

# ЕВОЛЮЦІЯ СТРУКТУРИ І ФУНКЦІЙ БІБЛІОТЕКОЗНАВСТВА ПІД ВПЛИВОМ СУЧАСНИХ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

~\* [027.021:061.12]:004.89

**Денис Соловяненко,**  
наук. співробітник НБУВ,  
канд. іст. наук

## Академічні бібліотеки у новому соціотехнічному вимірі: Частина друга. Інфраструктура зберігання електронних науково-інформаційних ресурсів

У другій частині дослідження авторська увага зосереджена на проблематиці створення академічними бібліотеками світу інфраструктури зберігання електронних науково-інформаційних ресурсів. Висвітлено діяльність бібліотек зі створення та підтримки сховищ даних наукових і освітніх установ. Детально розглянуто програмний пакет Dspace, проблемні аспекти зберігання бібліотеками онлайн-науково-інформаційних ресурсів, технологічну парадигму зберігання електронних LOCKSS.

**К л ю ч о в і с л о в а:** академічні бібліотеки, відкритий доступ, інфраструктура, сховища даних, цифрове архівування.

The paper is devoted to an overview of key trends in academic libraries of late 20th – early 21st centuries. The second part of the research is concentrated on issues relating to creation library infrastructure for digital preservation of scholarly information resources. The experience of academic libraries with institutional data repositories is presented. The software package DSpace is described in details. Furthermore, it is examined issues of digital preservation initiatives of academic libraries. The LOCKSS, a technological paradigm for digital preservation, is examined.

**К е у в о р д s:** academic libraries, open access, infrastructure, data repositories, digital archiving.

**П**ерспективний напрям інноваційного розвитку академічних бібліотек пов'язаний із розробленням і впровадженням спеціалізованих технологій архівування та зберігання електронних фондів. Тут є ціла низка абсолютно різних за своєю природою рішень, але всі вони об'єднані спільною метою – бібліотека прагне забезпечити своїм користувачам режим відкритого доступу до онлайн-інформаційних першоджерел, їх довготривалого та ефективного використання. У першу чергу, зупинимося на питанні розвитку академічними бібліотеками мережі відкритих сховищ даних науково-інформаційних ресурсів.

Інформаційна криза академічної галузі, спричинена процесами глобалізації та значного подорожчання передплатних наукових ресурсів, у першу чергу, вдарила по пересічних вчених. Парадокс ситуації полягає у тому, що видавничі монополії безкоштовно отримують зміст від вчених, а потім їм же (через бібліотеки або ж напряму) цей зміст продають за завищеними цінами [7]. Але якщо бібліотека не здатна придбати певний інформаційний ресурс, для пе-

ресічного вченого це невігідно одразу вдвоє: він не може читати праці колег-науковців, але й колеги не можуть читати (та цитувати) його праці. За таких умов вчені провідних країн світу в альянсі з академічними бібліотеками єдиним фронтом почали вдаватися до масових антимонопольних дій, головною серед яких став рух за відкритий доступ.

Останній започаткували ще у 1980-х рр., але системних рис він набув лише у 2002 р., з оприлюдненням фондом «Відкрите суспільство» (Угорщина) Будапештської ініціативи відкритого доступу (Budapest Open Access Initiative). Згідно з Ініціативою, під відкритим доступом до науково-інформаційних ресурсів розуміється надання засобами мережі Інтернет необмежених фінансовими, правовими та технічними перешкодами можливостей читання, завантаження, копіювання, розповсюдження, друку, переробки та використання з будь-якою законною метою літератури, котру дослідники виробляють і безкоштовно надають суспільству [1]. Будапештська ініціатива відкритого доступу пропонує два основні варіанти забезпечення від-

критого доступу: публікація у електронних журналах відкритого доступу (так званий «золотий шлях» відкритого доступу) та депонування (самоархівування) вченими у відкритих сховищах даних (особистих, тематичних, установи тощо) власних матеріалів, опублікованих у комерційних наукових часописах («зелений шлях»). Хоча тут існують й інші шляхи: публікація результатів досліджень у гібридних журналах, які надають вільний доступ до повних текстів статей через певний проміжок часу після виходу статті, або у комерційних журналах, котрі дають змогу автору сплатити за редакційну обробку його рукопису видавництвом і таким чином «спонсорувати» відкритий доступ до повного тексту своєї роботи.

Про методи забезпечення «золотого шляху» і роль у цьому процесі академічних бібліотек йшлося у попередній статті (Бібл. вісн. – 2010. – № 4. – С. 3–14). Тепер торкнемося юридичного аспекту даного питання. Правова передумова виникнення бар'єра відкритого доступу виникає на етапі передачі вченим свого рукопису видавцю для публікації. Традиційна видавнича модель передбачає, що автор передає видавцю і зміст, і усі права на публікацію, копіювання та розповсюдження цього змісту в оригінальному або модифікованому (редакторами та рецензентами) вигляді (копірайт). Таким чином, у подальшому видавець може застосовувати будь-яку бізнес-модель щодо поширення наукових робіт і автор не має засобів впливу на нього, оскільки переданий видавцю копірайт залишає автору лише винятково особисті права на опубліковане. Серед іншого, рух за відкритий доступ прагне усунути дану перешкоду відкритому доступу шляхом запровадження у журналах відкритого доступу нових моделей правовідносин авторів і видавців та нових видів ліцензій на зміст, які автори можуть передавати видавцям замість копірайту. Найбільш поширеними ліцензіями у журналах відкритого доступу є ліцензії неприбуткової організації Creative Commons.

За даними каталогу журналів відкритого доступу (DOAJ), що підтримується бібліотекою Лундського університету (Швеція), станом на початок 2010 р. у світі нараховується близько 4,8 тис. рецензованих наукових часописів, які забезпечують відкритий доступ до опублікованих статей [11]. Експерти відзначають [18], що серед наукових журналів частка тих, що забезпечують відкритий доступ, становить приблизно 10–15 %. Таким чином, для радикального покращення ситуації у науково-інформаційному забезпеченні важливим є також «зелений шлях» відкритого доступу.

Якщо дослідники безкоштовно віддають створений ними зміст видавництвам, то вони, звичайно ж, повинні мати право забезпечити безкоштовний доступ до цього змісту принаймні бібліотекам тих установ, у яких вчені працюють, або опублікувати цей зміст на власних веб-сайтах. Видавці витрачають гроші на редагування та верстку наукових робіт, але права на оригінали авторських рукописів, логічно, повинні належати дослідникам. Йдеться про публікацію авторами препринтів власних робіт шляхом їх депонування у зовнішні сховища даних. Під препринтом, як правило, розуміється «фінальна рецензована, редагована та прийнята до друку версія статті», хоча іноді під цим терміном розуміється «версія праці до її рецензування» [5].

Історія публікації дослідниками препринтів наукових робіт в Інтернеті веде свій відлік з початку 1990-х рр., коли тільки з'являлося всесвітнє павутиння. У 1991 р. було створене найстаріше і, станом на 2007 р., найбільше публічне онлайн-сховище препринтів [18] – проєкт arXiv.org бібліотеки Корнельського університету (США). Наразі сервер arXiv.org вміщує сховище обсягом до 600 тис. робіт у галузі фізико-математичних наук, схвалених до друку у наукових часописах, та має 15 дзеркал у різних країнах світу. Даний архів є дуже важливим джерелом отримання наукової інформації для вчених в усьому світі, він співпрацює не лише з авторами, а й з видавцями, завдяки чому праці з таких напрямів, як, наприклад, фізика високих енергій, депонуються у arXiv.org з вичерпністю майже 100 % [19].

