

Голубєв Ю.,
наук. співроб. НБУВ

АВТОМАТИЗОВАНА ТЕХНОЛОГІЯ ФОРМУВАННЯ ІНФОРМАЦІЙНОГО МАСИВУ НАЙМЕНУВАНЬ ОРГАНІЗАЦІЇ ДЛЯ АВТОРИТЕТНИХ ЗАПИСІВ

У статті наведено детальний опис та особливості використання автоматизованої технології формування інформаційного масиву найменувань організацій на першому етапі формування авторитетних записів. Представлений перелік функцій технології та кнопок керування ними.

Ключові слова: авторитетні записи, найменування організацій, автоматизована технологія, формування інформаційного масиву.

The article contain detailed description and the features for using of automated technology for forming of organizations names information array on the first stage of authority record formation. The list of technology functions and control buttons management is given.

Keywords: authority records, organization name, automated technology, information array.

В статье приведено подробное описание и особенности использования автоматизированной технологии формирования информационного массива наименований организаций на первом этапе формирования авторитетных записей. Представлен перечень функций технологии и кнопок управления ими.

Ключевые слова: авторитетные записи, наименование организации, автоматизированная технология, формирование информационного массива.

Одним зі складних та довготривалих процесів у технології формування авторитетного запису є етап пошуку варіантів попередніх форм найменувань організацій, зібраних із бібліографічних записів електронного каталогу. За період існування організація могла змінювати найменування та підпорядкування декілька разів. У поєднанні з варіантами форм найменувань головної установи або підпорядкованих підрозділів такі зміни спричиняють створення великого масиву варіантів найменувань, подібних до найменувань інших організацій. Використовуючи певну технологію, каталогізатор має знайти в масиві всі

можливі форми найменування організації або підрозділу та об'єднати їх у групу під одним ідентифікаційним номером. Об'єднані найменування організації складають інформаційний масив, умовно названий «корпорація». Масив використовується на етапах формування авторитетних записів для вибору авторитетного заголовку, варіантних форм, зв'язків з іншими авторитетними записами тощо.

Варіанти форм найменувань організацій, що були відібрані з бібліографічних записів електронного каталогу НБУВ, зберігаються в записах бази даних під назвою «CORP». Модель бази даних містить відповідні поля для зберігання найменування у UNIMARC-форматі, найменування у форматі «як на документі», ідентифікаційного номера запису найменування, ідентифікаційного номера корпорації, під яким гуртуються варіанти найменувань однієї організації, рік видання документа, з якого було отримано це найменування, та деякі інші поля для організації функціонування бази. По завершенню формування масиву всі записи поточної теми, наприклад національні вищі навчальні заклади, експортуються з бази даних. Під час експорту формуються записи у структурі бази даних системи – для «IRBIS64» подальшого імпорту до робочої бази даних авторитетних записів.

Технологія базується на автоматичній побудові гіпертекстових списків з кнопками керування у формах, монтованих на мові гіпертекстової розмітки (HTML) та посиланнями між елементами списків. Відповідно інтерфейс користувача здійснюється встановленим браузером на локальному комп'ютері. Керування програмами побудови здійснюється скриптами, реалізованими мовою PYTHON для локальної версії веб-сервера ENDELS. Налаштування технології зберігаються в папках локального сервера.

Програми передпобудовою гіпертекстових списків виконують задане користувачем добування записів з бази даних. Інтерфейс користувача дозволяє складати список пошукових термінів вибірки та встановлювати значення булевої логічної операції AND, NOR, OR для кожного рядка списку пошукових термінів. Елемент рядка можна задавати одним словом або декількома цілими чи скороченими словами. Кількість рядків не обмежена, будь-який рядок можна відключити, встановивши для нього режим OFF.

Склад гіпертекстових списків добірки пошукових термінів такий:

- абетковий список найменувань (Alist);
- список корпорацій, до яких входять найменування (Clist);
- абетковий список слів з найменувань (Wlist);
- список пошукових термінів складений із довільних цілих або відсічених справа слів для побудови наступної вибірки (Tlist);
- список пошукових термінів складений тільки зі слів поточної вибірки для побудови наступної вибірки (Slist);

- докладний склад окремої корпорації (Ccorp);
- список складу поточної корпорації та обраних найменувань (VReC);
- список історії виборок (Hisel).

