

# Інформаційно-логічні та архітектурні засади універсальних систем керування web-контентом

*Титенко С.В., к.т.н.*

*Національний технічний університет України «КПІ», м. Київ  
lab@setlab.net*

Робота присвячена проблемам моделювання та розробки універсальних систем керування контентом в веб-середовищі. Подаються ключові концептуальні та архітектурні передумови побудови таких систем.

## Вступ

Керування контентом – галузь, з якою користувачі всесвітньої мережі WWW зустрічаються щодня. Системи керування контентом (*англ.* content management systems, CMS) стали базовим і невід’ємним програмним забезпеченням інформаційних веб-ресурсів. Проблема інтелектуалізації та автоматизації процесів керування веб-контентом не втрачає своєї актуальності, незважаючи на безліч промислових CMS-систем.

## Універсальні CMS-системи та керування знаннями в Web-середовищі

Детальний огляд та класифікація CMS-систем здійснені в роботі українських авторів [1]. Тут широко висвітлено сукупність технічних аспектів з акцентом на функціональні можливості систем керування контентом. Спроба узагальнити структуру CMS-системи подані в роботі [2]. Дана структура покриває сукупність низькорівневих технічних задач CMS. Натомість керування веб-контентом в широкому сенсі фактично передбачає керування знаннями в веб-середовищі, що вимагає застосування методів інформаційного аналізу, моделювання і формалізації знань та інформації [3-7].

*Універсальні CMS-системи* – це такі системи керування контентом, які дозволяють подавати в веб-середовищі довільні інформаційні структури предметних областей, надаючи зручні механізми конструювання, публікації та навігації по інформаційним сутностям предметної області

веб-ресурса. Такі задачі зумовлюють потребу в особливій увазі до проблем побудови інформаційних моделей контенту CMS.

Питанням інформаційної моделі CMS та перспективам семантичного моделювання контенту присвячена роботи Олещького О.В., зокрема [8]. В роботі наголошується, що CMS-системи є документо-орієнтованими, пропонується формалізувати відношення документів системи та понять предметної області, а також ввести структуру навігаційних посилань у вигляді графа. У даному графі окрім динамічних відношень, пов'язаних з моделлю предметної області, пропонується використовувати ієрархічні відношення між елементами контенту. Натомість, у роботі не подано конкретних шляхів реалізації програмної системи та прикладів дослідної апробації запропонованих ідей. У подальших роботах розвинуто формальний теоретико-множинний апарат для запропонованого підходу. З іншого боку значна кількість робіт з галузі побудови навчальних веб-систем також детально розглядають питання формалізації контенту у поєднанні з моделюванням предметної області [9,10]. Узагальнюючи, можна підсумувати, що склався підхід до моделювання контенту інтелектуальних веб-систем, що передбачає семантичну структуру концептів предметної області, структуру елементів контенту та зв'язки між концептами та контентом. Актуальними є роботи, що разом з концептуальними пропозиціями акцентують увагу на працюючих проектних рішеннях та практичних апробаціях іноваційних методів керування контентом в універсальних CMS.

В попередніх роботах було подано модель контенту та перспективи керування знаннями в межах керування контентом в веб-середовищі [3-7]. Завданням даної роботи є узагальнення інформаційно-логічних та архітектурних вимог до універсальних CMS-систем та їх програмна реалізація.

## **Інформаційно-логічні та архітектурні засади CMS**

Результатом досліджень, проектування та програмної реалізації CMS став перелік інженерно-технічних та модельних рішень, що повинні бути покладені в основу системи керування контентом універсального типу (рис.1).

Об'єктна модель сторінки.

Ключовою інформаційною сутністю систем керування контентом є сторінка або елемент контенту. Об'єкт сторінки інкапсулює в собі усю логіку роботи із сторінкою, надаючи інтерфейс іншим програмним компонентам CMS для отримання доступу до інформації даного елемента контенту. На рівні БД CMS-системи існує спеціальна таблиця, що пред-

ставляє дану сутність, характеризуючи її певним набором полів. Основними полями об'єкта сторінки є заголовок, контент, анотація та ін. Так в розповсюджених CMS сторінка також моделюється і подається певним чином: в Drupal загальний елемент контенту позначається терміном Node (вузол), у Wordpress – Post (допис), в Joomla – матеріал.

