

УДК 378:004

В.М. Саух, Г.В. Велікжанін

Черкаський державний технологічний університет, Черкаси

ОСВІТНІ ПОРТАЛИ І ПРОБЛЕМИ ІНФОРМАЦІЙНОГО ВЗАЄМОБМІНУ У ФОРМІ МЕТАОПИСІВ

У статті висловлюється методологія створення типового освітнього порталу і описується набір перспективних програмних технологій для його створення. Приділено увагу навігаційній та інформаційній моделям опису інформаційних ресурсів, що забезпечують побудову ефективних механізмів введення даних, пошуку, класифікації ресурсів і комплексної обробки даних.

Ключові слова: освітній портал, інформаційна модель, метадані, колекція інформаційних ресурсів.

Вступ

Одним з найбільш перспективних напрямів розвитку індустрії інформаційних технологій, у тому числі і в області освіти, є створення порталів. Виникнення системи освітніх порталів, базовими елементами якої в даний час стають регіональні портали, призвело до необхідності стандартизованої реєстрації і каталогізації інформаційних ресурсів (ІР), придатних для використання в учбових цілях. Причому основними завданнями в цій сфері ставали: пошук і оцінка якості ІР, метаопис ІР, формування колекцій ІР по певних критеріях, пошук ІР по цілому ряду параметрів і ін.

У зв'язку з необхідністю створення єдиної системи освітніх порталів, стандартизація при створенні регіонального освітнього порталу стає особливо актуальною [1]. Система управління змістом порталу повинна мати функціональну можливість автоматизованого обміну описами ресурсів з іншими порталами (як регіонального, так і національного рівня), формування колекцій ІР, що доповнюють учбову програму.

Виходячи з рішення IEEE (Institute of Electrical and Electronics), модель метаопису освітніх ресурсів LOM була офіційно визнана стандартом, і вона ж є стандартом «де-факто» в системі українських освітніх порталів. Тому саме модель LOM вибрана як базова модель метаописів ресурсів в регіональному освітньому порталі.

Розробники освітніх порталів зіткнулися з необхідністю використання уніфікованої системи метаопису освітніх ресурсів та формування основних напрямків розробки системного проекту по створенню типового освітнього порталу.

Метою цієї публікації є визначення окремих напрямків типізації і стандартизації освітніх порталів, а саме, формування розподіленої бібліотеки навчально-методичних матеріалів, навігаційної моделі і інтерфейсів пошуку в зовнішніх класифікаторах та формування метаданих про просторові дані та послуги.

1. Освітній портал: проблема типізації і стандартизації

Будь-який освітній портал завжди розрахований на певні категорії користувачів (студентів, викладачів, абітурієнтів, батьків учнів, працевлаштувачів, фахівців в даній наочній області, працівників департаментів освіти і т.д.). Інформація, що міститься на його сторінках, повинна бути персоналізована відповідно до кожної групи користувачів і відображати їх специфічні інтереси. Портал повинен володіти здатністю налаштувати свій зовнішній вигляд, зміст і інтерфейси для кожної групи користувачів і кожного користувача індивідуально. До типового освітнього порталу пред'являється ряд організаційних, методичних, технічних і програмних вимог при проектуванні його програмної та інформаційної інфраструктури. Ці вимоги повинні бути пов'язані з основною технологічною вимогою, що визначає портал, як «інтерактивну, багату змістом інформаційно-сервісну структуру, що постійно функціонує в Інтернет, ефективно забезпечує освітнє співтовариство України якісними інформаційними ресурсами для відповідних рівнів освіти, функціонально і інформаційно пов'язаною зі всією системою загальноукраїнських освітніх порталів».

У типовому освітньому порталі з погляду його функціональної структури можна виділити три макрорівні:

- **Презентаційна частина** (структура учбового закладу, кадровий склад і персоналії, новинний блок).
- **Учбова частина** (учбові плани, учбові програми, навчально-методичні матеріали).
- **Сервісна частина** (реєстрація і аутентифікація, пошук на сайті, в порталі, мегапорталі та в Інтернет, засоби експорту/імпорту інформації, центр тестування, освітні послуги, виставки і заходи).

