

Технологічна реалізація об'єктно-тематичного доступу до бібліотечних електронних ресурсів

І. П. Перенесієнко

Національна бібліотека України імені В. І. Вернадського, Київ, Україна

Paper received 24.05.18; Accepted for publication 02.06.18.

<https://doi.org/10.31174/SEND-HS2018-169VI28-08>

Анотація. У статті розглянуто технологічні аспекти наукового представлення бібліотечних електронних ресурсів. Обґрунтовано важливість та актуальність системності об'єктно-тематичної організації інформації для сучасної аудиторії. Узагальнено особливості інтеграції бібліотечних електронних ресурсів. На прикладі низки діючих бібліотечних проектів та профільних нормативно-методичних документів репрезентовано реалізацію потенціалу технологічних рішень у процесі об'єктно-тематичного розкриття змісту документів.

Ключові слова: електронні бібліотеки, семантичний веб, організація знання, індексування.

Вступ. Поява терміну *електронна бібліотека* зумовлена потребою фіксації нових реалій інформаційно-комунікаційних технологій – розповсюдження локальних та широкосмугових комп'ютерних мереж, прискорення швидкості передачі даних, зростання гіпертекстового середовища та збільшення доступності баз даних. У 1989 р. Т. Бернерс-Лі запропонував глобальний проект, котрий передбачав публікацію гіпертекстових документів, пов'язаних між собою посиланнями. Протягом 1991–1993 рр. він створив перший сервер, гіпертекстовий браузер та сайт. Ці технології заклали основу Всесвітньої павутини, аудиторія котрої у 2017 р., згідно даними Internet World Stats, склала 51.7%. З огляду на стрімку зміну інформаційного середовища та природи наукової комунікації і переорієнтації читачів на електронні формати документів, бібліотеки постали перед завданням розробки технологій об'єктно-тематичного доступу до електронних ресурсів власного генерування.

Стислий огляд публікацій за темою. Технологічні аспекти реалізації доступу до бібліотечних електронних ресурсів розглядали низка науковців (О. Б. Антопольський, С. Ф. Бакан, Н. В. Вараксіна, М. Віллер (M. Willer), К. Вільямсон (C. Williamson), О. М. Волохін, Т. А. Калюжная, М. Криницька (M. Krynicka), О. Л. Лаврик, П. М. Лапо, В. П. Леонов, К. В. Лобузін, О. Ю. Мар'їна, Р. С. Мотульський, Л. В. Найханова, Б. Пайм (B. Paim), Н. П. Пасмор, Я. Пацек (J. Pasek), Г. В. Шемаєва та ін.). Наразі ця проблема вважається однією з ключових у вивченні розвитку семантичного веб. Не зважаючи на значну дослідницьку увагу, без належного висвітлення залишається чимало питань – зокрема, реалізація об'єктно-тематичного доступу.

Метою статті є вивчення та узагальнення прикладів технологічної реалізації об'єктно-тематичного доступу до бібліотечних електронних ресурсів, покликаних сприяти інтеграції документів.

Матеріалом дослідження є профільні нормативно-методичні документи Функціональні вимоги до бібліографічних записів (Functional Requirements for Bibliographic Records, FRBR), Міжнародний ідентифікатор стандартних найменувань (International Standard Name Identifier, ISNI), Віртуальний міжнародний авторитетний файл (Virtual International Authority File, VIAF), «Ключові проблеми розвитку колекцій елект-

ронних ресурсів» («Key issues for e-resource collection development: a guide for libraries»), а також європейська електронна бібліотека Європіана (Europeana), нідерландський проект Delpher та електронна бібліотека «Україніка» НБУВ.

Результати та їх обговорення. Для користувача участь бібліотекарів у функціонуванні електронної бібліотеки менш помітна, ніж у традиційній. Проте низка проектів припинили існування через відсутність заходів з організації доступу до документів. Їхні автори припускали, що для електронної бібліотеки достатньо оцифрувати документи та розмістити їх у веб-мережі.