Рух за відкритий доступ передбачає популяризацію депонування вченими не лише препринтів, але і «постпринтів» наукових статей. Поняття «постпринт» також має різні визначення (електронні комунікації змінили структуру видавничого процесу так, що з'явилися деякі нові поняття, яких у традиційному видавничому процесі просто не існувало). Юридично правильне визначення «постпринту» дають видавці – «версія роботи з правками, зробленими за результатами проходження нею процесу наукового рецензування», хоча самі автори наукових робіт часто вважають «постпринтом» остаточну електронну версію статті, яка з'являється у журналі [5].

Станом на 2010 р. найбільшим відкритим сховищем «постпринтів» наукових статей є проєкт PubMedCentral Національної медичної бібліотеки (США) [18], започаткований 2000 р. PubMedCentral – електронний архів англійської наукової періодики у галузі біомедичних наук та наук про життя, зв'язаний через ідентифікацію PMID з реферативною базою даних Medline. Специфіка формування даного сховища передбачає співпрацю бібліотеки

не з авторами, а з видавцями. Майже 800 галузевих часописів постачають (одразу після публікації або з затримкою) повні тексти опублікованих статей у PubMedCentral. Крім того, Національна бібліотека медицини має проект з оцифрування та представлення у PubMedCentral архівів ключових журналів, починаючи з їх першого випуску. Обсяг сховища у 2007 р. становив 1 млн повних текстів статей [22]. PubMedCentral конвертує статті, отримані від видавців, у спеціальний архівний формат на основі XML, тому у сховищі науковий зміст зберігається у компактному текстовому форматі. Деякі видавці додатково постачають для публікації у PubMedCentral також остаточні PDF-версії власних статей.

Популярність даного сховища серед вчених всього світу сприяла зростанню інтересу інших держав до створення аналогічних національних сховищ даних. Станом на квітень 2010 р. Національною медичною бібліотекою було авторизовано два національні проекти: UK PubMedCentral (підтримується Британською бібліотекою з 2007 р.); PubMedCentral Canada (спільний проект Канадського інституту медичних досліджень та Канадського інституту науково-технічної інформації (Національної наукової бібліотеки), запущений у жовтні 2009 р.) [21].

Названі сховища даних є тематичними, вони спеціалізуються на збиранні наукових робіт з певних галузей знання, опублікованих вченими з різних країн світу. Такі сховища підтримуються потужними установами і мають загальносвітову значущість у питаннях поширення ідей відкритого доступу. Але часто для науковців оптимальним варіантом депонування робіт є розміщення їх у локальних сховищах даних (так званих репозиторіях \* відкритого доступу) власних установ. Така ситуація зумовлена багатьма причинами. По-перше, потужні тематичні сховища існують не з усіх дисциплін і не завжди вони співпрацюють безпосередньо з авторами. По-друге, багато журналів обмежують політику депонування робіт вченими пунктом «...лише на особистому сайті або сайті своєї установи» [5]. По-третє, у локальному сховищі колегам-науковцям знайти роботу легше, а самим авторам легше отримати допомогу та фахову консультацію щодо передачі ними робіт у локальні архіви.

Розглядаючи інший аспект функціонування сховищ даних установ, слід вказати на їхній величезний потенціал щодо оприлюднення вченими власних «сірих архівів». Більшість науковців зберігають на своїх локальних комп'ютерах колосальні обсяги корисної наукової інформації, і лише нез-

начну частину власного доробку вони оприлюднюють у тому чи іншому вигляді. Таким чином, тисячі файлів з первинними даними досліджень, цікавими нереалізованими задумами, неопублікованими доповідями тощо залишаються «сірими архівами», вони старіють, «зачинені» на локальних комп'ютерах, і рано чи пізно остаточно втрачаються суспільством. Вчені мають повне право на приватність своїх архівів, але якщо вони свідомо прагнуть частково їх оприлюднювати, наукові установи повинні усіляко сприяти цьому. Отже, потреба у локальних сховищах наукових матеріалів наявна як у окремих вчених, так і в установ, де вони працюють. У більшості наукових організацій розроблення та підтримку таких локальних сховищ взяли на себе академічні бібліотеки.

Комерційні видавці у своїй більшості критично ставляться як до ідей відкритого доступу, так і до процесів самоархівування [10, с. 24]. Видавцям, зрозуміло, вигідно, щоб науковці максимально повно використовувати саме їх технологічні платформи. Проте, рух за відкритий доступ став надто поширеним у академічному середовищі, щоб монополії могли повністю його ігнорувати. Експерти відзначають, що комерційним монополіям загрожують не стільки нові журнали відкритого доступу, скільки саме практика депонування вченими власних робіт. Досить часто на рівні керівництва установ та дослідних фондів, які надають гранти, вчені заохочуються (або навіть примушуються) до підтримки ідей відкритого доступу; за таких умов переважна більшість (81 %) вчених надає перевагу депонуванню, замість того, щоб шукати для своїх робіт нові журнали, які підтримують політику відкритого доступу [4]. З метою мінімізації втрати свого впливу, комерційні видавці пішли шляхом часткових дозволів депонування у тому його вигляді, у якому воно найменше їм загрожує. За даними каталогу RoMEO [14], наразі абсолютна більшість рецензованих наукових журналів (з тих, які не є журналами відкритого доступу) підтримують політику депонування у тому чи іншому вигляді: 63,2 % дозволяють депонування «постпринтів», 31,68 % – депонування препринтів. Що, як правило, не дозволяється видавцями, так це депонування вченими остаточних версій статей у форматах PDF та HTML. Деякі видавці, наприклад, Elsevier [2], взагалі не дозволяють при депонуванні конвертацію вченими робіт у формати PDF та HTML. Таким чином, при дотриманні цієї вимоги, сховища мають наповнюватись рукописами вчених у стандартних текстових форматах (DOC, ODT, RTF). У такому вигляді довіра вчених до інформації, отриманої з

\* Репозиторій (англ. repository) – сховище, склад.

цих сховищ, буде найменшою (психологічно файл у текстовому форматі, на відміну від файлу у форматі PDF або HTML, сприймається не як готове, належно оформлене дослідження, а як робота, що перебуває ще на етапі доопрацювання і не готова до публікації). Хоча деякі дослідження доводять [5], що насправді вчені часто не виконують умов ліцензійних угод видавців (фактично, вони їх просто не читають); автори практикують депонування статей незалежно від того, чи дозволяють це видавці (архівуються вченими 36 % статей з журналів, які забороняють депонування та 27 % – з тих, які його дозволяють), у 90 % випадків вчені конвертують свої роботи у формат PDF перед виставленням на сайті, а часто депонують навіть остаточні PDF-версії статей із сайтів видавців.

Підтримка сховищ даних установ – системний напрям діяльності академічних бібліотек. Він, серед іншого, зумовив переосмислення тих технологічних засобів, які використовуються бібліотеками, та появу нових типів бібліотечного програмного забезпечення. Для того, щоб ефективно адмініструвати процеси депонування наукових робіт, бібліотека повинна забезпечити науковцю простий онлайн-інструментарій для передачі робіт у сховище, а також відповідні засоби для подальшого ефективного використання (пошуку, завантаження, друку, отримання статистики тощо) електронного архівного фонду установи. Бібліотеками та іншими заінтересованими установами була створена ціла низка програмних систем для адміністрування таких сховищ. Серед найпоширеніших програмних рішень тут можна назвати дві системи: DSpace, яка, за даними OpenDOAR, використовується у 31 % проектів та EPrints, яка використовується, відповідно, у 17 % проектів [20]. Детально розглянемо першу систему.

Система DSpace – програмний продукт з відкритим вихідним кодом, створений у рамках дворічного проекту співпраці компанії Hewlett-Packard та бібліотечної системи Массачусетського технологічного інституту (США) [31]. Перша версія даного продукту була випущена 2002 р.

