

Ірина Павлуша,
аспірантка НБУВ

ЕЛЕКТРОННІ БІБЛІОТЕКИ: ПИТАННЯ КОМПЛЕКТУВАННЯ Й ОБРОБКИ ВХІДНОЇ ІНФОРМАЦІЇ

Створення електронних бібліотечно-інформаційних ресурсів – один із нових важливих та актуальних напрямів роботи провідних бібліотек України. Для прискорення динаміки цього процесу необхідно розробити ряд теоретичних і методичних засад, вирішити питання термінологічного забезпечення цих робіт, чітко визначити базове поняття «одиниця збереження для фонду електронних документів».

У статті доводиться доцільність документної концепції розвитку книгозбірень, в яких має нагромаджуватися фонд електронних документів на носіях довготривалого збереження. Розглядаються передумови для організації комплектування електронних бібліотек, специфіка технологічного циклу «Шлях документа» для вхідного потоку електронної інформації.

Поняття «одиниця збереження» для фонду електронних документів

Невизначеність поняття «одиниця збереження» для фонду електронних документів зумовила дискусію про інформаційну та документну концепції розвитку бібліотек. Оскільки бібліотека з «книгосховища трансформується в банк даних», поняття «документ» в контексті «одиниця збереження» поступово відмиратиме, стверджують прихильники першої концепції [1, 2]. Аргументованішою є позиція Ю.М.Столярова. Він зазначає: «документу альтернативи нема», бо банк даних також є сховищем документів [3], які мають розглядатись і як «одиниці збереження».

Фахівці НБУВ, котрі вже мають практичний досвід створення електронної наукової бібліотеки, дійшли висновку про необхідність подальшого розвитку документної концепції бібліотек.

Сьогодні вхідний документний потік наукових бібліотек включає широкий спектр документів на комп'ютерних носіях. Однак значна кількість з них не забезпечує довготривалого збереження зафіксованої інформації (наприклад, гнучкі дискети). Очевидно, останні не можуть виступати в ролі «одиниць збереження» бібліотечних фондів, бо їх призначення – транспортування даних. Для довготривалого ж збереження отриманої інформації її слід попередньо

зафіксувати на інших матеріальних носіях (CD-ROM, DVD-дисках та ін.), що вже можуть розглядатись як одиниці збереження бібліотечних фондів.

Але існуючі в бібліотеках електронні каталоги та бази даних не можна вважати документними інформаційними ресурсами до тих пір, поки вони не будуть зафіксовані на електронних носіях довготривалого збереження й передані до загальних бібліотечних фондів з реєстрацією в книзі сумарного обліку документів, а також відображені у довідково-пошуковому апараті. Однак це не здійснюється в Україні навіть у національних книгозбірнях, які мають електронні каталоги обсягом понад 100 тис. бібліографічних записів. Враховуючи незначну вартість «чистих» CD-ROM дисків (приблизно 10 гривень) при великій ємності (650 MB), було б доцільно фіксувати на них у національних бібліотеках усі наявні електронні ресурси з квартальною періодичністю для забезпечення їх надійного збереження. Крім того, це дасть можливість отримати об'єктивні кількісні показники щоквартального збільшення обсягу електронних ресурсів бібліотек для оптимізації на цій основі управлінських рішень.

У цілому ж викладене є доказом на користь документної концепції розвитку книгозбірень. Навіть якщо бібліотека являтиме собою банк комп'ютерної інформації, в ній має існувати фонд електронних документів на носіях дов-

готривалого збереження, де відбиватимуться наявні інформаційні ресурси.

Отже, електронними ресурсами, на наш погляд, можна вважати документи, що відповідають вимогам:

- матеріальний носій інформації повинен забезпечувати її довгострокове збереження (CD-ROM, DVD диски та ін.);

- електронна інформація має бути зафіксована на носії у вигляді, що передбачає її подальше ефективне використання;

- електронний документ, як і традиційний, повинен мати вихідні дані для його реєстрації в книзі сумарного обліку документів і відображення в довідково-бібліографічному апараті бібліотеки.