Хоча процеси технологічної інформаційної адаптації бібліотек розвиваються більше тридцять років, формування суспільства глобального знання в умовах активного використання та створення інновацій в останні роки вимагає нового рівня семантичного узагальнення інформаційних послуг, освоєння нових технологій, в першу чергу, пов'язаних з обробкою, наданням доступу і виявленням нових знань [5, С.471].

Наразі доречно говорити про низку прикладів впровадження інформаційно-комунікаційних технологій до бібліотечної практики: електронні видання; поява нових технологій обслуговування та самообслуговування; віртуальна довідка, електронна доставка документів, онлайнове замовлення документів та ін.; створення електронних колекцій та бібліотек традиційними бібліотеками.

Головні завдання електронних бібліотек у свою чергу полягають у: веденні онтологічної бази знань, для чого необхідно розробити відповідні програмні інструменти; ануванні електронних ресурсів; веденні ієрархії каталогів, яка має на меті створення спеціальних словників для розподілу ресурсів за типами, видами та особливостями; забезпеченні можливості переходу між елементами опису електронних документів; сприянні нових технологій точному та повному пошуку документів [9, С.38].

Семантичний веб реалізує пошук за допомогою одного діалогового вікна – сучасні користувачі звикли до евристичної моделі пошукової системи Google, тож віддають перевагу подібним технологіям. Читачі при цьому часто навіть не розуміють, до якого джерела звертаються: чи це каталог, бібліографія, електронна бібліотека, архіви чи зовнішні пошукові систе-

ми; з одного місця чи відразу кількох отримано ці дані. Для них головне – швидкий доступ до інформації [12, С.5].

Поняття «контрольована точка доступу», згідно з FRBR, демонструє спрямованість на відображення зокрема об'єктно-тематичної складової. Організувати тематичний доступ до всіх ресурсів неможливо – навіть найрозвиненіші інформаційно-пошукові системи світу охоплюють у своїх індексах не більш ніж 30–40% наявних ресурсів. Необхідний для подолання проблеми консенсус власників систем подібного типу міг би забезпечити індексування «вертикальних» фрагментів мережі та якісну навігацію, проте цього наразі не відбувається – ефективними залишаються лише вузькотематичні каталоги та пошукові системи [4, С.23].

При побудові певної концептуальної моделі об'єктно-тематичного доступу до електронного ресурсу ключовим є рішення про визначення об'єктів та їхніх атрибутів. Розробники FRBR бачили суттєву перевагу в позначенні персоналій та організацій як окремих об'єктів, які могли б бути пов'язані з іншими, не включеними до відповідної моделі.

Практичний потенціал такої інформації можна простежити на прикладі ISNI та VIAF. Предметні рубрики, що відповідають персоналіям чи установам, повинні не лише розв'язувати питання тематичної організації локальної колекції. Вони мають потенціал сприяти налагодженню технологічної сумісності останніх з міжнародними авторитетними файлами – міжнародні ідентифікатори надають можливість продовжити пошук в каталогах інших бібліотек світу [6].

Стандарт ISNI регламентує унікальну ідентифікацію автора в вигляді номера. Ідентифікатор важливий для уніфікації представлення імен авторів, подолання невизначеності пошуку. Ця технологія відіграє значну роль у процесах зміни та розвитку сучасної наукової комунікації – на сьогодні вона об'єднує понад 9.8 мільйонів ідентифікаторів. Вирішується проблема неоднозначності назв та імен, мовних відмінностей. Наприклад, «William Shakespear» і «William Shakespeare» є альтернативними визначеннями однієї особи, тому належать до одного унікального ISNI [11]. Завдяки розв'язанню подібних проблем ISNI, ідентифікатори виступають у ролі сполучної ланки між різними інформаційними сегментами, та так само як критичним компонентом пов'язаних даних та семантичного веб.