DSpace є програмною платформою, яка дає змогу налагодити відносно простий процес розподіленого адміністрування, поповнення та використання локального сховища індексованих файлів у різних форматах: текстові документи, мультимедійні матеріали, електронні таблиці тощо. Глобально система виконує три основні функції [16]: полегшення процесу поповнення сховища новими ресурсами та індексації нових надходжень; забезпечення простого доступу до ресурсів сховища, їх пошуку та перегля-

ду; забезпечення довготривалого зберігання електронних інформаційних ресурсів.

Програмний пакет DSpace – це цілісний сервер програмних додатків, призначений для встановлення на локальному апаратному веб-сервері установи, що підтримує сховище даних. Система DSpace є операційно незалежним рішенням, вона може функціонувати під різними операційними платформами (Linux, Windows, OS X тощо), вимоги до серверного оточення включають Oracle або PostgreSQL у якості серверу баз даних та Java у якості програмної платформи. Користувацький інтерфейс DSpace, станом на початок 2010 р., локалізований 23 мовами (в т. ч. і українською) [13].

Платформа DSpace створена для того, щоб максимально спростити для пересічних науковців установи (або групи установ) процеси роботи зі спільним сховищем даних безпосередньо з їх робочих місць: поповнення сховища власними ресурсами та використання ресурсів, вже наявних у сховищі. Тому, для ефективної роботи сховища на базі DSpace, важливо, перш за все, забезпечити правильну архітектуру майбутнього сховища, «змодельовати» у ньому організаційну або іншу структуру установи та виокремити в межах єдиного сукупного фонду колекції окремих структурних одиниць установи.

Всі користувачі взаємодіють з платформою віддалено (крім окремих функцій системних адміністраторів) через спільний для всіх веб-інтерфейс (принцип «веб як платформа»). Користувачі DSpace поділяються на три основні типи: анонімні користувачі, зареєстровані користувачі та адміністратори. Базова політика доступу передбачає, що анонімні користувачі (тобто всі відвідувачі сайта сховища) мають права на перегляд змісту (всього або обмеженої частини), зареєстровані користувачі можуть мати більшу або меншу кількість прав на поповнення сховища новими об'єктами та модифікацію існуючих об'єктів, адміністратори мають повні права доступу до всіх об'єктів сховища. Зареєстровані користувачі можуть об'єднуватись адміністраторами у групи для полегшення їх адміністрування (наприклад, можна створити групи науковців, технічних фахівців, викладачів тощо). Зареєстровані користувачі у системі наслідують ті права, які має група, до котрої вони входять, якщо інше не вказано додатково.

Платформа DSpace має цілу низку засобів багаторівневої аутентифікації користувачів: за логінами / паролями та за IP-адресами. Можливе налаштування аутентифікації за логінами / паролями через LDAP або Shibboleth. Це доволі корисно, як-

що єдина платформа установи передбачає наявність кількох окремих програмних серверів для обслуговування різних інформаційних процесів, і кожен з цих серверів передбачає аутентифікацію користувачів. У даному випадку доцільно запропонувати користувачам глобальну «наскрізну» аутентифікацію на платформі в цілому. Наприклад, у першій частині цього дослідження ми відзначали, що багато бібліотек використовують систему OJS у якості журнального серверу та DSpace – у якості сховища даних. Зрозуміло, користувацький веб-інтерфейс цих серверів може бути приведений до схожого вигляду. В результаті користувач «безшовно» переходить між серверами, не замислюючись, де «закінчується» функціональність журнального серверу та починається функціональність сховища даних». За такої ситуації доцільно інтегрувати механізми аутентифікації та зберігати дані облікових записів користувачів не у базах даних конкретних систем (OJS, DSpace), а у глобальному LDAP-сховищі паролів.

При створенні сховища даних адміністратор DSpace насамперед має створити *Розділи* (спільноти верхнього рівня), що репрезентуватимуть, наприклад, великі організаційні підрозділи установи, працівники яких будуть поповнювати сховище інформаційними ресурсами. Так, *Розділи* можуть створюватися для відділів установи, лабораторій, дослідних центрів тощо, або, якщо сховище є спільним для кількох установ, *Розділами* можуть бути окремі установи. Якщо існує потреба створити більш складну структуру сховища, у межах *Розділів* можуть створюватися *Підрозділи*. Кожний *Розділ* або *Підрозділ* сховища має власні *Колекції*. Крім того, окремі *Колекції* можуть бути спільними для кількох *Розділів* (корисно для депонування вченими матеріалів спільних дослідних робіт). У разі необхідності, права доступу до *Колекції* можуть обмежуватись і надаватись лише окремим користувачам. Для кожної *Колекції* з числа користувачів призначається адміністратор (один або декілька), який матиме повні права доступу до всіх *Об'єктів Колекції*.

Поповнення сховища новим *Об'єктом* може виконуватись у два або три етапи: передача, технологічний процес та архівування. Технологічний процес налаштовується за бажанням заінтересованих учасників взаємодії. Цей процес надає окремим, заздалегідь визначеним, користувачам сховища можливість нагляду за тими *Об'єктами*, які передаються іншими користувачами для додавання у сховище.

Щоб розпочати процедуру передачі *Об'єкта*, користувач повинен, насамперед, зайти на платформу

як зареєстрований користувач системи, обрати *Колекцію*, у яку він бажає (та має право) передавати ресурс, ввести базові метадані (за замовчуванням: автор, назва, анотація, бібліографічний запис першоджерела, дата публікації), вивантажити на сервер файл (або файли) роботи та прийняти запропоновану ліцензію, з якою зберігатиметься робота у сховищі. В межах єдиного *Об'єкта* сховища користувач може вивантажити на сервер декілька файлів; це корисно, наприклад, якщо вчений депонує матеріали власного виступу на конференції: в межах єдиного *Об'єкта* сховища він може вивантажити файли з тезами своєї доповіді, її повним текстом та презентацією. Або, якщо користувач передає у сховище гіпертекстовий документ, він може в межах єдиного *Об'єкта* завантажити сам HTML-файл та графічні файли ілюстрацій цього документа.

Якщо *Об'єкт* передається у *Колекцію*, для якої налаштований технологічний процес, користувач, котрий відповідає за опрацювання нових *Об'єктів*, на цьому етапі отримує повідомлення електронною поштою про те, що у його *Колекцію* передано новий *Об'єкт*. Система DSpace пропонує три кроки (типи) технологічного процесу: «дозволити / відмовити», «дозволити / відмовити / редагувати метадані» та «редагувати метадані». Можна налаштувати проходження *Об'єктом* будь-якої комбінації з цих трьох кроків технологічного процесу. Наприклад, *Об'єкт* може спочатку потрапити на опрацювання представникові керівництва установи для контролю за тим, щоб сховище не поповнювалося неякісними ресурсами (крок «дозволити / відмовити»: у разі відмови користувачу, який передав *Об'єкт* до *Колекції* може бути надісланий лист з причиною відмови). Потім новий *Об'єкт* може потрапити представнику бібліотеки для контролю за тим, щоб *Об'єкти* архівувалися у сховищі з коректними комплектами метаданих (крок «редагувати метадані»). Налаштування даних кроків відмінне у різних установах і залежить від політики наповнення сховища даних.

Для забезпечення ефективного використання ресурсів сховища даних, особливе значення мають метадані. На платформі DSpace власні комплекти метаданих мають *Розділи*, *Колекції* та *Об'єкти*. Ці дані використовуються для організації ресурсів у сховищі та забезпечення можливостей їх пошуку. За замовчуванням, DSpace налаштований лише на підтримку формату метаданих Dublin Core, хоча, при потребі, ці налаштування можуть бути доповнені. Таким чином, сховище на базі DSpace може використовувати будь-який формат метаданих (які не суперечать загальним правилам синтаксису метаданих DSpace).