Склад елемента абеткового списку назв (Alist):

- порядковий номер у списку;
- посилання <Clist> до ідентифікаційного номера найменування в списку корпорацій (Clist);
- найменування організації;
- посилання <"слово"> від кожного слова найменування до відповідного елемента у списку слів з найменувань (Wlist);
- посилання <GO> до ідентифікаційного номера корпорації у списку корпорацій (Clist);
- код належності до теми.

Склад елемента списку корпорацій, до яких входять найменування (Clist):

- кнопка [NexC] здвигу та запам'ятовування позиції вказівника на наступний елемент у списку корпорацій (Clist);
- кнопка [C] повернення позиції вказівника на поточний елемент у списку корпорацій (Clist), встановлено кнопкою [NexC];
- кнопка [CORP] побудови докладного списку складу окремої корпорації (Ccorp);
- кнопка [RECR] запам'ятовування позиції вказівника на обрану поточну корпорацію у списку корпорацій (Clist);
- кнопка [V] повернення позиції вказівника на поточну корпорацію, обрану кнопкою [RECR] у списку корпорацій (Clist);
- кнопка [PART] відмітки корпорації як складової частини, раніше обраної кнопкою [RECR] у списку корпорацій (Clist);
- кнопка [PROL] відмітки корпорації як обробленої у базі даних CORP та у списку корпорацій (Clist);
- кнопка [ТЕМА] відмітки належності корпорації до тематичного списку у базі даних CORP та у списку корпорацій (Clist);
- ідентифікаційний номер корпорації;
- порядковий номер у складі корпорації;
- найменування організації;
- кнопка [NewC] включення або видалення обраного найменування до списку претендентів на приєднання до поточної корпорації (VReC);
- кнопка [BR] побудови бібліографічного опису та списку полів джерела найменування (Brfi);
- посилання <Alist> на відповідну назву в абетковому списку найменувань (Alist).

Склад елемента абеткового списку слів з найменувань (Wlist):

- віконце вводу пошукового терміна (in);
- кнопка [WW] введення пошукового терміна до першого рядка нового списку (Tlist);
- кнопка [N] введення пошукового терміна до нового рядка списку (Tlist);
- кнопка [A] введення пошукового терміна до останнього рядка списку (Tlist);
- кнопка [vWW] побудови списку введених пошукових термінів до списку (Tlist);
- віконця (all), (nf) вводу ознаки режиму перебудови списку слів з найменувань (List Short);
- кнопка [LisW] перебудови в режимі (List Short) списку слів з найменувань (Wlist);
- кнопка [Hi] побудови списку історії вибірок (Hisel);
- кнопка [VReC] побудови списку поточної корпорації та претендентів на приєднання до неї (VReC);
- кнопка [NexA] здвигу та запам'ятовування позиції вказівника на наступний елемент у абетковому списку найменувань (Alist);
- кнопка [A] повернення позиції вказівника на поточний елемент у абетковому списку найменувань (Alist);
- віконце вводу і кнопка [NewA] введення та запам'ятовування довільної позиції вказівника елемента в абетковому списку найменувань (Alist);
- кнопка [NexC] здвигу та запам'ятовування позиції вказівника на наступний елемент у абетковому списку корпорацій (Clist);
- кнопка [C] повернення позиції вказівника на поточний елемент у списку корпорацій (Clist);
- посилання <h> до шапки списку (Wlist);
- порядковий номер у списку;
- слово з кількістю найменувань в базі даних;
- кнопка [NexW] здвигу та запам'ятовування позиції вказівника на наступний елемент у абетковому списку слів (Wlist);
- кнопка [W] повернення позиції вказівника на поточний елемент у абетковому списку слів (Wlist);
- кнопка [M] запам'ятовування позиції елемента в абетковому списку слів (Wlist);
- кнопка [v] повернення вказівника на поточний елемент, що був помічений кнопкою [M] в абетковому списку слів (Wlist);
- кнопка [LN] включення поточного слова як пошукового терміна до першого рядка нового списку (Slist);

- кнопка [R] включення поточного слова як пошукового терміна до нового рядка списку (Slist);
- кнопка [A] включення поточного слова як пошукового терміна до останнього рядка списку (Slist);
- кнопка [Lv] побудови списку слів як пошукових термінів, включених до списку (Slist);
- список посилань <00123> до ідентифікаційного номера корпорації у списку корпорацій (Clist), до якої входить найменування з поточним словом.

