

Титенко С.В.

МОДЕЛЮВАННЯ СПЕЦІАЛІЗОВАНИХ ІНФОРМАЦІЙНИХ ОБ'ЄКТІВ В УНІВЕРСАЛЬНИХ СИСТЕМАХ КЕРУВАННЯ WEB-КОНТЕНТОМ

У роботі розглянуто задачу конструювання інформаційних об'єктів для Web-ресурсів під керуванням універсальних CMS-систем. Подано ключові реляційні моделі для реалізації спеціалізованих об'єктів.

Ключові слова: керування контентом, CMS, Web-сайт, спеціалізовані об'єкти контенту.

*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.
www.setlab.net*

Одним з напрямків розвитку Всесвітньої мережі WWW, що потребує комплексних наукових досліджень, є керування даними та контентом. Ця галузь містить такі напрямки як керування великими об'ємами даних, керування даними на основі хмарних обчислень, керування мультимедійними даними, а також є дотичною до Web-mining, задач кластеризації, класифікації та аналізу даних в Web, моделювання Web-контенту, Semantic Web тощо [1]. Задача ефективного керування Web-контентом набула великої значущості для багатьох галузей, що виникли на базі інфраструктури WWW, серед яких інтернет-комерція, дистанційне навчання та освітні Web-ресурси, керування великими інформаційними порталами, розробка та підтримка корпоративних порталів, створення Web-ресурсів електронного урядування, підтримка персональних сайтів та блогів тощо.

Серед популярних програмних рішень, що використовуються для керування контентом слід зазначити такі CMS-системи: Drupal, Joomla, Wordpress, Plone та ін. Для дистанційного навчання часто використовуються системи Moodle та aTutor, а також MediaWiki як засіб створення освітніх вікі-проектів. Перевагою цих систем, що зумовило їх популярність, є відкритий програмний код, наявність великої кількості додаків та безкоштовна ліцензія на використання. Серед пропріетарних систем слід зазначити такі: Magento, 1-C Бітрікс (електронна комерція); Microsoft SharePoint, Adobe Business Catalyst (комплексні SaaS-системи); Blackboard (система дистанційного навчання).

Незважаючи на велике розмаїття програмних систем, що застосовуються для керування контентом, актуальним завданням залишається дослідження, розробка та вдосконалення засобів універсального керування інформаційними об'єктами Web-ресурсів, що дозволить спростити створення нових ресурсів різного призначення, а також забезпечить ефективні механізми їх супроводження та налаштування. Зокрема актуальним є розв'язання наступних задач: 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. Серед систем з відкритим кодом Drupal надає найширші можливості по виконанню вказаних задач. Недоліком

системи є така організація структури даних та програмного коду, яка зумовлює велике навантаження на сервер баз даних. Це в свою чергу сповільнює швидкість генерації Web-сторінок ресурсу та вимагає великого ступеня кешування, що не завжди доцільно. Крім того в даній системі ускладнена робота по налаштуванню відображення інформаційних об'єктів, а також робота по загальному адмініструванню сайту. Загалом проблема подання спеціалізованих інформаційних об'єктів в контексті CMS є дотичною до напрямку об'єктно-реляційного відображення ORM (Object-relational mapping) [2], об'єктно-реляційних баз даних [3] та інших напрямів, пов'язаних із моделюванням даних та знань. Тому подальші дослідження у цій галузі мають перспективу вдосконалити процес керування контентом сучасних інформаційних Web-систем та підвищити ефективність створення та підтримки Web-ресурсів різного призначення.

Постановка задачі

Метою даної роботи є розгляд задачі створення спеціалізованих об'єктів контенту як засобу підвищення ефективності створення інформаційних Web-ресурсів різного призначення на базі універсальних CMS-систем. Для цього необхідно здійснити огляд методів реалізації засобів створення спеціалізованих об'єктів контенту в межах реляційних моделей бази даних CMS.

