УДК 004.9

ОРГАНІЗАЦІЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ КОРПОРАТИВНОГО ДОКУМЕНТО-ОРІЄНТОВАНОГО СХОВИЩА ДАНИХ ОСВІТНЬОЇ УСТАНОВИ

Д. С. Замятін, О. С. Кебкало, Л. І. Лукашевська, А. Ю. Михайлюк, О.С. Михайлюк, А. В. Петрашенко, В.П.Тарасенко

Вступ

Однією з характерних рис освітньої діяльності в суспільстві знань є виражена тенденція до використання переважно електронної форми подання навчальної інформації [1, 2]. Очевидно, це ставить перед системою освіти взагалі, і перед окремими навчальними закладами зокрема, ряд нових завдань, які потребують нагального вирішення. До числа таких завдань безумовно належить формування та упорядкування власного електронного інформаційного ресурсу (ІР) [3]. Частково ця задача розв'язується за рахунок використання специфічного програмного забезпечення (ПЗ) освітнього призначення [4]. Так, у системах дистанційного навчання, електронних підручниках, електронних бібліотеках тощо звичайно передбачають вбудований інструментарій для підтримки власного ІР. Однак, використання широкого переліку спеціалізованих систем породжує загалом негативне, з точки зору упорядкування інформації, явище технологічної та логічної гетерогенності ресурсу [5, 6]. Крім того, існує надзвичайно великий прошарок інформаційних об'єктів, що залишаються поза сферою впливу традиційних освітніх та офісних комп'ютерних систем [3, 7]. До числа таких об'єктів можна віднести електронні версії учнівських (студентських) робіт, різноманітні презентації, методичні матеріали, наукові розробки, будь-які електронні документи (Д), які з тих чи інших причин не вписуються у технологічний цикл традиційних систем електронного документообігу, і т.п. Ці дані формують окремий, також гетерогенний, інформаційний простір, який є надзвичайно важливим в плані використання у освітній та науковій діяльності учбового закладу, але не забезпечений конкретними ПЗ. Гетерогенність простору обумовлюється не тільки різноманітністю типів об'єктів, а й розмаїттям самих об'єктів одного типу. Об'єкти, що розглядаються, є слабоструктурованими, оскільки їх внутрішня організація залежить від багатьох факторів (тип Д, авторство, час створення тощо). Слід визнати, що перелік таких даних інтенсивно змінюється (в основному, збільшується), у зв'язку з чим використання для їх упорядкування деякої бази з наперед заданою сталою схемою даних не є реальним. Підтримка ж у повній мірі власного динамічного електронного ІР, передусім текстового, засобами стандартних систем керування базами даних сьогодні не під силу пересічним освітнім закладам, принаймні вітчизняним, зокрема з економічних причин [8].

Постановка задачі

Тому в контексті глобальної інформатизації освіти актуальною задачею є розробка та впровадження спеціалізованих, орієнтованих на використання в навчальних закладах, сховищ слабоструктурованих даних [9-12]. Далі пропонується підхід і концепція організації універсального програмного середовища, призначеного для формування відомчого документо-орієнтованого сховища даних навчального закладу, на базі якого існує можливість індивідуального конструювання засобів бізнес-логіки, характерної для різних ланок освітньої та наукової діяльності.

Загальна концепція спеціалізованого документо-орієнтованого сховища даних

В основу концепції спеціалізованого сховища слабоструктурованих даних покладена логічна модель «об'єкт-атрибут-значення», в якій кожна сутність або об'єкт подається певною сукупністю характеристик або атрибутів. Ця модель дозволяє не лише зберігати Д з різними наборами атрибутів, але й підтримувати динамічність цих наборів. Тому інформаційна база системи дозволяє накопичувати та обробляти сутності довільної структури (з довільною кількістю і типами атрибутів) [13].

Окремий ІР у межах сховища розглядається як підмножина документів, де окремий Д — довільний набір атрибутів і їх значень.

Розмежування прав доступу для роботи з ресурсами здійснюється на основі шаблонів. Шаблон — це набір умов для значень атрибутів Д, який дозволяє здійснити відбір потрібного сегменту ресурсів. Завдяки шаблонам групі користувачів можна призначити права в одному режимі роботи (у даному випадку і надалі маються на увазі режими роботи з Д) з одним набором Д, а у другому режимі — з іншим набором Д (певного ІР). Для ПЗ сховища даних (надалі - системи) визначено рівні доступу до ІР (гість, модератор тощо). Кожному рівню доступу відповідає набір режимів роботи, кожному з яких поставлено у відповідність шаблон. Можна призначити різні рівні доступу для різних ІР. Завдяки включенню користувача у довільний набір груп можна вибрати і призначити будь-яку комбінацію рівнів доступу до різних ІР Відповідність Д певному шаблону визначається динамічно при роботі у кожному з режимів.

Тобто, при зміні Д можливе включення або виключення його доступності у будь-якому з режимів для групи користувачів з певним рівнем доступу.

Для користувачів визначено наступні режими функціонування системи:

- додавання і редагування Д;
- перегляду Д;
- адаптивної каталогізації;
- адміністрування.

Режим додавання і редагування Д передбачає:

- створення, редагування, видалення, групування, розгруповування атрибутів;
- додавання, редагування Д (включення і вилучення з них атрибутів, встановлення і редагування значень, збереження змін).

Режим перегляду Д:

- перегляд повного набору атрибутів Д і їх значень;
- перегляд та друк Д у вигляді однієї з форм всіх шаблонів, яким відповідає Д;
- перегляд, друк і завантаження файлів, пов'язаних із атрибутами відповідного типу;
- модерування, закриття, перехід до редагування, вилучення Д;
- перегляд Д, з яким даний пов'язаний асоціативним зв'язком.

Режим адаптивної каталогізації:

- вибір Д для подальшого перегляду;
- пошук за значеннями і вмістом атрибутів;
- перегляд звітів обробки Д;
- систематизація, друк Д.