Іншим важливим елементом забезпечення ефективного використання ресурсів сховища є постійні ідентифікатори *Об'єктів*. У системі DSpace використовується ідентифікація Handle. Характерною рисою системи Handle є унікальність та стійкість кожного ідентифікатора у межах системи загалом. Це дає змогу ідентифікувати, власне, зміст *Об'єкта*, а не режим доступу до нього за поточною URL-адресою або його характеристики (контекст його опису певною схемою метаданих). Кожний *Об'єкт*, на етапі включення до сховища даних, автоматично отримує власний унікальний ідентифікатор Handle, який у подальшому ідентифікує *Об'єкт* у межах глобального інформаційного поля (Інтернету в цілому). Навіть якщо адреса сховища даних в Інтернеті зміниться або декілька сховищ будуть злиті, ідентифікатор Handle залишиться працездатним, він вказуватиме на *Об'єкт* за його новою адресою. Для того, щоб дана функціональність працювала коректно, сховище даних на етапі налаштування повинно бути зареєстроване у Corporation for National Research Initiatives (CNRI) – організації, яка підтримує глобальний реєстр Handle. Реєструючи сховище, установа отримує від CNRI глобальний префікс ідентифікаторів Handle, який ідентифікуватиме сховище загалом через його наявність у структурі унікального ідентифікатора кожного *Об'єкта* цього сховища\*.

Нині систему DSpace використовують більше 700 установ по всьому світу [32]. Типово DSpace використовується для створення сховищ інформаційних ресурсів наукових та освітніх установ [16; 32]. Деякі установи використовують систему для публікації електронних версій наукових журналів [12]. Фактично, онлайнове сховище рецензованих статей, якщо воно супроводжується інструментарієм для проведення інформаційного пошуку, саме по собі починає виконувати функцію наукового журналу [29]. Тобто, з точки зору читача, складно провести межу між сховищем статей та журнальною платформою, оскільки сховища також можна поповнювати з певною періодичністю, групувати підбірки статей у випуски та номери або систематизувати їх іншим чином.

Більше того, розвиток сховищ відкритого доступу спричинив виникнення істотно нового типу наукових часописів – оверлейних\*\* журналів (overlay journal). Оверлейні журнали – це наукові видан-

ня відкритого доступу, майже завжди електронні, які публікують зміст, відібраний зі сховищ даних або інших відкритих джерел інформації. Такі журнали, як правило, не практикують проведення процедури повномасштабного наукового рецензування статей. Деякі оверлейні журнали є тематичними, вони відбирають у власні випуски наукові статті з вузьких тем, які вже публікувалися в інших часописах; інші – відбирають зі сховищ цікавий зміст, який ще не оприлюднювався. Оверлейні журнали відрізняються за своїм призначенням та редакційною політикою. Зокрема, одним з найбільш успішних [8] журналів такого зразка є український науковий часопис SIGMA відділу прикладних досліджень Інституту математики НАН України, який оголосив себе оверлейним журналом сховища arXiv.org. До друку у цьому часописі приймаються лише ті роботи, препринти яких розміщені у сховищі arXiv.org. Таким чином, процедура подання авторами рукописів у SIGMA завжди повинна розпочинатися з їх депонування у arXiv.org [30].

Наразі всесвітня мережа сховищ даних наукових та освітніх установ є дуже великою. Всього, за даними OpenDOAR [20], станом на квітень 2010 р. у світі нараховується понад півтори тисячі сховищ даних (81 % з них – сховища даних установ), в яких сукупно зберігається понад 20 млн об'єктів. Переважна частина сховищ сконцентрована у Європі (48 %) та Північній Америці (27 %). Серед існуючих сховищ 62 % наповнюються журнальними статтями, 50 % – текстами дисертацій, 41 % – неопублікованими науковими матеріалами, 35 % – матеріалами конференцій, 31 % – монографіями та їх розділами, 23 % – мультимедійними матеріалами тощо.

Для академічних бібліотек підтримка сховищ даних власних установ є важливим напрямом діяльності з декількох міркувань. Безперечно важливість цієї діяльності у контексті зберігання наукового доробку вчених власної установи; академічна бібліотека стає збирачем, хранителем та розпорядником «колективного інтелектуального капіталу» установи, розпорошеного по тисячах різних джерел. У цій якості вона робить суттєвий крок на шляху до підвищення статусу та репутації бібліотеки як академічного інституту. Крім того, розвиток сховищ даних – важливий компонент «нової видавничої моделі» [10, с. 4], він надає бібліотекам арсенал засобів у сфері забезпечення авторських прав вчених, забезпечує позитивне міжнародне сприйняття наукового, соціального та фінансового потенціалу установи, надає додаткові аргументи для діалогу з видавничими монополіями, розширює коло взаємодії бібліотеки з ученими як доно-

\* Детальніше про принципи організації Handle-взаємодії див. [3].

\*\* Оверлейний (англ. overlay) – такий, що ґрунтується на чомусь іншому, утворює новий шар чогось іншого.

рами та акцепторами в процесах інформаційного обміну. Більше того, сховища даних, як говорилося вище, самі, у певному сенсі, виконують функції періодичних видань: якщо вчений передає до сховища ще неопублікований зміст, акт його прийняття на зберігання у сховище можна вважати актом публікації матеріалу; якщо зміст вже публікувався – актом його перевидання. Разом з тим, залежно від політики наповнення сховища даних, його повністю або частково можна розглядати як електронний архів установи, тобто як стале джерело отримання відомостей стосовно історії установи та діяльності окремих вчених. Нарешті, отримуючи дані щодо публікаційної активності вчених, академічна бібліотека, фактично, виходить на рівень аудитора не тільки наукового процесу загалом, а й окремих ланок науки; стає можливим застосування кількісних методів оцінки наукового потенціалу установи та вироблення на цій основі системи індикаторів розвитку її академічного середовища.

Інтерес бібліотек до розвитку сховищ даних науково-інформаційних ресурсів, як вже говорилося, був спричинений, насамперед, системною кризою галузі та необхідністю застосування інноваційних підходів у питаннях інформаційного супроводу академічного процесу. Ціла низка таких інноваційних високотехнологічних підходів пов'язана з питаннями архівування та зберігання ресурсів. Розвиток сховищ даних – потужний, але не єдиний підхід академічних бібліотек до реалізації даного завдання. Розглянемо питання зберігання академічними бібліотеками онлайнних науково-інформаційних ресурсів, доступних на умовах передплати.

Поступове перенесення процесів науково-інформаційного забезпечення у електронне середовище поставило перед бібліотеками низку додаткових проблем. Однією з найважливіших з них є проблема контролю [6, с. 221]. Контроль, у даному аспекті, ми розуміємо як упевненість у тому, що певний ресурс буде доступний бібліотечним користувачам у майбутньому. Бібліотека може контролювати ресурси, якими вона володіє фізично, але вона не завжди здатна забезпечити контроль за тими ресурсами свого фонду, які зберігаються поза межами бібліотеки, зокрема, за передплатними онлайнними ресурсами, що зберігаються на зовнішніх серверах видавців та інформаційних агрегаторів і до яких бібліотека має «лише доступ». Цей «лише доступ» докорінно змінює роль бібліотек у суспільстві, яку вони відігравали протягом століть – були довіреними хранителями наукового та культурного надбання [23]. З розповсюдженням онлайнних технологій, традиційну для бібліотек функцію зберігання

інформаційних ресурсів взяли на себе видавці [10, с. 12], бібліотеки ж перетворилися на орендарів інформаційних ресурсів, повністю залежних від поточної та майбутньої кон'юнктури ринку.

