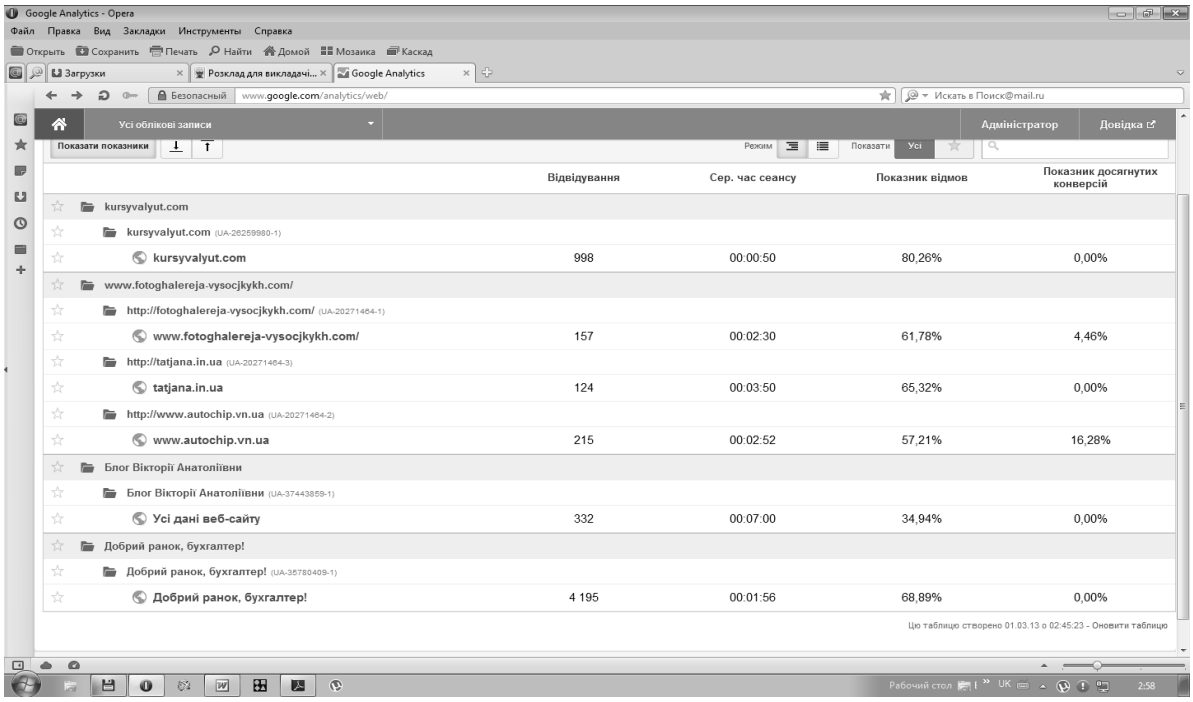


Рис. 12. Приклади показників відмов

Наприклад, відвідувачі можуть залишити сайт на сторінці входу через дизайн сайту чи проблеми зі зручністю використання. Або може бути, що лише певні сторінки на сайті мають високий показник відмов через дуже вагомі причини. Нижче описано деякі проблеми, які можуть спричинити появу високого показника відмов. Щоб зрозуміти різницю між показником виходів і показником відмов для певної сторінки сайту, слід взяти до уваги такі три аспекти.

1. Для всіх переглядів сторінки показником виходів є відсоток останніх за сеанс переглядів.
2. Для всіх розпочатих зі сторінки сеансів показником відмов є відсоток єдиних за сеанс відвідувань.
3. Показник відмов для сторінки розраховується на основі даних лише про ті відвідування, які почалися з цієї сторінки.

Розгляньмо цей останній аспект на простому прикладі. На сайті є сторінки від А до В. Існує лише один на день сеанс із наведеним нижче порядком перегляду сторінок.

Понеділок: сторінка А ® сторінка Б ® сторінка В
Вівторок: сторінка Б ® сторінка А ® сторінка В
Середа: сторінка А ® вихід

Звіт про вміст для Сторінки А покаже три перегляди сторінки, а також показник відмов, який становить 50 %. Показник відмов не становитиме 33 %, бо перегляд сторінки А за четвер не було враховано в розрахунку показника відмов. Відмова – це сеанс, за який відбулася лише одна взаємодія з відвідувачем. Показник відмов для сторінки має значення лише тоді, коли з цієї сторінки починається сеанс. Тепер дослідимо показники виходів і показники відмов для групи днів з одним сеансом на сайті.

Понеділок: сторінка Б ® сторінка А ® сторінка В
Вівторок: сторінка Б ® вихід
Середа: сторінка А ® сторінка В ® сторінка Б
Четвер: сторінка В ® вихід
П'ятниця: сторінка Б ® сторінка В ® сторінка А

Нижче подано розраховані відсоткові показники виходів і відмов.

Показник виходів:

- Сторінка А: 33 % (лише 3 з 5 сеансів містили сторінку А)
- Сторінка Б: 50 % (лише 4 з 5 сеансів містили сторінку Б)
- Сторінка В: 50 % (лише 4 з 5 сеансів містили сторінку В)

Показник відмов

- Сторінка А: 0 % (жоден сеанс не почався зі сторінки А, тому для неї немає показника відмов)
- Сторінка Б: 33 % (показник відмов вищий за показник виходів, оскільки зі сторінки Б почалися 3 сеанси й один із них привів до відмови)
- Сторінка В: 100 % (один сеанс розпочався зі сторінки В й привів до відмови)

Показник досягнутих конверсій

Для обраної цілі i, де $i = \overline{1,20}$, цей показник відображає відсоток відвідувань, результатом яких стала конверсія (досягнення цієї цілі). Конверсія відбувається, коли відвідувач досягає цілі. Існує три типи цілей:

1. *Цільова URL-адреса* – це сторінка, яку бачать відвідувачі одразу після виконання дії. Для реєстрації облікового запису це може бути сторінка *Дякуємо за реєстрацію* або *Дякуємо за покупку,* або сторінка квитанції. Ця ціль ініціює конверсію, коли відвідувач переглядає вказану сторінку.

2. *Час перебування на сайті* – це граничний показник часу, який визначає модератор сайту. Коли відвідувач провів більше чи менше часу на сайті, ніж визначений граничний показник часу, здійснюється конверсія.

3. *Кількість переглянутих сторінок за відвідування* дає змогу визначити граничний показник переглянутих сторінок. Коли відвідувач переглянув більше чи менше сторінок, ніж установлений граничний показник, здійснюється конверсія.

