

УДК 025.4; 004.738.5:02; 004.78:025.4.03

І. В. Лобузін

Національна бібліотека України імені В.І. Вернадського
Проспект 40-річчя Жовтня, 3, 03039 Київ, Україна

Створення, представлення та перспективи використання електронного фонду цифрових копій документів бібліотеки

Розглянуто проблеми формування та впорядкування електронного фонду цифрових копій документів бібліотеки. Викладено основні науково-методичні засади створення електронного архіву цифрових копій документів. Детально проаналізовано структуру та підсистеми цифрового фонду наукової бібліотеки: поповнення, зберігання, облік, представлення, пошук, захист і розділення доступу до інформації. Обґрунтовано необхідність і доцільність інтеграції створеного електронного архіву з автоматизованою бібліотечною інформаційною системою.

Ключові слова: електронний архів, цифровий фонд, електронний документ, автоматизована бібліотечна інформаційна система, цифрова копія документа.

Вступ

У сучасному світі відбуваються процеси конвергенції технологій, що ведуть за собою великі зміни у ході розвитку суспільства. Змінюється природа інформації та форми її подання. Ці процеси значною мірою торкаються питань збереження світової культурної спадщини як основи всієї людської цивілізації в цілому. Як зазначили підсумки міжнародної конференції «EVA» (Електронні зображення та візуальні мистецтва), що відбулась у 2002 р. під керівництвом ЮНЕСКО [10]: в Україні, незважаючи на труднощі та проблеми, процеси електронної готовності, інформатизації, становлення і розвитку інформаційного суспільства в останні роки стійко прогресують. Створюються розподілені бази даних і знань, удосконалюється телекомунікаційна інфраструктура, в різноманітних сферах (науки, культури, освіти) формується середовище цифрової спадщини. Створюються мережі інформаційних веб-порталів; розвиваються електронні бібліотеки, музеї, архіви та експозиції; удосконалюється система використання світових інформаційних ресурсів для надання високоякісних і високоефективних освітніх та інших послуг. Однак, слід зауважити, що в Україні не створена цілісна система нормативно-мето-

дичного забезпечення процесів архівного зберігання електронних документів, гармонізована з нормами міжнародного права з питань розвитку інформаційного суспільства. Тому єдиних правил щодо ведення електронних архівів відомчих державних підприємств та установ немає. Електронні архіви створюються залежно від вирішуваних архівом задач зі збереження даних та економічних можливостей установи [11].

Світове співтовариство приділяє велику увагу збереженню культурної спадщини. За ініціативою ЮНЕСКО підготовлено проект «Хартії про збереження цифрової спадщини», який є крупною міжнародною ініціативою, що має значний інтерес для усіх країн світової спільноти. За своєю природою Хартія — потужний каталізатор, що активно сприяє вирішенню однієї з найважливіших задач діяльності ЮНЕСКО — розвитку культури у світі та взаєморозуміння між країнами та народами.

Наслідком здійснення різноманітних програм і проектів з оцифрування історико-культурної спадщини стало усвідомлення того, що необхідна кооперативна взаємодія на національному та міждержавному рівнях для створення повноцінних цифрових колекцій, що будуть відображати як національну самобутність окремих регіонів, так і все розмаїття світового наукового й культурного досвіду. Така кооперація зусиль допоможе віртуальній реконструкції розпорошених між різними державами колекцій, буде запобігати зайвому дублюванню інформації, надасть єдину точку доступу користувачам до необхідної їм інформації. Усвідомлення всіх цих реалій призвело до розгортання міжнародних проектів зі збереження та надання доступу до об'єктів національного надбання: Європейська Електронна Бібліотека (The European Library, www.theeuropeanlibrary.org), Європеана (Europeana, <http://www.europeana.eu>); Світова Цифрова Бібліотека (World Digital Library, <http://www.wdl.org>); Золота Колекція Євразії (Проект «Золотая коллекция Евразии», <http://bae.rsl.ru/programs/golden-collection>). Метою всіх цих проектів є представлення на єдиній технологічній основі всього етнічного та національного розмаїття світової культури.

Як справедливо зазначає директор Інституту корпоративних бібліотечно-інформаційних систем Санкт-Петербурзького державного політехнічного університету Н.В. Соколова (2010), проекти створення цифрових бібліотек (електронних бібліотек оцифрованих документів) можна умовно поділити на два основні напрями: книжкові проекти в Інтернеті та цифрові проекти саме бібліотек. Найбільш відомим книжковим проектом Інтернету є «GOOGLE Книги». Свого часу запуск цього проекту в Інтернеті у 2004 р. справив шокове враження на бібліотечних спеціалістів, однак час виявив, що бібліотеки та GOOGLE можуть співіснувати та співпрацювати. Основна місія цього проекту була сформульована як необхідність систематизувати світову книжкову інформацію та зробити її загальнокорисною, засоби цього упорядкування базуються на бібліотечних підходах. Бібліотеки привнесли зі свого боку в електронне інформаційне середовище свої сильні риси: якісну каталогізацію ресурсів, стандарти опису та систематизації документів, турботу про збереження цифрового твору. Бібліотеки представили світу якісно упорядковані та описані електронні колекції, що пов'язані зі збереженням культурної спадщини та наданням доступу до неї широких верств населення.