Режим адміністрування:

- створення, редагування, видалення інформаційних об'єктів системи (атрибути, шаблони, профілі користувачів та ін.);
- встановлення порядку відображення атрибутів Д при перегляді і редагуванні набору атрибутів шаблону;
- конструювання формул, конструювання форм перегляду Д і форм звітів.

Адміністрування організоване таким чином, що для кожного типу інформаційних об'єктів (атрибут, шаблон, форма тощо), які використовуються у системі, можна додавати дозволи на виконання дій над ними. Кожному користувачеві або групі користувачів можна надавати ці дозволи або позбавляти їх. Для груп є можливість надання і позбавлення дозволів і прав на режими роботи з сегментами ІР. Шляхом включення певної категорії користувачів у групу можна автоматично надати їм дозволи і права, які визначені для даної групи. При роботі у кожному з режимів будь-якому користувачеві надаються лише ті можливості, які доступні йому згідно його прав.

Архітектура програмного забезпечення сховища даних

До складу системи входять наступні підсистеми (ПС) (рис. 1):

- *інформаційна база (ПС збереження ІР),* відповідає за збереження ІР, які використовуються іншими модулями системи;
- *ПС адміністрування,* забезпечує підтримку сесій, налаштування шаблонів, розмежування прав у кожному з режимів для різних ІР на основі шаблонів;
- *ПС введення даних,* відповідає за наповнення і оновлення інформаційної бази;
- *пошукова ПС,* здійснює каталогізацію, реалізує операції пошуку і встановлення асоціативних зв'язків;
- *ПС обробки даних,* реалізує аналітичну обробку текстів, обслуговує користувацькі режими обробки Д і статистичну обробку даних;
- *ПС візуалізації,* забезпечує формування інтерактивного інтерфейсу для користувача.

Далі подано опис складових частин вказаних підсистем.

Інформаційна база включає наступні сегменти даних.

1) *Сегмент даних про атрибути* — властивості атрибутів (назва, тип, текст підказки, можливість (не)відображення підказки, значення за замовчанням, порядок у списку атрибутів при перегляді, параметр для задання групових атрибутів та ін.), з яких складаються Д.

2) *Сегмент Д* - Д, які можуть належати до різних їх категорій.

3) *Сегмент освітніх ІР різного формату* — файли, які можна пов'язувати зі значеннями атрибутів типу «Файл» Д (навчально-методичні матеріали, фото і т.д.) і які можуть бути довільного формату.

4) *Сегмент адміністративних даних* – інформація про користувачів (профілі користувачів): логін, пароль, прізвище, ім'я, e-mail, привілеї (активність, статус персоналу, статус суперкористувача, дозволи на виконання дій над інформаційними об'єктами), останній вхід, дата приєднання, входження у групи тощо. Статус суперкористувача визначає, що такий користувач має всі дозволи без їх точного

зазначення. Тут також міститься інформація: про групи користувачів, для яких можна визначити права на виконання дій над інформаційними об’єктами системи (наприклад, на додавання, редагування і видалення атрибутів) та права на режими роботи з відповідними ІР (ресурс публікацій, ресурс педагогічного досвіду тощо), які визначаються умовами шаблонів; про підтримку шаблонів — умови

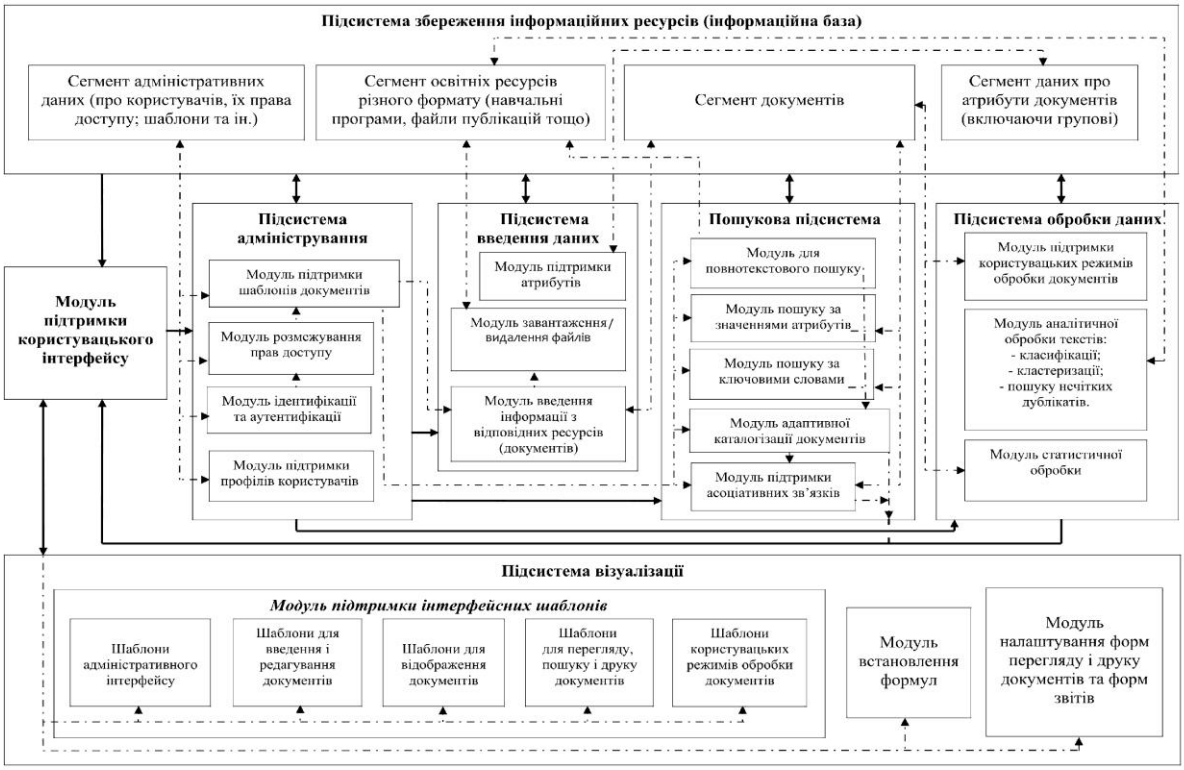

Рисунок 1 - Архітектура програмного забезпечення сховища даних навчального закладу.

