

В.П. Дуплій¹,
Ю.М. Бабич²,
Б.В. Моргун¹

¹ Інститут клітинної біології
та генетичної інженерії
НАН України

² Інститут проблем міцності
ім. Г.С. Писаренка НАН
України

СТВОРЕННЯ БАЗИ ДАНИХ ПУБЛІКАЦІЙ ТА ІНФОРМАЦІЙНОГО САЙТА ЖУРНАЛУ "ЦИТОЛОГІЯ І ГЕНЕТИКА"

Завдяки розвитку сучасних технологій доступ до інформації про матеріали, опубліковані в періодичних виданнях, і до самих матеріалів стає швидшим і доступнішим. Зручність пошуку наукових статей спеціалізованими засобами в Інтернеті витісняє на другий план традиційні підходи та методи. Виданням, що не мають можливості розміщуватися на власному сайті, який знайомить з останніми публікаціями, залишається сподіватися лише на лояльність своїх читачів і бажання сторонніх організацій включити до своїх публічних ресурсів інформацію про ці матеріали. Проте редакція журналу також має потребу в базі даних своїх публікацій. Робота над системою, описаною в цій статті, була розпочата майже 10 років тому, але й сьогодні продовжується, триває її удосконалення, адаптація до вимог сьогодення.

Вибір програмного забезпечення. У наш час програмне забезпечення з відкритим вихідним (програмним) кодом, або відкрите програмне забезпечення, за якістю та інформаційною підтримкою користувача не поступається комерційному. У галузі веб-технологій такі розробки, як сервер *Apache* (Хокинс С., 2001), мова програмування *PHP* (Котеров Д.В., Костарев А.Ф., 2006) та система управління базами даних *MySQL* (*DuBois P.*, 2008) стали стандартами *de facto* і надаються майже всіма компаніями, що забезпечують хостинг сайтів. Усі три пакети мають версії для більшості операційних систем і можуть використовуватися як в Інтернеті, так і в локальній мережі, або навіть на окремому комп'ютері. Разом ці програмні продукти дозволяють створити єдину базу даних і використовувати її як для власних потреб, так і для сторонніх користувачів, відкриваючи їм тільки частину інформації. Для того, щоб система була гнучкою і водночас швидкою, було вирішено не використовувати СКВ (Система керування вмістом) сторонніх розробників, а створити власну.

Структура даних публікації. Дані про публікації зберігаються лише в одній таблиці *MySQL*, що забезпечує швидку роботу системи та зручне адміністрування. Треба зазначити, що журнал "Цитология и генетика" є російськомовним, перекладається і перевидается англійською мовою, але в оригінальній версії статті публікуються однією з трьох мов (українською, російською чи англійською) за вибором автора.

В.П. ДУПЛІЙ, Ю.М. БАБИЧ, Б.В. МОРГУН

Як видно з табл. 1, деякі поля парні і заповнюються як мовою оригіналу, так і англійською мовою. Поле "rubric" ("огляд" або "оригінальна робота") представлено однозначним числом, що при використанні лише кількох варіантів не важко заповнити вручну, але при потребі можна запрограмувати підбір із довідкової таблиці. Електронна пошта вноситься до бази даних, але користувачеві сайта вона не доступна, спілкування з авторами організовано через спеціальну форму. Це зроблено для захисту від спаму.

Структура даних загальних сторінок. Сайт журналу має кілька загальних сторінок, на яких наведено склад редколегії, правила для авторів тощо. Структура таблиці сторінок побудована за тими ж принципами, що й таблиці публікацій. Хотілося б тільки зупинитися на полях *file* і *dir*, що містять ім'я файлу та назву папки відповідно (табл. 2). Для прискорення видачі сторінки користувачу та зменшення навантаження на сервер була створена система попередньої генерації загальних сторінок і сторінок публікацій. Крім того, сукупність утворених сторінок можна використовувати на веб-сервері, що не має *PHP* і *MySQL*, або розповсюджувати на електронних носіях. Назви файлів публікацій складаються із номерів сторінок, тому та номеру журналу і розміщуються в

Таблиця 1. Поля таблиці публікацій

| Поле | Опис | Тип |
|-------------|------------------------------------|-------------|
| id | Ідентифікатор статті | int(4) |
| year | Рік | varchar(4) |
| volume | Том | char(2) |
| number | Номер журналу | char(1) |
| pages | Сторінки публікації | varchar(10) |
| rubric | Рубрика | int(1) |
| lang | Мова статті | char(2) |
| author_ru | Автори мовою оригіналу | tinytext |
| title_ru | Назва мовою оригіналу | tinytext |
| origin_ru | Установи авторів мовою оригіналу | text |
| author_en | Автори англійською мовою | tinytext |
| title_en | Назва англійською мовою | tinytext |
| origin_en | Установи авторів англійською мовою | text |
| email | Електронна пошта авторів | varchar(40) |
| keywords_ru | Ключові слова мовою оригіналу | text |
| summary_ru | Резюме мовою оригіналу | text |
| keywords_en | Ключові слова англійською мовою | text |
| summary_en | Резюме англійською мовою | text |

папці року публікації. Для загальних сторінок назва файла та папка зберігаються в базі даних.