Цілі й послідовності – універсальний спосіб визначення успішності інформаційного ресурсу або програми з огляду на поставлені завдання. Можна налаштувати окремі цілі для відстеження різноманітних дій, наприклад, трансакцій із мінімальною сумою покупки або часу, проведеного на певному екрані. За допомогою послідовності можна вказати очікуваний шлях проходження трафіку для досягнення цілі. Поєднання цілей і послідовностей дає змогу аналізувати, наскільки ефективно сайт або програма спрямовує користувачів до поставленої мети. Кожного разу, коли дія користувача відповідає цілі, в обліковому записі Google Analytics реєструється конверсія. Якщо значення цілі задано в грошовому еквіваленті, дані про конверсії також міститимуть відповідні значення. Рейтинги конверсій із виконанням цілі (тобто рівні виконання цілей) можна переглянути у звітах щодо цілей. Проаналізувати конверсії з виконанням цілі також можна за допомогою інших звітів, включаючи звіти про відвідувачів, трафік, звіти “Пошук на сайті”, а також звіти про події.

Доступно чотири типи цілей на вибір. Коли відвідувач виконує вибрану дію, конверсія ініціюється й фіксується у звітах про цілі. Потрібний тип цілі для відстеження можна вибрати зі списку під час налаштування цілей в обліковому записі.

1. URL-адреса призначення: адреса певного розташування, наприклад веб-сторінки (або віртуальної сторінки) чи екрана програми, які було завантажено. Так, веб-сторінка або екран програми з подякою за реєстрацію можуть бути призначенням для кампанії е-комерції, орієнтованої на визначення кола потенційних клієнтів. Ця ціль чудово поєднується з послідовностями.

2. Тривалість відвідування: відвідування, які тривають протягом установленого часу або довше. Цю ціль можна використовувати для визначення кількості відвідувачів, які залишаються на сторінці чи екрані покупки довше 2 хв.

3. Сторінки/відвідування (для веб-сторінок) Екрани/відвідування (для програм): відвідувач переглядає певну кількість сторінок або екранів під час сеансу. Використовують цей тип цілі, коли потрібно фіксувати відвідувачів, які переглянули, наприклад, не менше 3 сторінок.

4. Подія: відвідувач ініціює дію, яку модератор визначив як подію (наприклад, соціальну рекомендацію або клік оголошення). Перш ніж установити ціль такого типу, потрібно спочатку налаштувати відстеження подій.

Цілі автоматично групуються в набори, проте модератор визначає, які цілі мають належати до кожного з наборів. Використовують набори для класифікації різних типів цілей для сайту. Наприклад, можна відстежувати завантаження, реєстрації та сторінки з квитанціями в окремих наборах цілей. Цілі встановлюються на рівні профілю (не більше 20 в одному профілі). У кожному профілі можна створити чотири набори з максимум п'ятьма цілями. Щоб відстежувати понад 20 цілей для веб-сайту або програми, створюють новий профіль для відповідної власності.

Під час налаштування цілі призначають для конверсії певну вартість у доларах. Щоразу, коли цілі буде досягнуто, це число реєструватиметься. Потім усі випадки зареєстрованого числа додаватимуться і відображатимуться як *Вартість цілі*. Будь-яку дію, виконану користувачем на веб-сайті чи в програмі, можна перенести в суму в доларах. Один зі способів визначити вартість цілі – це підрахувати, як часто користувачі, що досягнули цілі, стають клієнтами. Наприклад, якщо до команди з продажів звертається 10 % користувачів, які підписалися на інформаційні листи, а середня сума трансакції становить 500, то для цілі *Підписка на отримання інформаційних листів* можна призначити вартість 50 (тобто 10 % від 500). Відвідувачі досягають цілі, коли переходять на останню сторінку реєстрації. Якщо ж покупку здійснює лише 1 % з усіх підписаних користувачів, то для цілі *Підписка на інформаційний лист* можна призначити вартість лише 5. Хоча визначати вартість під час налаштування цілей зовсім не обов'язково, користуються цією опцією, щоб мати змогу оцінити вартість і отримати прибуток від будь-яких взаємодій відвідувачів із сайтом. Система Google Analytics також використовує значення вартості цілі для обчислення інших показників, таких як рентабельність інвестицій (ROI) та середній показник.

За допомогою *послідовності* можна вказати очікуваний шлях проходження трафіку для досягнення цільового призначення. Орієнтуючись на вказані модератором етапи послідовності, система Analytics відстежує моменти, коли відвідувачі виходять на шлях до цілі й залишають його. Завдяки цьому можна отримати цінну інформацію про сайт. Так, наприклад, можна визначити сторінку в послідовності, на якій фіксується велика кількість виходів на шляху до досягнення цілі, що вказує на проблеми з цією сторінкою. Якщо є багато пропущених етапів, це вказуватиме на складність навігації або надто довгий шлях до конверсії з великою кількістю зайвих кроків. Послідовності можна використовувати лише в поєднанні з цільовими призначеннями. Остання сторінка в послідовності є сторінкою цілі (введеною як цільове призначення), тоді як попередні сторінки утворюють послідовність. Якщо ціль послідовності – привабити потенційних клієнтів, то першу сторінку послідовності можна призначити URL-адресою форми запиту на контакт, а цільову сторінку – URL-адресою для сторінки подяки за запит, що відображається після того, як користувач надсилає запит на контакт.

Обчислення показників

1. Показник "Порівняно з початковою сторінкою". Цей показник порівнює рейтинги конверсій експериментального варіанта сторінки та початкової сторінки.

Порівняння здійснюється за такою формулою:

$$\left(\frac{\textit{рейтинг конверсій експериментального варіанта}}{\textit{рейтинг конверсій початкового варіанта}} - 1\right) \cdot 100 .$$

Наприклад, якщо експериментальний варіант ефективніший за початковий:

- рейтинг конверсій експериментального варіанта = 35 %;
- рейтинг конверсій початкового варіанта = 25 %;

Тоді значення показника порівняно з початковою сторінкою $\left(\frac{35}{25} - 1\right) \cdot 100 = 40\,\%$.

Якщо експериментальний варіант менш ефективний, ніж початковий:

- рейтинг конверсій експериментального варіанта = 20 %;
- рейтинг конверсій початкового варіанта = 40 %;

Значення показника Порівняно з початковою сторінкою $\left(\frac{20}{40} - 1\right) \cdot 100 = -50\,\%$.