VIAF є віртуальним каталогом міжнародного нормативного контролю під загальною егідою OCLC. Кожен запис отримує унікальний номер, містить набір первинних записів, посилається на оригінальні джерела. Дані перебувають у вільному доступі і можуть бути використані для досліджень і обміну ними. Типами записів є «Назви організацій», «Географічні назви», «Особи», «Твори», «Вирази», «Прийняті форми імені», «Гочна форма форма заголовків», «Заголовки» [14]. Наприклад, за запитом «НБУВ» у полі «Назви організацій» VIAF містить назви: «Національна бібліотека України імені В. І. Вернадського», «Національна академія наук України», «НБУВ, Інститут рукопису», «НБУВ, Інститут архівознавства», «НБУВ, Інститут біографічних досліджень» тощо.

Бібліотека Конгресу США підтримує серверний протокол Z39.50, створений для застосування у міжбібліотечних пошукових каталогах, та який часто включають до автоматизованих інформаційних систем. Для інтеграції бібліотечних каталогів у репозиторії відкритого доступу в Бібліотеці Конгресу США розроблено протокол SRU/SRW, який дає змогу здійснювати пошук інформації як на бібліотечних серверах Z39.50, так і в репозитаріях підтримуваних протоколом OAI-PMH. Таке технологічне рішення забезпечує можливість інтегрувати власні інформаційні ресурси в сучасній інформаційні сервіси відкритого доступу не змінюючи бібліотечну технологію, де реалізується можливість створення єдиного вікна доступу незалежно від того, який протокол обміну даними підтримує АБІС конкретної національної бібліотеки [7, С.133]. З погляду тематичного доступу інтеграція метаданих передбачає наявність таких уніфікованих даних, як стандартизовані предметні рубрики та систематичні індекси (наприклад, Предметні рубрики Бібліотеки Конгресу США, індекси УДК).

Відповідно до керівництва «Ключові проблеми розвитку колекцій електронних ресурсів», процес технологічної організації електронних ресурсів має забезпечувати: а) функціональність пошуку та видачі результатів, наприклад, усечення, перегляд, історія пошуку, транслітерація; б) експорт та завантаження, наприклад: друк, відправлення електронною поштою, завантаження на комп'ютер чи інший електронний пристрій; в) можливості сортування та ранжирування результатів пошуку, наприклад, за автором, заголовком, датою, релевантністю, фасетами і т.д.; г) інтуїтивну зрозумілість інтерфейсу, матеріальні навігації, допоміжні тексти на навчальні матеріали; д) інтеграція; е) надійність та доступність – наприклад, мінімальний час очікування, режим доступу 24/7 [3, С.12–13].

Більше 10 мільйонів текстових сторінок доступні в Інтернеті завдяки платформі Delpher, проекту Королівської бібліотеки Нідерландів. Електронна бібліотека має на меті об'єднати установи, дотичні до збереження історико-культурної спадщини Нідерландів: «Перевага платформи Delpher полягає в тому, що ви можете вести перехресний пошук у всіх голландських бібліотеках та установах-зберігачах культурної спадщини на одному онлайн-порталі» [10]. Тематичну організацію електронних колекцій технологічно здійснено на базі каталогу основного фонду, у самих колекціях є лише невеликий список предметних рубрик для оперативної роботи читачів. Кожну тематичну електронну колекцію організовано шляхом об'єднання бібліографічних записів відповідної тематики з уже діючого каталогу основного фонду. Наприклад, при виборі першої у списку колекції «Ахтерберг Герріт», гіперпосилання скеровує читачів до головного каталогу. Виняток з технології представлення документів у Delpher є лише стародруки та книжкові пам'ятки. Тематичний доступ до цієї колекції у вигляді предметних рубрик здійснено на окремому сайті «Medieval illuminated manuscripts».

