

УДК 004.67

ЛАНДЕ Д.В., доктор технічних наук
БРАЙЧЕВСЬКИЙ С.М., кандидат фізико-математичних наук

МОЖЛИВОСТІ ДОВІДКОВИХ МЕРЕЖЕВИХ РЕСУРСІВ ДЛЯ СТВОРЕННЯ ЕЛЕКТРОННОЇ ЕНЦИКЛОПЕДІЇ ЗАКОНОДАВСТВА УКРАЇНИ

Анотація. *Визначення можливостей і переваг технології MediaWiki. Обґрунтовується можливість її застосування для створення електронної енциклопедії законодавства України.*

Ключові слова: *електронна енциклопедія, технологія, Інтернет, веб-сайт, MediaWiki.*

Аннотация. *Определение возможностей и преимуществ технологии MediaWiki. Обосновывается возможность ее применения для создания электронной энциклопедии законодательства Украины.*

Ключевые слова: *электронная энциклопедия, технология, Интернет, веб-сайт, MediaWiki.*

Summary. *Decision of possibilities and advantages of technology of the MediaWiki technology is provided. Possibility of its application for creation of electronic encyclopedia of legislation of Ukraine is grounded.*

Keywords: *electronic encyclopedia, technology, Internet, website, MediaWiki.*

Постановка проблеми. Одним з важливих напрямків розбудови інформаційного суспільства є створення умов для гарантованого надання його членам доступу до багатопрофільної інформації. Ця задача не може бути ефективно вирішена без тієї чи іншої системи розподілу інформації по широкій мережі незалежних інформаційних ресурсів. Такі розподілені системи, як правило, відзначаються високою стабільністю по відношенню до зовнішніх впливів [1].

Центральною проблемою створення розподілених систем незалежних інформаційних ресурсів є їх наповнення та організація обміну наборами даних великого обсягу. Тут слід зазначити два суттєві моменти. По-перше, необхідно створювати відповідні програмні комплекси, що є достатньо складним процесом, який вимагає не лише високої кваліфікації розробників, але й значного часу на створення придатного для експлуатації продукту. По-друге, такі програмні продукти мають бути належним чином уніфіковані в плані пакетної обробки інформаційних масивів, що необхідно для реалізації швидкого перенесення даних з одного ресурсу до іншого.

На цей час в Національній академії правових наук України створюється електронна енциклопедія законодавства (далі – ЕлЕнЗ), яка має охоплювати понятійно-категоріальний апарат, визначений законами, підзаконними актами та міжнародно-правовими документами. ЕлЕнЗ міститиме термінологію, зв'язки з нормативно-правовими документами, аналітичними та оглядовими матеріалами, практикою застосування. При цьому передбачається можливість розширення проекту за рахунок нових інформаційних компонент. З цієї точки зору перспективним, на погляд авторів, є використання технології MediaWiki, добре відомої широкому загалу на прикладі Вікіпедії (точніше, сімейства вікіпедій) [2, 3]. Ця технологія поширюється вільно і є доступною широкому колу організацій та окремих осіб, що працюють в сфері інформаційних технологій. Метою цієї статті є обґрунтування можливості застосування мережевої технології MediaWiki для створення ЕлЕнЗ.

Метою статті є визначення можливостей застосування технології MediaWiki для створення електронної енциклопедії законодавства України.

Виклад основних положень. З самого початку технологія MediaWiki призначалася для створення енциклопедій, статті яких формувались безпосередньо в інтерактивному режимі за допомогою спеціальних вбудованих засобів. Власне, ці засоби інтерактивного наповнення системи практично необмеженою кількістю користувачів і розглядалися як ядро wiki-технології. Але згодом виявилось, що отримана технологія може застосовуватись в значно ширшому спектрі проблем, що стимулювало швидкий розвиток початкових технологічних ідей. Тому проект MediaWiki має свою історію, яка добре ілюструє основні аспекти створення і розвитку розподілених інформаційно-технологічних систем, здатних акумулювати в собі великі обсяги структурованої інформації за допомогою інтерактивних інструментальних засобів.