З погляду архітектурно-програмної реалізації типового освітнього порталу [2] необхідно застосувати відкриті стандарти і протоколи для забезпечення кризової інформаційної взаємодії. Для цих цілей необхідно мати:

- Інформаційну модель опису інформаційного ресурсу (IMS/LOM, Dublin Core).

- Загальнодоступні тематичні рубрикатори (Інформіка, Стек, ГРНТІ, УДК).
- Словники і класифікатори (типи ресурсів, рівні освіти, аудиторія і так далі).
- Реплікаційні шаблони описів у вигляді XML-схем.
- Програмне забезпечення для проведення процедури каталогізації і формування XML-карток описуваних ІР.
- Засоби доставки і розповсюдження.

При такому підході надзвичайно легко маніпулювати структурно-організованою інформацією і експортувати її в будь-які різнорідні системи. Більш того, інформацію, наприклад, про структуру учбового закладу (XML-file) можна представляти абсолютно по-різному (для цього необхідно мати декілька шаблонів XSL) для різних категорій користувачів: формат дерева для викладачів, наочна таблиця для абітурієнтів, ієрархічні списки ключових показників для працівників департаментів освіти і т.д.

При цьому необхідно відзначити, що завдання збору і накопичення в своїх базах величезних масивів освітньої інформації не є першорядним завданням національних освітніх порталів. Освітня інформація виникає в регіонах на місцях, тому на національному порталі її необхідно виявити і відшукати. Тому важливо створити кореневі структури і реплікаційні шаблони для периферійної освітньої інформації, по яких вона одноманітно формуватиметься і розповсюджуватиметься.

2. Формування розподіленої бібліотеки навчально-методичних матеріалів

У освітньому порталі повинні бути сформована бібліотека навчально-методичних матеріалів і комплекс універсальних програмно-технологічних засобів для її обслуговування, колективного ведення і розповсюдження. Для того, щоб успішно вирішити ці завдання, необхідно розробити і використовувати:

- метадані про освітні інформаційні ресурси;
- використання стандартних тематичних класифікаторів (ГРНТІ, УДК);
- XML-орієнтовані програмні засоби для універсальної каталогізації.

Модель метаданих повинна точно відображати специфіку наочної області і всесторонньо описувати інформаційний ресурс. Специфікація метаданих про освітні ІР будується як розширення схеми LOM.

Як базові класифікатори і словники, використовувані в метаданих, можна використовувати:

- словник цільових призначень ІР (на основі ГОСТ 7.60-90 і ГОСТ 7.83-2001);
- класифікатор видів ІР (на основі типології видів видань);
- наочні класифікатори (УДК, ГРНТІ);
- словник рівнів освіти (відповідно до Закону України про освіту);
- словник цільової аудиторії (розширений словник LOM).

Для запропонованої інформаційної моделі метаданих розробляються: профілі описів ІР різних класів, обмеження цілісності, XML-прив'язка.

Для даної інформаційної моделі розробляється програмне забезпечення, що дозволяє формувати XML-карточки описів створюваних в Інтернет нових освітніх об'єктів, тобто проводити універсальну процедуру каталогізації інформаційних ресурсів для формування розподілених по мережі навчально-методичних матеріалів і процесів. У плані постановки цієї проблеми для всієї мережі освітніх установ України необхідно вирішити наступні основні завдання:

- розробити оптимальну схему XML, що реалізує основні позиції інформаційної моделі опису ІР. У описах елементів і атрибутів проектованої XSD-схеми повинні бути описані всі основні сервіси системи. Структура схеми повинна враховувати зв'язок із зовнішніми класифікаторами, а також бути відкритою для подальших розширень при майбутній реалізації можливостей OLAP-аналізу і моніторингу ІР;

- розробити навігаційні схеми наочної класифікації і створити модель навігатора, що існує у формі кроссплатформного Web-додатку з прописаним в ньому інтерфейсом пошуку в зовнішніх класифікаторах, таких як УДК і ГРНТІ;

- розробити структуру генератора Web-форм на базі XSD-схем для забезпечення призначених для користувача інтерфейсів при реєстрації нових інформаційних ресурсів;

- розробити серверні сценарії, що реалізують заповнення користувачем Web-форм (з підключенням інструментів навігації) і автоматичне формування XML-файлів (XML-карточки) описуваних ІР.

Загальна схема архітектури програмного забезпечення формування інформаційних ресурсів на будь-якому освітньому порталі приведена на рис. 1.