Загальні елементи контенту

Ключовою інформаційною сутністю систем керування вмістом сайтів (CMS) є сторінка або елемент контенту [4]. Як правило, в базі даних CMS-системи існує спеціальна таблиця, що представляє дану сутність, характеризуючи її певним набором полів. Основними полями загального елемента контенту є заголовок, html-текст, анотація та ін. В залежності від системи по-різному можуть бути реалізовані ієрархічні та структурні зв'язки між елементами контенту. В Drupal загальний елемент контенту позначається терміном Node (вузол), у Wordpress – Post (допис), в Joomla – матеріал. У попередній роботі було запропоновану ієрархічно-мережеву модель контенту [5], де сутністю, що відповідає за сторінки є множина елементів контенту $V=\{v_i\}$.

Інтеграція і розмежування спеціалізованих інформаційних об'єктів із загальними елементами контенту

Спеціалізовані інформаційні об'єкти веб-ресурсу можуть реалізовуватись двома шляхами: (1) як доповнення до набору полів загального елемента контенту – приєднання до контенту; (2) як окрема незалежна сутність, не пов'язана напряду з елементами контенту загального типу – розмежування контенту та спеціалізованих об'єктів (рис.1).

Приєднання до загального контенту вигідно використовувати у випадках, коли сутність, яку потрібно представити, є незначним розширенням сторінки загального типу і повинна грати роль деякого інформаційного елемента контенту, аналогічного сторінкам сайту. В Drupal це може бути реалізовано за допомогою функціональності Content types, що стала стандартною у версії 7 даної CMS [6-7]. Wordpress реалізує згадані функції за допомогою додатка Custom Post Type [6]. Для Joomla існує цілий набір розширень, що реалізують дану функціональність, найбільш відомим з яких є

K2 [8]. K2 додає більше десятка таблиць в базу даних і фактично підмінює систему створення контенту в Joomla.

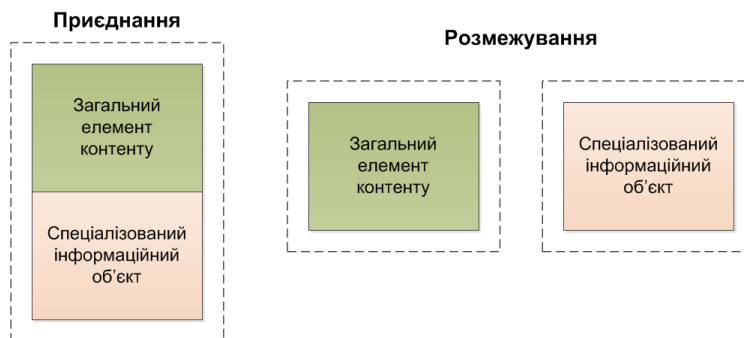


Рис. 1. Способи реалізації спеціалізованого інформаційного об'єкта у відношенні до загальних елементів контенту.

Розмежування загального контенту та спеціалізованих об'єктів в CMS передбачає наявність незалежних сутностей в БД системи, що відповідають за подання таких об'єктів. При цьому редагування даних об'єктів відбувається за допомогою окремого інтерфейсу, а набір полів не обтяжується стандартним набором полів загального елемента контенту. Спосіб відображення даних об'єктів на сайті повинен налаштовуватись за допомогою спеціальних засобів. Таким чином відбувається керування виглядом таких об'єктів, налаштування способів навігації по їх сховищу та інші спеціалізовані функціональні можливості.

Розмежування загального контенту та спеціалізованих об'єктів у згаданих CMS можна порівняти із розширенням стандартного набору полів для такої сутності CMS як елемент таксономії. Таксономія в математиці – це деревоподібна структура класифікацій деякого набору об'єктів. Таксономія в CMS використовується для різних типів класифікацій елементів контенту: для створення ієрархії розділів сайту, для організації міток (тегів), для організації меню та навігації, для тематичної класифікації контенту.

У попередній роботі [5] запропоновано структуру тематичних груп контенту як засіб моделювання предметних областей інформаційного наповнення сайту: $G=\{g_i\}$ – множина тематичних груп; $FG: F \rightarrow F$ – батьківські зв'язки в ієрархії груп. Даний підхід можна розглядати також як засіб створення таксономії.