Звіти про багатоканальні послідовності створюються на основі шляхів до конверсії – послідовностей взаємодій (тобто кліків або переходів із каналів), які призвели до кожної конверсії та трансакції протягом 30 днів. Дані про багатоканальні послідовності збираються з даних без вибірки. Якщо у звітах не вказано інше, дані про багатоканальні послідовності враховують усі конверсії, трансакції та шляхи до конверсії, починаючи із січня 2011 року. Збирання даних про багатоканальні послідовності здійснюється із затримкою до двох днів, тому відомості за два минулі дні не відображатимуться у звітах. Google Analytics записує до 5000 взаємодій на один шлях до конверсії. Немає обмеження на кількість унікальних шляхів до конверсії, які можна записати. У звітах про багатоканальні послідовності дані відображаються за умови здійснення принаймні однієї конверсії. Шлях до конверсії створюється для кожної конверсії та трансакції, які реєструються службою Google Analytics. Повторні конверсії одного відвідувача відображаються як різні шляхи. Визначити й аналізувати окремі підмножини шляхів до конверсії можна за допомогою сегментів конверсій. У звітах про багатоканальні послідовності, якщо конверсія відбувається під час "прямого" відвідування (тобто з використанням закладки або введення URL-адреси сайту у веб-переглядачі), конверсія зараховується до "прямого" каналу. Це відрізняється від інших звітів Google Analytics, у яких конверсія зараховується до попередньої непрямої кампанії або джерела за їх наявності. Наприклад, в інших звітах Google Analytics, якщо відвідувач переспрямовується на сайт, що призводить до "прямої" конверсії, "пряме" джерело ігнорується. Натомість конверсія зараховується до переспрямування. У звітах про багатоканальні послідовності "пряме" джерело не ігнорується. Конверсія приписується "прямому" джерелу як останній взаємодії перед конверсією, а переспрямування враховується як допоміжна взаємодія.

Дані про шлях до конверсії створюються для кожної досягнутої конверсії та трансакції е-комерції, зареєстрованої службою Google Analytics. Тому кількість конверсій і трансакцій, яка відображається у звіті про багатоканальні послідовності, відповідатиме кількості конверсій і трансакцій, яка відображається в інших звітах Google Analytics.

2. Збирають дані про багатоканальні послідовності із затримкою до 2 днів. З цієї причини кількість конверсій за останні 2 дні у звіті про багатоканальні послідовності не відповідатиме даним, які відображаються в інших звітах служби Google Analytics.

3. Загальна кількість конверсій у звітах про багатоканальні послідовності – це сума загальної кількості досягнутих конверсій і загальної кількості трансакцій е-комерції.

4. У звітах Google Analytics, відмінних від звітів про багатоканальні послідовності, якщо з відвідуванням не пов'язано кампанію, використовується кампанія з попереднього відвідування. Проте для звітів про багатоканальні послідовності служба Google Analytics вважатиме це відвідування "прямим".

Наприклад, якщо особа переходить на сайт із сайту google.com, конверсія зараховується "прямому" відвідувачу й у звітах Google Analytics відображатиметься 1 конверсія для джерела трафіку "google.com / звичайний". У звітах про багатоканальні послідовності відображатиметься 1 конверсія зі шляхом "google.com / звичайний > прямий / (немає)" у звіті "Найпопулярніші послідовності", а також "прямий / (немає)" з 1 конверсією за останньою взаємодією та "google.com / звичайний" з 1 асоційованою конверсією. Звіти про багатоканальні послідовності обмежені 30-денним періодом ретроспективного аналізу конверсій. Для інших звітів Google Analytics немає періоду ретроспективного аналізу.

Метод обчислення відношення конверсій за допоміжними й останніми взаємодіями

Відношення конверсій за допоміжними й останніми взаємодіями – це результат ділення кількості конверсій, для яких канал виконував допоміжну роль, на кількість конверсій, для яких канал був останньою взаємодією. Якщо канал кілька разів виконував допоміжну роль на шляху до однієї конверсії, як асоційована конверсія для коефіцієнта враховується лише одна конверсія.

Асоційовані конверсії для різних каналів не виключають одна однієї. Для всіх каналів, які виконували допоміжну роль на одному шляху до конверсії, враховується асоційована конверсія.

Тому загальна кількість асоційованих конверсій для окремих каналів може перевищувати загальну кількість асоційованих конверсій для всіх каналів.

Сегментація даних про шлях до конверсії

Сегменти конверсій дають змогу визначити й аналізувати окремі набори шляхів до конверсії у звітах про багатоканальні послідовності. Наприклад, можна створити сегмент конверсії, який містить лише послідовності, першою взаємодією в яких був перехід із певного сайту. Потім можна здійснювати пошук у звіті про багатоканальні послідовності, переглядаючи дані лише для цього сегмента конверсії. Можна також порівнювати дані в цьому сегменті з інформацією в інших сегментах. Щоб переглянути підмножину послідовностей конверсій, необхідно застосувати сегменти конверсій замість профілів із фільтрами. Використання профілю з фільтрами може негативно вплинути на точність звітів про багатоканальні послідовності. Використовують засіб створення умов, щоб указати шляхи до конверсій, які потрібно включити в сегмент. Наприклад, щоб налаштувати сегмент конверсії, який містить лише шляхи до конверсій, що починаються переходом із сайту example.com, необхідно вказати наведені нижче умови.

Атрибутивна модель – це правило або набір правил, що визначає спосіб призначення кредитів за продажі й конверсії точкам дотику з аудиторією в шляхах до конверсій. Наприклад, атрибут "Остання взаємодія" призначає 100 % кредиту кінцевим точкам дотику з аудиторією (тобто, клікам), які передують продажам або конверсіям. Атрибут "Перша взаємодія" призначає 100 % кредиту точкам дотику до аудиторії, які починають шляхи до конверсій. Нижче подано приклад атрибутивних моделей. Розраховане значення конверсії (а також кількість конверсій) для кожного маркетингового каналу змінюватиметься відповідно до використовуваної атрибутивної моделі. Канал, який здебільшого починає шляхи до конверсій, завдяки атрибутивній моделі "Перша взаємодія" матиме вищу цінність конверсії, ніж за використання моделі "Остання взаємодія".

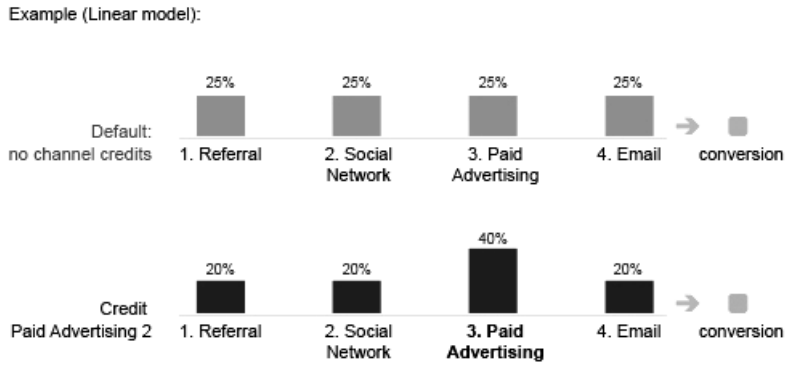
Клієнт знайшов сайт, натиснувши одне з оплачених оголошень. Він повертається за тиждень через соціальну мережу. Того ж дня він повертається втретє, щоб придбати товар – цього разу через одну з кампаній на основі е-пошти. В атрибутивній моделі :