шаблонів, порядок відображення атрибутів для кожного шаблону; про параметри відображення форм перегляду Д та форм звітів, параметри звітів (формули). Для кожного шаблону або виду звіту можна створювати форми, у параметрах відображення яких задається інформація про координати розміщення і розміри блоків для атрибутів і їх значень (для форм Д) або для результатів звітів, а також для блоків міток.

ПС адміністрування має наступні складові.

1) *Модуль підтримки профілів користувачів* — дозволяє створювати, редагувати і видаляти профілі користувачів.

2) *Модуль ідентифікації та аутентифікації* — забезпечує можливість введення користувачем логіну і паролю та встановлення відповідності користувача пред’явленому ним ідентифікатору шляхом співставлення введених значень з інформацією, наявною у сегменті адміністративних даних.

3) *Модуль розмежування прав доступу* —у відповідності з правами доступу надає користувачеві можливість роботи у доступних режимах з доступними ІР системи, які визначаються умовами шаблонів. Для системи визначені наступні режими роботи з Д: *додавання, редагування, модерування, перегляд*. Режим модерування призначений для контролю за правильністю введених даних. Право на модерування надає користувачеві можливість обмежити доступність редагування певного набору Д для групи інших користувачів, після перевірки на коректність введених ними значень атрибутів. Таким чином, користувач системи, у обов'язки якого входить лише введення даних для певного сегменту ІР, може редагувати доступні йому Д доти, доки їх не перевірить і не зафіксує (промодерує) відповідальна особа.

4) *Модуль підтримки шаблонів* — дозволяє додавати, редагувати і видаляти шаблони і умови для них, здійснювати відбір Д, відповідних певному шаблону, а також визначати порядок відображення атрибутів для наборів Д, які відповідають певному шаблону. Для шаблону можна задавати умови не лише для значень атрибутів, які доступні для режиму редагування, але й атрибутів, які мають адміністративний

характер («Модерування», «Власник»). Завдяки цьому забезпечується функціонування режиму модерування і фільтрація Д, створених певним користувачем, з метою надання йому прав лише на роботу із сегментом власних Д у визначеному режимі.

ПС введення даних включає наступні модулі.

1) *Модуль підтримки атрибутів* — дозволяє створювати, редагувати і вилучати атрибути з усіма їх властивостями, групувати і розгруповувати атрибути для користувачів, у яких є дозвіл на зміну інформаційних об'єктів даного типу. Аналогічні можливості надаються у підсистемі адміністрування, але на відміну від підсистеми введення даних з дещо іншою логікою діалогу і ширшими можливостями та зручнішими інструментами для відбору необхідних атрибутів (сортування за певними полями, пошук за назвою, фільтрація за типом).

2) *Модуль введення Д* — дозволяє вводити і редагувати Д для різних ІР. Конкретний користувач може вводити (редагувати) лише ті Д, які відповідають шаблону, визначеному для режиму введення (редагування) для груп, у які включений даний користувач. При введенні даних користувачеві пропонується обрати шаблон, після чого відображається певний набір атрибутів, який динамічно формується у базі даних у процесі перебору всіх існуючих Д. Таким чином, набір атрибутів шаблону — це сукупність всіх існуючих атрибутів збережених Д, котрі відповідають обраному шаблону. Додавання нового атрибуту в Д призводить до появи даного атрибуту у наборах атрибутів всіх шаблонів, яким відповідає вказаний Д. Якщо при зміні значення одного з атрибутів Д атрибут втрачає відповідність певному шаблону, то атрибути даного Д не враховуються при формуванні набору атрибутів даного шаблону. До Д можна включати атрибути, які уже присутні, для введення декількох значень для одного атрибуту.

3) *Модуль завантаження/видалення файлів* — дозволяє завантажувати з файлової системи на сервер і видаляти зі сховища освітні ІР або довідкові файли, які пов'язуються зі значеннями атрибутів типу «Файл», при введенні і редагуванні Д. У режимі перегляду модуль забезпечує завантаження з серверу необхідних файлів для детальнішого ознайомлення з їх змістом.

ПС пошуку містить наступні модулі.

1) *Модуль адаптивної каталогізації* — систематизує Д сховища, надає структуровану інформацію про Д і можливість її застосування для подальшого пошуку, забезпечує доступ до Д за атрибутами. Перегляд Д здійснюється через каталог, у якому відображається той набір Д, який доступний користувачу у цьому режимі згідно його прав. Будь-який з Д, доступних для перегляду, може мати довільний набір атрибутів. Користувачеві надається список всіх атрибутів, що присутні у доступному для перегляду наборі Д. При виборі атрибуту відображається список всіх його можливих значень у алфавітному порядку серед всіх Д набору, які можуть бути використані для подальшого пошуку. Список атрибутів змінюється відповідно до зміни набору Д.

2) *Модуль для повнотекстового пошуку* — дозволяє здійснити пошук серед файлів, які пов'язуються з Д. Даний модуль також має реалізувати пошук за ключовими словами у тексті файлів Д і використовувати порядок відображення результатів, який залежить не лише від частоти входження вказаних слів, але і від інших критеріїв, які підвищують релевантність пошуку, за рахунок впровадження засобів Text Mining.

3) *Модуль пошуку за значеннями атрибутів* — дозволяє відфільтрувати Д згідно заданих значень атрибутів. Даний вид пошуку може бути ієрархічним. Таким чином, реалізується операція перетину для множин Д з обраними значеннями атрибутів.

4) *Модуль пошуку за ключовими словами* — дозволяє знайти Д, у значеннях відповідних атрибутів яких містяться введені користувачем ключові слова (регістронезалежний пошук). Пошук за ключовими словами може здійснюватись для значень довільного набору атрибутів. Таким чином, будуть відібрані ті Д, які мають вказані користувачем атрибути, і у відповідних значеннях яких містяться введені ключові слова. Пошук може бути комбінований (взаємодіючий). Можна також паралельно використовувати ці два види пошуку (за значеннями атрибутів і за ключовими словами), що забезпечує фільтрацію Д за значеннями і вмістом відповідних атрибутів.