Передумови організації комплектування електронних бібліотек

Основна завада на шляху формування електронних бібліотечно-інформаційних ресурсів - авторське право, точніше - відсутність у законодавстві чіткого визначення прав книгозбірень стосовно отримання, зберігання і надання читачам електронних документів, зокрема комп'ютерних версій друкованих видань у наукових, освітянських і культурологічних цілях [4]. Офіційна позиція ІФЛА з цього питання полягає в тому, що виготовлення книгозбірнями електронних копій матеріалів у згаданих напрямках не слід вважати порушенням авторських прав [5]. Однак у законодавствах країн, зокрема в чинному Законі України «Про обов'язковий примірник документів» [6], цю позицію не втілено.

Подолання завади, пов'язаної з авторським правом, можна здійснити насамперед, використовуючи позитивний досвід в організації бібліотечної справи у нашій державі (вітчизняні наукові бібліотеки, як правило, є складовими елементами не лише системи соціальних комунікацій, а й певних наукових об'єднань, скажімо, національної чи галузевих академій наук, університетів тощо. Завдяки цьому вказані бібліотеки мають організаційні передумови для створення електронних депозитаріїв наукових матеріалів організацій, до складу яких вони входять. Потрібно лише розробити нормативне забезпечення для реалізації такої функції. Вирішенню питання могло б сприяти прийняття відповідних постанов президій академій наук, рішень учених рад університетів та ін. Як першочергове завдання при цьому слід розглядати організацію інформаційної співпраці між центральними бібліотеками академій наук і видавництвами, що входять до їх складу.

Однак передача центральним бібліотекам науково-інформаційних матеріалів академічних

установ не означає передачу бібліотекам прав на їх довільне використання і, зокрема, на організацію он-лайнного доступу до повних текстів усіх матеріалів. На ряд документів мають поширюватися архівні правила використання. Але, безперечно, слід забезпечити кумуляцію і постійне збереження всієї наявної сьогодні електронної наукової інформації. Існуюча практика вилучення комп'ютерних версток журналів та книг у «корзину» після отримання їх твердих копій свідчить про наявність прогалин у вирішенні питань інформатизації системи документних комунікацій.

Перший крок у напрямі створення загальноакадемічної електронної бібліотеки зроблено в НАН України. Ним став семінар з питань інформаційного забезпечення наукових досліджень у НАН України, проведений у квітні 1999 р. Інформаційно-бібліотечною радою Президії НАН України. Учасники заходу прийняли рішення, в якому доручили НБУВ організувати передачу в науково-дослідні установи Академії копій ЕК і рекомендували цим установам організувати передачу до книгозбірні своїх загальнодоступних електронних інформаційних ресурсів для створення електронної наукової бібліотеки НАН України.

Другим підходом до вирішення питання авторського права може стати практика укладання типових договорів про інформаційне співробітництво між авторами (фізичними особами) й бібліотеками, в яких кожна зі сторін братиме на себе певні зобов'язання. Бібліотека в рамках такого договору здійснюватиме обробку, постійне збереження та пропаганду електронного документа. Книгозбірня, за бажанням автора, може надавати його працю користувачам на некомерційних засадах або виступати посередником при комерційному використанні. Цей підхід особливо актуальний для України, де при наявності значного інтелектуального потенціалу (знань науковців і фахівців, які мають кумулюватися й поширюватися в часі та просторі) науково-видавничу діяльність у зв'язку з економічними негараздами належно не підтримується.

Не менш важливим питанням при організації комплектування електронними матеріалами є визначення єдиних вимог до подання на комп'ютерних носіях наукової інформації (статей, монографій, звітів про НДР тощо) для її наступного включення до фондів електронних бібліотек. Такі вимоги мають виключити або звести до мінімуму додаткові трудовитрати на конвертування вхідних даних у формат зберігання інформації у фондах цих бібліотек.

Дотримання даних вимог є запорукою до-

сягнення успіху, про що свідчить функціонування російської «Національної електронної бібліотеки» [7]. Це - найбільший у світі електронний архів російськомовних документів (понад 3,5 млн.) - матеріалів російських газет, журналів, радіо- і телестанцій, інформаційно-аналітичних агенцій. Щодня електронна бібліотека обробляє близько 4,5 тис. нових документів. Опрацювання такого вхідного потоку було б неможливе без чіткого дотримання постачальниками інформації єдиних правил її подання, в яких визначаються вимоги до кодування символів, подання тексту (неприпустимість переносів, псевдографіки та інших нетекстових символів), перелік обов'язкових полів-заголовків, довжини рядків тощо.