Drupal та Wordpress містять засоби створення таксономії (Wordpress використовує термін «категорії»). У цих системах на основі згаданих вище модулів для елементів таксономії можуть створюватись додаткові поля, у такому випадку окремі розділи таксономії можуть розглядатися як набори спеціалізованих інформаційних об'єктів.

Порівняно із стандартними елементами контенту, елементи таксономії не будуть містити значного переліку стандартних полів, тому такий метод опису спеціалізованих об'єктів можна розглядати як варіант розмежування контенту та відповідних об'єктів. Натомість слід зазначити, що такий спосіб не є строгим розмежуванням, і маніпуляції з новоствореними об'єктами будуть обтяжуватись стандартним шаром обробки, характерним для кожного елемента таксономії.

Реляційна реалізація спеціалізованих інформаційних об'єктів CMS

Робота по керуванню спеціалізованими інформаційними об'єктами в контексті CMS передбачає опис нового типу контенту як нової сутності в інформаційній системі, що повинно включати:

- *загальний опис* створюваної сутності, що включає такі характеристики як ім'я для програмного опрацювання та підпис для відображення користувачам;
- *опис набору полів*, які повинні характеризувати відповідні інформаційні об'єкти, що включає власне перелік полів, а також опис кожного окремого поля із зазначенням його імені, підпису, типу даних та ін. інформації;
- *створення інфраструктури для збереження* екземплярів новостворюваної сутності, тобто безпосередніх об'єктів із значеннями їх полів;
- *налаштування способу відображення об'єктів на сайті*, їх адміністрування, а також будь-яких інших засобів роботи з ними.

Таким чином спеціалізовані об'єкти в БД CMS-системи декомпонуються на такі сутності, які повинні певним чином обов'язково відображатися в системі:

- «тип об'єкту»;
- «поля об'єкту»;
- «об'єкти» (або екземпляри);
- «значення полів» об'єктів.

У випадку, коли для створення об'єкта застосовується спосіб *приєднання до контенту*, роль «об'єкту» виконуватиме сторінка загального типу, до якої буде приєднано спеціалізовані «значення полів» відповідно до «типу об'єкту». У випадку розмежування спеціалізованих об'єктів та контенту, «об'єкт» буде представлено окремою сутністю в базі даних.

Реляційна логіка та можливості сучасних SQL-систем допускають велике розмаїття способів реалізації спеціалізованих інформаційних об'єктів в контексті бази даних CMS. Основні способи подані на рис. 2 у вигляді прототипів інформаційно-логічних моделей реляційної бази даних CMS-системи. Деякі моделі вимагають застосування інструкцій DDL (створення таблиць) на етапі додавання нових типів об'єктів. Розглянемо кожну із поданих моделей.

1. *Готова об'єктно-реляційна надбудова* (рис.2., модель 1). Без застосування DDL-інструкцій. Недоліком моделі є використання однієї таблиці для збереження значень різних спеціалізованих полів, що вимагає або використовувати різні колонки для збереження значень різних типів, або накладати обмеження на тип даних спеціалізованих полів. Іншим варіантом є застосування окремих таблиць для кожного з типів даних.

2. *Об'єктно-реляційна надбудова «Нове поле – нова таблиця»* (рис.2., модель 2). Потребує застосування DDL з метою створення окремих таблиць для зберігання «значень полів». Дана модель є природним розвитком попередньої – тут вирішується згадана проблема подання полів різних типів. Відбувається це за рахунок створення спеціалізованих таблиць для кожного з полів. Недоліком даної моделі є складність «збирання» об'єкта, так як об'єкт виявляється розподіленим між різними таблицями, і його композиція вимагає значних SQL-витрат.