Технологія MediaWiki (від гавайського wiki-wiki – “якнайшвидше”, автор концепції дізнався про нього в аеропорту Гонолулу, де місцеве населення так називає автобуси) була створена ще у 1995 році Уордом Каннингемом (Ward Cunningham), і призначалася для колективної розробки документації (текстової і гіпертекстової). Ще MediaWiki розшифровують як “технологію швидкої правки”, проте частіше говорять про неї як про “технологію живих документів”.

Перша реалізація MediaWiki була здійснена Каннингемом для веб-вузла Pattern Languages Community з метою спростити спільне створення і ведення програмних документів. Вільне доступне програмне забезпечення MediaWiki на PHP доступно за адресою <http://wikipedia.sourceforge.net> було розроблено некомерційною організацією WikiMedia. На цьому ж програмному забезпеченні працюють проекти WikiMedia, серед яких можна виділити енциклопедію Вікіпедія, Вікі-підручник, Вікі-словник, Вікі-новини тощо.

Технологія MediaWiki – це основа побудови веб-систем, призначених для колективної розробки, зберігання, структуризації тексту, гіпертексту, файлів, у тому числі мультимедіа. Усі сторінки MediaWiki-сайту є статтями, вміст яких являє собою текст, в якому можна використовувати просту Wiki-розмітку або теги HTML.

На цей час загальновідомий проект Вікі (на базі технології MediaWiki) виявився настільки успішним, що вийшов за рамки програмного продукту і перетворився на концепцію.

Саме технологію MediaWiki було вибрано як базову для створення ЕлЕнЗ, виходячи з таких 7 головних причин:

1. З самого початку мережева технологія, успішність підтверджена практикою.
2. Вільна ліцензія, не потрібна власна розробка ПЗ.
3. Можливість включення великої кількості інформаційних адміністраторів.
4. Можливість організації гіперпосилань як на зовнішні ресурси, так і на окремі свої статті.
5. Можливість створення різних категорій, нарощування видів інформаційних ресурсів.
6. Можливість “підключення” мультимедійних матеріалів.
7. Можливість пакетного завантаження вже накопичених матеріалів.

Головною перевагою технології MediaWiki є, безперечно, комплекс вбудованих інструментальних засобів, які дозволяють виконувати штатні операції з даними в інтерактивному режимі через стандартний веб-інтерфейс.

Насправді в сучасному стані технологія MediaWiki дозволяє виконувати набагато ширший спектр робіт з даними, що зберігаються в системі [4]. Перш за все, це розвинені операції експорту та імпорту даних без суттєвих обмежень обсягу, що дозволяє

тиражувати комплекси та утворювати “клони” (якщо є така потреба). Обмін даними здійснюється в форматі XML, який допускає створення потрібних наборів даних різноманітними технічними засобами від пакетних конверторів до звичайних XML-редакторів. Формуючи набір тегів, маємо можливість безпосередньо керувати складом та структурою даних, що завантажуються в wiki-систему.

Слід відзначити, що використання формату XML створює умови для зручної обробки досить складної структурованої інформації великих обсягів. Головна перевага такої технології полягає в тому, що формат XML дозволяє точно описати та відобразити будь-яку, навіть дуже складну структуру даних. Опис конкретної wiki-системи виконується прозоро і містить в собі всю наявну інформацію – від характеристик форматування головної сторінки до параметрів ревізій кожної статті. Такий підхід дозволяє створювати та модифікувати контент wiki-системи шляхом незалежної обробки кожного елемента загального набору даних. При цьому текст кожної статті являє собою окремий тег, який безпосередньо містить всі елементи внутрішньої розмітки wiki, що спрощує створення нових документів, оскільки не вимагає використання додаткових засобів.

Таким чином, маємо можливість безпосередньо керувати складом та структурою даних, що завантажуються в wiki-систему, просто формуючи набори відповідних тегів.

Технологія MediaWiki дозволяє здійснювати різноманітні операції з модифікації даних, в тому числі з ревізіями статей. Вони можуть виконуватися в двох основних режимах: інтерактивними засобами безпосередньо в системі, та автоматично шляхом підготовки даних у вигляді файлу для пакетного завантаження.