Склад списку поточної корпорації та обраних найменувань (VReC):

- порядковий номер у складі корпорації;
- склад поточної корпорації;
- кнопка [AppC] запуску процесу приєднання;
- кнопка [DelC] видалення обраного найменування зі списку претендентів на приєднання до поточної корпорації;
- кнопка [MovC] видалення обраного найменування зі складу корпорації (нова корпорація).

Склад елемента списку пошукових термінів для побудови вибірки (Tlist):

- порядковий номер у списку;
- кнопка [Lnew] запуску процесу побудови вибірки;
- кнопка [Тема] встановлення режиму роботи фільтра належності корпорації до тематичних списків;
- кнопка [Proc] встановлення режиму роботи фільтра належності до списку вже оброблених корпорацій;
- порядковий номер у списку;
- склад рядка пошукових термінів;
- кнопка [OAN] встановлення режиму роботи фільтра OR, AND, NOT для кожного рядка пошукових термінів.

Склад елемента списку пошукових термінів для побудови наступної вибірки, складений зі слів поточної вибірки (Slist):

- кнопка [Lnew] запуску процесу побудови вибірки;
- кнопка [Тема] встановлення режиму роботи фільтра належності корпорації до тематичних списків;
- кнопка [Proc] встановлення режиму роботи фільтра належності до списку вже оброблених корпорацій;
- порядковий номер у списку;
- склад рядка пошукових термінів;
- кнопка [OAN] встановлення режиму роботи фільтра OR, AND, NOT для кожного рядка пошукових термінів;

- кнопка [CS] встановлення режиму роботи фільтра C, S для кожного рядка пошукових термінів.

Виконання технології починається з побудови списку пошукових термінів. Він складений із характерних слів і вводить їх до віконця (in), перегляду списку та запуску вибірки. Критеріями для вибору характерних слів, що відокремлюють організацію серед інших, є вид та характеристика діяльності, географічна ознака (місцезнаходження), найменування організації.

З метою уникнення граматичної залежності під час пошуку слова можна вводити усіченими справа. Наприклад, замість «аграрного», «аграрному», «аграрний» вводимо «аграрн». У зв'язку з існуванням найменувань організацій російською мовою треба вводити аналогічні російськомовні перекладені та скорочені назви. Наприклад, «харчовиц» можна вводити до одного рядка списку пошукових термінів. Пошукові терміни, до яких мають бути застосовані логічні операції AND, NOT, OR, повинні вводитись до різних рядків списку. Наприклад, «універ універ ун-т» вводимо до одного а «Шевчен» до наступного рядка. У результаті застосування AND до обох рядків на виході добірки отримуюмо перетин множин першого та другого рядків.

Керування формуванням списку довільних пошукових термінів здійснюється кнопками:

- [WW] – слово відкриває перший рядок нового списку;
- [N] – слово відкриває наступний рядок у списку;
- [A] – слово додаємо як наступне до останнього рядка;
- [vWW] – побудова списку пошукових термінів для перегляду перед використанням у добірці.

Після формування списку для кожного рядка треба встановити необхідні булеві логічні операції. Автоматично для кожного нового рядка пошукових термінів програма встановлює для роботи фільтру режим AND, який можна змінити кнопкою [OAN]. Якщо потрібно, можна встановити режими додаткових фільтрів корпорацій добірки відповідними кнопками. Кнопка [Тема] установки режиму роботи фільтра належності корпорації до тематичних списків дозволяє циклічно встановлювати наступні режими:

- «+T» – додати до складу вибірки тільки корпорації, що мають відмітку поточної теми;
- «-T» – додати до складу вибірки тільки корпорації, що не мають відмітки поточної теми;
- «+A» – додати до складу вибірки тільки корпорації, що мають відмітку будь-якої зі вказаних тем;
- «-A» – додати до складу вибірки тільки корпорації, що не мають відмітки будь-якої зі вказаних тем;

- «OFF» – ігнорувати наявність відмітки.

Кнопка [Proc] установки режиму роботи фільтра належності корпорації до списку оброблених дозволяє циклічно встановлювати такі режими:

- «ON» – додати до складу вибірки тільки корпорації, що мають відмітку обробленої

- «OFF» – ігнорувати наявність відмітки.

Використання кнопок [Тема] та [Proc] полегшує пошук за рахунок зменшення кількості елементів для перегляду. Це дозволяє швидше знаходити під час повторного пошуку пропущені споріднені найменування та не відмічені оброблені чи не віднесені до тематичного списку. Для цього потрібно встановити режим для теми «-Т» або «-А» та режим «ON» для обробки й запустити побудову вибірки кнопкою [Lnew] зі списку (VReC).