5) *Модуль підтримки асоціативних зв'язків* — дозволяє знайти за допомогою засобів адаптивної каталогізації ті Д, які пов'язані з обраним, і встановити асоціативні зв'язки за рахунок атрибутів відповідного типу. Якщо у користувача немає прав на перегляд Д, які пов'язані асоціативним зв'язком з доступним, то відповідні атрибути, які встановлюють вказаний зв'язок, взагалі не відображаються.

ПС обробки даних включає наступні модулі.

1) *Модуль підтримки користувацьких режимів обробки Д* — забезпечує реалізацію необхідних користувачеві задач шляхом здійснення операцій (вибірка, фільтрація та ін.) над необхідними наборами Д. В основі лежить використання синтаксичного аналізатора формул і математичних виразів, завдяки чому забезпечується підтримка змінних вимог до проведення обробки інформації. При необхідності здійснення операцій над новоствореним сегментом Д з новими атрибутами відповідальна особа вводить формули з відповідними параметрами, які визначаються логікою обробки.

2) *Модуль аналітичної обробки текстів* забезпечує наступні процеси.

- Класифікацію — групування вихідних понять, наявних у тексті, за будь-якими суттєвими ознаками і розташування їх у певному порядку, що відображає ступінь схожості.
- Кластеризацію — логічне продовження класифікації, але класи набору даних заздалегідь невідомі. Призначена для розбиття сукупності об'єктів на однорідні групи.
- Пошук нечітких дублікатів (для виявлення плагіату). Забезпечує визначення того, чи фрагменти даного тексту взяті із уже існуючих, тобто наявних у сегменті освітніх ІР інформаційної бази системи.

3) *Модуль статистичної обробки* забезпечує застосування засобів статистики (статистичних показників, критеріїв) для формування даних, необхідних для звітної документації. *ПС візуалізації* містить наступні складові частини.

1) *Модуль підтримки інтерфейсних шаблонів* — забезпечує генерацію веб-сторінок з необхідною інформацією і включає наступні групи шаблонів.

- *Шаблони адміністративного інтерфейсу* призначені для привілейованих користувачів (адміністраторів). Забезпечують інтерфейс з елементами управління і навігацією, які дозволяють виконувати дії адміністративного характеру. Для кожного типу об'єктів існує форма редагування, яка дозволяє додавати, редагувати і видаляти конкретні записи і список всіх доступних у базі даних об'єктів. У формі редагування для кожного типу значень визначено відповідні елементи відображення (календар для типу «Дата», перемикач для булевого типу тощо). У списку об'єктів забезпечуються зручні візуальні засоби для здійснення пошуку, фільтрації, сортування змінюваних інформаційних об'єктів. При встановленні порядку для атрибутів шаблону здійснюється динамічне відображення набору атрибутів даного шаблону у порядку, який був збережений у базу. Для зміни цього порядку використовується технологія drag-and-drop (перетягування за допомогою миші).
- *Шаблони для введення і редагування Д*. Забезпечується відображення атрибутів у встановленому порядку у відповідному вікні з метою надання користувачеві можливості вибору і включення певного атрибуту в Д. Елементи інтерфейсу забезпечують зручний режим введення (редагування) Д за рахунок:
- використання динамічного відображення атрибутів, які відповідають обраному шаблону;
- спливаючих підказок для введення значень;
- різних форм відображення елементів для введення значень атрибутів (перемикачі, прапорці, текстові поля, поля введення, випадаючі списки, поля типу "Файл", спливаючі календарики для введення значень типу "Дата");
- діалогових вікон з адаптивною каталогізацією (для атрибутів типу "Асоціативний зв'язок") та ін.

Для кожної з форм відображення елементів введення, якщо значення для доданого в Д атрибуту не було встановлено, атрибут зберігається у Д і доступний при наступному редагуванні без додаткових включень.

- *Шаблони для відображення Д*. Атрибути Д відображаються у встановленому модулем підтримки шаблонів порядку. Для певних форматів завантажені освітні ІР можна переглядати у відповідному вікні і здійснювати друк. Всі файли можна завантажувати з серверу, користуючись посиланням. У даному шаблоні перегляду можна відобразити Д у вигляді будь-якої форми інших шаблонів, яким відповідає даний Д, а також роздрукувати форму. Для атрибутів типу «Асоціативний зв'язок» значення відображаються у вигляді гіперпосилання, за допомогою якого можна перейти до перегляду відповідного Д.
- *Шаблони для перегляду, пошуку і друку Д*. Наявна посторінкова навігація доступних для перегляду Д. Пошук за ключовими словами полегшується за рахунок використання спливаючих підказок, а пошук за значеннями атрибутів — за рахунок динамічного відображення значень у алфавітному порядку. Для друку обраних Д можна обирати атрибути, значення яких відображені у таблиці для виведення інформації. У таблиці для відображення вмісту Д значення складних атрибутів групуються так само, як і декілька значень одного й того ж атрибуту в одному Д.
- *Шаблони користувацьких режимів обробки Д*. Користувач може переглядати у відповідних формах звіти, які містять інформацію про результати обробки певних ІР.

2) *Модуль налаштування форм перегляду і друку Д і форм звітів* забезпечує можливість створення, редагування і видалення форм перегляду і друку Д та форм звітів, а також налаштування параметрів відображення форм. Доступні такі режими роботи з формами: створення нової форми, створення нової форми на основі існуючої, редагування форми. У конструкторі форм Д при виборі шаблону здійснюється динамічне відображення списку набору атрибутів даного шаблону, списку форм, визначених для нього, і списку міток, наявних у вже створених формах. Користувач має можливість здійснювати такі операції: перетягування атрибутів на графічне поле, встановлення та зміна координат розміщення та розмірів для блоків відображення назв і значень атрибутів, видалення їх із області видимості графічного поля. Аналогічні дії можна виконувати над блоками для міток. У конструкторі форм звітів надається перелік існуючих форм, список параметрів звітів (з назвами і відповідними формулами) та список міток, що використовуються. При виборі однієї з форм відмічаються наявні в обраній формі параметри та мітки.