Операції над даними можуть здійснюватись на кількох рівнях: штатними засобами Web-інтерфейсу системи; штатними засобами з використанням спеціальних утиліт з командного рядка та зовнішніми утилітами, які містяться в бібліотеках на мовах Perl, PHP, Python. Зовнішні утиліти поділяються на дві категорії. Перша категорія передбачає безпосереднє використання API MediaWiki, а друга працює на рівні ботів.

Стандартні операції обміну даними можуть бути виконані штатними утилітами, що входять до комплекту MediaWiki. Вивантаження даних в файл у форматі XML здійснюється за допомогою утиліти `dumpBackup.php`, яка знаходиться в директорії `/maintenance`. Залежно від потреб можна вивести в вихідний файл всі ревізії статей, або лише останні. Відповідно, для завантаження даних необхідно конвертувати їх у формат XML, що має таку ж структуру, як і у випадку вивантаження. Створений XML-файл завантажується утилітою `importDump.php`, що знаходиться в тій самій директорії.

Для досягнення максимальної ефективності така система повинна бути розрахована на отримання інформаційних наборів даних, що надходять з різних джерел у різних форматах. Отже, система має забезпечувати можливість уніфікованого конвертування поширених форматів у обумовлений формат, який використовується в ній для пакетного завантаження. З цієї точки зору технологія MediaWiki може застосовуватись також для створення різноманітних інтеграторів довідкової інформації.

Слід також зазначити, що технологія MediaWiki, дозволяє використовувати різні інформаційно-пошукові системи, що забезпечує додаткову гнучкість створюваних продуктів.

Особливий інтерес, на нашу думку, становить можливість формувати ієрархічний гіпертекст, що значно спрощує доступ до потрібної користувачеві інформації. Забезпечення інформаційних потреб користувача в даному випадку складається з двох фаз. Перша фаза полягає в звичайному інформаційному пошуку за ключовими словами (статистика свідчить, що принаймні 80 % запитів складаються з одного слова). Друга фаза – це розширення інформаційного поля шляхом отримання додаткових статей за гіперпосиланнями, які можуть міститися в кожній статті.

На цей час створено діючу модель ЕлЕнЗ, яка містить 6000 статей, доповнених як зовнішніми посиланнями на законодавчі акти, кодекси України, так і внутрішніми – посиланнями з одних статей із ЕлЕнЗ на інші (Рис. 1).