- *Остання взаємодія* – остання точка дотику до аудиторії – у цьому випадку канал е-пошти – отримає 100 % кредиту за продаж.
- *Перша взаємодія* – перша точка дотику до аудиторії – у цьому випадку канал оплачених оголошень – отримає 100 % кредиту за продаж.
- *Лінійна* кожна точка дотику до аудиторії на шляху до конверсій – у цьому випадку канал оплачених оголошень, соціальна мережа й канали е-пошти – отримає рівну частину (33,3 % кожному) кредиту за продаж.
- *Ослаблення з часом* точки дотику до аудиторії, найближчі за часом до продажу або конверсії, отримують найбільшу частину кредиту. У цьому випадку канали е-пошти й соціальної мережі отримають найбільшу частину кредиту, оскільки клієнт взаємодіяв із ними в день конверсії. Оскільки взаємодія з оплаченим оголошенням сталася на тиждень раніше, цей канал отримає значно менший кредит.
- *На основі позиції* 40 % кредиту призначається першій взаємодії, 20 % – проміжним взаємодіям і 40 % – останній взаємодії. Канали оплачених оголошень й е-пошти отримають по 40 %кредиту, а соціальна мережа – 20 %.

Засіб атрибутивного моделювання дає змогу порівняти вплив різних атрибутивних моделей на оцінювання маркетингових каналів. Виберіть у засобі атрибутивну модель (наприклад, Остання *взаємодія*), після чого в таблиці з'явиться кількість конверсій (або залежно від вибору ціна) для кожного каналу, обчислена за вибраною моделлю. Можна вибрати до трьох атрибутивних моделей одночасно й порівняти результати кожної з них у таблиці. Після імпорту даних про вартість в AdWords або іншої подібної інформації з'являться стовпці, за допомогою яких можна проаналізувати та порівняти відповідні показники. Окрім базових моделей, можна використати засіб атрибутивного моделювання для створення, збереження й застосування користувацької моделі, яка використовує вказані модератором правила. Це дає змогу пристосувати моделі

спеціально до набору припущень, які потрібно оцінити в даних шляху до конверсії. У розділі “Застосування спеціальних правил кредитування” можна вказати умови, які визначають точки дотику на шляху до конверсії відповідно до таких характеристик, як позиція (перша, остання, середня, допоміжна), тип точок дотику (клік, показ, пряме відвідування), а також тип кампанії або джерела трафіку (кампанія, ключове слово чи інші параметри). Після визначення потрібних точок дотику можна вказати спосіб розподілення кредиту конверсії для них відносно інших точок.

У всіх правилах визначається відносний розподіл кредитів. Наприклад, лінійна модель поділяє кредит конверсії порівну між точками дотику. Тому на шляху до конверсії з чотирма взаємодіями кожна точка дотику отримає 25 % кредиту. Однак якщо до каналу оплачених оголошень застосується кредит 2 точок, а третя точка на шляху – “Оплачене оголошення”, то кредит застосується так:



Після застосування кількох правил до однієї точки дотику вагу кредиту для правил, що перекриваються, буде помножено. Як у прикладі вище, якщо за другим правилом визначено, що “Пошук за брендом” отримує, як і інші точки дотику, 0,1 кредиту, точки дотику типу “Пошук за брендом”, які відповідають правилу, визначеному як “Оплачене оголошення”, отримають 2×0,1=0,2 кредиту. Для оцінювання ефективності своїх каналів використовують атрибутивні моделі, які відображають рекламні цілі й бізнес-моделі. Незалежно від використовуваних моделей, можна перевірити свої припущення за допомогою експерименту. Збільште або зменште інвестиції в канал за результатами моделювання, а потім перегляньте результати в даних.

Модель *Остання взаємодія* передає 100 % ціни конверсії останньому каналу, з яким взаємодіяв користувач перед покупкою або здійсненням конверсії. Google Analytics використовує цю модель за умовчанням для передавання ціни конверсії у звіти, які не належать до багатоканальних послідовностей. Оскільки модель “Остання взаємодія” – модель за замовчуванням для звітів, які не належать до багатоканальних послідовностей, вона дає корисні результати, які можна порівняти з результатами інших моделей. Також якщо оголошення й кампанії розроблено для приваблення людей у момент покупки або якщо компанія працює в основному на трансакціях і має цикли продажу, який не містить фази обмірковування, модель “Остання взаємодія” може бути прийнятною.

Модель *Перша взаємодія* передає 100 % ціни конверсії першому каналу, з яким взаємодіяв користувач. Ця модель прийнятна, якщо покази оголошень або кампанії виконуються для створення початкової поінформованості. Наприклад, якщо бренд не дуже відомий, можна зосередити зусилля на ключових словах або каналах, які першими представляють бренд клієнту.

Модель *Лінійна* розподіляє порівну кредит кожному каналу взаємодії на шляху до конверсії. Модель корисно використовувати, якщо кампанії розроблено для підтримання контакту з клієнтом і його поінформованості під час усього циклу продажу. У такому випадку кожна точка важлива для процесу обмірковування.

Якщо цикл продажів містить коротку фазу обмірковування, можна використати модель *Ослаблення з часом*. Ця модель передає найбільшу частину кредиту точкам дотику, які сталися в часі якнайближче до конверсії. Якщо виконуються одно- або дводенні рекламні кампанії, можна

передати більше кредиту взаємодіям протягом рекламних днів. У цьому випадку взаємодії, що сталися за тиждень до просування, отримають лише невелику частку кредиту порівняно з точками дотику біля конверсії. Модель *Ослаблення з часом* дає змогу відповідним чином кредитувати точки дотику під час одного або двох днів, які ведуть до конверсії.

На рис. 12 подано контекстну діаграму потоків даних типового регіонального інформаційного ресурсу, а на рис. 13 – деталізовану діаграму потоків даних типового регіонального інформаційного ресурсу.