Швидкість роботи системи виявилась достатньо високою, щоб перейти до динамічної генерації сторінок на запит користувача. Для того, щоб перехід відбувся непомітно для пошукових машин, ми використовуємо можливості модуля *Apache Rewrite*, який підмінює шлях до сторінки відповідним запитом до скрипта системи. Крім того, акуратна традиційна адреса сторінки краще сприймається користувачами та деякими пошуковими системами.

Підготовка та введення даних. Вихідний файл даних, що надається у форматі *Microsoft Word*, за допомогою текстового процесора з відкритими вихідними кодами *OpenOffice.org Writer* (Weber J.H., 2004) перезаписується в форматі *html*. Створена нами система регулярних виразів (Фридл Дж., 2003) дозволяє позбутися зайвих елементів розмітки тексту. В поля потрапляє тільки виділення курсивом, верхнім і нижнім індексом. Ми не прагнули більшої автоматизації процесу перенесення інформації до бази даних через намагання забезпечити більш ретельний її контроль.

Форми редагування генеруються автоматично, за однаковою алгоритмом для публікацій та для загальних сторінок. Для полів бази даних типу *"text"* виводиться чотирирядкове поле, для решти – однорядкове. Для зручнішого редагування даних типу *"text"*, їхні поля мали б бути більшими, але потреба в редагуванні виникає дуже рідко.

Інтерфейс користувача системи керування вмістом. Систему керування вмістом ми намагалися створити якомога універсальнішою, придатною для будь-якого періодичного видання. В таблиці публікацій обов'язковими є поля, що містять рік, номер і сторінки публікації. Також наявні поля

Таблиця 2. Поля таблиці загальних сторінок

| Поле | Опис | Тип |
|-------------|----------------------------------|--------------|
| id | Ідентифікатор сторінки | int(4) |
| file | Назва файла | varchar(40) |
| dir | Назва папки | varchar(40) |
| picture | Малюнок | varchar(60) |
| title_ru | Назва сторінки російською мовою | varchar(255) |
| keywords_ru | Ключові слова російською мовою | text |
| menu_ru | Пункт меню російською мовою | varchar(40) |
| article_ru | Текст сторінки російською мовою | text |
| title_en | Назва сторінки англійською мовою | varchar(255) |
| keywords_en | Ключові слова англійською мовою | text |
| menu_en | Пункт меню англійською мовою | varchar(40) |
| article_en | Текст сторінки англійською мовою | text |



Рис. 1. Фрагмент вікна браузера

авторів і назви статті мовою оригіналу, але вони можуть бути незаповненими. Решта полів є варіативною та використовується лише на етапі генерації сторінки. При адаптації системи до іншого періодичного видання, крім коригування структури полів таблиці публікацій, залишається лише переробити шаблони сторінок сайта і файл стилів, щоб вони відповідали потребам видання.

В інтерфейсі користувача СКВ ми прагнули лаконічності та зручності. Вікно браузера ділиться на дві частини, у лівій завжди доступне головне меню (рис. 1). Угорі автоматично виводиться доменне ім'я сайта, є кнопки відображення списку публікацій і загальних сторінок. Є можливість отримати список публікацій відмічених номерів за вибраний рік. Якщо при цьому буде позначено поле "Generate", будуть згенеровані сторінки вибраних номерів.

У заголовок списку публікацій (рис. 2) автоматично виводиться назва таблиці бази даних, у якій він зберігається. Для кожної статті показується рік, номер і сторінки журналу, а також перелік авторів і назва. Виводяться кнопки для їхнього показу, редагування і видалення. Список також має кнопку введення нової статті, яка викликає незаповнену форму редагування, описану вище. Список загальних сторінок подібний.

Публічні сторінки сайта. Сайт ділиться на англomовну і російськомовну частини, кожна сторінка має посилання на відповідну їй у певній частині сайта. Доступ до загальних сторінок здійснюється через меню. Опис публікації (рис. 3) складається з переліку авторів, назви статті, відомостей про місце її розміщення в журналі, резюме та перелік установ, де працюють автори, позначається "оригінальна робота" чи "огляд". Для авторів, що надали адресу електронної пошти, виводиться іконка для виклику форми зв'язку з авторами. Планується також розміщувати посилання на файл статті (pdf), якщо вона є у вільному доступі.