Національна бібліотека України імені В.І. Вернадського (НБУВ) також активно приймає участь у міжнародних і міжнаціональних процесах створення історико-культурної та наукової цифрової спадщини. У НБУВ у 2010 р. створено групу з оцифрування документів на традиційних носіях, перед якою було поставлено два основні завдання: створення страхового фонду документів із фондів НБУВ і створення електронного фонду користування (переважно на основі укладання та організації цифрових колекцій). З давніх часів людству доводилося вирішувати проблему зберігання інформації, накопичених багатьма поколіннями знань, культурних, технічних, наукових та інших цінностей. Вирішення цих проблем тісно пов'язано з технологіями зберігання, обліку, розмежування доступу, введенням в обіг нових одиниць зберігання, створенням нових знань на основі вже наявних. Отже, природно, що під час організації роботи зі створення фонду цифрових копій документів НБУВ постало ціле коло питань, яке перекликається з тими основними проблемами, що пов'язані зі створенням електронних архівів і бібліотек, а саме:

- організація поповнення інформації;
- організація зберігання інформації;
- організація обліку інформації;
- організація представлення інформації та системи пошуку;
- організація системи захисту та розділення доступу до інформації;
- створення на основі перерахованих пунктів єдиної інформаційної системи.

З точки зору вдосконалення програмно-технологічних рішень для проектів оцифрування книжкових фондів сьогодні пропонуються комплексні рішення провідними розробниками світового ринку інформаційних послуг, такими як АBBY, XEROX, CANON. Рішення компаній XEROX, CANON пропонують комплексну автоматизацію процесів створення цифрових архівів: від постачання обладнання до програмно-технологічних комплексів. Технологічний комплекс передбачає: ретроконверсію довідково-пошукового апарату бібліотеки (сканування й розпізнавання каталожних карток та описів фондів); виготовлення цифрових копій паперових документів; підготовку електронних версій документів; введення та редагування метаданих; формування електронних колекцій; онлайн або локальну публікацію створеного цифрового контенту; виготовлення поліграфічних аналогів унікальних паперових документів, що максимально наближені до оригіналу; підтримку пошукового апарату колекцій і системи захисту інформації від розповсюдження та копіювання. Компанія XEROX з успіхом працює на бібліотечному ринку Росії, сьогодні вона підписала угоду з ДПНТБ Росії з метою розробки нового програмного продукту, який спрямований на створення електронних архівів і бібліотек вузів і наукових установ [25]. Рішення компанії АBBY зорієнтовані, у першу чергу, на оптичне розпізнавання символів (Optical Character Recognition — OCR). Програмне забезпечення компанії АBBY для розпізнавання старовинних текстів відіграє важливу роль у міжнародному проекті ІМРАСТ (ІМProving АCcess to Text, «покращення доступу до тексту») [19].

Розглянемо більш детально стан проблеми та рішення, що були прийняті в НБУВ стосовно окресленого кола питань і організації основних підсистем інформаційної системи цифрового фонду НБУВ.

Організація поповнення інформації

Для цифрових архівів поповнення інформації відбувається за рахунок переведення паперових документів у електронний вигляд. Для цього використовуються, залежно від типу та формату документів, сканери, цифрове фотообладнання та програмне забезпечення цифрової обробки зображень. Сканери, які пропонуються на сучасному ринку комп'ютерної техніки, дуже відрізняються за своїми характеристиками: вузькоформатні (до формату А3 включно); широкоформатні сканери («більше» формату А3); контактні та безконтактні сканери; фотосканери — обладнані цифровими камерами; швидкісні поточкові сканери — для масового опрацювання картотек та офісних документів; спеціалізовані книжкові сканери; роботизовані сканери — з автоматичним перегортанням сторінок. Бібліотеки, які у своїх фондах мають різноформатні документи, що відрізняються за багатьма характеристиками: матеріальною основою (зброшуровані, неконтрастні, пошкоджені, крихкі), хронологічними межами, фізичними розмірами, знаковою основою (текст, ноти, карти, рукопис, образотворчі матеріали), стоять перед непростим вибором під час придбання обладнання для оцифрування фондів — практично їм можуть знадобитися сканери будь-якого типу. Одним із вдалих прикладів «все в одному», який враховує практично будь-які особливості бібліотечних документів і має адаптоване для автоматичного опрацювання зображень програмне забезпечення є сучасний планетарний сканер фірми Zeutschel [12]. Сканери компанії Zeutschel мають у підрозділах електронної обробки фондів Національна бібліотека Білорусі [15] та Російська державна бібліотека [7], придбання яких значно розширило спектр оцифрованих документів. Для масового користувача сьогодні пропонується ручний книжковий сканер «Book Saver Book Scanner» компанії Ion Audio, який дозволяє за 15 хвилин оцифрувати книгу у 200 сторінок та перетворити її у формат, зручний для перегляду будь-яким е-рідером [6]. Однак за своїми ціннісними характеристиками таке обладнання виявляється мало доступним для бібліотек України.