Як вітчизняний приклад можливості отримання вагомих результатів по опрацюванню значних масивів вхідної інформації у стислі строки завдяки дотриманню єдиних вимог подання її може слугувати «Електронна наукова бібліотека НБУВ» [8].

Основний інформаційний ресурс цієї електронної бібліотеки становлять комп'ютерні версії авторефератів дисертаційних робіт, які отримуються з ВАК України (згідно з чинними правилами до переліку матеріалів, які дисертант подає в атестаційну комісію, входить дискета з текстом автореферату, формат DOC). Після розгляду дисертаційних справ інформація з цих дискет передається в НБУВ на депозитарне збереження. «Транспортними» носіями інформації виступають ZIP- та гнучкі дискети. Єдиний формат подання інформації дозволив упродовж одного місяця включити до фонду електронної бібліотеки тисячу авторефератів дисертаційних праць. Порівняємо: у фонді найвідомішого проекту створення електронної бібліотеки - «Гутенберг» [4], започаткованого майже 30 років тому, сьогодні нараховується тисяча книжок.

Таким чином, передумовами для організації комплектування електронних бібліотек слід вважати:

- розробку нормативного забезпечення для передачі до наукових бібліотек комп'ютерної інформації установ і організацій, до складу яких вони входять;

- розробку типових договорів про інформаційну співпрацю між авторами (фізичними особами) й бібліотеками з делегуванням останнім на певних умовах прав на використання електронних документів;

- визначення єдиних вимог до подання на комп'ютерних носіях наукової інформації для її наступного включення до фондів електронних бібліотек з мінімальними додатковими трудовит-

ратами на конвертування вхідних даних у формат зберігання інформації у фондах цих бібліотек.

Особливості технологічного циклу «Шлях документа» для вхідного потоку електронної інформації

Як відзначалося, основна частина вхідного потоку електронної інформації (комп'ютерні версії книг, статей тощо) надходить до бібліотек комп'ютерними мережами або на «транспортних» носіях типу ZIP- та гнучких дискетах. Перед включенням цієї інформації до бібліотечного фонду її потрібно зафіксувати на носіях, що можуть виступати в статусі «одиниці збереження» (CD-ROM, DVD-диски та ін.). Оскільки останні мають велику ємність, на них можна записати електронні версії кількох сотень книжок чи кількох тисяч статей. Має сенс структурувати вхідний потік спочатку за видовою ознакою (книги, статті, автореферати дисертацій тощо), а потім за тематичною (природничі та гуманітарні науки, медицина, сільське господарство та ін.).

При цьому кожній інформаційній одиниці (книзі, статті) має бути надане унікальне в межах «одиниці збереження» (диску) ім'я.

Пропонується для визначення такого імені використовувати алгоритмічний код, що включає вісім символів: два перших символи - дві останні цифри року видання; три наступних символи - перші літери прізвища, імені та по батькові автора документа; три заключних символи - перші літери трьох останніх слів з назви документа (без урахування службових). Викладене правило формування алгоритмічного коду для визначення імені комп'ютерної версії документа використовується, зокрема, в НБУВ при формуванні електронної наукової бібліотеки авторефератів дисертацій [9], пошуковий інтерфейс якої подано на рис. 1. Для прикладу в пошукові вікна введені терміни «Методика бібліографування» (слова з тексту), «07.00.08» (шифр спеціальності «Книгознавство, бібліотекознавство, бібліографознавство») і «1998» (рік видання). Результат пошуку наведено на рис. 2.

Цей результат являє собою бібліографічний опис автореферату кандидатської дисертації Н.С.Дяченко з теми «Формування системи рекомендаційних бібліографічних посібників для дітей в Україні» (1998).

Звернемо увагу на розширений опис автореферату, що містить, крім обов'язкових елементів, анотацію дисертаційної роботи. Вона дає змогу отримати попереднє уявлення про зміст автореферату й на цій підставі прийняти рішення про доцільність ознайомлення з його

Рис. 1.

Рис. 2.