В умовах, коли національні бібліотеки, видавничі монополії беруть на себе обов'язки по збереженню наукових ресурсів на електронних носіях, для академічних бібліотек ситуація не була б такою критичною, якби не кризові явища у науково-видавничій галузі. Проте, у реаліях, що склалися, поняття «зберігання для суспільства» і «зберігання для користувачів окремої бібліотеки» істотно відрізняються. Так, наукова бібліотека, оформлюючи на поточний рік передплату на певний онлайнний ресурс, не впевнена, що у наступному році вона цю передплату зможе продовжити. Більше того, втрачаючи доступ до платформи видавця або агрегатора, бібліотека втрачає не лише доступ до поточних ресурсів, але й доступ до архіву, до тих ресурсів, які донедавна були доступні її користувачам. Здавалося б, закономірний шлях вирішення даної проблеми для бібліотек – фізичне переміщення передплатених ресурсів на власні сервери та створення локальних колекцій цих ресурсів на бібліотечних сайтах. Але це абсолютно не вигідно видавцям, вони бажають мінімізувати процеси безконтрольного розповсюдження власного змісту, а тому створення бібліотеками локальних повнотекстових колекцій ресурсів, що передплачуються, у більшості випадків, суворо забороняється ліцензійними угодами. Існує інший шлях: якщо бібліотека купує у видавців «лише доступ», вона могла б вимагати, принаймні, щоб цей «лише доступ» був «вічним» (тобто умови, за яких бібліотеці гарантуються права «вічного» доступу до ресурсів видавця, опублікованих у роки передплати з мінімальною платною лише за технічний супровід цього доступу). Видавці навіть самі пропонують такі умови доступу, певною мірою це вигідно їм (нижче ми поговоримо про це). Але гарантії «вічності» цього доступу в будь-якому разі сумнівні для бібліотек [25]. Тому бібліотеки шукають альтернативні (технологічні) шляхи забезпечення власним вченим режиму сталого доступу до онлайнних наукових ресурсів. Ті рішення, які вони знаходять, іноді стимулюють виникнення цілих напрямів розвитку інформаційних технологій.

З огляду на вищесказане розглянемо доволі популярну сьогодні технологічну парадигму зберігання електронних ресурсів LOCKSS (Lots Of Copies Keper Stuff Safe). В ядрі цієї концепції лежить реалізація елементарних принципів Інтернет-взаємодії. Але, перш ніж розглядати функціональність систе-

ми LOCKSS як програмної інфраструктури, згадаємо базові принципи взаємодії користувачів всесвітнього павутиння з віддаленими серверами.

Базова працездатність всесвітнього павутиння передбачає, що перегляд користувачем документа з віддаленого веб-сайта можливий лише після завантаження цього документа з віддаленого серверу у спеціальну теку тимчасових файлів (кеш-теку) на локальному комп'ютері користувача. Але після перегляду файл не знищується на локальному комп'ютері одразу. Він ще певний час зберігається у кеш-теці, на випадок, якщо користувач захоче повернутися до документа, з яким він нещодавно працював (щоб знову не завантажувати його з віддаленого серверу). Таким чином, якщо вчений читає статті наукових журналів з комп'ютера, то на ньому завжди, хоче він того чи ні, зберігатиметься (у прихованому вигляді), так би мовити, «локальна колекція журнальних статей». Якщо ж протягом короткого проміжку часу вчений двічі захоче прочитати одну й ту саму статтю, то перший раз він знайде на платформу видавця або агрегатора, знайде статтю, вона завантажиться з віддаленого серверу і відкриється. Другим разом все буде так само, але стаття вже не завантажуватиметься, а відкриватиметься безпосередньо з локального комп'ютера (з теки тимчасових файлів). Звичайно ж, у межах локальної мережі дану функціональність можна розширити – створити проміжний кеш-сервер між веб-сайтами видавців та комп'ютерами у читальних залах бібліотеки, на якому вічно зберігатимуться як кеш-файли ресурси, до яких бібліотека має (або мала в минулому) авторизований доступ. Отже, бібліотека буде мати власну локальну копію того змісту, який вона купила у видавців, ця копія змісту перебуватиме цілковито під контролем бібліотеки, але доступ до цього змісту читачі отримуватимуть за місцем його публікації (тобто з технологічних платформ видавців).

Ми зупинилися на описі цього нескладного алгоритму для того, щоб у подальшому була більш зрозумілою функціональність системи LOCKSS як спеціалізованого програмного рішення для створення, підтримки та використання електронних архівних колекцій бібліотек. Скринька LOCKSS виконує три основні функції: збирає ресурси, зберігає їх та видає, в разі потреби, користувачам. На відміну від низки інших програмних засобів для створення архівів електронних ресурсів вона не збагачує ресурси метаданими і не має публічного інтерфейсу для кінцевих користувачів архіву. На етапі збору та видачі ресурсів LOCKSS працює як кеш-сервер файлів з видавничих веб-сайтів.

Щоб уникнути плутанини, відзначимо: LOCKSS – це спеціалізоване програмне рішення саме для журнального (та іншого статичного) веб-змісту. Ця система не призначена для архівування змісту веб-середовища загалом. Гарантування збереженості глобальних веб-сегментів (електронного документного надбання націй) – завдання національних бібліотек, і у них для цього є свої відмінні технологічні рішення. Статті у наукових часописах вирізняються тим, що вони публікуються одноразово, їх зміст не змінюється з часом; крім того, архіви часописів завжди мають чітку ієрархію; LOCKSS пропонує рішення саме для такого типу веб-змісту [25].

Отже, згідно з самовизначенням розробників, LOCKSS (Lots Of Copies Keep Stuff Safe) – це міжнародна ініціатива бібліотечної системи Стенфордського університету (США), в рамках якої заінтересованим бібліотекам надається інструментарій та необхідний супровід щодо запровадження легкого та економічного збору і зберігання їх власних копій авторизованого веб-змісту [17]. Як технологічне рішення, LOCKSS є програмним пакетом з відкритим вихідним кодом, який дає змогу організувати збір веб-ресурсів у всіх форматах, налаштувати однорангову децентралізовану інфраструктуру зберігання цих ресурсів та забезпечити доставку зібраних ресурсів на комп'ютери в локальній мережі бібліотеки. Програму LOCKSS започаткували у 1998 р., протягом наступних кількох років були випущені перші версії програмного пакета та проведена серія тестувань нової технології. До неї долучалися Бібліотека Конгресу, Національна сільськогосподарська бібліотека (США), Британська бібліотека, 40 потужних академічних бібліотек та 53 видавці науково-інформаційних ресурсів [27].

LOCKSS – програмний пакет, який працює на одному з локальних комп'ютерів бібліотеки, що має власну IP-адресу (з діапазону тих, які авторизовані ліцензійними угодами з видавцями). З міркувань захисту системи пакет LOCKSS не встановлюється на комп'ютер як звичайна програма, а записується на захищений носій. Тут потрібно нагадати, що будь-яка система, підключена до глобальної мережі, досить вразлива щодо зовнішніх атак та вірусів. Тому система зберігання електронних ресурсів завжди повинна бути або надійно захищена від небажаних зовнішніх впливів, або повністю ізольована від мережі [26]. У випадку LOCKSS застосовується перший підхід; після конфігурації система записується на компакт-диск без можливості перезапису й надалі працює у незмінному вигляді. LOCKSS



має власну захищену операційну систему та власний комплект драйверів для комп'ютерного обладнання, тому для роботи їй не потрібно жодного програмного оточення, лише персональний комп'ютер, підключений до Інтернету та локальної мережі, на якому є пристрій для читання компакт-дисків та носій з місткістю, достатньою для створення на ньому архіву веб-ресурсів. Комп'ютер, який працює під керуванням системи LOCKSS, називається скринькою LOCKSS (LOCKSS Box).