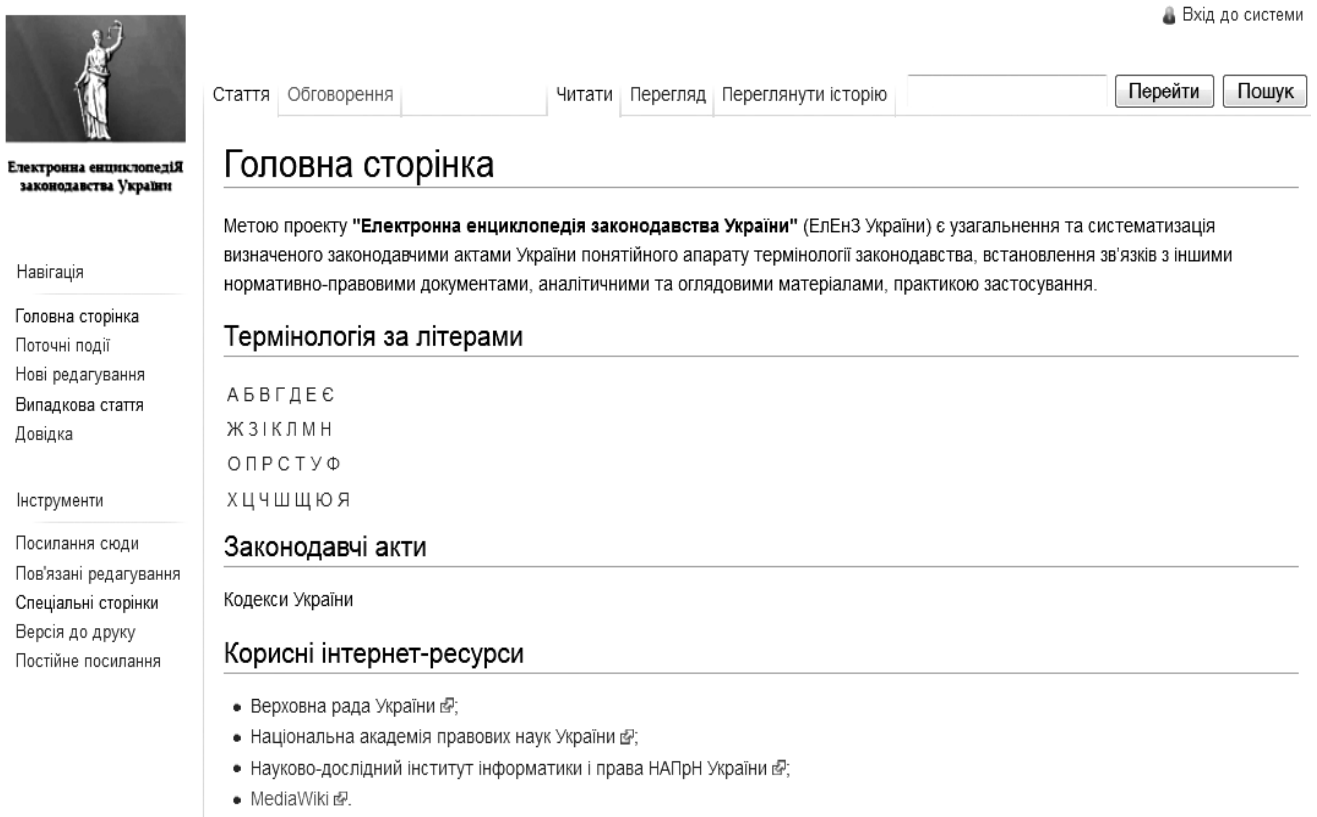


Рис. 1 – Інтерфейс моделі ЕлЕнЗ

Можливості технології MediaWiki дозволяють організувати колективну роботу по веденню/супроводженню ЕлЕнЗ у модернованому режимі (Рис. 2).



Рис. 2 – Схема організації колективної роботи

У відповідності з наведеною схемою головний адміністратор ЕлЕнЗ має повноваження призначення інформаційних адміністраторів, управління ботами – програмами, що здійснюють пакетне завантаження даних у базу даних системи, та отримання статистики роботи користувачів і наповнення бази даних. Інформаційні адміністратори мають здійснювати введення та коригування даних, зареєстровані корпоративні користувачі також мають можливість введення та коригування введених ними даних, відвідувачі з мережі Інтернет мають право доступу до баз даних ЕлЕнЗ у режимі “читання” та пошуку.

Висновки.

При створенні електронної енциклопедії законодавства України на базі технології MediaWiki вважаємо за необхідне врахувати наступне:

1. Необхідність синхронізації із поточними змінами у законодавстві. Це потребує технологічного зв’язку із відповідними БД.

2. Необхідність “ручної” роботи, яка має виконуватися за рахунок підключення зацікавлених осіб: студентів, правників.

3. Застосування де-факто стандарту MediaWiki здешевлює проект, робить його доступним у методичному плані.

4. Для підвищення якості ЕлЕнЗ необхідно підключення до проекту фахівців-лексикографів.

5. Термінологічна БД ЕлЕнЗ має доповнюватися повними документами, гіперпосиланнями, мультимедійними матеріалами.

Використання технології MediaWiki відкриває широкі перспективи розширення набору інструментальних засобів для розробки уніфікованих систем, що забезпечують можливість накопичення та обміну даних без розробки складних програмних комплексів.

Використана література

1. Горбулін В.П. Інформаційні операції та безпека суспільства: загрози, протидія, моделювання : монографія / В.П. Горбулін, О.Г. Додонов, Д.В. Ланде. – К. : Інтертехнологія, 2009. – 164 с.

2. Ландэ Д.В. За знаниями - к Википедии. – Ч. 1 // Телеком, 2005. – № 9. – С. 60-64.

3. Ландэ Д.В. За знаниями - к Википедии. – Ч. 2 // Телеком, 2005. – № 11. – С. 60-64.

4. Daniel J. Barrett. MediaWiki. – O’Reilly Media Inc., 2009. – 376 p.

~~~~~ \* \* \* ~~~~~