повною електронною версією (передача повних текстів по глобальних комп'ютерних мережах потребує значних витрат часу). Така можливість дозволяє зберігати й передавати матеріали в заархівованому вигляді, що забезпечує підтримку на одному й тому ж дисковому просторі сервера майже втричі більшої кількості документів і суттєво прискорює їх передачу комп'ютерними мережами.

У наведеному прикладі іменем документа, сформованого у відповідності з вищенаведеним правилом алгоритмічного кодування, є 98dnspdu.doc.

Дане ім'я знаходиться у новій для практики каталогізації області бібліографічного опису, що визначає електронну адресу документа та спосіб доступу до нього. Принципи занесення даних у цю область розроблені IFLA в 1997 р. і викладені в проекті ISBD (ER) - Міжнародний стандартний бібліографічний опис електронних ресурсів - схваленому на 64-й сесії федерації у 1998 р. [10, 11]. Рекомендації IFLA відображено в UNIMARC, зокрема в «Блоці міжнародного використання», де з'явилося нове поле 856 - «Електронна адреса документа» [12]. Воно містить інформацію, яка дає можливість визначити місцезнаходження електронної версії документа і отримати доступ до нього. Наприклад, для електронної версії автореферату дисертації Н.С.Дяченко, що зберігається в електронній науковій бібліотеці НБУВ, така адреса має вигляд: URL:

<http://www.nbu.gov.ua/ard/1998/98dnspdu.zip>,

де <http://www.nbu.gov.ua/> - Internet-адреса серверу НБУВ;

ard - каталог, у якому зберігаються повні тексти авторефератів дисертацій;

1998 - підкаталог авторефератів дисертаційних робіт, захищених у 1998 р.;

98dnspdu.zip - ім'я файлу з текстом автореферату.

Структуровані масиви вхідної електронної інформації мають записуватися на носії довготривалого збереження і фіксуватися в книзі сумарного обліку документів. Ці операції повинні мати місце і для власної інформаційної продукції бібліотеки (баз даних, електронних копій раритетів, текстових матеріалів тощо). Потім отримані «одиниці збереження» проходять каталогізацію, яка принципово не відрізняється від аналогічних процесів при обробці матеріалів на паперових носіях.

Однак при обробці електронних одиниць збереження бібліотечного фонду необхідне неухильне дотримання міжнародних правил БО, оскільки їх ігнорування унеможливує доступ до електронних бібліотечно-інформаційних ресурсів по глобальних комп'ютерних мережах.

Особливість індексування електронних матеріалів впливає з можливості створення автоматизованим способом їх розширених пошукових образів, які включали б, крім класифікаційних індексів або предметних рубрик, і реферати чи анотації та ключові фрагменти документів. Як такі фрагменти для праць монографічного рівня доцільно використати їх зміст (перелік розділів і підрозділів), що дає наочне уявлення про сутність роботи. При цьому зміст не набирається клавіатурно, а копіюється з електронного документа.

Для зменшення трудовитрат на розширену індексацію слід використати результати існуючої в бібліотеці аналітико-синтетичної переробки вхідного документного потоку. Так, на-

приклад, у НБУВ українська наукова література не лише бібліографується, а й проходить процес реферування. Тому при надходженні комп'ютерних версій публікацій їх досить лише за певними правилами «підключити» до існуючої реферативної бази про наукові видання України й отримати якісно новий продукт - електронну наукову бібліотеку з розвиненим довідково-пошуковим апаратом. Така можливість уже сьогодні реалізується в НБУВ у електронній науковій бібліотеці авторефератів дисертацій.

Якщо ж електронні ресурси формуються не в бібліотеці з налагодженою технологією аналітико-синтетичної переробки документів, а в організації, де цей процес здійснюється аматорами, створення розвинених лінгвістичних засобів вбачається малоімовірним. Тому, наприклад, віддаючи належне електронному зібранню Максима Мошкова [4] та зібранням інших ентузіастів, усе ж відзначимо, що їх колекції не мають лінгвістичного забезпечення в класичному розумінні цього терміна.