Кнопка [OAN] установки режиму роботи фільтра операції для слів у рядку дозволяє циклічно встановлювати такі режими:

- «OR» – застосувати для всіх слів у рядку логічну операцію OR під час добірки;

- «AND» – застосувати для всіх слів у рядку логічну операцію AND під час добірки;

- «NOT» – застосувати для всіх слів у рядку логічну операцію NOT під час добірки;

- «OFF» – ігнорувати весь рядок.

Кнопка [CS] установки режиму роботи фільтра слів у рядку дозволяє циклічно встановлювати такі режими:

- «S» – додати до складу добірки тільки корпорації, вивантажені з бази даних CORP та фільтровані поточною добіркою перед фільтром операції [OAN];

- «C» – додати до складу добірки всі корпорації, вивантажені з бази даних CORP перед фільтром операції [OAN].

Кнопка [CS] присутня тільки у списку пошукових термінів зі слів поточної добірки.

У будь-який момент можна зробити нову дочірню добірку і побудувати гіпертекстові списки та продовжити обробку корпорації у новому вікні браузера тільки з окремими словами і тільки з поточної вибірки (режим S). Такий підхід зменшує кількість найменувань для перегляду під час пошуку споріднених складових та сприяє зменшенню помилок. Для цього потрібно визначити і знайти важливі слова у списку слів та послідовно сформувати з них список пошукових термінів кнопками (рис. 1):

- [LN] – слово відкриває перший рядок нового списку;

- [R] – слово відкриває наступний рядок у списку;

- [A] – слово додаємо як наступне до останнього рядка;

- [Lv] – побудова списку слів для перегляду та використання у вибірці;

- [OAN] – встановити один з режимів роботи фільтра OR, AND, NOT або OFF для кожного рядка пошукових термінів. OFF ігнорує рядок у фільтрі;

- [CS] – встановити режим C або S роботи фільтра для кожного рядка пошукових термінів;

- [Lnew] – запустити процес побудови вибірки.

Приклад екрану відразу після вибірки

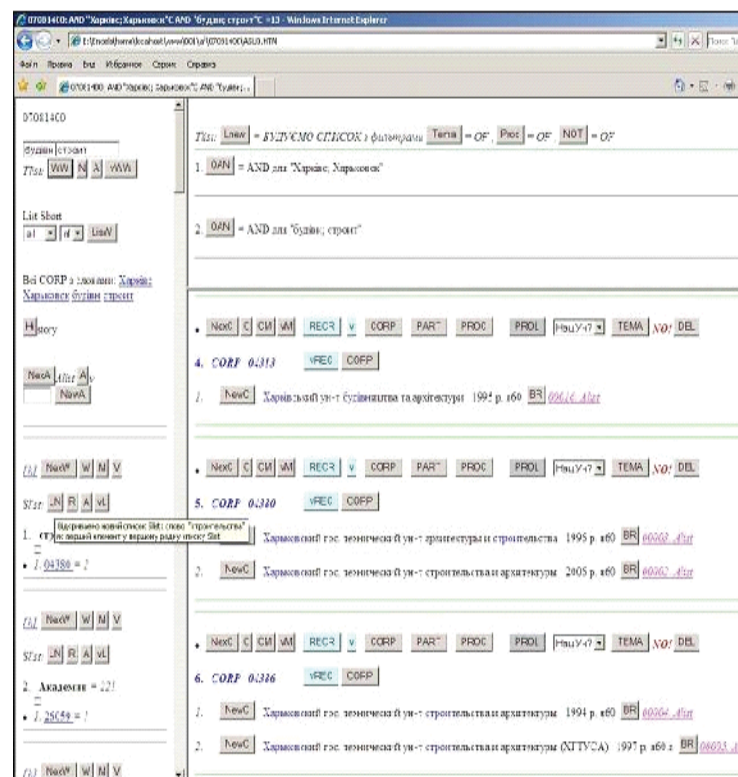


Рис. 1. Вигляд екрану відразу після вибірки

Пошук наступної корпорації для обробки може здійснюватись у списках корпорацій, найменувань, слів. Переключення між ними можна

здійснювати за допомогою посилань <<слово>> до слова, <Clist> до корпорації, <00123> до корпорації, <Alist> до найменування. Використання кнопок [NexA], [NexW], [NexC] робить елементи поточними та «підсвічує» їх у списках для пошуку. Під час мандрів по списку слів корисно використовувати кнопку [M] та [V] для повторного повернення до важливого слова. Повернутися до поточного елемента після мандрів по списках можна відповідно кнопками [A], [W], [C].