Насамперед, видавець (або розробник іншого веб-сайта) повинен дати дозвіл системі LOCKSS на збирання та зберігання веб-змісту. Тут потрібно зауважити, що серед архівістів немає наразі єдиної думки щодо того, чи потрібний дозвіл на архівування веб-змісту. Наприклад, фахівці проекту Internet Archive сповідають підхід «питати дозволу – заохочувати відмови» і архівують зміст, захищений авторським правом, без отримання дозволів на це у правовласників [26]. Розробники LOCKSS керуються іншою філософією, тому системою архівується лише той зміст, який дозволяють архівувати його власники. Дозволом на архівування для LOCKSS-скриньки є або відповідна ліцензія Creative Commons на сайті, або текст, де вказано, що система LOCKSS може збирати і зберігати веб-зміст ресурсу [24].

Скринька LOCKSS за допомогою спеціального кравлера, схожого на той який використовують пошукові машини Інтернету, з певною частотою збирає новий зміст певних розділів визначених веб-сайтів і завантажує його як кеш на локальний носій даних (жорсткий диск комп'ютера). Оскільки кравлер заходить на сайт видавця з IP-адреси бібліотеки (передплатника), він має доступ до повнотекстового змісту, доступного на умовах передплати. Інформація про веб-сайти видавців зберігається у спеціальних XML-файлах системи LOCKSS. В цих файлах вказано місце на сайті, де зберігається файл з дозволом на архівування, які розділи сайту потрібно архівувати, який час та які дні є оптимальними для збирання даних тощо. Директорію цих XML-файлів підтримує бібліотечна система Стенфордського університету. Кожна скринька LOCKSS працює повністю автономно і збирає веб-зміст незалежно від інших.

Оскільки інформаційні ресурси в електронних архівах традиційно дуже рідко використовуються, жодна система зберігання електронних ресурсів не може у питаннях аудиту цілісності та автентичності архіву розраховувати на дані щодо взаємодії користувачів з архівом. Окремі дані можуть бути пошкодженими і у такому вигляді зберігатися де-

сятками років, поки кінцевий користувач не спробує до них звернутися. Таким чином, необхідно забезпечити регулярний аудит цілісності всіх даних у архіві за допомогою спеціальних програмних рішень [26]. У системі LOCKSS аудит цілісності архіву проводиться засобами однорангової децентралізованої інфраструктури. Час від часу LOCKSS-скриньки різних бібліотек з'єднуються через особливий протокол LCAP (Library Cache Auditing Protocol) і за допомогою спеціального алгоритму голосування порівнюють цілісність та автентичність веб-змісту, наявного у них на зберіганні. Якщо у одній зі скриньок зміст виявляється неповним або пошкодженим, він відновлюється з веб-сайта видавця або з інших скриньок LOCKSS. Чим більше скриньок LOCKSS зберігають певний веб-зміст, тим надійнішим є механізм його зберігання. Для забезпечення надійності зберігання веб-змісту рекомендується, щоб зберігалися принаймні сім його реплік (копій) [24].

Зберігання численних копій змісту у територіально розподілених сховищах – базова стратегія побудови систем цифрового зберігання. Такий підхід застосовується в архівній системі Британської бібліотеки, системі DAITSS Центру бібліотечної автоматизації у Флориді (США) та багатьох інших проектах [26]. Якщо сховища друкованих ресурсів потерпають від пожеж, стихійних лих, крадіжок та інших непрогнозованих подій, то у випадку з електронними ресурсами тим більше не можна гарантувати збереженості документа, який має єдину копію. Крім небезпек, які загрожують друкованим зібранням, виникають й інші. Зокрема, у випадку «світлич архівів», до яких мають доступ широкі кола громадськості, існує загроза нападу агресивно налаштованих хакерів-фанатиків тієї чи іншої ідеї. У системі LOCKSS механізмом запобігання даної небезпеки є зберігання реплік у численних територіально та організаційно розподілених LOCKSS-скриньках. Скриньки мають інформацію одна про одну, вона необхідна для проведення голосувань щодо автентичності змісту. Теоретично потенційний зловмисник може, напавши на одну зі скриньок, використати списки LOCKSS-скриньок для цілеспрямованого нападу на всі місця зберігання певного змісту. Тому в системі LOCKSS передбачено спеціальний механізм постійного оновлення цих списків таким чином, щоб жодна LOCKSS-скринька ніколи не мала повного списку інших скриньок, у яких зберігаються репліки певного змісту. Це дає змогу гарантувати, навіть у випадку вдалого цілеспрямованого нападу на чисельні LOCKSS-скриньки, що зміст все одно буде автома-

тично відновлений у всіх скриньках, де він зазнав пошкоджень [28].

Якщо LOCKSS-скринька правильно сконфігурована у локальній мережі установи, вона забезпечує доставку необхідних файлів на комп'ютери кінцевих користувачів, у випадках, коли сайти видавців з певних причин виявляються недоступними. Така доставка можлива за допомогою двох основних методів: прозорого проксі-серверу або обслуговування змісту. У першому випадку скринька LOCKSS інтегрується з проксі-сервером установи. Коли читач із комп'ютера у локальній мережі робить запит до сайту видавця (вводить адресу URL або переходить за посиланням – не важливо), проксі-сервер перехоплює цей запит і перенаправляє його скриньці LOCKSS, яка, в свою чергу, перенаправляє цей запит знову до сайту видавця. Якщо сайт видавця надає зміст, скринька LOCKSS персилає його кінцевому користувачу. Якщо сайт видавця не надає змісту (він тимчасово недоступний, термін передплати на зміст сплив, сталася помилка на сервері тощо), скринька LOCKSS доставляє кінцевому користувачеві зміст із власного архіву. При цьому кінцевий користувач не повинен навіть знати, що в локальній мережі існує LOCKSS-функціональність. Другий метод (обслуговування змісту) більш потужний, він пов'язаний із застосуванням OpenURL-функціональності. Детально питання використання технології OpenURL розглядатиметься у наступній частині дослідження, коли йтиметься про інтеграцію гетерогенних інформаційних потоків академічними бібліотеками, тому тут не будемо зупинятися на цьому питанні.

Для забезпечення можливостей сталого доступу до збереженого веб-змісту протягом багатьох років у системі LOCKSS передбачено механізм міграції форматів. Всі документи LOCKSS-скринька зберігає у їх оригінальних форматах, але якщо веб-браузер користувача виявляється не здатним коректно прочитати зміст певного веб-ресурсу (документ або його частина у старому форматі, який вже не підтримується сучасними веб-браузерами), LOCKSS-скринька може, за запитом, створити тимчасову копію документа в тому форматі, який здатен прочитати браузер користувача.

Для бібліотек система LOCKSS цікава з багатьох позицій. Не остання з них – економічність даного рішення. До появи LOCKSS бібліотеки, які мали бюджети, достатні для впровадження повноцінних систем зберігання електронних ресурсів, можна було перелічити на пальцях [26]. Якісне зберігання електронних ресурсів, традиційно, дуже витратний напрям діяльності установ. LOCKSS зробив мож-

ливим навіть для невеликих бібліотек, які не мають коштів зберігання електронних ресурсів, забезпечення формування власних електронних архівів [24]. LOCKSS-скринька не вимагає ані дорогого обладнання, ані постійної людської підтримки. У питаннях обладнання, розробники навіть рекомендують використовувати для створення LOCKSS-скриньок застарілі комп'ютери (оскільки новітні апаратні засоби, частіше всього, пропонують збільшену потужність за рахунок зниженої надійності). Що стосується підтримки, сконфігурований LOCKSS вимагає втручання системного адміністратора у еквіваленті від 15 хвилин до однієї робочої години на місяць [17].

Система LOCKSS відтворює для академічних бібліотек традиційну для них модель передплати, з якою вони працювали сотні років, тобто можливість передплачувати ресурси (купувати копії, а не брати їх в оренду) не лише для нинішніх, але й для майбутніх користувачів. Отже, передплачені онлайнові ресурси стають такою самою сталою частиною бібліотечного фонду, як і друковані видання, до них можуть застосовуватися всі традиційні технологічні процеси обробки.