Спираючись на об'єкт сторінки, відбувається процес публікації (рис.2).

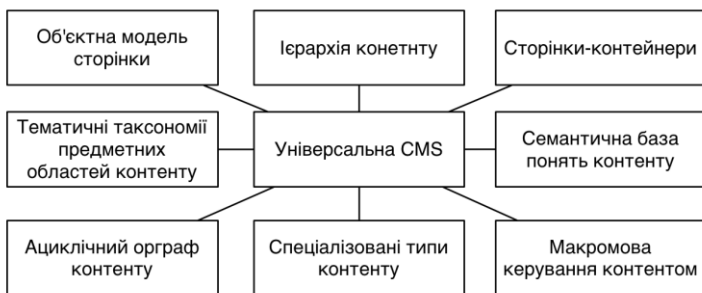


Рис.1. Архітектурні та модельні компоненти CMS універсального типу

## Ієрархія контенту

Ієрархія – це розташування частин або елементів цілого в певному порядку від вищого до нижчого. Ієрархія – найзручніший і природно зрозумілий спосіб структурування речей, це ж справедливо і для елементів контенту. Незважаючи на те, що ієрархічне структурування в Web викликає труднощі у випадках великої кількості інформації, тим не менше ієрархія, як базова структура, для організації контенту є зручним і практично доцільним способом структурування інформації веб-ресурсу.

## Сторінки-контейнери

Сторінки-контейнери – сторінки, головним призначенням яких є подання списку анотованих посилань на дочірні сторінки. Власний контент таких сторінок, як правило, не має великого об'єму і є короткою анотацією до розділу або набору статей. Сторінки-контейнери займають друге місце після простих інформаційних сторінок по частоті застосування в веб-ресурсах. Наявність сторінок-контейнерів як стандартної функціональності системи є необхідною передумовою для універсальності CMS-системи. За допомогою сторінок-контейнерів моделюються найрізноманітніші розділи веб-ресурсу – добірки статей, розділи новин, категорії товарів та будь-які інші набори інформаційних об'єктів предметної області веб-ресурсу.