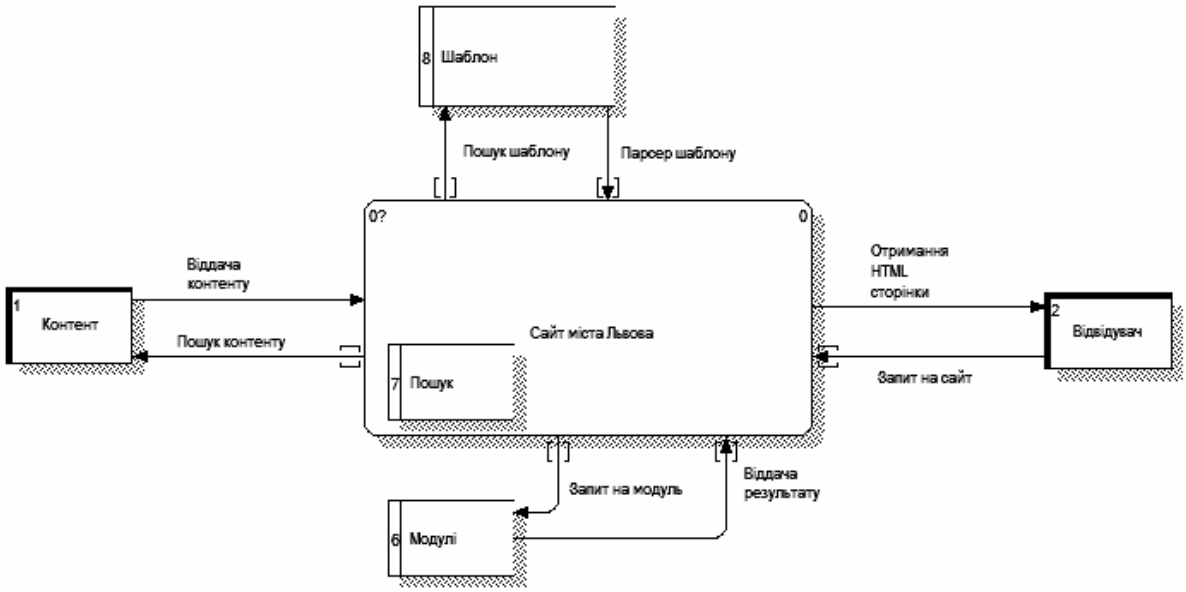


Рис. 12. Контекстна DFD типового регіонального інформаційного ресурсу

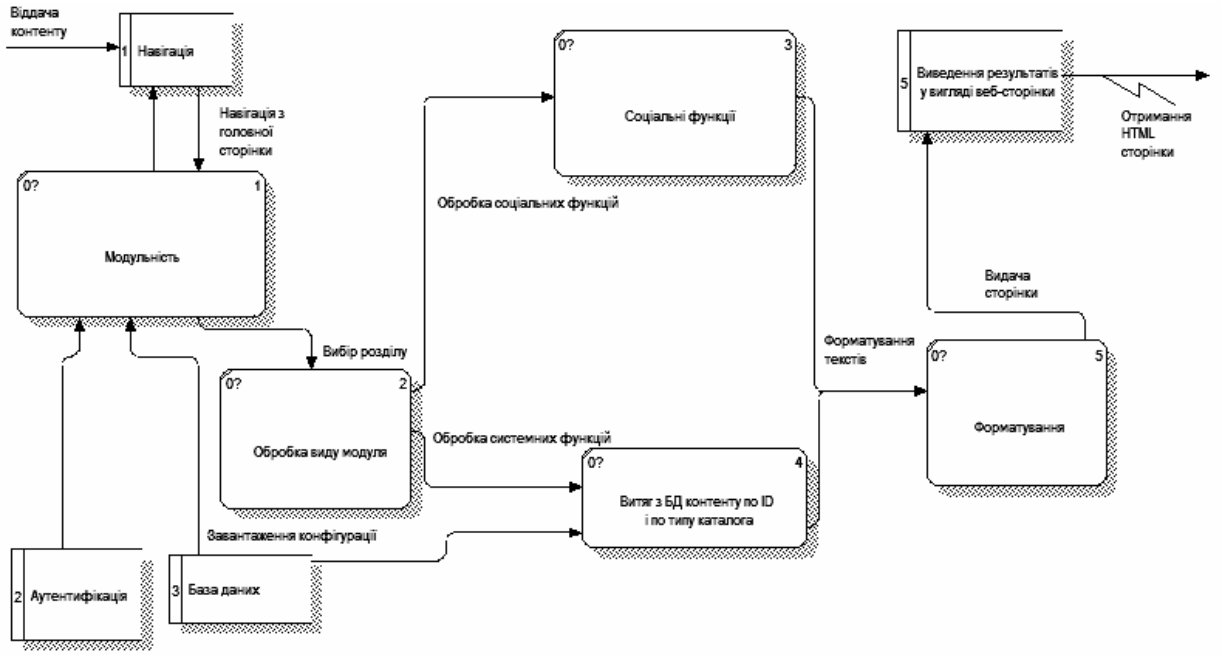
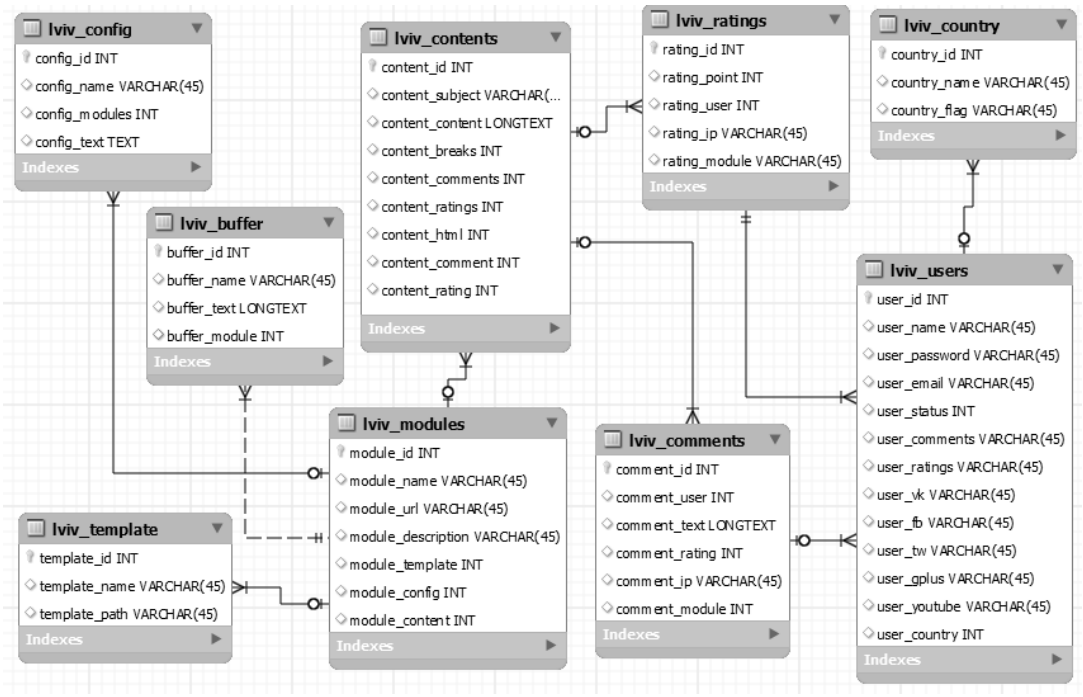


Рис. 12. Деталізована DFD типового регіонального інформаційного ресурсу

Модель *На основі позиції* дає змогу поєднати моделі *Остання взаємодія* й *Перша взаємодія*. Замість того, щоб віддавати весь кредит першій або останній взаємодії, його можна поділити між ними. Зазвичай 40 % кредиту призначають першій та останній взаємодії, а 20 % – проміжним взаємодіям. Якщо цінуєте більше ті точки дотику, які познайомили користувача з брендом, а також точки, які призвели до продажу, використовують модель *На основі позиції*.