Спосіб реалізації LOM-специфікації моделі опису інформаційних освітніх ресурсів [3] на базі мови XML вибраний по наступних причинах:

- Структура мови XML ідеально підходить для опису ієрархічних моделей.
- Мова XML є розширюваною на рівні мета-описів.
- Специфікація на основі XML є незалежною від платформи розгортання і мови програмування.
- Схема XML дозволяє вносити до опису документа обмеження на дані, що містяться, надаючи тим самим розвинену систему варіації даних.

На даний момент в специфікації XML існують два способи завдання схеми XML документу:

1. Document Type Definition (DTD) – опис типів документа.

2. XML Schema language – мова опису XML схем.

Хоча обидва ці способи дозволяють описати структуру XML документу, проте XML схеми дозволяють описати структуру і обмеження гнучкіше і широко. Тому для реалізації інформаційної моделі опису інформаційного ресурсу із застосуванням мови XML доцільно вибрати XML-схему.

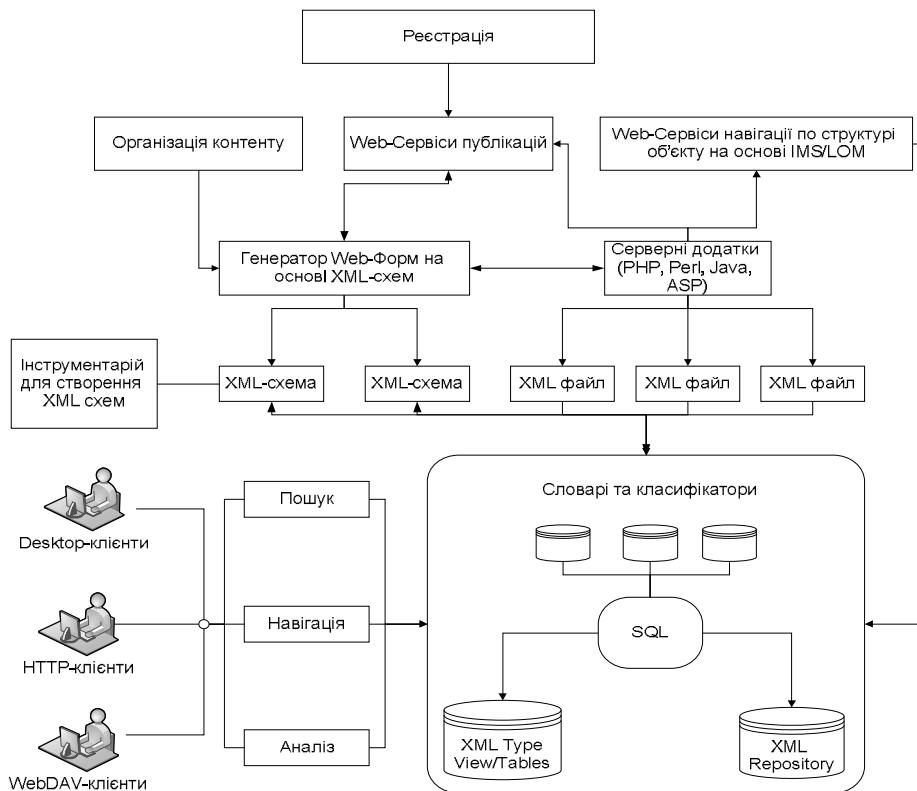


Рис. 1. Архітектура сервісів, що забезпечують процедури універсальної каталогізації зберігання і розповсюдження інформації про інформаційні ресурси

3. Навігаційна модель і інтерфейси пошуку в зовнішніх класифікаторах

Інформаційна модель опису інформаційних ресурсів припускає накопичення великих об'ємів метаописів інформаційних ресурсів, для яких повинні бути побудовані ефективні механізми введення даних, пошуку, класифікації ресурсів і комплексної обробки зібраних матеріалів. Відсутність зручних і надійних механізмів комплексної обробки великих об'ємів накопичених даних не дозволяє забезпечити пошук необхідної інформації. Використання в описі інформаційних ресурсів стандартних класифікаційних кодів істотно полегшує це завдання, забезпечуючи надійний стійкий базис для рубрикації інформаційних ресурсів. Проте проста наявність описових класифікаторів (ГРНТІ, УДК, RUS_LOM і т.д.) не вирішує проблеми осяжності великих об'ємів інформації, а тільки робить це рішення можливим завдяки маркіровці об'єктів, що класифікуються, і набору типових методів, що забезпечують зручність обробки масивів елементів інформації, яка будується на технологіях:

- збору даних з джерел первинної інформації, реалізованої в загальному випадку в онлайн-режимі за допомогою будівника форм, що володіє виходом на інтерфейс навігації по класифікаторах;
- реалізації інтерактивного пошуку з можливістю навігації по стандартних класифікаторах, що мають, в загальному випадку, деревовидну структуру і нормалізовані по власних додаткових ознаках;

- використання масивів інформаційних об'єктів для комплексного аналізу їх сукупності, наприклад, за допомогою багатовимірного аналізу з використанням стандартних класифікаторів як вимірювання інформаційних кубів.

Ці технології зв'язані одна з одною, оскільки у всіх випадках для їх вирішення застосовується принцип нормалізації структури даних.

Нормалізація метаописів IP [4] можлива по наступних напрямках:

- по опису інформаційного ресурсу (r1:Resource), в якому передбачена наявність ряду класифікаційних ознак (r1:Classification), кожен з яких позначає місце IP в межах вказаного класифікатора;
 - по відомостях про юридичну або фізичну особу, представлені у формі vCard і нормалізовані по ряду ознак, які необхідні для подальшої якісної реалізації механізмів навігації і аналітичної обробки. Такими ознаками є регіон і країна. Тільки за умови нормалізації описів ресурсів по територіальних кодах виявляється можливим просторовий аналіз територіально-прив'язаної інформації з використанням багатовимірної обробки і ГІС-ТЕХНОЛОГІЙ.
 - по тимчасовій прив'язці описів інформаційних ресурсів забезпечується входженням елементів r1:Date в структуру схеми документа. Стандартні тимчасові класифікатори років, кварталів і місяців можуть використовуватися для завдання тимчасових інтервалів пошуку і комплексного аналізу.
- Сам навігаційний сервіс реалізується у вигляді мережевого додатку, що працює з реляційною базою

даних через інтерфейс метаданих. Використання класифікаційних кодів є загальним способом формального опису будь-яких об'єктів. У освітній сфері інформаційні ресурси прийнято класифікувати за допомогою тематичних довідників, причому можна підтримуватися ГРНТІ, УДК, `RUS_LOM` і інших класифікаторів.

Простим навігаційним підходом при роботі з класифікатором є використання QBE-запиту (`Query By Example`). Простота і природність підходу QBE очевидна. Проте робота з великими багаторівневими довідниками важка через декілька причин.

Побудова великої ієрархії при зверненні до довідника займає тривалий час і вимагає великих обчислювальних ресурсів. Це особливо критично при реалізації мережевих сервісів, коли додаткове навантаження на сервер за рахунок кожного нового з'єднання повинне бути мінімальним. Проглядання великого деревовидного класифікатора і пошук в ньому досить важка задача для користувача.

Отже, для вирішення завдання навігації по базі даних з використанням довідників необхідно вирішити більш спеціалізовану задачу – завдання навігації по самих довідниках. Тобто при пошуку і обробці даних користувачеві повинні бути надані додаткові сервісні засоби полегшення огляду класифікаторів і вибору записів з них. Найбільш простим способом навігації по ієрархічному класифікаторі є послідовний спуск по рівнях ієрархії від узагальнювальних вузлів до термінальних елементів. Тобто для вибору певної коди користувачеві надається список вузлів класифікатора верхнього рівня, потім – список вузлів наступного рівня, що є спадкоємцями вузла, вибраного на першому кроці, і так далі до виконання користувачем остаточного вибору запису класифікатора. Безпосередня навігація по рівнях дерева проста для розуміння і підходить для вибору одного відповідного коду, особливо коли користувач не знайомий з повною структурою класифікатора, а логічна структура ієрархії інтуїтивно зрозуміла. Але в декількох випадках такий підхід недостатній, наприклад, коли потрібно:

- вибрати декілька записів класифікатора з різних розділів;
- сумістити пошук по логічному контексту з пошуком по інших критеріях (наприклад, вибрати всі записи класифікатора, назви яких містять слово «система»);
- здійснити пошук по критеріях, заснованих на суті і атрибутах, пов'язаних з класифікатором відносинами посилальної цілісності.