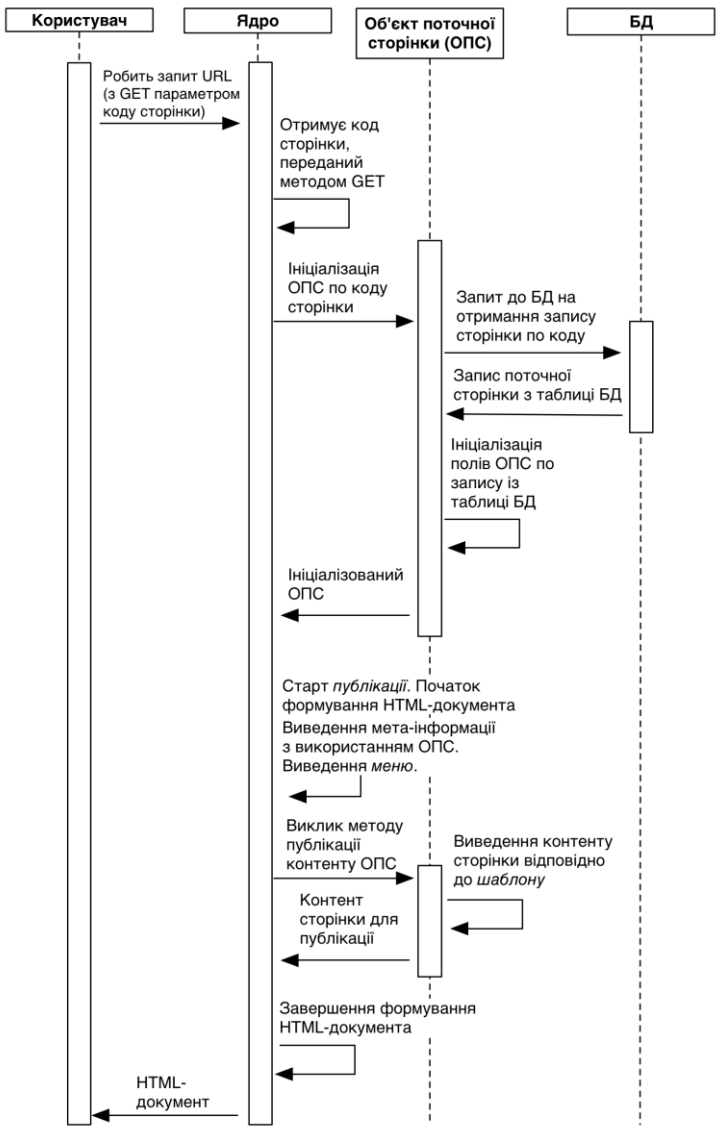


Рис.2. Діаграма послідовності CMS

## Ациклічний оргграф контенту

Тенденція примноження інформації ускладнює структурування інформаційних елементів у строгій деревоподібній формі. Крім цього строга ієрархія часто не задовольняє навігаційним вимогам веб-ресурсу. Потреба одночасної присутності одного і того ж елемента контенту в різних розділах сайту – досить розповсюджене і актуальне завдання. Прикладом цього може служити новина, яка по своєму змісту одночасно задовольняє різним розділам новин, товар, що підходить до різних категорій, деяка службова інформаційна сторінка сайту, яку доцільно розмістити в різних частинах сайту.

Подібні об'єктивні неоднозначності вимагають від системи функціональної можливості розташування сторінки сайту в різних місцях ієрархії. Доцільною структурою для такого завдання є ациклічний орієнтований граф. Ациклічний оргграф є узагальненням дерева. Для оптимізації інформаційно-логічної структури CMS-системи ациклічний оргграф доцільно реалізувати на базі відношень псевдонімів [4] (рис. 3).

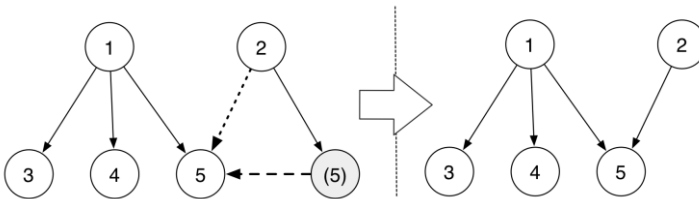


Рис. 3. Принцип роботи відношення псевдонімів

## Спеціалізовані типи контенту

Актуальним завданням розвинених CMS-систем залишається дослідження, розробка та вдосконалення засобів універсального керування інформаційними об'єктами веб-ресурсів, що дозволить спростити створення нових ресурсів різного призначення, а також забезпечить ефективні механізми їх супроводження та налаштування. В межах даної задачі потрібно вирішувати наступні завдання: створення спеціалізованих інформаційних об'єктів із користувацьким набором полів, керування відображенням таких об'єктів, сортування, фільтрація та навігація по сховищу створених об'єктів, інтеграція сховища інформаційних об'єктів із загальним сховищем контенту, керування ієрархією об'єктів, класифікація та групування об'єктів тощо. Огляд інформаційно-логічних

методів моделювання спеціалізованих типів контенту подано в попередній роботі [5]

### Макромова керування контентом.

Макромова керування контентом в CMS – це спеціальна мова логічного керування контентом, оператори якої вбудовуються безпосередньо в інформаційний html-текст веб-сторінки. Такі команди можуть використовуватись на рівні контент-менеджера або архітектора веб-сайту.

Робота з контентом часто передбачає багаторазове виконання однотипних задач оформлення інформації. Виконання подібних задач вручну призводить до необхідності повторного формування html-розмітки, додаткового програмування та нагромадження в контенті сайту технічно-залежного коду. Ефективним розв'язком проблеми є створення системи *макрокоманд контенту*, що дозволять зосередити логіку оформлення повторюваних фрагментів контенту в одному місці [6].

### Тематичні таксономії предметних областей контенту

Тематичні групи служать для організації різноманітних міжпредметних і внутрішньопредметних зв'язків між елементами контенту. Тематичні групи використовуються для моделювання предметних областей, каталогізації, групування і вибірки асоціативного контенту. Організація тематичних груп відбувається у ієрархічній структурі. Це дозволяє вбудовувати таксономію предметних областей. Кожен елемент контенту може брати участь у довільній кількості тематичних груп [4].