Аналогічно уже згаданому конструктору, користувач може перетягувати нові параметри і мітки на графічне поле, встановлювати необхідні координати і розміри шляхом перетягування та видаляти їх.

3) *Модуль встановлення формул* (конструктор формул) — дозволяє вибрати необхідні параметри для формул та автоматизує процес їх створення. Надається перелік створених формул, при виборі однієї з яких вона відображається у вікні редагування, де можна змінити її параметри. Підбір параметрів при створенні і редагуванні формул встановлюється за рахунок: вибору виду формули; вибору шаблону(ів), атрибути якого необхідні для формул, або опції для перегляду всіх атрибутів; вибору необхідних атрибутів (включаючи групові і складові групових) з можливим вказуванням їх значення із використанням підказки. Вказані дії можна виконувати у довільному порядку. При збереженні формул здійснюється їх перевірка на коректність.

Модуль підтримки користувацького інтерфейсу забезпечує динамічну взаємодію підсистем, яким ставиться задача, з інтерфейсними шаблонами, які відображаються для користувача, тобто забезпечує підстановку результуючих даних у відповідні шаблони, взаємодію шаблонів (вставку одних шаблонів у інші) та ін.

Порядок функціонування програмного забезпечення сховища даних

На рис.2 наведено діаграму основних станів системи в процесі її експлуатації

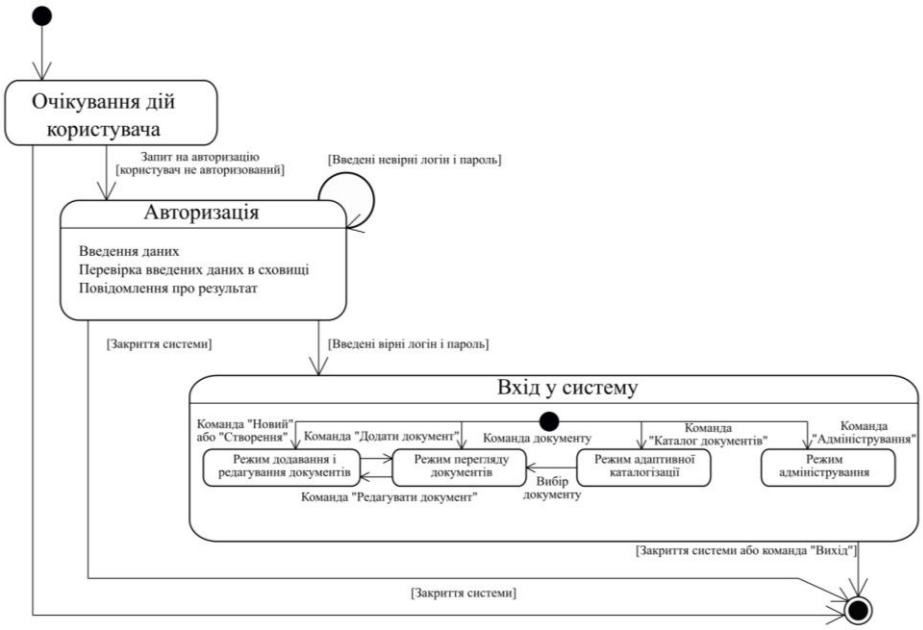

Рисунок 2 - Діаграма основних станів системи.

користувачами. Користувач не буде допущений до роботи з системою, доки не виконає авторизацію. У кожен з режимів можна перейти шляхом вибору відповідної вкладки. Існують додаткові можливості переходів між режимами. Після вибору потрібного Д режимі адаптивної каталогізації відкривається режим перегляду для детального ознайомлення із вмістом Д. З режиму перегляду в разі необхідності можливий перехід у режим редагування відкритого Д. Після внесення необхідних змін або додавання до ІР нового Д користувач має можливість перейти в режим перегляду для остаточної перевірки внесеної інформації. Вихід здійснюється або при закритті системи або при виконанні команди «Вихід».

Далі подано діаграми станів для основних режимів системи, які доступні звичайним користувачам. Режиму додавання і редагування Д відповідає діаграма на рис. 3.

У згаданому режимі забезпечується паралельність виконання дій з атрибутами і роботи з Д. За наявності необхідних прав користувачеві доступні такі операції: додавання, вилучення, редагування, групування і розгруповування атрибутів. Внесені зміни динамічно відображаються за допомогою відповідних запитів, при успішному виконанні яких здійснюється перезавантаження списку атрибутів.

Користувач має змогу перейти до стану роботи з Д безпосередньо шляхом включення необхідних атрибутів, після чого у формі Д відображаються елементи для заповнення значень, або шляхом вибору шаблону і запуску відповідної команди, після чого здійснюється відбір набору атрибутів шаблону і відображення їх у формі. На будь-якому кроці можна включати нові і вилучати наявні в Д атрибути. У екранній формі Д для атрибутів можливий один з п’яти станів для встановлення значень: введення значення, вибір значення із запропонованих, введення значення з використанням підказки, вибір файлу з відповідного діалогового вікна, вибір Д із вікна каталогу. При збереженні або оновленні Д здійснюється перехід у режим перегляду.

Діаграма станів системи в режимі перегляду Д подана на рис.4.