На рис. 14 подано ERD типового регіонального інформаційного ресурсу.



Математичне співвідношення для детермінованих систем у загальному випадку:

$$\frac{dy}{dt} = f(y,t), y(t_0) = y_0, y = \begin{pmatrix} y_1 \\ \mathbf{L} \\ y_n \end{pmatrix}$$

Для сайта міста Львова розглянемо модель адміністратор-відвідувач:

$$\begin{cases} \dfrac{dy_1}{dt} = \dfrac{(y_2^{\text{Cavg}})}{\text{ALL}} y_1 \\ \dfrac{dy_2}{dt} = \left(\dfrac{y_2}{\text{ALL}} \right)^{\text{Call}} \end{cases}, \quad \text{ALL} \in (0;1]$$

де y_1 – кількість адміністраторів; y_2 – кількість відвідувачів; Cavg – кількість контенту; Call – загальна кількість контенту; ALL – корисність інформації, що визначається в шкалі від 0 до 1.

Керована система Лоткі–Вольтерра має вигляд:

$$\begin{cases} \dfrac{dy_1}{dt} = \dfrac{(y_2^{\text{Cavg}})}{\text{ALL}} y_1 + f_1 u_1(t) \\ \dfrac{dy_2}{dt} = \left(\dfrac{y_2}{\text{ALL}} \right)^{\text{Call}} + f_2 u_2(t) \end{cases}, \quad \text{ALL} \in (0;1]$$

$$I = \int_0^T \left[\frac{1}{2} \left(y - y^* \right)^T R \left(y - y^* \right) + \frac{1}{2} u^T Q u \right] dt \to \min_y$$

Лінеарізуємо функцію:

$$\frac{dy_1}{dt} = \frac{(y_2^{\text{Call}})}{\text{ALL}} y_1 = 0\left(y - y^* \right) + \frac{\ln \text{Call}}{\text{ALL}} \cdot y_1^* \left(y_2 - y_2^* \right) + o\left(y - y^* \right)$$

$$\frac{dy_2}{dt} = \left(\frac{y_2}{\text{ALL}} \right)^{\text{Cavg}} y_1 = \left(\frac{1}{\text{ALL}} \right) y_2^* \left(y_1 - y_1^* \right) + 0\left(y_2 - y_2^* \right) + o\left(y - y^* \right)$$

$$A=\begin{pmatrix} 0 & \frac{\ln \mathrm{Call}}{\mathrm{ALL}} y_1^* \\ \left(\frac{1}{\mathrm{ALL}}\right) y_2^* & 0 \end{pmatrix}$$

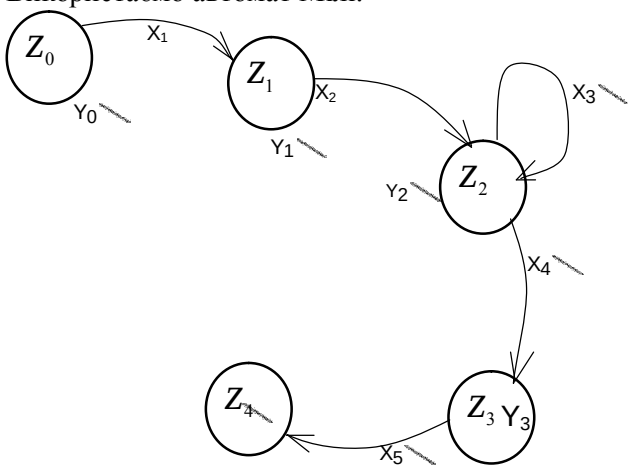
Дискретно-детермінована модель інформаційної системи. Для реалізації цієї системи можна використовувати 1 з 2-х автоматів: Мілі або Мура.

Автомат 1-го роду – автомат Мілі: $\begin{cases} z(t+1)=j\left[z(t),x(t)\right], t=0,1,2,...; \\ y(t)=y\left[z(t),x(t)\right], t=0,1,2,...; \end{cases}$

Автомат 2-го роду – автомат Мура: $\begin{cases} z(t+1)=j\left[z(t),x(t)\right], t=0,1,2,...; \\ y(t)=y\left[z(t-1),x(t)\right], t=0,1,2,...; \end{cases}$

Для котрого: $y(t)=y\left[z(t)\right], t=0,1,2,...;$

Використаємо автомат Мілі:



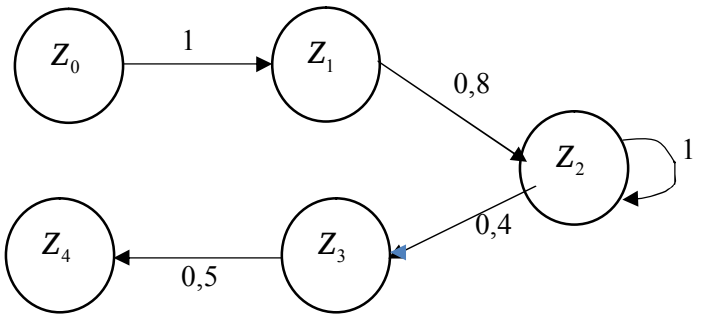
де Z_0 – запит на сервер; Z_1 – обробка запиту; Z_2 – запуск скрипта; Z_3 – з'єднання з БД; Z_4 – генерація і віддача сторінки відвідувачу.