Одним з ефективних способів подолання проблем, що виникають через специфічні характеристики описів документів, є просте запозичення даних у різних постачальників. Тематичні збірки колекції Європіани,

наприклад, –набори фільтрів. Обираючи елементи на основі предмета, дисципліни та інших критеріїв, користувач може узгоджувати та спрямовувати пошук. Незважаючи на простоту механізму, він виявився надзвичайно ефективним – статистичні показники свідчать, що відвідувачі об'єктно-тематичних колекцій переглядають більше сторінок та довше залишаються на кожній із них.

З погляду технічної інфраструктури, архітектура Європіани побудована навколо пошукової платформи. Можливості останньої включають повнотекстовий та фасетний пошук, індексування у реальному часі, динамічну кластеризацію, інтеграцію з базами даних, можливості обробки складних форматів, розподілений пошук та реплікацію індексу. Крім того, в системі адаптовано плагін BM25 для визначення релевантності документів пошуковому запиту, що використовується для ранжування документів. Використання машинного навчання покращує пошукові результати на 24,1% в порівнянні з незбалансованим пошуком [13].

Модель інтегрованого веб-простору передбачає глобальну уніфікацію структур даних та організацію кількох взаємозалежних централізованих реєстрів. Формуючи середовище на рівні даних, необхідно чітко визначити соціальний потенціал взаємодії з тими чи іншими проектами і на цій основі передбачити необхідні технологічні компоненти для побудови каналів ефективної співпраці [2, С.165–166].

Електронні бібліотеки вирішують проблеми невеликої кількості примірників чи збереження книжкового фонду. Це надає можливість відновлення втраченого чи пошкодженого видання, дозволяє отримувати інформацію незалежно від часу та місця перебування, а також виконувати операції з електронними документами, які виходять за межі простого читання тексту чи перегляду зображення. Попри прагматичну природу своєї появи, електронним бібліотекам вдалося вплинути у розвиток інформаційно-комунікаційних технологій. Певною мірою вони стали їхньою квінтесенцією – «концентрованою формою утвердження електронних технологій у бібліотечній сфері» [7, С.60].

Мета проекту НБУВ електронна бібліотека «Україніка» – акумулювати в електронному форматі твори –

незалежно від мови та місця видання – про українців, територію України та про всі народи, які жили або живуть на ній. Для наукової організації тематичного доступу у електронній бібліотеці «Україніка» застосовано певну структурну модель. Принциповим є впровадження єдиної системи упорядкування знань на основі наукового довідково-пошукового апарату. Основні поняття представлені галузі знань супроводжено іконографічною та довідково-енциклопедичною інформацією, що надає змогу визначити для кожного ресурсу його тематичні, територіальні, часові й інтелектуальні атрибути. Все це створює для дослідників умови ефективного пошуку джерел наукової інформації [1].

Надалі передбачається створення єдиної інформаційної системи із засобами розподіленого пошуку в електронних ресурсах україніки. Веб-ресурси українознавчої тематики будуть представлені як окремим блоком (інтернет-навігатор з питань українознавства), так і систематизовані разом із іншими цифровими об'єктами, відповідно до тематичного наповнення [9].

Висновки. Підсумовуючи аналіз технологічної реалізації об'єктно-тематичної організації бібліотечних електронних ресурсів, можна сказати наступне. З одного боку, функції традиційних та електронних бібліотек подібні – формування фонду, його зберігання, опис джерел, наукова систематизація, обслуговування читачів. З іншого, електронний документ з реалізованими зв'язками не зовсім тотожний паперовому відповіднику – це вже інший об'єкт, що має інтерактивні можливості. Виникає відмінна від практики традиційних бібліотек потреба забезпечення взаємоузгодженої системи посилення, результатом реалізації якої стає електронний об'єкт як елемент організованого сховища електронного контенту. Сучасні технології створюють умови для забезпечення уніфікованого тематичного представлення об'єднаних територіально розподілених ресурсів. Найявний досвід демонструє, що головним пріоритетом є не нагромадження фондів, а політика отримання метайнформації – інформації про інформацію в електронній формі та сумісність форматів машиночитаних бібліографічних описів у нових умовах.

