

## **ФОРМУВАННЯ НЕЗАЛЕЖНИХ ІНФОРМАЦІЙНИХ РЕСУРСІВ НА ОСНОВІ ТЕХНОЛОГІЇ MEDIAWIKI**

***Ланде Д.В., Брайчевський С.М.,  
ІПРІ НАН України, НДІП НАПрН України***

Одним з важливих напрямків забезпечення інформаційної безпеки суспільства є створення умов для гарантованого надання його членам доступу до багатопрофільної інформації. Оскільки, нажаль, завжди є реальна загроза кібератак на потужні загальновідомі інформаційні ресурси, дана 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 надзвичайно високою резистентією 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 (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.

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.

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 являли собою ядро wiki-технології.

Саме технологія 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, нарачування видів інформаційних ресурсів.
6. Можливість “підключення” мультимедійних матеріалів.
7. Можливість пакетного завантаження вже накопичених матеріалів.

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

склад та структуру даних, що завантажуються в wiki-систему.

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 поступають 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.

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 як 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. Схеми 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.

При розробленні цієї моделі Електронної енциклопедії законодавства на базі технології MediaWiki враховувалось:

1. Необхідність синхронізації із поточними змінами у законодавстві. Це потребує технологічного зв'язу 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.

## Література

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

2. Ланде Д.В. За знаннями - к Вікіпедии. Часть 1. История и идеология // Телеком, 2005. - № 9. - С. 60-64.

3. Ланде Д.В. За знаннями - к Вікіпедии. Часть 2. Страна и ее обитатели // Телеком, 2005. - № 11. - С. 60-64.

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