Процес об'єднання. Обрана поточна корпорація встановлюється кнопкою [PECR] (рецептор) у списку корпорацій (Clist) та підсвічується кольором. Користувач розшукує споріднені найменування для поточної корпорації в списках найменувань (Alist) або корпорацій (Clist), аналізує та вибирає рішення на користь приєднання знайденого найменування до поточної корпорації. Знайдене найменування помічається кнопкою [NewC] як претендент на приєднання до поточної корпорації. Помічати можна послідовно декілька найменувань. Повторне натискання кнопки [NewC] видаляє найменування з претендентів на приєднання. Переглянути склад поточної корпорації та зібрані претендентів на приєднання до неї можна кнопкою [VReC], яка запускає процес побудови списку (VReC). Можна використати у цьому списку кнопку [DelC] для видалення найменування з претендентів на приєднання. Після аналізу та прийняття рішення можна запустити процес приєднання у базі даних кнопкою [AppC]. Після закінчення процесу список оновлюється з підсвіченими приєднаними найменуваннями організації. Після процесу об'єднання поточну корпорацію можна помітити як оброблену кнопкою [PROL] та внесену до тематичного списку кнопкою [TEMA].

Особливості пошуку та аналізу у списку слів (Wlist). Елемент списку має порядковий номер, слово, кнопки керування та нумерований внутрішній список корпорацій з числом найменувань, що мають поточне слово та входять до корпорацій, а також посиланням до ідентифікаційного номера у списку корпорацій (Clist). Кнопкою [NexW] встановлюється поточне слово для аналізу. Для швидкого повернення до слова використовується кнопка [W]. Послідовно обираючи посилання із внутрішнього списку посилань, користувач переглядає та розшукує споріднені між найменувань обраної корпорації у списку корпорацій (Clist). У разі необхідності уточнення за допомогою посилання (Alist) користувач встановлює позицію вказівника на поточний елемент у алфавітному списку найменувань (Alist) і за допомогою посилання <<слово>> послідовно переглядає внутрішні списки для всіх обраних ним слів з найменування. Під час цієї подорожі по списку слів позицію окремого слова можна запам'ятати кнопкою [M] для подальшого швидкого повернення до нього кнопкою [V] та перегляду його корпорацій. Після знаходження споріднених найменувань кнопкою [PECR] починається процес об'єднання, описаний вище.

Використання кнопки [LisW]. Перед використанням треба встановити режими фільтрації (List Short) перебудови для списку слів:

- «all» – включити у список всі слова, що мають у внутрішньому списку одну або більше корпорацій;

- «one» – включити до списку тільки слова, що мають у внутрішньому списку одну корпорацію;

- «som» – включити до списку тільки слова, що мають у внутрішньому списку більше одної корпорації;

- «af» – включити до списку тільки слова, що мають у внутрішньому списку корпорацію, яка не помічена як оброблена (фільтр оброблених);

- «nf» – не фільтрувати корпорацію, що не помічена як оброблена.

Різні поєднання режимів (List Short) дають змогу побудувати декілька варіантів скорочених списків слів і не переглядати повторно вже оброблені та підсвічені червоним корпорації. Корпорація, що була складена з одного найменування, приєданого до іншої корпорації, залишається у списку, але не має числа найменувань після знака «=>» і не потребує перегляду. Рекомендуємо спочатку перебудувати список у режимі «som» та «af», щоб отримати список слів, які мають у внутрішньому списку більше одної корпорації, переглянути його, провести можливі приєднання та відмітити оброблені. Після приєднань деякі корпорації зникнуть внаслідок переходу найменувань до інших, де зросте кількість найменувань. Деякі корпорації почнуть відображатись у списку слів у режимі «one». Повторна побудова списку слів у режимі «som» та «af» відобразить проведені зміни. Побудова списку слів у режимі «one» включити до списку тільки слова, що мають у внутрішньому списку одну корпорацію. Користуючись цим списком, можна швидко відмітити опрацьовані корпорації.