### Семантична база понять контенту

Для реалізації керування контентом на рівні семантики предметної області необхідно створювати сховище понять предметної області контенту та формалізувати відношення між ділянками контенту і поняттями предметної області. Така задача вирішується на основі онтологій та семантичних моделей предметних областей [7,8,10].

## Висновок

У роботі представлено сукупність архітектурних, інженерно-технічних та інформаційно-логічних рішень для побудови на їх ґрунті універсальної системи керування контентом. Запропонований комплекс дозволяє на програмному рівні реалізувати ключові завдання універсальних CMS щодо конструювання, публікації та навігації по інформаційним сутностям предметної області веб-ресурсу. Прототип системи апробований на порталі [znannya.org](http://znannya.org). Публікації та лекційний відеоматеріал щодо теми дослідження розміщено на сайті [setlab.net](http://setlab.net).

## Література

1. *Курилов М. А.* Классификация систем управления содержимым web-ресурсов и их использование для разработки сайта дистанционного обучения / М.А. Курилов, С.В. Терещенко // Штучний інтелект. — 2010. — № 3. — С. 648-654.
2. *Любченко С. С.* Формирование обобщенной структуры системы управления содержимым web-сайта / С. С. Любченко // Искусств. интелект. - 2005. - № 2. - С. 109-115.
3. *Титенко С. В.* Моделювання області знань в системі безперервного навчання на основі інтеграції моделі контенту Tree-Net і понятійно-тезисної моделі / С. В. Титенко, О. О. Гагарін // VIII международная конференция «Интеллектуальный анализ информации ИАИ-2008», Киев, 14-17 мая 2008г. : Сб. тр./ Ред. кол. : С.В. Сирота (гл.ред.) и др. – К.: Просвіта, 2008. – С. 475-484.
4. *Титенко С. В.* Модель навчального Web-контенту Tree-Net як основа для інтеграції керування знаннями і безперервним навчанням / С. В. Титенко, О. О. Гагарін // Системні дослідження та інформаційні технології. – 2009. – № 1. – С. 74–86.
5. *Титенко С. В.* Моделювання спеціалізованих інформаційних об'єктів в універсальних системах керування Web-контентом / С. В. Титенко // Вісник Східноукраїнського національного університету імені Володимира Даля – 2012. – №8 (179). Ч.2. — С. 235-239.
6. *Титенко С. В.* FreshKnowledge – система управління навчальним Веб-контентом на семантичному рівні / С. В. Титенко, О.О. Гагарін // VII международная конференция «Интеллектуальный анализ информации ИАИ-2007», Киев, 15-18 мая 2007г. : Сб. тр./ Ред. кол. : С.В. Сирота (гл.ред.) и др. – К.: Просвіта, 2007. – С. 342-352
7. *Титенко С. В.* Структурные основы онтологически-ориентированной системы управления информационно-учебным Web-контентом / С. В. Титенко // Управляющие системы и машины: информационные технологии: междунар. науч. журн. - 2012. - № 2. - С. 35-42. ISSN 0130-5395
8. *Олецький О. В.* Застосування формальних моделей онтологій для формалізації інформаційних потоків у системах управління контентом / О. В. Олецький // Теоретичні та прикладні аспекти побудови програмних систем. Матеріали міжнародної конференції ТАAPSD'2005, Київ, 7-9 грудня 2005 р. — С. 26-29.
9. *Семикин, В. А.* Семантическая модель контента образовательных электронных зданий: Автореф. дис. ... канд. тех. наук: 05.13.18 / Семикин Виктор Алексеевич ; Тюменск. гос. ун-т. – Тюмень, 2004. – 21 с.
10. *Brusilovsky, P.* Adaptive hypermedia, an attempt to analyze and generalize / P. Brusilovsky // Multimedia, Hypermedia, and Virtual Reality (Lecture Notes in Computer Science). [P. Brusilovsky, P. Kommers, N. Streitz (Eds.)]. – Berlin: Springer-Verlag, 1996. – Vol. 1077. – P. 288-304.