У НБУВ для цифрового опрацювання фондів створено два спеціалізованих комп'ютерних комплекси. Перший комплекс включає: сканер книжковий АТІЗ, сканер А4 Mustek (вузькоформатний), сканер А3 Mustek (багатокольоровий), мікрофільмуючу камеру «Senator»; другий комплекс має таке обладнання: сканер книжковий АТІЗ, книжковий скануючий комплект SMA21, сканер А2 Minolta (чорно-білий), сканер А4 Epson (вузькоформатний), установку «Докуматор», цифровий фотоапарат Olympus. Наявне обладнання дозволяє створити цифрові копії практично будь-яких документів із фондів НБУВ, але потребує значного постопрацювання цифрових зображень, особливо це стосується зображень, що отримані за допомогою цифрових фотокамер (книжкового сканера АТІЗ і фотоапарату Olympus). У спеціалізованих фондах НБУВ (рукописному, стародруків і рідкісних видань, історичних зібрань і бібліотечних колекцій) зберігаються раритетні документи, копії яких неможливо отримати іншим способом ніж фотографування (широка та жорстка оправа, крихкий папір тощо), для цього використовуються модернізовані мікрофільмуючі установки («Senator» та «Докуматор»), які пристосовані для цифрової фотографії (особливо цінними є система освітлення та позиціонування документів, книжкові опори різної конфігурації). Цифрове обладнання

забезпечено програмно-апаратним комплексом комп'ютерної техніки, на базі якого створено автоматизовані робочі місця спеціалістів з оцифрування документів. Програмне забезпечення включає комплекс засобів мережевої взаємодії та цифрової обробки зображень: FileZila (FTP-сервер), Photoshop (програма для професійного опрацювання зображень), Book Drive Capture (програмне забезпечення сканера ATIZ), Scan Kromsator та XnView (універсальні програми для пакетного опрацювання зображень).

Остаточні опрацьовані цифрові копії документів з робочих місць фотографів за допомогою FTP-сервера передаються до файлового сервера, де перевіряються відповідальною особою на повноту і якість, описуються та обліковуються.

Введення даних у систему відбувається у двох основних видах: поповнення електронного страхового фонду НБУВ і формування фонду користування. Для страхового фонду документи зберігаються у вигляді зображень високої роздільної здатності, а до фонду користування вони надходять опрацьованими у вигляді електронних версій документів. Електронні версії документів виготовляються різними способами, у різному дизайні та форматі, залежно від електронної колекції, до якої вони належать. Це може бути формат PDF для сучасних наукових видань, HTML-формат для швидкого перегляду через веб-інтерфейс, FLASH-формат із гарним оформленням, перегортанням сторінок і засобами захисту від копіювання для обслуговування користувачів у віртуальному читальному залі особливо цінними та популярними документами.

Організація зберігання інформації

Досвід створення електронних архівів свідчить про те, що обсяг архіву буде збільшуватись і швидко досягне терабайтних розмірів [11, 16], проте в інтенсивному оперативному доступі знаходиться лише 5–10 % інформації. Отже, підсистема зберігання інформації повинна мати щонайменше два основних розділи: область оперативно доступної інформації з високою інтенсивністю використання та область довготривалого зберігання з високими вимогами до надійності, цілісності та безпеки даних.

Оперативний архів реалізують розміщенням на сервері у спеціальному розділі жорсткого диска (для невеликих обсягів інформації), на окремому жорсткому диску або RAID-масиві. Для підвищення надійності функціонування подібних систем використовують засоби резервного копіювання, дублювання інформації та «дзеркальних дисків». RAID-масиви також використовують для зберігання метаданих та облікової бази даних архіву. За рекомендацією Національного архіву Сполучених Штатів Америки [U.S. National Archives] ці дані мають назву Master References File (еталонний файл) [23]. Найбільш перспективним рішенням, що одночасно задовольняє вимогам швидкості та надійності, є RAID-5 або -6.

Для організації області довгострокового зберігання використовуються змінні носії інформації CD/DVD/UDO та роботизовані бібліотеки на їхній основі [8, 11, 16]. Найбільш досконалі рішення підтримують надщільний оптичний формат (Ultra Density Optical — UDO) та технології Blue-Ray. Популярність технології UDO ґрунтується на можливості записати на один носій 30–60 Гбайт інформації з терміном зберігання інформації 50 років. Широку перспективу також має техно-

логія Blue-Ray, прийнята в якості кращого архівного носія. Довговічність BD (Blue-Ray Disk) (більш 50 років), їхня ємність (50 ГБ, у перспективі до 200 ГБ) і швидкісні характеристики (від 9 МБ/с) дозволяють цим дискам забезпечити створення унікальних за конкурентоздатністю систем архівування. Дисководи Blue-Ray сумісні за читанням/записом з носіями CD/DVD і підтримуються файловою системою UDF.

Роботизована бібліотека є наступним етапом створення електронного архіву довгострокового зберігання, вона представляє собою масив змінних дисків, що розміщений у спеціальному корпусі. Усередині такої дискової бібліотеки розміщені приводи, що забезпечують подачу необхідного диска до слоту для зчитування інформації. Сама роботизована бібліотека є SCSI-пристроєм, який підключається до комп'ютера, де сприймається операційною системою як єдиний логічний ресурс. Рішення на основі роботизованої бібліотеки зі змінними носіями інформації також дозволяє досить легко вирішити питання технічного захисту інформації і забезпечити її надійне зберігання [8].