Тому логічнішим рішенням представляється максимальна нормалізація таблиці класифікатора в реляційній базі даних по можливих ознаках, а також реалізація можливості багатокритерійного пошуку по будь-якому атрибуту структури зв'язаної суті.

Для зручності навігації слід максимально нормалізувати структуру даних, тому доцільно виконати категоризацію класифікаторів. Частиною навігаційного підходу є перетворення моделі «суть-зв'язок» в ієрархічну структуру суті і атрибутів, що

представляються у формі навігаційних метаданих. Такі метадані залежно від способу представлення інформації можуть зберігатися у вигляді реляційних таблиць або структурованих документів (XML). На цій основі пропонується навігаційний інтерфейс багатокритерійного пошуку, що використовує можливість текстового пошуку засобами вибраної мови запитів (`SQL`, `XPath`, `XQuery`).

Що стосується пошуку інформації в сховищі даних метаописів IP, він може бути організований як безпосередньо по колекції зібраних документів, що заповнені в будівнику форм і задовольняють заданій схемі, так і в паралельно організованій аналітичній вітрині даних по реєстрації інформаційних ресурсів. У першому випадку метаданими навігаційної моделі є сама XML схема описів ресурсів, де як класифікаційні довідники можуть використовуватися типи, що перераховують, схеми, явно задані в описі форми (`xs:appinfo` базового елементу), масиви можливих значень і зовнішні довідники, що імпортуються з реляційних баз даних. У другому випадку навігаційні метадані включаються до складу бази даних аналітичної вітрини у вигляді зарезервованих таблиць.

4. Бази метаданих — як елемент інфраструктури просторових даних

Об'єми накопичуваних, оброблюваних і поширюваних просторових даних стрімко зростають, розширюється номенклатура продукції і послуг, заснованих на просторово-прив'язаній інформації. Під інфраструктурою просторових даних (ІПД) розуміється сукупність технологій, стандартів і людських ресурсів, необхідних для збору, обробки, накопичення, зберігання, розподілу і покращуваного використання просторових даних.

ІПД складається з трьох основних блоків:

- Блок «дані» включає всі види просторових даних, що створюються, накопичуваних, оброблюваних і передаваних в рамках ІПД.
- Блок «метадані» є сукупність баз метаданих, що містять зведення про характеристики просторових даних. Метадані (дані про дані) є важливим елементом інфраструктури, оскільки дозволяють з необхідним і достатнім ступенем подробиць описати набір просторових даних або послуг для забезпечення взаєморозуміння між споживачами і виробниками.
- Блок «сервіси» є сукупність технологій пошуку, доступу, уявлення, обробки, публікації і аналізу просторових даних і метаданих. Саме тут міститься весь функціонал, необхідний споживачам просторових даних.

Роль метаданих в процесах інформаційного обміну переоцінити важко: вони використовуються при каталогізації, обліку, статистичній обробці і аналізі. Саме за допомогою метаданих потенційний споживач продукції або послуги може оцінити її придатність для використання.

Формування метаданих про просторові дані і послуги ведеться відповідно до стандарту ISO/TC

№19115 «Географічна інформація. Метадані». (ГОСТ Р 52573-2006) і створення бази метаданих на основі цього стандарту. У стандарті визначена термінологія, методологія і універсальний набір елементів метаданих, за допомогою яких ці метадані описуються.

Стандарт ISO/TC211 №19115 (як і багато інших стандартів ISO) розроблений із застосуванням мови UML (Unified Modeling Language) – універсальної мови моделювання, де метадані представляються у вигляді окремих тематичних блоків (UML пакетів), кожен з яких характеризує той або інший аспект метаданих.

Стандартом визначений конкретний набір елементів метаданих, проте далеко не всі з них використовуються при описі тієї або іншої просторової інформації. Як основа для формування метаданих визначений наступний базовий набір елементів метаданих (так зване «ядро»), необхідний для основного документування просторових даних: загальний опис даних, місцезнаходження даних, час створення даних, організація-виробник.