Станом на кінець 2009 р. до Публічної мережі LOCKSS входить більше 200 бібліотек та інформаційних центрів з усього світу [24]. Крім академічних бібліотек, серед членів мережі також Бібліотека Конгресу, Британська бібліотека, Баварська державна бібліотека, Італійський, Канадський та Французький інститути науково-технічної інформації, Національна бібліотека ПАР, Національна науково-технічна бібліотека Німеччини, Нью-Йоркська публічна бібліотека, OCLC тощо. Приєднання до Публічної мережі для бібліотек безкоштовне. Воно дає змогу брати участь у спільному архівуванні вільно доступного журнального веб-змісту. Але більшість бібліотек входять не лише до Публічної мережі, а й до Альянсу LOCKSS. Альянс LOCKSS (LOCKSS Alliance) – це об'єднання академічних бібліотек – передплатників науково-інформаційних ресурсів, які спільно працюють над удосконаленням програмної інфраструктури LOCKSS та виробляють стратегічну політику ініціативи. Членство в Альянсі дає змогу архівувати також зміст, який бібліотеки мають на умовах передплати. Більшість бібліотек через членство в Альянсі мають змогу архівувати більшу частину змісту, який вони передплачують. Середній показник реплікації (кількості архівних копій певного змісту) у Публічній мережі LOCKSS станом на 2009 р. становить приблизно 40 [24]. Академічні бібліотеки Великої Британії заснували власний Альянс

(UK LOCKSS Alliance), вони проводять власну політику у переговорах з видавцями та мають змогу архівувати додатковий передплатний веб-зміст.

Крім Публічної мережі існують також Приватні мережі LOCKSS, закриті лише для бібліотек-членів (типово від семи до п'ятнадцяти). Це, переважно, «темні архіви», до яких немає публічного доступу. У таких мережах бібліотеки забезпечують за допомогою технології LOCKSS колективне зберігання особливих типів ресурсів: оцифрованих бібліотечних фондів, власних веб-сайтів, урядових документів тощо. Крім того, всі члени Приватних мереж LOCKSS є також членами Альянсу LOCKSS [24], вони беруть участь у колективному архівуванні безкоштовних та передплатних науково-інформаційних ресурсів.

Технологія LOCKSS вигідна не лише бібліотекам. У певному сенсі вона також вигідна видавцям. Система LOCKSS нічого не порушує у видавничих механізмах розповсюдження змісту, вона лише зберігає файли веб-сайта видавця на локальному комп'ютері бібліотеки як кеш-файли та надає ці файли користувачеві з локальної мережі бібліотеки, працюючи як проміжний проксі-сервер. Для видавців це важливо в кількох аспектах. По-перше, видавці зацікавлені в тому, щоб бібліотеки передплачували саме онлайнві науково-інформаційні ресурси, а не друковані. Модель друкованого видання витратна, вона не подобається науковцям, а самим видавцям не дає змоги отримати очікуваних прибутків. Багато потужних видавців із задоволенням взагалі б відмовилися від друкованих видань [27]. По-друге, видавці не зацікавлені у безконтрольному розповсюдженні власних науково-інформаційних ресурсів. Найгірший варіант такого безконтрольного розповсюдження – це створення повнотекстових колекцій ресурсів на локальних веб-сайтах. Потужні видавці тримаються на іміджі та престижі, який вони мають серед пересічних науковців [27]. Левову частку своїх прибутків вони витрачають саме на підтримку цього іміджу. Веб-платформи престижних видавців завжди мають професійний веб-дизайн, якісну графіку, зручний інтерфейс та, що найважливіше, потужний пошуковий та навігаційний апарат. Вони створюються для того, щоб жоден науковець не пішов з такої платформи незадоволеним. Якщо ресурс видавця наявний у локальній колекції, зрозуміло, імідж видавця втрачає від традиційно невисокої якості оформлення такої колекції. Тому, вивчаючи потреби науковців та оптимальні шляхи їх задоволення, потужні видавничі корпорації впроваджують цілу низку високотехнологічних засобів, відслідкову-

ють «кожний рух» користувачів на своїх видавничих платформах, досліджують залежності між тими ресурсами, до яких має доступ науковець, та тими, які він цитує в своїх працях, і створюють типові моделі поведінки науковців. Це серйозні дослідження, вони важливі для стратегічного розвитку видавничих платформ. Але безконтрольне розповсюдження ресурсів завжди вносить серйозну похибку в розроблення цих моделей. Видавці прагнуть мати доступ до вичерпної статистики використання своїх ресурсів, у т. ч. до даних щодо процесів взаємодії користувачів з ресурсами [27]. Вони забороняють бібліотекам-передплатникам створення локальних повнотекстових колекцій, але дослідники, невпевнені у фінансовій стабільності своїх бібліотек, все одно їх створюють, і з цим важко боротися. Технологія LOCKSS вирішує багато проблем видавців: ресурси не можуть безконтрольно розповсюджуватись зі скриньок LOCKSS; система побудована таким чином, що дані з кеш можуть видаватися за запитом лише дружнім скринькам LOCKSS для перевірки їх цілісності та користувачам з локальної мережі бібліотеки. При цьому взаємодія користувачів з ресурсами залишається повністю прозорою для видавців [27]; поведінка науковців на видавничій платформі ніяк не змінюється, вони так само шукають ресурси за допомогою пошукових та навігаційних засобів платформи, читають певні ресурси та роблять інші дії – всі ці дані доступні видавцю.

Станом на листопад 2009 р. до ініціативи LOCKSS приєдналося понад 400 видавців, які сукупно забезпечують бібліотекам дозволи на архівування 2 тис. назв наукових видань [28], у т. ч. 250 видавців дозволяють архівувати зміст усіх або окремих їхніх видань всім зацікавленим установам, для архівування змісту інших видавців вимагається членство в Альянсі LOCKSS.

Розвиток технологічної парадигми LOCKSS не залишився поза увагою потужних видавничих монополій. Як вже говорилося, LOCKSS має цілу низку вигідних для видавців рішень, до яких можна додати й таке – видавці, напевно, більше, ніж всі інші соціальні групи, зацікавлені у гарантуванні збереженості власного веб-змісту, і LOCKSS, як прикладне рішення для створення архівів, є прекрасним інструментом забезпечення такої гарантії. Проте деякі ланки цієї технології монополістам не завжди вигідні, зокрема, у частині повної відкритості LOCKSS-скриньок для кінцевих користувачів. Гіганти видавничої індустрії беруть участь у ініціативі LOCKSS, але, в основному, надають дозволи на архівування лише окремих назв своїх видань (ви-

давничий концерн Elsevier станом на 2010 р. не представлений взагалі). Але LOCKSS надає видавцям значно більші можливості щодо укріплення власної репутабельності. Як вже говорилося, видавці взагалі зацікавлені перейти до суто онлайн-моделі видання. Але тоді вони втрачуть свою репутабельність в очах значної частини вчених. У найбільш загальному вигляді, це «проблема геніїв», які пишуть свої праці «не для сучасників, а для нащадків». Протягом століть передачу наукових знань від покоління до покоління забезпечували бібліотеки як хранителі друкованої спадщини, і в цій ролі вони заслужили довіру академічної спільноти. З приходом онлайн-технологій все змінилося, тепер у вченого немає гарантії, що навіть через 10 років колеги зможуть прочитати його статтю в науковому журналі, якщо вже сьогодні лише окремі бібліотеки можуть надати до неї доступ. Гарантії з боку самих видавничих монополій ненадійні, оскільки вони є комерційними підприємствами, які можуть банкрутувати та припинити доступ до певної частини змісту з фінансових міркувань. Найпотужніші видавці (Elsevier, Springer, Wiley-Blackwell, Taylor & Francis та ін.) зробили ставку на престиж Королівської бібліотеки (Нідерланди) та поклали на неї функції вичерпного зберігання всього власного змісту (проект e-Depot [15]). Проте, якою б потужною й репутабельною не була Королівська бібліотека, це лише одна інституція в світі, і цього мало для повного гарантування збереженості наукових ресурсів. Таким чином, видавцям вигідно, щоб їх зміст як у електронній, так і у друкованій формах зберігався в академічних бібліотеках по всьому світу, це додає репутабельності їхнім виданням в очах академічної спільноти. З цією метою у 2006 р. провідні світові постачальники науково-інформаційного змісту об'єдналися з потужними науковими бібліотеками задля створення глобального «темного архіву» науково-інформаційних ресурсів. Ця ініціатива отримала назву CLOCKSS (Контрольований LOCKSS).