З точки зору зберігання самих файлів електронних документів є також два основних підходи [16]: за умов використання першого з них електронні документи зберігаються у бінарному вигляді у відповідних полях самої бази даних архіву, другий підхід базується на тому, що файли можуть зберігатися окремо, а в базі даних є тільки електронні адреси файлів (вказується шлях доступу до файлу). Перший із способів швидко призводить до цілої низки проблем: ускладнення резервування даних, перехід до іншої інформаційної платформи, зниження швидкодії системи. Він може бути рекомендований лише для архівів з невеликим фізичним обсягом інформації.

Крім надійності та продуктивності електронного архіву або сховища даних важливою також є логічна структура сховища [11]. За базову структуру рекомендується прийняти ієрархію папок з відповідними документами, яка буде відображати логічний зв'язок основних вузлів системи. Така організація, сприяє інтуїтивно зрозумілій навігації: користувач може легко зорієнтуватися де можна шукати необхідний документ, полегшується інтеграція даних електронного архіву зі спеціалізованим програмним забезпеченням з опрацювання даних.

У НБУВ рішення щодо зберігання інформації цифрового фонду були обґрунтовані поставленими завданнями: створення страхового фонду документів та оперативного фонду користування.

Для формування страхового фонду використовуються UDO-диски, на які переносяться цифрові копії високої роздільної здатності. Надалі передбачається налагодження та підключення роботизованої бібліотеки на основі системи автоматичного накопичення та пошуку інформації Plasmon, яка дозволяє керувати створеним UDO-архівом й отримувати доступ до необхідних документів.

Для формування оперативного електронного фонду користування виготовляються електронні версії документів зниженої роздільної здатності, яка достатня для ознайомлення зі змістом документа, його читання та перегляду. Електронні версії документів каталогізуються (атрибутується їхня належність до певних колекцій) та розміщуються на сервері у спеціальному розділі жорсткого диска, де

формується електронний архів. Відповідно до визначених правил регламентується доступ до документів: службовий, локальний або онлайнний.

Організація обліку інформації

Необхідність обліку інформації не викликає сумнівів, правильно організований облік значно полегшує процеси доступу та пошуку інформації, запобігає її втраті та унеможливує повторне виготовлення електронних аналогів уже опрацьованих паперових документів. Організація обліку потребує вирішення двох першочергових питань: вибору інформаційної платформи та вибору структури метаданих. Розглянемо більш детально ці основні завдання.

Системи обліку інформації для електронних архівів використовують системи керування базами даних, які представлені на сучасному ринку цілим спектром популярних програмних продуктів від універсальних рішень (Access, MsSQL, Oracle) до спеціалізованих платформ (QStar HSM, Saperion) — класу ECM (Enterprise Content Management), що призначені для керування електронним документообігом підприємства. Як показує практика експлуатації програмного забезпечення для створення електронних архівів [11, 14, 16], на певному етапі перед підприємством постає питання інтеграції створеного архіву з інформаційною системою, яка прийнята для керування документами. Тому передбачити цю ситуацію слід уже на перших етапах проектування цифрового фонду.

Для організації обліку створених цифрових копій у НБУВ зупинилися на платформі, яка є сьогодні базовою основою інформаційних сервісів НБУВ — системі автоматизації бібліотек (САБ) «ІРБІС64» [4]. Таке рішення організації обліку на основі автоматизованої бібліотечної інформаційної системи (АБІС) мало безперечні переваги та було оптимальним в умовах бібліотеки: система має можливість для розробки власних прикладних рішень; сучасні АБІС працюють на основі клієнт-серверної архітектури, що дає змогу організувати мережеву взаємодію користувачів багатьох різних підрозділів; для опису документів (стародруки, рукописи, ноти, карти, образотворчі матеріали) необхідно залучати фахових спеціалістів різних структурних підрозділів; для великої кількості документів уже створено записи в електронному каталозі та спеціалізованих базах даних; система підтримує зберігання цифрових об'єктів у бінарному вигляді та гіпертекстові посилання до електронних документів; створений електронний фонд легко інтегрувати з іншими інформаційними сервісами бібліотеки.

Ідея спільного проекту Херох і ДПНТБ Росії народилась у результаті успішної співпраці компаній під час створення Наукової електронної бібліотеки ДПНТБ Росії: Херох реалізувала в бібліотеці проект «Електронний архів науково-технічної рідкісної книги» (<http://www.vlibrary.ru/>) на базі промислової платформи для комплексного управління документами і даними Херох «DocuShare» та програмного забезпечення ДПНТДБ Росії «ІРБІС». У 2009 році було запущено електронне сховище і забезпечено доступ до унікальних видань з віртуальних читальних залів і мережі Інтернет. Зараз компанії працюють над створенням інтегрованого рішення «ІРБІС» + «DocuShare» під назвою «БІАР», яке має забезпечити комплексну автоматизацію електронних бібліотек і архівів [25].