ЛІТЕРАТУРА

1. Дубровіна Л. А., Лобузін К. В., Попик В.І. Історія та розвиток поняття «Україніка» як національного ресурсу документальної пам'яті [Електронний ресурс] // Україніка – електронна бібліотека. URL: <http://irbis-nbuv.gov.ua/cgi-bin/ua/elib.exe?C21COM=F&I21DBN=UKRAINICA&P21DBN=UKRLIB>.
2. Інтеграція України у світове співтовариство в контексті розвитку бібліотечних інформаційних технологій [О. С. Онищенко, Л. А. Дубровіна, В. М. Горвий та ін.]; НАН України, Нац. б-ка України ім. В. І. Вернадського. К.: НБУВ, 2011. 232 с.
3. Ключевые проблемы развития коллекций электронных ресурсов: руководство для библиотек; Междунар. федерация библиотечных ассоциаций и учреждений. Секция по комплектованию и развитию коллекций, Рос. библиотечная ассоциация. – Санкт-Петербург: Российская национальная библиотека, 2012. 52 с.
4. Ландэ Д. В. Поиск знаний в Internet. Профессиональная работа / пер. с англ. М.: Издательский дом «Вильямс», 2005. – 272 с.
5. Лобузін К. В. Библиотечные технологии организации знаний в электронной научно-образовательной среде // Образовательные технологии и общество. 2014. Т. 17. № 1. С. 469–486.
6. Лобузін К. В. Авторитетний файл імен осіб – розширення можливостей пошуку в електронному каталозі НБУВ [Електронний ресурс] // 2018. Національна бібліотека України імені В. І. Вернадського URL: <http://www.nbuv.gov.ua/node/3919>.
7. Лобузін К. В. Технології організації знань у бібліотечно-інформаційній діяльності: монографія; НАН України, Нац. б-ка України ім. В. І. Вернадського. – К., 2012. 252 с.
8. Лобузін К. В. Фундаментальна електронна бібліотека «Україніка»: технологічна організація та основні принципи управління інформаційними ресурсами [Електронний ресурс] // «Міжнародна наукова конференція «Бібліотека. Наука. Комунікація», 2018. Національна бібліотека України імені В. І. Вернадського. URL: <http://conference.nbuv.gov.ua/report/view/id/656>.

9. Хоай Л., Тузовский А. Ф. Использование онтологии в электронных библиотеках // Известия Томского политехнического университета, 2012. Т. 320. № 5. С. 36–42.
10. Delpher.nl – Dutch cultural heritage at your fingertips [Electronic resource] // The European library. Connecting knowledge. URL: <http://www.theeuropeanlibrary.org/tel4/newsitem/3900>.
11. International Standard Name Identifier (ISO 27729) [Electronic resource] // 2012. URL: <http://www.isni.org/>.
12. Pacek J., Krynicka M. «Are you trying to kill me ?» – the national bibliography asks Internet [Electronic resource] URL: <https://www.ifla.org/past-wlic/2012/215-pacek-en.pdf>.
13. Petras V. Europeana – a Search Engine for Digitised Cultural Heritage Material / V Petras, T. Hill, J. Stiller, M. Gäde [Electronic resource] // Datenbank-Spektrum. 2017. 17(1). P. 41–46. URL: <https://doi.org/10.1007/s13222-016-0238-1>.
14. VIAF: Виртуальный международный авторитетный файл [Electronic resource] // URL: <https://viaf.org/>.