Повернення до повного списку кнопкою [LisW] у режимі «all» та «nf».

Особливості пошуку та аналізу в алфавітному списку найменувань (Alist). Елементи списку мають посилання до інших списків:

- посилання <Clist> до ідентифікаційного номера найменування в списку корпорацій (Clist);

- посилання <GO> до ідентифікаційного номера корпорації у списку корпорацій (Clist);

- посиланням <<слово>> від кожного слова назви до відповідного елемента у списку слів з найменувань (Wlist).

Сортування списку проводиться особливим способом: як сортувальний ключ використовується найменування організації, але якщо в найменуванні присутня крапка, то частина найменування від правого краю тексту до першої крапки додається до ключа з лівого краю. Завдяки цьому під час сортування у списку відбувається групування найменувань структурних

підрозділів організації, які у процесі каталогізації відділялися від основної частини назви крапкою. Таке групування дозволяє зібрати в одному місці списку близькі найменування підрозділу та головної організації для з'ясування приналежності до корпорації.

Наприклад:

14. <Clist>Сумський держ. аграрний ун-т. Кафедра іноземних мов
CORP_06358 <GO>.

15. <Clist>Сумський держ. аграрний ун-т. Кафедра кібернетики
CORP_13457 <GO>.

16. <Clist>Сумський національний аграрний ун-т. Кафедра
кібернетики CORP_13457 <GO>.

17. <Clist>Сумський аграрний ун-т. Кафедра кібернетики та
інформатики CORP_13457 <GO>.

18. <Clist>Сумський національний аграрний ун-т. Кафедра
кібернетики та інформатики CORP_13457 <GO>.

19. <Clist>Сумський національний аграрний ун-т. Кафедра філософії
та соціології CORP_28733 <GO>.

Використання посилання від кожного слова найменування до списку (Wlist) дозволяє швидко переглядати та зіставляти склад корпорацій для знаходження розбіжностей між ними. Серед розбіжностей можуть бути споріднені найменування, які треба переглянути в списку корпорацій (Clist) та, можливо, приєднати. Склад корпорації можна також швидко переглянути, використовуючи посилання <GO> або <Clist>, яке встановлює вказівник на поточне найменування у складі корпорації.

Особливості пошуку та аналізу в списку корпорацій (Clist).

Сортування списку проводиться за ідентифікаційним номером корпорації, а склад корпорації сортований за роком видання документа. Окремі слова найменування, через які корпорація потрапила до вибірки, підсвічені синім кольором. Список показує користувачу склад корпорації та призначений для пошуку споріднених (та, можливо, чужих) найменувань і відмітки претендентів на приєднання. Останні підсвічуються жовтим кольором кнопкою [NewC]. Претендента можна додатково подивитися в алфавітному списку найменувань через посилання <Alist>, а вже звідти перейти до списку слів (Wlist) та діяти, як описано вище (рис. 2).

Приклад екрану перед об'єднанням корпорацій



Рис. 2. Екран перед об'єднанням корпорацій

Після закінчення об'єднання найменувань у корпорацію в поточній вибірці користувач може побудувати контрольні списки для неї. Кнопкою [PROC] запускаємо процес побудови. Програма відкриває в окремих вікнах браузера список складу «Корпорація 12345» та список усіх слів з найменувань тільки обраної корпорації «Ключі 12345». Вибираючи тільки характерні слова кнопкою [KeyW], складаємо список ключів, за якими буде зроблена вибірка. Кожен ключ кнопками [CatR], [CatL] може відсікатись справа або зліва для ліквідації граматичної залежності та формування пошукового терміна. Ці скорочення можна відновити кнопками [ResR] та [ResL].

Керування фільтрами [Тема] та [Proc]. Кнопка [Тема] установки режиму роботи фільтра приналежності корпорації до тематичних списків дозволяє циклічно встановлювати такі режими:

- «ON» – додати до складу вибірки тільки корпорації, що не мають відмітки поточної теми;
- «AL» – додати до складу вибірки тільки корпорації, що не мають відмітки будь-якої зі вказаних тем;
- «OFF» – ігнорувати наявність відмітки.

Кнопка [Proc] установки режиму роботи фільтра приналежності корпорації до списку оброблених дозволяє циклічно встановлювати такі режими:

- «ON» – додати до складу вибірки тільки корпорації, що не мають відмітки обробленої;

- «OFF» – ігнорувати наявність відмітки.