Стосовно метаданих на сьогодні є два найбільш значимі підходи до опису електронних інформаційних ресурсів: розробка ІФЛА «Міжнародний стандартний бібліографічний опис для електронних ресурсів» — International Standard Bibliographic Description for Electronic Resources ISBD (ER) [1] і проект метаданих «Дублінське ядро» (Dublin Core), ініціатором якого є OCLC, точніша назва DCMІ (Dublin Core Metadata Initiative — Ініціатива Дублінського ядра метаданих) [21].

Стандарт DCMІ для опису електронних інформаційних ресурсів є оптимальним з погляду на простоту його застосування й надання можливості побудови метаописів необхідної повноти. Стандарт містить 15 метаданих із відповідними множинами кваліфікаторів і пропонує правила опису загальних характеристик електронних інформаційних ресурсів, що відбивають вміст ресурсу (Title — Назва, Subject — Тема, Description — Опис, Type — Тип, Source — Джерело, Relation — Відношення), інтелектуальну власність (Creator — Створювач, Publisher — Видавець, Contributor — Співавтор, Rights management — Права, Coverage — Охоплення) та його стан (Date — Дата, Format — Формат, Identifier — Ідентифікатор, Language — Мова) [5, 13].

Бібліографічний стандарт ISBD відповідає основним правилам укладання бібліографічних записів документів у бібліотеках UNIMARC [24], є більш складним і менш формалізованим. Однак, більшість документів у бібліотеці описано саме за цими правилами й тому для бібліотечних документів природнім є обрання саме цього стандарту. Крім того, можна побачити, що в основній описовій частині ці два стандарти мають спільні елементи: автори, назва, тема (предметна рубрика), мова, дата та місце видання (створення), опис (анотація).

Провідні фахівці зі створення електронних архівів, такі як Національний архів Сполучених Штатів Америки (U.S. National Archives, www.archives.gov) [23], рекомендують під час формування фонду цифрових копій вводити не тільки описові метадані документів, а обов'язково передбачити технічні та правові метадані. Технічні метадані мають відображати основні характеристики зображення (колір, роздільну здатність, розмір, обсяг), обладнання та програмне забезпечення, засобами яких створено зображення та електронна версія документа. Правові метадані особливо важливі для документів архівного фонду, бо за правилами на копії розповсюджуються ті самі обмеження, що стосуються їхніх паперових аналогів.

У НБУВ за основу укладання метаданих цифрових об'єктів була обрана модель метаданих WDL — Світової Цифрової Бібліотеки (World Digital Library), яка розроблена на основі досвіду Бібліотеки Конгресу США в процесі створення цифрового порталу Пам'ять Америки (American Memory) [20]. Базовий комплект метаданих Світової Цифрової Бібліотеки, що заснований на Dublin Core (DC) форматі, включає такі елементи: назву, опис об'єкта, автора, видавця, місце видання, дату створення, мову, місце, тему, тип документа, додаткові рубрики, примітки, опис фізичних властивостей, назву колекції, опис колекції, організацію, адресу сайту. Можливості системи підтримки електронного каталогу НБУВ дозволяють конвертувати дані як у MARC-формат, так і DC-формат, експортувати їх у табличні форми, необхідні для реєстрації оцифрованих об'єктів національного надбання у світових цифрових бібліотеках.

З метою обліку проведених робіт зі створення цифрових копій паперових документів було розроблено на платформі САБ «ІРБІС64» спеціалізовану базу даних, яка передбачає ведення записів двох основних типів: «Замовлення» та «Страховий фонд».

Запис «Замовлення» дозволяє слідкувати за всіма етапами поповнення електронного фонду: від виготовлення зображень до запису на диск UDO, створення та публікації електронної версії видання. Основними структурними елементами запису є: *замовлення* (стан виконання, дати (початок – кінець), № диска UDO, підрозділ НБУВ або установа/організація; замовник, примітка, тип замовлення, текст замовлення, підстава виконання, призначення копії, найменування проекту або колекції, тип документа, обсяг сторінок); *опис* (стислий бібліографічний опис документа з необхідними ідентифікаторами: місце зберігання, фонд/колекція, шифр, інвентарний номер, посилання до електронної версії видання); *виконання замовлення* (обладнання, проведені роботи, дата виконання, виконавець); *обсяг робіт* (назва папки/файлу, кількість файлів, обсяг у байтах, формат файлів, роздільна здатність, колір, глибина кольору).

Запис «Страховий фонд» має на меті отримання опису документів страхового фонду, записаних на той чи інший диск UDO. Запис має дві частини: опис сторони *A* та сторони *B*. Основними структурними елементами запису є: № диска UDO, дата (початок – кінець), підрозділ НБУВ, колекція, проект, виконавці, опис документа (№ замовлення, ідентифікатори, стислий опис, назва папки/файлу, кількість файлів, обсяг у байтах, формат файлів, роздільна здатність, колір, глибина кольору). За введеними даними на екрані легко побачити обсяг накопиченої інформації та роздрукувати за необхідності обліковий акт.

Застосовані рішення щодо обліку робіт дали змогу оцінити продуктивність роботи створених цифрових комплексів, автоматизувати звітність, контролювати ефективність роботи обладнання та окремих працівників.