Рисунок 3 - Діаграма станів системи в режимі додавання і редагування документів

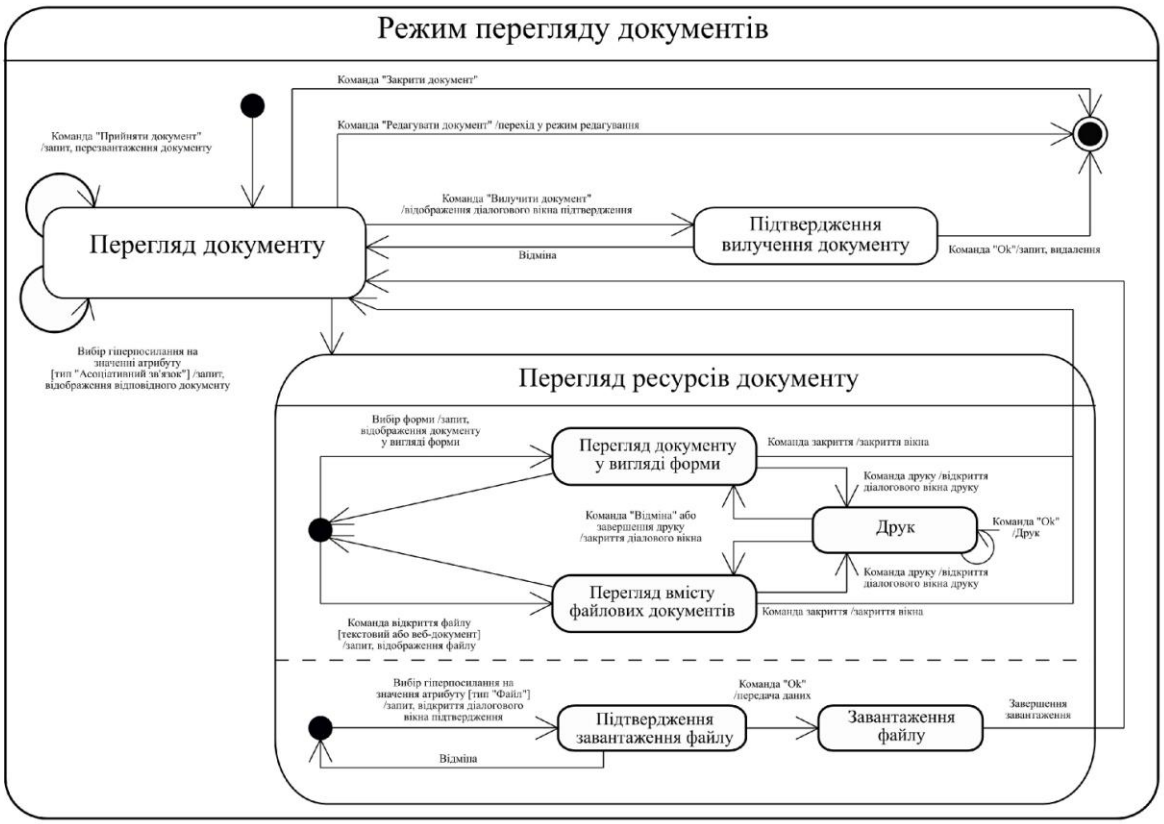

Рисунок 4 - Діаграма станів системи в режимі перегляду документів.

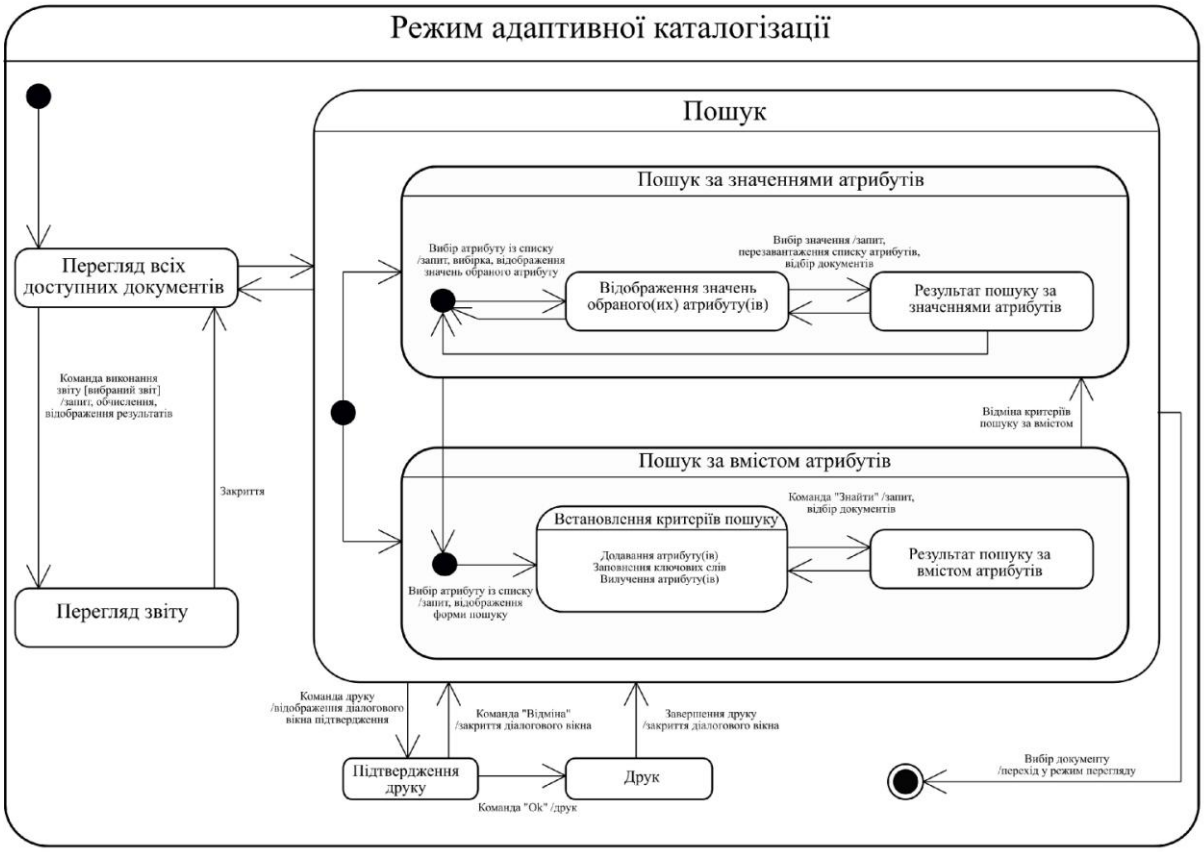

Рисунок 5 - Діаграма станів системи в режимі адаптивної каталогізації.