3. *Природна реляційна модель «Новий тип об'єктів – нова таблиця»* (рис.2., модель 3). Потребує застосування DDL для створення таблиць, що

представляють об'єкти стандартним для реляційних БД чином – кожен запис таблиці представляє об'єкт, кожне поле – значення поля об'єкта. Додаткова мета-інформація про тип об'єктів та налаштування полів можуть зберігатися у відповідних таблицях. Якщо застосовується приєднання об'єктів до контенту, тоді між таблицею сторінок та таблицею певного типу об'єктів існуватиме зв'язок «один-до-одного» (1–0..1). З іншого боку дана модель є найбільш природним рішенням у випадку розмежування контенту і інформаційних об'єктів. При цьому для реалізації спеціалізованої бізнес-логіки для маніпуляції об'єктами доцільним є наявність в CMS відповідного API для роботи із спеціалізованими об'єктами. Подібні функціональні можливості реалізовані в системі, що пропонується автором [9].

4. *Модель тегів* (рис.2., модель 4). Без використання DDL. Застосовується принцип тегування: сторінці, що є об'єктом, призначається набір міток, кожна з яких представляє деяку характеристику об'єкта, тобто значення деякого поля. Тут кожна мітка в таблиці тегів відповідає деякому одному значенню поля. Значення полів (тобто теги) можуть групуватися в словники, тоді словник представлятиме деяке поле, а теги – область його значень. Даний підхід має значні обмеження у виразності опису об'єктів, проте він є дуже зручним у випадку, коли спеціалізовані об'єкти використовуються для фільтрації та групування контенту за деякими ознаками. Кожне значення поля отримує свій власний ідентифікатор, тому формування URL та відповідні SQL-запити по видобуванню об'єктів виконуються технічно просто.

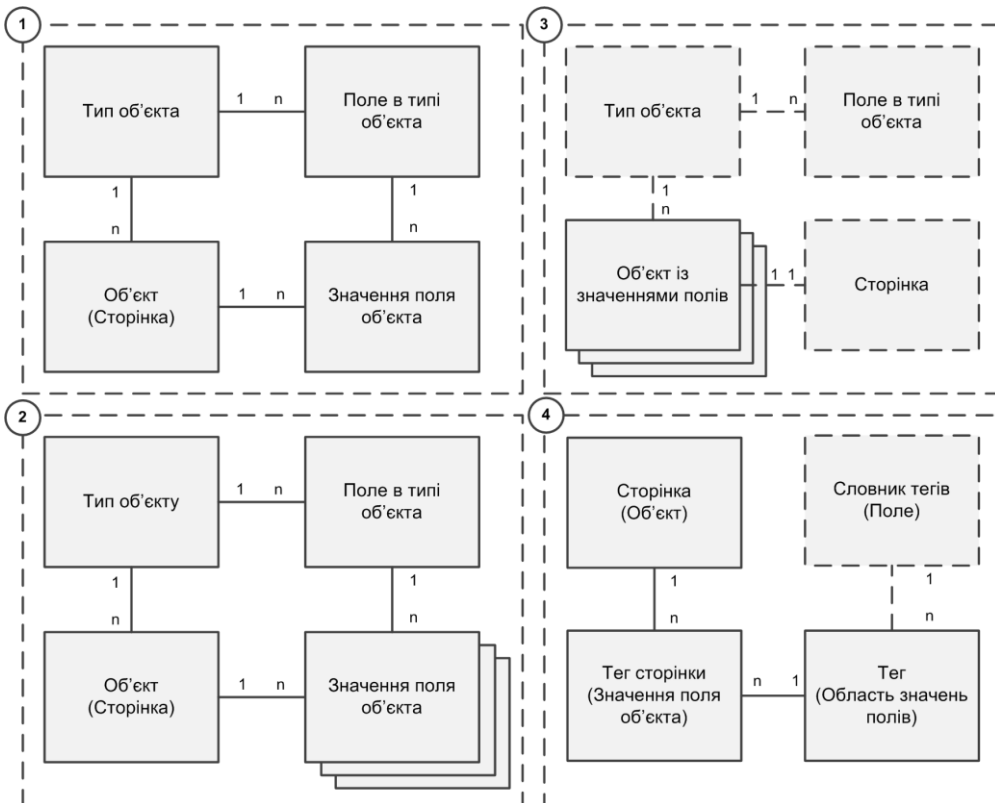


Рис.2. Способи реляційної реалізації спеціалізованих інформаційних об'єктів CMS

Висновки

У роботі розглянуто задачу моделювання спеціалізованих інформаційних об'єктів в універсальних CMS-системах. До ключових задач в контексті керування спеціалізованими об'єктами відноситься створення та зберігання об'єкта, публікація об'єкта, навігація по сховищу об'єктів, фільтрація та групування об'єктів за їх характеристиками.