Організація представлення інформації та системи пошуку

Функціональні вимоги щодо організації пошуку та представлення інформації в електронному архіві великого розміру або сховищі даних (Data Warehouse) детально проаналізували R. Kimball (2000) [22] та Г. Асеев (2009) [2]. З точки зору користувача основними вимогами є: структура даних сховища має бути зрозумілою користувачам; повнота і достовірність збережених даних; підтримка внутрішньої несуперечності даних; зручні, прості механізми формування запитів; підтримка високої швидкості одержання даних зі сховища; підтримка складних запитів для великого обсягу даних (тисяч або мільйонів записів); наявність зручних утиліт перегляду даних у сховищі даних.

У НБУВ облік робіт з оцифрування документів ґрунтується на платформі САБ «ІРБІС64», яка має можливості для ефективного представлення бібліотечних інформаційних ресурсів будь-якої природи та обсягу, що дало змогу легко інтегрувати ресурси сформованого цифрового фонду в систему обслуговування НБУВ. Спеціальні модулі САБ «ІРБІС64» дають змогу надати читачам/користувачам бібліотеки доступ до електронних інформаційних ресурсів як у локальному режимі в читальних залах НБУВ, так і у віддаленому доступі через веб-інтерфейс. Інфор-

маційно-пошукова бібліотечна система у звичному інтерфейсі надає користувачу необхідні навігаційні засоби за стандартними атрибутами: автор, назва, ключові слова, дата видання (створення), предметні рубрики, колекції тощо. Після остаточного опрацювання — записування страхової копії та виготовлення електронної версії видання — електронний документ публікується у фонді користування: створюється опис документа (або експортується вже введений) з посиланням до повного тексту; сам документ розміщується на сервері у спеціальному розділі жорсткого диску; за необхідності виготовляється зображення обкладинки видання. Документи, які призначені для онлайнної публікації, розміщуються на веб-сервері НБУВ і засобами WEB ІРБІС представляються віддаленим користувачам. Записи електронного фонду наділені необхідним комплектом полів для опису цифрових ресурсів, який відповідає сучасним стандартам The European Library (www.theeuropeanlibrary.org) та World Digital Library (<http://www.wdl.org>), і можуть бути експортовані для інтеграції електронних колекцій НБУВ у міжнародні цифрові бібліотеки. Отриманий результат можна побачити на сайті НБУВ, переглянувши «Електронний фонд оцифрованих видань» (<http://irbis-nbuv.gov.ua/fond/>). На основі цифрових інформаційних ресурсів для повноти розкриття змісту особливо цінних і популярних фондів НБУВ формуються окремі тематичні електронні повнотекстові колекції (<http://irbis-nbuv.gov.ua/fond/ires.htm>): «Почаївські видання», «Колекція інкунаблів», «Золота колекція Євразії: стародруки, рукописи, прижиттєві видання класиків української літератури». Оцифровані матеріали використовуються для укладання ілюстрованого фонду електронних виставок, присвячених знаменним і пам'ятним датам (<http://irbis-nbuv.gov.ua/fond/show.html>).

Організація системи захисту та розділення доступу до інформації

Доступ користувачів до електронного архіву необхідно регулювати залежно від виду, призначення і ступеня важливості інформації, способу обробки даних тощо. Інформацію можна розділити на чотири класи [9]: життєво важлива, модифікація якої призведе до невиправної втрати; важлива, доступна невеликій групі користувачів; інформація, постійний несанкціонований доступ до якої може привести до витоку більш цінних даних; не представляє конкретного інтересу для зловмисників, але вимагає захисту від випадкових порушень через безвідповідальність виконавців. Найбільш часто на практиці застосовують системи захисту у вигляді трьох складових частин: основного контуру безпеки, засобів протидії випадковому несанкціонованому доступу та засобів управління системою захисту. Основний контур виконує функції захисту носіїв даних (як змінних, так і незмінних), захисту технічних засобів від несанкціонованого проникнення, контролю введення апаратури в режим виконання регламентних робіт. Випадковому несанкціонованому доступу можна запобігти постійним контролем за якістю та цілісністю інформації (контрольною сумою файлів), застосуванням мережевого програмного забезпечення із засобами аутентифікації. Центральною ланкою системи захисту є адміністрування системи безпеки, що включає реєстрацію користувачів, регламентацію їхнього доступу, реєстрацію паролів, перевірку журнальних записів (записів логфайлів), за якими можна прослідити за зверненням до інформації, її зміною та передаванням у межах локальної (корпоративної) мережі.

У НБУВ для забезпечення захисту інформації в процесі створення цифрових копій було застосовано наступні організаційно-технічні заходи.

1. Основний контур безпеки підтримують такі заходи:

- розташування обладнання для оцифрування, комп'ютерів і проміжного сервера в окремому приміщенні з обмеженим доступом навіть для співробітників;
- виділення окремого сегмента мережі для роботи виключно групи оцифрування документів, у якого відсутній безпосередній доступ до загальної мережі підприємства (дані передаються виключно відповідальною особою до сервера);
- відсутність фізичної можливості підключення змінних носіїв інформації до комп'ютерів, що задіяні в процесі створення цифрових копій та відсутність можливості несанкціонованої передачі даних до локальної мережі або Інтернет;
- доступ до запису дисків UDO фізично обмежений, їхній запис здійснюється відповідальною особою, диски захищені паролем, стійким до зламу.