При перегляді неприйнятих Д уповноважений відповідним чином користувач може промодерувати один з них, після чого буде здійснено збереження змін і перезавантаження Д. Якщо наявні значення для атрибутів типу «Асоціативний зв'язок», то вибір відповідного гіперпосилання супроводжується відкриттям обраного Д у режимі перегляду. Можна паралельно здійснювати такі дії:

- завантаження файлу;
- перегляд Д у вигляді певної форми або перегляд вмісту текстових і веб-документів.

Вийти з даного режиму можна шляхом переходу у режим редагування, закриття або вилучення відкритого Д. При вході у режим адаптивної каталогізації надається можливість перегляду всіх доступних Д. Користувач може перейти у стан пошуку (рис. 5). При виборі атрибуту зі списку здійснюється запит і відображення значень обраного атрибуту. Після вибору значення здійснюється перезавантаження списку атрибутів і вибірка Д, які відповідають умовам пошуку.

Вказані операції можна повторювати, додаючи або відміняючи умови для фільтрації Д. На будь-якому кроці пошуку за значеннями атрибутів можна перейти до задавання критеріїв для пошуку за вмістом. При цьому відібраний сегмент Д буде відповідати критеріям обох видів пошуку. Для пошуку за вмістом атрибутів аналогічно можна додавати або відміняти умови пошуку шляхом додавання, вилучення атрибутів і встановлення ключових слів для значень атрибутів. При переході до стану пошуку за значеннями атрибутів здійснюється відміна критеріїв пошуку за вмістом.

У кожному зі згаданих вище станів можна здійснити друк списку Д, по відміні або завершенні якого закривається діалогове вікно друку і здійснюється перехід у попередній стан.

На довільному кроці пошуку можна перейти в стан перегляду всіх доступних Д шляхом вибору команди «Документи». Після вибору одного з доступних користувачу звітів і запуску команди його формування здійснюється запит і парсинг відповідних формул для обчислення передбачених у звіті параметрів та для вибірки необхідних значень атрибутів.

Особливості програмної реалізації системи

Для розробки і проектування системи було обрано перспективні сучасні інструментальні засоби, які підтримуються широким колом відкритого співтовариства, — високорівнева об'єктно-орієнтована мова Python і фреймворк Django (з відкритим вихідним кодом) [14, 15].

Обраний програмний інструментарій (Django) має свої переваги:

- вбудований ORM (Object-relational mapper) — забезпечує проекціювання реляційний даних в об'єкти;
- підтримка MVC (Model-View-Controller) — дозволяє розділити логіку доступу до даних, бізнес-логіку (управління) і логіку відображення (інтерфейс користувача), внаслідок чого забезпечується зручність оновлення;
- висока швидкість роботи (може витримувати велике навантаження і має вбудовані можливості для кешування і розподілу навантаження);
- підвищення рівня безпеки за рахунок спеціальних пакетів для захисту від SQL-ін'єкцій, атак типу «підробка HTTP запитів» (Cross-Site Request Forgery), підробки сесій та ін.

Програмування інтерфейсу здійснено з використанням динамічної слаботипізованої мови Javascript з об'єктно-орієнтованими можливостями і новітніх веб-технологій, які забезпечують кросбраузерність (за рахунок бібліотеки jQuery) і динамічність (за рахунок технології AJAX) інтерфейсу [16]. Враховуючи орієнтованість можливість широкого застосування системи у навчальних закладах різних рівнів, інтерфейс реалізовано із застосуванням стандартних засобів управління у інтуїтивно зрозумілій пересічному користувачеві формі (рис. 6).

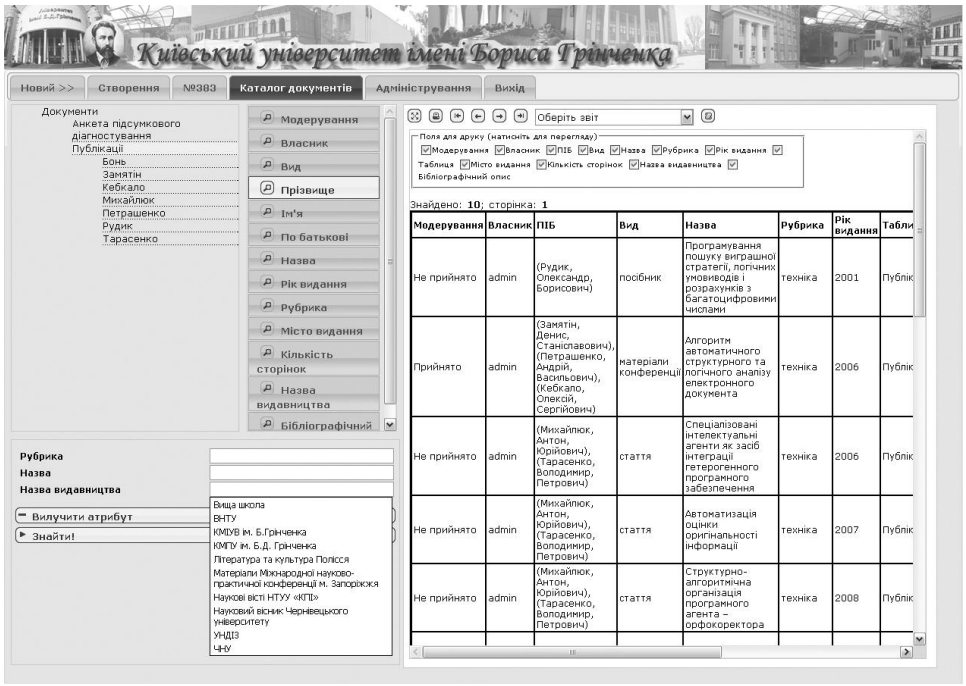

Рисунок 6 - Інтерфейс програмного забезпечення сховища даних (режим адаптивної каталогізації).

Висновки

Запропоновано концепцію архітектури ПЗ для формування та підтримки сховища слабоструктурованих даних довільного навчального закладу з метою впорядкування його відомчих електронних ІР.