Особливістю цього списку є наявність кнопок [G], [S] встановлення рівня шуму від 0 до 9. До складу вибірки попадають тільки ті корпорації, де кількість входжень зі встановленого списку ключів до слів із будь-якого найменування перевищує встановлений рівень або є хоча б одне входження будь-якого ключа, поміченого кнопкою [NotW]. Кнопкою [NotW] можна встановити ознаку «не шумові» для обраних ключів у списку перед побудовою вибірки. Різні поєднання фільтрів, рівня шуму та ознака «не шумові» дозволяє зменшувати кількість найменувань у вибірці та необхідність вичитувати та аналізувати їх. Кнопкою [LIST] запускається побудова нової вибірки в новому вікні.

Список історії вибірок (Hisel). Елементи списку містять номер папки, адже з папки починалась побудова вибірки. Якщо з однієї вибірки будували декілька «дочірніх» вибірок, усі вони будуть відображені нижче, а «батьківські» – вище у списку історії. Цей список потрібен для нагадування користувачу всіх пошукових термінів та номера папки знаходження вибірки, використаних під час сеансу роботи. Список історії будується як новий під час побудови списку вибірки (Tlist).

Представлену технологію було застосовано під час опрацювання понад 300 найменувань організацій та установ Національної академії наук України. Всі записи додані до баз даних електронного каталогу для подальшої обробки та використання.

Список використаних джерел

1. Антоненко И. П. Авторитетные записи и функциональный поход к их созданию / И. П. Антоненко // Библиотеки национальных академий наук: проблемы функционирования, тенденций развития : науч.-практ. и теорет. сб. / Междунар. асоц. акад. наук, Совет директоров науч. б-к и информ. центров. – К., 2006. – Вып. 4. – С. 197–211.

2. Лобузінa К. Технології організації знанневих ресурсів у бібліотечно-інформаційній діяльності : монографія / Катерина Лобузінa ; відп. ред. О. С. Онищенко ; НАН України, Нац. б-ка України ім. В. І. Вернадського. – К., 2012. – 252 с.

3. Лутц М. Изучаем Python : пер. с англ. / М. Лутц. – 3-е изд. – СПб. : Символ-Плюс, 2009. – 848 с.

4. Саммерфильд М. Программирование на Питон 3 : подробное руководство : пер. с англ. / М. Саммерфильд. – СПб. : Символ-Плюс, 2009. – 608 с.

5. Система автоматизации библиотек ИРБИС64. Общее описание системы / ГПНТБ России. – М., 2004. – 259 с.

6. RUSMARC в примерах : учеб. пособие для каталогизаторов. Ч. 3. Авторитетные записи / Нац. информ.-библ. центр «ЛИБНЕТ». – М. : ФАИР-ПРЕСС : Центр «ЛИБНЕТ», 2003. – (Специальный издательский проект для библиотек). – 2005. – 496 с.

7. UNIMARC / Authorities. Международный коммуникативный формат UNIMARC для авторитетных/нормативных записей. – М. : ГПНТБ : РГБ : РКП, 1994. – 100 с.

Голубев Ю. Автоматизована технологія формування інформаційного масиву найменувань організації для авторитетних записів.

Ключові слова: авторитетні записи, найменування організації, автоматизована технологія, формування інформаційного масиву.

Технологія дозволяє каталогізатору знайти в базі даних всі форми найменування організації або підрозділу та об'єднати в групу під одним ідентифікаційним номером. Об'єднані найменування організації складають інформаційний масив, який має умовну назву *корпорація*.

Технологія основана на автоматичній побудові гіпертекстових списків із кнопками керування у формах, монтованих на мові гіпертекстової розмітки (HTML), посиланнями між елементами списків. Програми під час формування списків виконують добірку записів з бази даних, необхідні фільтрації та сортування, побудову посилань і кнопок, групування елементів та їх підсвічування. Програми формують три основні списки: найменувань організацій, склад корпорацій та слова, виділені з найменувань. Допоміжні списки призначені для побудови вибірок та маніпулювання основними списками.

Перший список відображає всі найменування, з яких складається вибірка та сортований за алфавітом. Кожний елемент списку має посилання до себе в складі корпорації або до ідентифікаційного номеру корпорації у списку корпорацій, посилання від кожного слова найменування до відповідного елементу у списку слів з найменувань.

Список складу корпорацій сортований за ідентифікаційним номером та роком видання документа. Кожний елемент списку корпорації має посилання на свій ідентифікаційний номер в алфавітному списку.