2. Випадковому несанкціонованому доступу протидіють такі засоби безпеки:

- необхідність авторизації співробітників НБУВ, які задіяні в процесі виготовлення цифрових копій, відсутність фізичної можливості у користувачів змінювати системні й мережеві налаштування комп'ютерів та іншого обладнання;
- цілісність та якість інформації забезпечує система обліку та контролю оцифрованого матеріалу (перевіряється якість зображень, їхня повнота та відповідність оригіналу, технічні метадані цифрових об'єктів містять необхідну інформацію для їхньої ідентифікації: кількість, розмір, формат тощо);
- записи страхового фонду на дисках UDO захищені стійким до зламу паролем.

3. Адміністрування системи безпеки забезпечено такими інструментами:

- наявність системи спостереження за діями користувачів і даними, що передаються у виділеному для створення цифрових копій сегменті локальної мережі;
- для кожної із груп співробітників, які приймають участь у формуванні цифрового ресурсу (фотографів, реставраторів, каталогізаторів, адміністраторів), налагоджено індивідуальні профілі з регламентованим доступом лише до необхідної для їхньої роботи інформації.

Однією із суттєвих проблем, що пов'язана з регламентацією доступу до інформації (поза службовими взаємовідносинами), з якою зустрічаються бібліотеки під час обслуговування електронними інформаційними ресурсами, є дотримання авторських прав і правил архівного обслуговування [3]. Не всі документи можуть бути представлені у мережі для вільного відкритого доступу, частина документів може бути доступною тільки в спеціалізованих читальних залах, де професійні працівники зможуть правильно розпорядитися цифровою копією документа: надати її тільки для перегляду або дозволити копіювання фрагментів документа згідно з установленими правилами.

Ці питання для публікації оцифрованих інформаційних ресурсів НБУВ були вирішені введенням спеціального поля для визначення статусу документа, яке має значення: службовий, локальний та онлайнвий. Залежно від визначеного статусу здійснюється публікація документа. До документів зі статусом «службовий» мають доступ лише співробітники, які наділені повноваженнями, доступ гнучко налагоджується для кожної із груп користувачів адміністратором баз даних і здійснюється через службовий модуль «Каталогізатор» після авторизації. До-

кументи зі статусом «локальний» доступні через модуль «Читач» для користувачів у спеціалізованих читальних залах НБУВ, доступ до них можна отримати після реєстрації у системі чергового бібліотекаря. Всі інші документи зі статусом «онлайн» публікуються для віддалених користувачів і доступні на сайті НБУВ через веб-інтерфейс бібліотечної інформаційно-пошукової системи (<http://irbis-nbuv.gov.ua/fond/>).

Висновки

1. Створення фонду цифрових копій документів бібліотеки потребує вирішення цілого кола питань з організації: поповнення, зберігання, обліку, представлення, пошуку, захисту та розділення доступу до інформації.

2. Підсистема зберігання інформації електронного архіву бібліотеки повинна мати два основні розділи: область оперативно доступної інформації та область довготривалого зберігання. Оперативно доступну інформацію рекомендується зберігати засобами RAID-масивів. Для організації довгострокового зберігання використовують змінні носії інформації (UDO або Blue-Ray диски) та роботизовані бібліотеки на їхній основі.

3. Особливо важливим для впорядкування електронного архіву є організація повноцінної підсистеми обліку, яка має ґрунтуватися на детально розробленому наборі метаданих: описових, технічних і правових.

4. Формування та облік фонду оцифрованих документів бібліотеки на основі автоматизованої бібліотечної інформаційної системи (АБІС) має безперечні переваги та є оптимальним рішенням з огляду на: клієнт-серверну архітектуру сучасних АБІС, що дає змогу організувати мережеву взаємодію користувачів багатьох різних підрозділів; створені на платформі АБІС записи для багатьох документів бібліотеки в електронному каталозі та спеціалізованих базах даних; підтримку зберігання цифрових об'єктів у бінарному вигляді та гіпертекстових посилань до електронних документів; легку інтеграцію створеного електронного фонду з іншими інформаційними сервісами бібліотеки.

5. Перспективою розвитку закладених науково-організаційних засад є удосконалення структури та розробка єдиної технологічної схеми формування, функціонування, використання та представлення користувачам оцифрованого фонду бібліотеки.

1. *Антоненко І.П.* Каталогізація електронних ресурсів: наук.-метод. посіб. / І.П. Антоненко, О.В. Баркова; наук. ред. О.В. Воскобойнікова-Гузєва; Національна бібліотека України ім. В.І. Вернадського НАН України. — К.: НБУВ, 2007. — 116 с.

2. *Асєєв Г.* Концепція електронного сховища даних / Г. Асєєв // Вісник Книжкової палати. — 2009. — № 2. — С. 28–30.

3. *Бобришева О.* Правові засади формування комплексної системи захисту інформації в бібліотеках / О. Бобришева // Вісник Книжкової палати. — 2009. — № 12. — С. 23–26.