Характерною рисою ПЗ є здатність підтримувати високу динаміку логічної схеми даних. Передусім ПЗ орієнтоване на оперування електронними Д, в першу чергу текстомісткими, однак на його базі можуть бути реалізовані також широкопрофільні банки даних, здатні підтримувати ІР практично для повного спектру інформаційної діяльності в галузі освіти.

ПЗ впроваджено на базі Київського університету імені Бориса Грінченка, де з його допомогою формувались і вводилися до експлуатації сховища студентських випускних робіт та навчальних презентацій, ресурсний центр методичного забезпечення, різноманітні репозиторії даних навчального спрямування. Подальше дослідження та розробка проводяться в напрямку створення в складі ПЗ сховища даних універсального інструментарію для інтерактивного формування розвиненої бізнес-логіки.

Література

1. Згуровський М.З., Родіонов М.К., Жиляєв І.Б. Розвиток інформаційного суспільства в Україні: Правове регулювання у сфері інформаційних відносин. — К.: НТУУ «КПІ», 2006. — 542 с.
2. Кюблер Х.-Д. Міфи про суспільство знань. — К.: Видавничий дім Дмитра Бурого, 2010. — 264 с.

3. Управління інформаційними технологіями вищих навчальних закладів: навчальний посібник. Видання третє, доповнене / За ред. О.В. Співаковського. — Херсон: Айлант, 2010. — 302 с.
4. Елочкин М.Е., Брановский Ю.С., Николаенко И.Д. Информационные технологии: Учебник. — М.: Оникс, 2007. — 256 с.
5. Ландэ Д.В., Снарский А.А., Безсуднов И.В. Интернетика. Навигация в сложных сетях: модели и алгоритмы. — М.: Либроком, 2009. — 264 с.
6. Ландэ Д.В. Поиск знаний в Internet. Профессиональная работа. — М.: Вильямс, 2005. — 272 с.
7. Саак А.Э., Пахомов Е.В., Тюшняков В.Н. Информационные технологии управления: Учебник для вузов. 2-е изд. — СПб.: Питер, 2008. — 320 с.
8. Тарасенко В.П., Михайлюк А.Ю., Сніжко М.В., Бігун Л.М. Функціональність спеціалізованих інформаційно-аналітичних систем для підтримки інформаційно-навчальної діяльності // Проблеми інформатизації та управління. — 2009. — № 3 (27). — С. 123-130.
9. Кебкало О.С., Михайлюк А.Ю., Тарасенко В.П. Тематична класифікація структурованих та частково структурованих текстових документів освітньої спрямованості // Збірник праць Четвертої міжнародної конференції «Нові інформаційні технології в освіті для всіх: інноваційні методи та моделі». — 2009. — С. 75-81.
10. Бігун Л.М., Замятін Д.С., Михайлюк А.Ю., Петрашенко А.В. Інформаційно-аналітична система для підтримки самостійної навчальної роботи // Збірник праць Четвертої міжнародної конференції «Нові інформаційні технології в освіті для всіх: інноваційні методи та моделі». — 2009. — С. 331-338.
11. Mykhailyuk A.Y., Zamiatin D.S., Petrashenko A.V. Unstructured Data Warehouse Processing System Based on an Uniform Set of Functions // Proceedings of the 4-th International Conference "Advanced Computer Systems and Networks: Design and Application" (ACSN-2009). — 2009. — P. 117-119.
12. Михайлюк А.Ю., Замятін Д.С., Петрашенко А.В. Проблеми систематизації даних у неструктурованих текстових сховищах // Інформатика, обчислювальна техніка та кібернетика: Вісник Університету «Україна». — 2010. — № 8. — С. 88 -91.
13. Mykhailyuk A.Y., Petrashenko A.V., Zamiatin D.S. Method of Unstructured Text-Based Warehouses Systematization Based on Composite Attributes // Збірник праць Ювілейної міжнародної науково-практичної конференції, що присвячена 50-річчю створення першої на Україні кафедри обчислювальної техніки РКС-2010 "Розподіленні комп'ютерні системи". — 2010. — Т. 1. — С. 86-88.
14. Python Documentation Index [Electronic resource]. — Access mode: http://python.org/doc/.
15. Django | Django documentation [Electronic resource]. — Access mode: http://docs.djangoproject.com/en/1.2/.
16. Бибо Б., Кац И. jQuery. Подробное руководство по продвинутому JavaScript. — Спб.: Символ-Плюс, 2009. — 384 с.

Відомості про авторів

Замятін Денис Станіславович – к.т.н., доцент кафедри системного програмування і спеціалізованих комп’ютерних систем (СП і СКС), НТУУ “Київський політехнічний інститут”, пр.Перемоги, 37, м. Київ, 03056, тел. 454-90-31, e-mail: dsz@ukr.net.

Кебкало О. С., науковий співробітник кафедри СП і СКС, НТУУ “Київський політехнічний інститут”, пр.Перемоги, 37, м. Київ, 03056, тел. 454-94-92, e-mail: kebka@mail.ru.

Лукашевська Л. І. – інженер, Київський університет імені Бориса Грінченка, м. Київ, тел. 068-127-86-96, lukashe@ukr.net.

Михайлюк Антон Юрійович – к.т.н., доцент кафедри інформатики, Київський університет імені Бориса Грінченка, м. Київ, тел. 461-02-43., e-mail: may-62@ukr.net.

Михайлюк Олена Станіславівна - науковий співробітник кафедри СП і СКС, НТУУ “Київський політехнічний інститут”, пр.Перемоги, 37, м. Київ, 03056, тел. 454-94-92, e-mail: mes@scs.ntu-kpi.kiev.ua.

Петрашенко Андрій Васильйович - к.т.н., доцент кафедри СП і СКС, НТУУ “Київський політехнічний інститут”, пр.Перемоги, 37, м. Київ, 03056, тел. 454-90-31, e-mail: petrashenko@yandex.ru.

Тарасенко Володимир Петрович – д.т.н., професор, завідувач кафедри СП і СКС, НТУУ “Київський політехнічний інститут”, пр.Перемоги, 37, м. Київ, 03056, тел. 236-32-02.