У ряді випадків (наприклад, проведення каталогізації за допомогою метаданих) цілком можна обмежитися вказаним базовим набором елементів. У інших випадках можна доповнити базовий набір іншими елементами метаданих і тим самим забезпечити необхідний ступінь деталізації залежно від вирішуваного завдання.

Обумовлена в стандарті методика представлення метаданих є добре структурованою і універсальною і при її допомозі можна описати будь-яку просторову інформацію з необхідним і достатнім рівнем деталізації.

Стандарт передбачає наступні основні етапи створення бази метаданих:

- проектування структури бази метаданих на основі стандарту ГОСТ Р 52573-2006 за допомогою CASE-технологій;
- фізична реалізація структури бази метаданих на платформі СУБД;
- розробка оболонки користувача для роботи з метаданими.

Як засоби розробки можна використовувати засоби MS VISIO, MS SQL SERVER і BORLAND DELPHI.

ВИСНОВОК

Для створення типового освітнього порталу необхідно розробити системний проект, що визначає набір рекомендованої інформації, перелік, склад і способи реалізації базових сервісів, шаблонів і структур для регіональних порталів і сайтів учбових закладів, що забезпечують можливість їх взаємодії. Весь блок організаційних, методичних, технічних і програмних вимог до освітніх порталів повинен бути тісно пов'язаний з основною технологічною вимогою що визначають портал, як інформаційно-сервісну структуру, що постійно функціонує в Інтернет, ефективно забезпечує освітнє співтовариство України якісними інформаційними ресурсами для відповідних рівнів освіти, функціонально і інформаційно пов'язаною зі всією системою загальноукраїнських освітніх порталів.

Список літератури

1. Про стандарт метаданих інформаційних освітніх ресурсів для інтернет-каталогів // В збірці наукових статей "Інтернет-портали: зміст і технології". Вип. 3 / Редкол.: А.Н. Тихонов (перед.) і др.; ФДМ ГНІ ІТТ "Інформіка". - М.: Освіта, 2005. - С. 26-47.

2. Антопольський А.Б. Звіт про результати робіт 2004 р. За проектом РФФІ 04-07-90087 «Дослідження і розробка системи метаданих для електронних інформаційних ресурсів і сервісів у фундаментальній науці», 2005 [Електронний ресурс] / А.Б. Антопольський, В.І. Ауссем, С.А. Блау, А.І. Жежель. – Режим доступу до ресурсу: www.inforeg.ru.

3. Соколовський В.В. Порівняльний аналіз стандартів метаданих для сфери науки, культури і освіти / В.В. Соколовський. – М., 2006. – 12 с. (Серія «Електронні бібліотеки: теорія і методика»; Препрінт № 06.004, березень 2006 р. / Російська асоціація електронних бібліотек — НП ЕЛБІ).

4. Метадані для інформаційних ресурсів сфери утворення. Керівництво по застосуванню інформаційної моделі і її XML-прив'язки. Версія 1.0. Москва, 2006. [Електронний ресурс]. – Режим доступу до ресурсу: <http://www.spec.edu.ru>.

Надійшла до редколегії 29.04.2009

Рецензент: д-р техн. наук, проф. В.М. Рудницький, Черкаський державний технологічний університет, Черкаси.

ОБРАЗОВАТЕЛЬНЫЙ ПОРТАЛ И ПРОБЛЕМЫ ИНФОРМАЦИОННОГО ВЗАИМООБМЕНА В ФОРМЕ МЕТАДАНЫХ

В.М. Саух, Г.В. Великжанин

В статье излагается методология создания типового образовательного портала и описывается набор перспективных программных технологий для его создания. Уделено внимание навигационной и информационной моделям описания информационных ресурсов, обеспечивающих построения эффективных механизмов ввода данных, поиска, классификации ресурсов и комплексной обработки данных.

Ключевые слова: образовательный портал, информационная модель, метаданные, коллекция информационных ресурсов.

EDUCATIONAL PORTALS AND INFORMATION EXCHANGE AS METADESCRIPTION PROBLEMS

V.M. Sauh, G.V. Vekikzhanin

In this article the typical educational portal creation methodology is expressed and a set of prospective software technologies for its creation is described. The attention is given to navigation and information models of information resources description which provide data input, data search, resources classification, and complex data processing effective technologies building.

Keywords: educational portal, information model, metadata, information resources collection.