REFERENCES

1. Dubrovina L. A., Lobuzina K. V., Popyk V. I. Istoriia ta rozvytok poniattia «Ukrainiky» yak natsionalnoho resursu dokumentalnoi pamiaty [Elektronnyi resurs] // Ukrainika – elektronna biblioteka. URL: <http://irbis-nbuv.gov.ua/cgi-bin/ua/elib.exe?C21COM=F&I21DBN=UKRAINICA&P21DBN=UKRLIB>.
2. Intehratsiia Ukrainy u svitove spivtovarystvo v konteksti rozvytku biblioteknykh informatsiinykh tekhnolohii [O. S. Onyshchenko, L. A. Dubrovina, V. M. Horovy ta in.]; NAN Ukrainy, Nats. b-ka Ukrainy im. V. I. Vernadskoho. K.: NBUV, 2011. 232 s.
3. Key issues for e-resource collection development: a guide for libraries; Mezhdunar. federaciya bibl. asociacij i uchrezhdenij. Sekciya po komplektovaniyu i razvitiyu kollekcij, Ros. bibl. assoc. – Sankt-Peterburg: Rossijskaya nacionalnaya biblioteka, 2012. 52 s.
4. Lande D. V. Poisk znaniy v Internet. Professionalnaya rabota / per. s angl. M.: Izdatelskij dom «Vilyams», 2005. – 272 c.
5. Lobuzina E. V. Biblioteknyye tehnologii organizacii znaniy v elektronnoj nauchno-obrazovatelnoj srede // Obrazovatelnye tehnologii i obshestvo. 2014. T. 17. № 1. S. 469–486.
6. Lobuzina K. V. Avtorytetnyi fail imen osib – rozshyrennia mozhlyvostei poshuku v elektronnomu katalogi NBUV [Elektronnyi resurs] // 2018. Natsionalna biblioteka Ukrainy imeni V. I. Vernadskoho URL: <http://www.nbuv.gov.ua/node/3919>.
7. Lobuzina K. Tekhnolohii orhanizatsii znannievkykh resursiv u bibliotekno-informatsiini diialnosti: monohrafiia; NAN Ukrainy, Nats. b-ka Ukrainy im. V. I. Vernadskoho. – K., 2012. 252 s.
8. Lobuzina K. V. Fundamentalna elektronna biblioteka «Ukrainika»: tekhnolohichna orhanizatsiia ta osnovni pryntsyipy upravlinnia informatsiinykh resursamy [Elektronnyi resurs] // «Mizhnarodna naukova konferentsiia «Biblioteka. Nauka. Komunikatsiia», 2018. Natsionalna biblioteka Ukrainy imeni V. I. Vernadskoho. URL: <http://conference.nbuv.gov.ua/report/view/id/656>.
9. Hoaj L., Tuzovskij A. F. Ispolzovanie ontologii v elektronnykh bibliotekah // Izvestiya Tomskogo politehnicheskogo universiteta, 2012. T. 320. № 5. S. 36–42.

Technological realization of object-thematic access to library electronic resources

I. P. Perenesienko

Abstract. The article deals with the technological aspects of the scientific presentation of library electronic resources. The importance and relevance of the systematization of the object-thematic organization of information for a modern audience are substantiated. Features of integration of library electronic resources are generalized. On the example of a number of existing library projects and profile normative and methodological documents, the implementation of the potential of technological solutions in the process of object-thematic disclosure of the contents of documents.

Keywords: *electronic library, semantic web, knowledge organization, indexing.*

Технологическая реализация объектно-тематического доступа к библиотечным электронным ресурсам

И. П. Перенесиенко

Аннотация. В статье рассмотрены технологические аспекты научного представления библиотечных электронных ресурсов. Обоснованы важность и актуальность системности объектно-тематической организации информации для современной аудитории. Обобщены особенности интеграции библиотечных электронных ресурсов. На примере ряда действующих библиотечных проектов и профильных нормативно-методических документов представлена реализация потенциала технологических решений в процессе объектно-тематического раскрытия содержания документов.

Ключевые слова: *электронные библиотеки, семантический веб, организация знаний, индексирование.*