Матриця функцій переходів:

X_i	Z_k				
	Z_0	Z_1	Z_2	Z_3	Z_4
Переходи					
X_1	Z_1	-	-	-	-
X_2	-	Z_2	-	-	-
X_3	-	-	Z_2	-	-
X_4	-	-	-	Z_3	-
X_5	-	-	-	-	Z_4
Виходи					
X_1	Y_1	-	-	-	-
X_2	-	Y_2	-	-	-
X_3	-	-	Y_2	-	-
X_4	-	-	-	Y_3	_
X_5	-	-	-	-	Y_4

Дискретно-стохастична модель інформаційної системи

1	0	0	0	0
0	0,8	1	0	0
0	0	0,4	0	0
0	0	0	0,5	1
0	0	0	0	0



Неперервно-стохастична модель інформаційної системи

Кожного дня у БД додаються нові статті, нова інформація. Отже, на початку робочого дня в накопичувачі немає нічого. Тобто стани підсистеми описуватимуться системою рівнянь:

$$\begin{cases} P_n(t+\mathrm{W}t) = P_n(t)\left[1-(l+m)\mathrm{W}t\right] + P_{n-1}(t)\, l\,\mathrm{W}t + P_{n+1}(t)\, m\mathrm{W}t, n \geq 1, \\ P_0(t+\mathrm{W}t) = P_0(t)(1-l\,\mathrm{W}t) + P_1(t)\, m\mathrm{W}t, \end{cases}$$

де $P_n(t)$ – ймовірність перебування системи у стані; $z_n(t) \in Z$ у момент часу *t,* тобто у ній *n* заявок. Диференціюємо систему, якщо у ній є *n* заявок:

$$\begin{cases} dP_n(t)/dt = -(l+m)P_n(t) + l\,P_{n-1}(t) + mP_{n+1}(t) \\ dP_0(t)/dt = -l\,P_0(t) + mP_1(t) \end{cases}$$

При прирівняні часу до нуля:

$$\begin{cases} (l+m)\,p_n = l\,p_{n-1} + m\,p_{n-1}, n \geq 1, \\ l\,p_0 = m\,p_0, \end{cases} \quad \begin{cases} (1+r)\,p_n = p_{n+1} + r\,p_{n-1}, n \geq 1, \\ p_1 = r\,p_0. \end{cases}$$

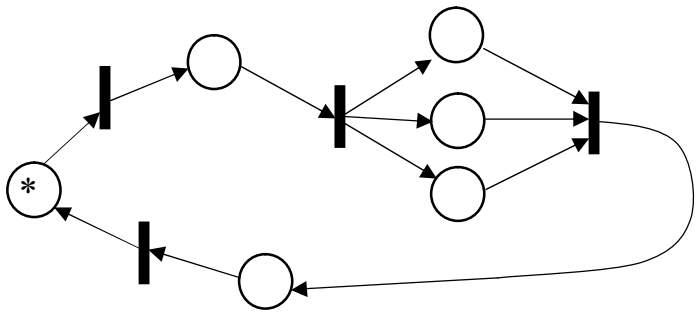
При $p_1 = 1$: $p_n = r^n(1-r)$

Математичне очікування кількості контенту, що перебуває у системі:

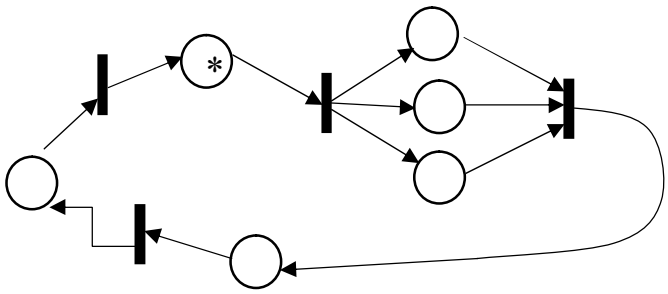
$$l_n = \sum_{n=0}^{\infty} np_n = (1-r)\sum_{n=0}^{\infty} nr^n = r(1-r)$$

Середній час перебування: $l_H = l_n/l = r^2/\left[l(1-r)\right]$

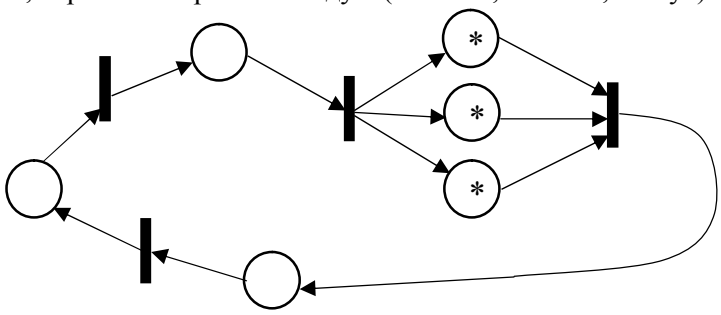
Мережна модель. За допомогою мережної моделі побудуємо принцип роботи CMS для проекта. Почнемо з того, що користувач задав запит на сервер:



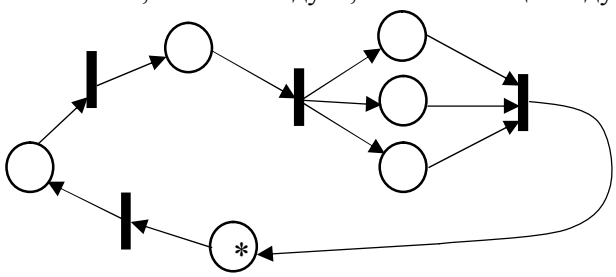
Далі сервер передає дані на обробник скриптів, а сам скрипт на роутеринг, і з ротутеринга на модульність, де відбувається пошук модуля, шаблона по GET запиту.



Після роутеринга, скрипт вибирає тип модуля(контент, каталог, пошук):



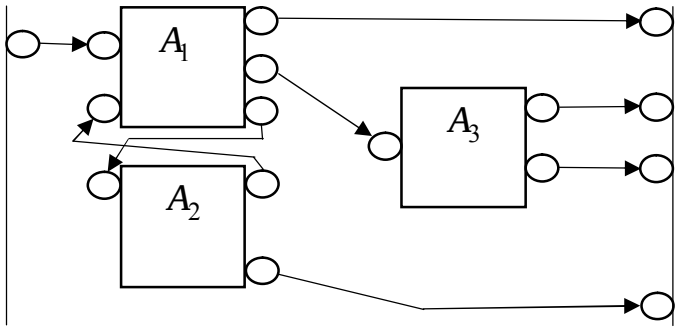
Завантажуються тип контента, шаблон модуля, системні опції модуля, контент і сам модуль:



Сервер виводить згенеровану сторінку користувачу в браузер.

Комбінована модель

Агрегативна система



A_1 – сервер, A_2 – CMS нашого проекту, A_3 – БД.

1 – запит на сервер.

2 – з’єднання з БД.

3 – помилки.

4 – видача клієнту згенерованої сторінки, завершення роботи.

5 – передавання даних методом POST, GET, HEAD.

6 – отримання даних методом POST, GET, HEAD.

7 – видача контенту за критеріями.

Висновки та пперспективи подальших наукових розвідок

1. Проаналізовано способи формування комерційного контенту та досліджено відомі моделі життєвого циклу контенту та стандартизовані сервіси управління контенту, що дає можливість визначити вимоги для створення оптимального життєвого циклу комерційного контенту.

2. Досліджено інтернет-технології для побудови сервіс-орієнтованих систем електронної комерції, що дало можливість класифікувати системи електронної комерції та системи електронної контент-комерції.