Показано, що реалізація спеціалізованих об'єктів контенту відбувається одним з двох шляхів – як приєднання до загального контенту, або як розмежування з сторінками загального типу. Доцільність кожного з цих способів обумовлюється конкретною задачею та ступенем складності структури спеціалізованого об'єкта.

У роботі розглянуто основні реляційні моделі реалізації спеціалізованих інформаційних об'єктів у межах системи керування контентом. Показано, що кожна з моделей має власні переваги та недоліки, і тому повинна застосовуватись в залежності від типу задач. У зв'язку з цим робиться висновок про те, що універсальна CMS-система повинна надавати редакторам різні способи створення спеціалізованих об'єктів контенту.

Подальші дослідження будуть зосереджені на пошуку оптимальних моделей в контексті практичної реалізації інформаційних об'єктів для різних задач. Іншими важливими напрямками є використання полів посилального типу, полів із визначеним набором допустимих значень, полів-колекцій, використання відношень між об'єктами тощо. Попередні роботи і поточний хід досліджень висвітлюються на сайті setlab.net.

Література

1. Scientific topics of World Wide Web Conference 2012 – Електронний ресурс, режим доступу: <http://www2012.wwwconference.org/program/scientific-topics/>
2. Семенов В.А., Морозов С.В., Порох С.А. Стратегии объектно-реляционного отображения: систематизация и анализ на основе паттернов – Електронний ресурс, режим доступу: <http://citforum.ru/database/articles/strategy/>
3. Кузнецов С.Д. Объектно-реляционные базы данных: прошедший этап или недооцененные возможности? – Електронний ресурс, режим доступу: <http://citforum.ru/database/articles/ordbms10/4.shtml>
4. Титенко С.В. CMS. Принципи розробки систем керування вмістом сайтів. Електронний ресурс, режим доступу: <http://www.znannya.org/?view=cms-development>
5. 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.
6. Prokopenko A. WordPress 3 vs Drupal ССК – Електронний ресурс, режим доступу: <http://justcoded.com/implementation/wordpress-3-vs-drupal-cck/>
7. Content Types, Drupal Documentation – Електронний ресурс, режим доступу: <http://drupal.org/node/21947>
8. K2 Joomla extension – Електронний ресурс, режим доступу: <http://getk2.org/>

9. Титенко С. В. FreshKnowledge – система управління навчальним Веб-контентом на семантичному рівні / С. В. Титенко, О. О. Гагарін // VII міжнародна конференція «Інтелектуальний аналіз інформації ІАІ-2007», Київ, 15-18 мая 2007г. : Сб. тр./ Ред. кол. : С.В. Сирота (гл.ред.) и др. – К.: Просвіта, 2007. – С. 342-352

Титенко С. В. Моделирование специализированных информационных объектов в универсальных системах управления Web-контентом.

В работе рассмотрена задача конструирования информационных объектов для web-ресурсов под управлением универсальных CMS-систем. Представлены ключевые реляционные модели для реализации специализированных объектов.

Ключевые слова: управление контентом, CMS, Web-сайт, специализированные объекты контента.

Tytenko S. V. Custom information objects modeling in universal web content management systems.

The task of information objects construction for web resources built with CMS was considered. Key relational models for custom objects building were represented.

Keywords: content management, CMS, web site, custom content objects.

Титенко Сергій Володимирович, к.т.н., старший викладач кафедри автоматизації проектування енергетичних процесів та систем Національного технічного університету України "Київський політехнічний інститут"

*Стаття подана
18.05.2012.*