4. *Возможности системы автоматизации библиотек «ИРБИС64» для организации работы специализированных фондов научных библиотек / К.В. Лобузина, И.В. Лобузин // Библиотеки на-*

циональных академий наук: проблемы функционирования, тенденции развития: научно-практический и теоретический сборник. — К., 2010. — Вып. 85. — С. 120–130.

5. *Волохін О.М.* Каталогізація цифрових ресурсів Інтернет: Дублінське ядро метаданих: посібник. — Кіровоград, 2003. — 70 с.

6. *Как быстро оцифровать ваши книги с Book Saver Book Scanner* [Электронный ресурс] // Micro-system.org: робототехника и новинки электроники. — URL: <http://www.micro-system.org>

7. *Каштаньер А.* Оцифровка газет в РНБ [Электронный ресурс]: настоящее и будущее // Диалог национальных библиотек: Россия – Великобритания, 15–16 октября 2003: [материалы конференции]. — РНБ, 2003. — URL: <http://www.nlr.ru:8101/tus/151003/kashtanier.htm>

8. *Меленець А.В.* Обрання типу сховища для зберігання документів страхового фонду в електронному вигляді // Системи обробки інформації. — Харків, ХУПС, 2010. — № 6(87). — С. 199–203.

9. *Мельников В.В.* Защита информации в компьютерных системах / В.В. Мельников. — М.: Финансы и статистика; Электронинформ, 1997. — 2005. — 368 с.

10. *Міжнародна наукова конференція «Електронні зображення та візуальні мистецтва»* — International Conference «Electronic Imaging & the Visual Arts»: EVA 2002 Київ, 22–24 травня 2005 р.: зб. праць української конференції серії EVA / Міжнародний науково-навчальний центр ЮНЕСКО інформаційних технологій та систем НАН і міністерства науки і освіти України. — [К.], 2005. — 260 с.

11. *Петров І.В.* Обґрунтування вибору програмно-апаратних засобів архівів електронних документів довготермінового збереження / І.В. Петров, А.М. Стеценко, Н.В. Солоніна // Реєстрація, зберігання і оброб. даних. — 2010. — Т. 12. — № 1. — С. 79–88.

12. [Обладнання для сканування компанії Zeutschel] [Электронный ресурс] — URL: http://www.zeutschel.com/products_scanners.html

13. *Захарова Е.Г.* Опис наукових електронних ресурсів метаданими дублінського ядра / Е.Г. Захарова, О.В. Захарова, В.А. Резніченко // Проблеми програмування. — 2008. — № 2–3. Спец. випуск. — С. 507–514.

14. *Журавский Г.Л.* Построение электронных хранилищ документации больших систем / В.Г. Журавский, В.В. Гольдин // Микропроцессорная и вычислительная техника. — 2007. — № 3. — С. 74–80.

15. *Проблемы оцифровки аудиовизуальных документов* [Электронный ресурс] / Голубев А.М., Прозорович Ю.Э., Неборская Т.В. // Междунар. науч.-практ. конф. «Миссия библиотеки в формировании интеллектуального потенциала общества», 4–6 декабря 2006 г., Минск: [материалы конференции]. — 3 с. — URL: old.nlb.by/press/upload/2883.pdf

16. *Рындин А.* Архив без пыльных полок или способы организации архива предприятия // Jet Info: информационный бюллетень. — 2002. — № 10 (113). — С. 3–40.

17. *Система автоматизации библиотек ИРБИС64.* Общее описание системы. — М.: ГПНТБ России, 2004. — 259 с.

18. *Соколова Н.В.* Электронные библиотеки как факт современного информационного ландшафта / Н.В. Соколова // Университетская книга. — 2010. — [№ 2] февраль. — С. 62–66.

19. *Технологии АBBYU* помогают оцифровать культурное наследие Европы России [Электронный ресурс]. — URL: <http://www.abbyu.ua/company/news/>

20. *American Memory from the Library of Congress* [Электронный ресурс]. — URL: <http://memory.loc.gov/ammem/index.html>

21. *Dublin Core Metadata Initiative* [Электронный ресурс]. — URL: <http://dublincore.org>

22. *Kimball R.* The Data Webhouse Toolkit: Building the Web-Enabled Data / Ralph Kimbal. — Warehouse: John Wiley & Sons, 2000. — 416 p.

23. *Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files — Raster Image* [Електронний ресурс] / S.T. Puglia, J. Reed, E. Rhodes; U.S. National Archives and Records Administration (NARA). — 2004. — 87 p. — URL: <http://www.archives.gov/preservation/technical/guidelines.html>

24. *Unimarc Manual. Bibliographic Format: Руководство по применению UNIMARC (ISO 2709) для библиографических данных / Международная федерация библиотечных ассоциаций и учреждений (ИФЛА); ГПНТБ России.* — М.: ГПНТБ России, 1998. — 52 с.

25. *Xerox* и ГПНТБ России выпускают новый программный продукт для создания электронных библиотек и архивов «БИАР» [Электронный ресурс]. — URL: www.xerox.ru/ru/press-centre/2011/

Надійшла до редакції 20.09.2011