3. Детально розглянуто інформаційні ресурси та виробничі процеси систем електронної комерції, що дає можливість розробити оптимальний життєвий цикл контенту та типову архітектуру систем електронної контент-комерції.

4. Проаналізовано технологія управління контентом в електронній комерції, що дає змогу розробити формальні моделі, уніфіковані методи та програмні засоби опрацювання інформаційних ресурсів у системах електронної контент-комерції.

5. З позиції системного підходу проаналізовано сучасні методи і засоби проектування, моделювання та реалізації систем електронної контент-комерції, а також обґрунтовано необхідність та доцільність створення уніфікованих методів та програмних засобів опрацювання інформаційних ресурсів.

1. Береза А. М. Електронна комерція / А. М. Береза. – К.: КНЕУ, 2002 р. 2. Большакова Е. И. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учеб. пособие / Е. И. Большакова, Э. С. Клышинский, Д. В. Ландэ, А. А. Носков, О. В. Пескова, Е. В. Ягунова. – М.: МИЭМ, 2011. – 272 с. 3. Джерк Н. Разработка приложений для электронной коммерции / Н. Джерк, СПб.: Питер, 2001 р. 4. Евдокимов Н. В. Раскрутка Web-сайтов. Эффективная Интернет-коммерция. / Н. В. Евдокимов. – М.: Вильямс, 2007. – 160 с. 5. Ермаков А. Е. Лингвистическая модель для компьютерного анализа тональности публикаций СМИ / А. Е. Ермаков, С. Л. Киселев // Компьютерная лингвистика и интеллектуальные технологии: труды Международной конференции Диалог'2005. – М.: Наука, 2005. – С. 282–285. 6. Іванов В. Ф. Контент-аналіз: Методологія і методика дослідження ЗМК: навч. посібник / Наук. ред. А. З. Москаленко / В.Ф. Іванов. – К.,1994. – 112 с. 7. Клифтон Б. Google Analytics: профессиональный анализ посещаемости веб-сайтов / Б. Клифтон. – М. : Вильямс, 2009. – 400 с. 8. Основы моделирования и оценки електронных информационных потоков / [Д. Ландэ, В. Фурашев, С. Брайчевский, О. Григорьев]. – К. : Інжиніринг, 2006. – 348 с. 9. Ландэ Д. Основы интеграции информационных потоков: монография / Д. Ландэ. – К. : Інжиніринг, 2006. – 240 с. 10. Поспелов Д. Ситуационное управление: теория и практика / Д. Поспелов. – М.: Наука. – 1986. – 288 с. 11. CM Lifecycle Poster [Electronic resource] / Content Management Professionals. – Retrieved 20 July 2010. – Access mode: http://www.cmprosold.org/resources/poster/. – Title from the screen. 12. EMC. Content Management Interoperability Services. Appendices. Version 0.5 / EMC, IBM, Microsoft. – Hopkinton : EMC, 2008. – 17 p. 13. EMC. Content Management Interoperability Services. Part I. Version 0.5 / EMC, IBM, Microsoft. – Hopkinton : EMC, 2008. – 76 p. 14. EMC. Content Management Interoperability Services. Part II – REST protocol binding. Version 0.5 / EMC, IBM, Microsoft. – Hopkinton : EMC, 2008. – 79 p. 15. EMC. Content Management Interoperability Services. Part II – SOAP protocol binding. Version 0.5 / EMC, IBM, Microsoft. – Hopkinton : EMC, 2008. – 37 p. 16. Hackos J. Content Management for Dynamic Web Delivery / J. Hackos. – Hoboken : Wiley, 2002. – 432 p. 17. Halvorson K. Content Strategy for the Web / K. Halvorson. – Reading : New Riders Press, 2009. – 192 p. 18. McGovern G. Content Critical / G. McGovern, R. Norton. – Upper Saddle River : FT Press, 2001. – 256 p. 19. McKeever S. Understanding Web content management systems: evolution, lifecycle and market / S. McKeever // Industrial Management & Data Systems (MCB UP), 2003. – № 103 (9). – P. 686–692. 20. Nakano R. Web content management: a collaborative approach / R. Nakano. – Boston: Addison Wesley

Professional, 2002. – 222 p. 21. Osgood C. The nature and measurement of meaning / C. Osgood // Psychological Bulletin, 49 (1952). – P. 197-237. 22. Papka R. On-line News Event Detection, Clustering, and Tracking : thesis for the degree doctor of philosophy / R. Papka. – Amherst : Massachusetts University, 1999. – 154 p. 23. Woods R. Defining a Model for Content [Electronic resource] / R. Woods. – 2010. – Access mode: http://www.contentmanager.net/magazine/article_785_defining_a_model_for_content_governance.html. – Title from the screen. 24. Rockley A. Managing Enterprise Content: A Unified Content Strategy / A. Rockley. – Reading : New Riders Press, 2002. – 592 p. 25. Stone W. R. Plagiarism, Duplicate Publication and Duplicate Submission: They Are All Wrong! / W.R. Stone // IEEE Antennas and Propagation, 2003. – Vol. 45. – № 4. – P. 47-49. 26. Sullivan D. Invisible Web Gets Deeper [Electronic resource] / D. Sullivan // Search Engine Report. – 2002. – Access mode: http://searchenginewatch.com/sereport/article.php/2162871. – Title from the screen. 27. The Content Management Possibilities Poster [Electronic resource] / Metatorial Services, Inc. – Retrieved 20 July 2010. – Access mode: http://metatorial.com/pagea.asp?id=poster. 28. В. В. Корнеев, А. Ф. Гареев, С. В. Васютин, В. В. Райх. Базы данных. Интеллектуальная обработка информации. Москва Издательство "Нолидж", 2000. 352 с. 29. Г. Г. Белоногов, В. И. Богатырев. Автоматизированные информационные системы. – М.: Сов. радио, 1973. 30. Э. В. Попов. Общение с ЭВМ на естественном языке. – М.: Наука, 1982. 31. V. N. Gudivada. Поиск информации в World Wide Web. ComputerWeek N35_97 с. 19–21, 26, 27. 32. P. J. Hayes and S. P. Weinstein. Construe/TIS: A system for content-based indexing of a database of news stories. In Innovative Applications of Artificial Intelligence 2, pages 49-64. The AAAI Press/The MIT Press, Cambridge, MA, 1991. 33. Н. В. Лукашевич. Автоматическое рубрицирование потоков текстов по общественно-политической тематике // НТИ. Серия 2. Информационные процессы и системы. 1996. №10. С.22–30. 34. Гареев А. Решение проблемы размерности словаря при использовании вероятностной нейронной сети для задач информационного поиска. Нейрокомпьютеры: разработка, применение. – № 1. – С. 60–63. 2000. 35. Цаленко М. Ш. Моделирование семантики в базах данных. – М.: Наука, 1989.