Список слів відображає слова, виділені з кожного найменування, з яких складається вибірка. Він сортований за алфавітом та номером корпорації, до складу

якої входить найменування з виділенням словом. Кожне слово має свій список посилань до відповідного ідентифікаційного номеру у списку корпорацій.

Для елементів списків є ряд кнопок керування у процесі пошуку та обробки споріднених найменувань, відмітки в базі даних властивостей корпорації, організації списку пошукових термінів для наступної вибірки.

За допомогою кнопок керування і посилань у основних та допоміжних списках користувач призначає поточну корпорацію, розшукує споріднені найменування, маніпулюючи списками, виконує приєднання знайдених найменувань корпорації до поточної.

Golubev Yu. An automated technology of the organizations names information array creation for authority records.

Keywords: authority records, organization name, automated technology, information array.

The technology allows the cataloguer of finding in the database all the forms of the organization name or the unit name, and of grouping them under a single identification number. The combined organization names arrange an information array, which calls conventionally «corporation».

The technology is based on the automatic construction of the hypertext lists with control buttons in the forms written in the hypertext markup language (HTML) and the links between the elements of the lists. The programs, during the lists producing, should discharge the selection of records from the database, the necessary filtering and sorting, links and buttons building, elements grouping and illumination implementing. The programs form the three main lists: organizations names, corporations structure and the words marked out from the names. The auxiliary lists have been intended to set selections and manipulate the main lists.

The first list can map all the names setting the selection and can be sorted alphabetically. Each element from the list forms the link to itself in the structure of the corporation or to the identification number of the corporation in the list of corporations. The links are generated from each word of the name to a suitable element in the list of words composed of names.

The list with the corporations structure has sorted by ID number and the year of the document publication. Each element of the corporation list forms the link to its identification number in the alphabetical list.

The list of words can map the words marked out from each name setting the selection. It has been sorted alphabetically and by the number of corporation, which includes the name with the selected word. Each word has its own link list to a suitable identification number in the corporations list.

For the elements of the lists there are a number of control buttons within a framework of searching and processing of the related names, the notes in the corporation property database, the organizations of the searching terms list for the next selection.

With the help of the control buttons and the links in the primary and the auxiliary lists user can assign the current corporation, manipulating the lists he can search the related names and connect the found names to the current.

Голубев Ю. Автоматизированная технология формирования информационного массива наименований организации для авторитетных записей.

Ключевые слова: авторитетные записи, наименование организации, автоматизированная технология, формирование информационного массива.

Технология позволяет каталогизатору найти в базе данных все формы наименования организации или подразделения и объединить в группу под одним идентификационным номером. Объединенные наименования организации составляют информационный массив, который имеет условное название «корпорация».

Технология основана на автоматическом построении гипертекстовых списков с кнопками управления в формах, монтируемых на языке гипертекстовой разметки (HTML), и ссылками между элементами списков. Программы при формировании списков выполняют подборку записей из базы данных, необходимые фильтрации и сортировки, построение ссылок и кнопок, группировки элементов и их подсветку. Программы формируют три основных списка: наименования организаций, состав корпораций и слова выделенные из наименований. Вспомогательные списки предназначены для построения выборки и манипулирования основными списками.

Первый список отображает все наименования, составляющие выборку, и сортирован по алфавиту. Каждый элемент списка имеет ссылку к себе в составе корпорации или к идентификационному номеру корпорации в списке корпораций, ссылки от каждого слова наименования к соответствующему элементу в списке слов из наименований.

Список состава корпораций сортирован по идентификационному номеру и году издания документа. Каждый элемент списка корпорации имеет ссылку на свой идентификационный номер в алфавитном списке.

Список слов отображает слова, выделенные из каждого наименования, составляющих выборку. Он сортирован по алфавиту и номеру корпорации в состав которой входит наименование с выделенным словом. Каждое слово имеет свой список ссылок к соответствующему идентификационному номеру в списке корпораций.

Для элементов списков есть ряд кнопок управления в процессе поиска и обработки родственных наименований, отметки в базе данных свойств корпорации, организации списка поисковых терминов для следующей выборки.

С помощью кнопок управления и ссылок в основных и вспомогательных списках пользователь назначает текущую корпорацию, разыскивает родственные наименования, манипулируя списками, и выполняет присоединение найденных к текущей корпорации.