Для кожного номера автоматично формується список статей, а для тому – алфавітний показчик, відсортований за першим автором. Доступ до цих списків можливий з головної сторінки, а також і зі сторінки "Публікації" обох частин сайта.

| paper | | | | | | | |
|------------|---|-------|--|--|------|------|--------|
| Add Record | | | | | | | |
| 2011 | 1 | 3-10 | Танасиенко И.В., Емец А.И., Пирко Я.В., Корховой В.И., Абумхадн Н., Блом Я.Б. | Получение трансгенных линий ячменя, продуцирующих лактоферрин человека, с использованием мутантного альфа-тубулина в качестве селективного маркерного гена | Show | Edit | Delete |
| 2011 | 1 | 11-17 | Матвеева Н.А., Василенко М.Ю., Шаховский А.М., Банникова М.А., Кваско О.Ю., Кучук Н.В. | Эффективная агробактериальная трансформация растений цикория (<i>Cichorium intybus</i> L.) вектором с геном туберкулезного антигена ESAT6 | Show | Edit | Delete |

Рис. 2. Зовнішній вигляд списку публікацій у СКВ

ISSN 0564-3783

ЦИТОЛОГИЯ И ГЕНЕТИКА

Главная страница
Контакты
Публикации
Тематика журнала
Подписка
Вниманию авторов
Редколлегия

Сахно Л.А. ✉

Вариабельность жирнокислотного состава рапсового масла: классическая селекция и биотехнология

Обзор

Институт клеточной биологии и генетической инженерии НАН Украины,
Киев

Цитология и генетика 2010, том 44, N 6, 70-80

Проанализированы проблемы и достижения в селекции масличного рапса *Brassica napus L. var. oleifera*, направленной на изменение состава жирных кислот в масле семян, с использованием традиционных и генноинженерных подходов. Отмечено, что для дальнейшего прогресса в этой области оптимальным является сочетание биотехнологических разработок и методов классической селекции.

| | | |
|-----------|---------|---------|
|1544 | РЕЙТИНГ | 379261 |
| Sp iG | 77 | mail.ru |
| | | 53 |
| | | 29 |

Copyright © ICBGE 2002-2011
Coded & Designed by Volodymyr Duplij
Modified 30.03.11

Рис. 3. Сторінка з описом статті

Система пошуку публікацій. Кожна сторінка має форму пошуку. Пошук ведеться на послідовність набраних символів у будь-яке поле таблиці публікацій. Такий підхід є простим і ефективним, але з ростом бази даних, як за рахунок нових статей, так і внаслідок розміщення давніх публікацій, постає завдання додатково надавати користувачу засоби розширеного пошуку. Планується дозволити пошук за окремими полями таблиці, одним або кільком одразу, обмежувати часовий інтервал публікацій, надати можливість пошуку по фразах і сукупності ключових слів. Окремим завданням стоїть формування списків статей по кожному автору.

Експорт бібліографічних даних. Однією з найрозповсюдженіших систем автоматизації бібліотек серед країн СНД є система ІРБІС. На нашу думку, до найцінніших її рис треба віднести: відповідність державним стандартам, можливість обміну даними з іншими бібліотеками та імпорту/експорту бібліографічної інформації у відкритому форматі.

В.П. ДУПЛИЙ, Ю.М. БАБИЧ, Б.В. МОРГУН

Для перенесення даних про публікації журналу до системи ІРБІС бібліотеки Інституту клітинної біології та генетичної інженерії у СКВ сайта була додана функція експорту бібліографічної інформації у форматі *UNIMARC*. Є можливість експортувати як дані про конкретну статтю, так і про будь-який номер або том журналу. В майбутньому планується зробити цю функцію доступною користувачам сайта, а також робити експорт в інші поширені формати.

Висновки. Створено систему керування базою даних публікацій та вмістом сайта (*www.cytgen.com*) журналу "Цитологія і генетика", що має достатні і необхідні ознаки для функціонування БД публікацій та інформаційного сайта журналу. А саме:

- просту і невибагливу до ресурсів (займає 256 КВ пам'яті на диску);
- не вимагає довгої підготовки персоналу для вводу даних і не обтяжує зайвими інтерфейсними надбудовами;
- дозволяє вести пошук одночасно по всіх полях бази даних;
- має можливість експорту сайта і бази даних публікації для розповсюдження на електронних носіях;
- має можливість експорту бібліографічних даних у форматі *UNIMARC* для систем автоматизації бібліотек, таких як ІРБІС.

Для цієї системи розроблено автоматизовану технологію підготовки даних.

1. *DuBois P.* MySQL. — Addison-Wesley, 2008. — 1200 p.
2. *Weber Jean Hollis.* OpenOffice.org Writer — O'Reilly Media, 2004. — 240 p.
3. *Воройский Ф.С.* Основы проектирования автоматизированных библиотечно-информационных систем. Изд. 2-е доп. и переработанное. — М.: ФИЗМАТЛИТ, 2008. — 512 с.
4. *Котеров Д.В., Костарев А.Ф.* РНР 5 В Подлиннике. — БХВ-Петербург, 2006. — 1120 с.
5. *Фридл Дж.* Регулярные выражения. 2-е изд. — СПб.: Питер, 2003. — 464 с.
6. *Хокинс С.* Администрирование Web-сервера Apache. — Вильямс, 2001. — 336 с.