У цілому особливості технологічного циклу «Шлях документа» для вхідного потоку електронної інформації полягають у наступному:

- в існуючий технологічний цикл вводиться новий етап, що включає надання кожній інформаційній одиниці (книзі, статті) унікального ідентифікаційного імені, структурування мате-

ріалів за видовою ознакою і записом отриманих масивів на носії довготривалого збереження;

- вбачається доцільним уже сьогодні фіксувати на CD-ROM чи DVD-дисках усі електронні інформаційні ресурси, створювані в бібліотеках з квартальною чи іншою періодичністю для забезпечення їх надійного збереження. Крім того, це дасть можливість отримати об'єктивні кількісні показники збільшення об'єму електронних інформаційних ресурсів для оптимізації на цій основі управлінських рішень;

- у книзі сумарного обліку документів реєструється не документ на «транспортному» носії інформації (ZIP- чи гнучкий диск), а створена безпосередньо в самій книгозбірні «одиниця збереження» (CD-ROM, DVD-диск тощо);

- у процесі каталогізації електронних ресурсів створюється розширений пошук образу, що включає анотацію, реферат чи ключові фрагменти матеріалів, і зазначення електронної адреси документа та способу доступу до нього;

- при обробці електронних одиниць збереження бібліотечного фонду необхідне неухильне дотримання міжнародних правил бібліографічного опису, оскільки їх ігнорування унеможливує доступ до електронних бібліотечно-інформаційних ресурсів по глобальних комп'ютерних мережах.

1. Скворцов В.В. Российская библиотечная наука: анализируя настоящее, вглядываясь в контуры будущего // Библиотечное дело и проблемы информатизации общества: Тез. докл. междунар. научн. конф.; Москва, 27-28 апр. 1999 г. - М., 1999. - Ч. 1. - С. 57-58.
2. Гуляревский Р.С. Электронные издания в библиотеке: Проблемы и пути их решения // Библиотечное дело и проблемы информатизации общества: Тез. докл. междунар. научн. конф.; Москва, 27-28 апр. 1999 г. - М., 1999. - Ч.1. - С. 17-19.
3. Столяров Ю.Н. Документу альтернативы нет // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества: Материалы междунар. конф. «Крым-99»; Судак, 5-13 июня 1999 г. - Симферополь: Таврида, 1999. - Т.1. - С.282-284.
4. Павлуша І.А. Електронна бібліотека: зарубіжний досвід: питання розробки української концепції // Бібл. вісн. - 1999. - № 4. - С. 13-24.
5. Заявление, отражающее официальную позицию ИФЛА по вопросам авторского права в электронной среде // Библиотеки за рубежом. - 1996. - Вып. 3 [Электрон. ресурс] / Способ доступа: URL: http://cs.libfl.ru/win/nzb96_3.html#8. - Загол. з екрана.
6. Про обов'язковий примірник документів: Закон України від 9 квітня 1999 р. № 595-XIV [Електрон. ресурс] / Способ доступа: URL: http://www.nbuv.gov.ua/law/99_oblik.html. - Загол. з екрана.
7. Национальная электронная библиотека [Электрон. ресурс] / Способ доступа: URL: <http://www.nns.ru/sources/index.html>. - Загол. з екрана.
8. Електронні тексти: електронна наукова бібліотека НБУВ [Електрон. ресурс] / Способ доступа: URL: <http://www.nbuv.gov.ua/eb/>. - Загол. з екрана.
9. Автореферати дисертацій: електронна наукова бібліотека НБУВ [Електрон. ресурс] / Способ доступа: URL: <http://www.nbuv.gov.ua/eb/ard.html>. - Загол. з екрана.
10. Каспарова Н.Н. Библиографическое описание электронных ресурсов в России: национальные аспекты в контексте международного опыта // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества. - С.304-305.
11. Библиографическое описание электронных ресурсов: метод. рекомендации [Электрон. ресурс] / Сост. Т.А.-Бахтурина, И.С.Дудник, Л.Д.Плохоцкая; Отв. ред. Н.Н.Каспарова. - М.: РГБ, 1998. - Способ доступа: URL: <http://www.rsl.ru/pub/erdes.htm>. - Загол. з екрана.
12. Российский коммуникативный формат представления библиографических записей в машиночитаемой форме [Электрон. ресурс] / Способ доступа: URL: <http://www.rba.ru/rusmarc/format.htm>. - Загол. з екрана.