Технічно CLOCKSS реалізується приблизно так само, як і LOCKSS, використовується той самий пакет програмного забезпечення. Відмінність полягає у тому, що, по-перше, членство у CLOCKSS платне як для видавців, так і для бібліотек, і зміст архіву є закритим, за звичайних обставин користувачі бібліотек не мають до нього доступу. Але у випадку, якщо виникає одна з чотирьох обставин (видавець припиняє своє існування, доступ до назви видання більше не пропонується на сайті видавця, видавець припиняє доступ до архівних випусків видання або виникає форс-мажорна обставина, в

результаті якої видавець не має змоги забезпечувати онлайн-доступ до видання), бібліотеки-учасниці забезпечують надання змісту у вільний доступ всім бажаючим користувачам Інтернету. Станом на квітень 2010 р. у проекті беруть участь 23 потужні видавці (Elsevier, Nature Publishing Group, SAGE Publications, Springer, Taylor & Francis, Wiley-Blackwell та ін.), 10 бібліотек та OCLC, які підтримують локальні CLOCKSS-скриньки та 35 бібліотек, котрі забезпечують супровід проекту. За час існування проекту у вільний доступ потрапило 2 назви журналів SAGE Publications та одна назва Oxford University Press [9].

Отже, у сфері забезпечення інфраструктури зберігання електронних фондів протягом останніх 10–15 років академічні бібліотеки зробили суттєвий крок вперед. Функція бібліотеки як хранителя документної спадщини завжди була дуже важливою для формування її інституційного статусу. Із упровадженням новітніх технологій ця функція зазнала істотного розширення. З одного боку, для бібліотек стали доступними нові засоби і методи забезпечення збереженості власних фондів на різних видах носіїв, а з другого – нові канали розповсюдження інформації породили для бібліотек цілу низку нових проблем, вирішення яких вимагає інноваційного бачення бібліотечної діяльності.

### Список використаних джерел

1. Будапештская инициатива «Открытый доступ» [Электронный ресурс]. – Режим доступа : URL : [www.soros.org/openaccess/ru/read.shtml](http://www.soros.org/openaccess/ru/read.shtml). – Загл. с экрана.
2. Захарова, Г. М. Интернет-технологии: открытый доступ и самоархивирование / Г. М. Захарова // Научные и технические библиотеки. – 2006. – № 4. – С. 58–66.
3. Соловяненко, Денис. Цифровий ідентифікатор об'єкта (DOI): «ISBN суспільства знань» / Д. Соловяненко // Бібл. вісн. – 2009. – № 4. – С. 3–15.
4. Albert, K. M. Open access: implications for scholarly publishing and medical libraries / Karen M. Albert // Journal of the Medical Library Association. – 2006. – Vol. 94. – N. 3. – P. 253–262.
5. Antelman, K. Self-archiving practice and the influence of publisher policies in the social sciences / Kristin Antelman // Learned Publishing. – 2006. – Vol. 19. – N. 2. – P. 85–95.
6. Budd, J. The changing academic library: operations, culture, environments / John M. Budd. – Chicago: Association of College and Research Libraries, 2005. – V, 323 p.
7. Case, M. M., John, N. R. Opening up scholarly information at the university of Illinois at chicao [Virtual Resource] / Mary M. Case, Nancy R. John // First Monday. – 2007. – Vol. 12. – Is. 10. – Access Mode : URL : [firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1956/1833](http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1956/1833).

8. Cassella, M., Calvi, L. New journal models and publishing perspectives in the evolving digital environment [Virtual Resource] / Maria Cassella, Licia Calvi // Proceedings of the World Library and Information Congress: 75th IFLA General Conference and Assembly «Libraries create futures: Building on cultural heritage» (Italy, Milan, 23–27 August 2009). – Access Mode : URL : [www.ifla.org/files/hq/papers/ifla75/179-calvi-en.pdf](http://www.ifla.org/files/hq/papers/ifla75/179-calvi-en.pdf).
9. CLOCKSS [Virtual Resource]. – Access Mode : URL : [www.clockss.org/clockss/Home](http://www.clockss.org/clockss/Home). – Title from Screen. – Date of Access: 6 April 2010.
10. Crow, R. The Case for Institutional Repositories: A SPARC Position Paper / Raym Crow; SPARC. – Washington: SPARC, 2002. – 37 p.
11. DOAJ: Directory of Open Access Journals [Virtual Resource]. – Access Mode : URL: [www.doaj.org/articles/questions](http://www.doaj.org/articles/questions). – Title from Screen. – Date of Access: 10 March 2010.
12. Hahn, K. L. Research Library Publishing Services: New Options for University Publishing / Karla L. Hahn; Association of Research Libraries. – Washington: Association of Research Libraries, 2008. – 40 p.
13. I18nSupport [Virtual Resource] // DSpace Wiki. – Access Mode : URL : [wiki.dspace.org/index.php/I18nSupport](http://wiki.dspace.org/index.php/I18nSupport). – Title from Screen. – Date of Access: 6 April 2009.
14. Journal Policies – Summary Statistics So Far [Virtual Resource] // RoMEO. – Access Mode : URL : [romeo.eprints.org/stats.php](http://romeo.eprints.org/stats.php). – Title from Screen. – Date of Access: 14 November 2009.
15. The KB's e-Depot: a trustworthy steward for the digital scholarly record / Koninklijke Bibliotheek – National library of the Netherlands. – Access Mode : URL : [www.kb.nl/dnp/e-depot/operational/suppliers/national\\_suppliers-en.html](http://www.kb.nl/dnp/e-depot/operational/suppliers/national_suppliers-en.html). – Date of Access: 14 November 2009.
16. Lewis, S., Yates, C. DSpace Course – An Introduction to DSpace [Virtual Resource] / Stuart Lewis, Chris Yates. – Access Mode : URL : [hdl.handle.net/2160/617](http://hdl.handle.net/2160/617). – Title from Screen. – Date of Access: 13 November 2009.
17. LOCKSS [Virtual Resource]. – Access Mode : URL : [www.lockss.org/lockss/Home](http://www.lockss.org/lockss/Home). – Title from Screen. – Date of Access: 15 November 2009.
18. Morrison, H. Rethinking collections – Libraries and librarians in an open age: A theoretical view [Virtual Resource] / Heather Morrison // First Monday. – 2007. – Vol. 12, Is. 10. – Access Mode : URL : [firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1965/1841](http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1965/1841).
19. Morrison, H. G. Scholarly Communication for Librarians / Heather Morrison. – Oxford: Chandos, 2009. – 264 p. – Access Mode : URL : [eprints.rclis.org/16282/1/Chapter\\_6openaccess.pdf](http://eprints.rclis.org/16282/1/Chapter_6openaccess.pdf).
20. OpenDOAR: Directory of Open Access Repositories [Virtual Resource]. – Access Mode : URL : [www.opendoar.org](http://www.opendoar.org). – Title from Screen. – Date of Access: 13 November 2009.
21. PubMed Central [Virtual Resource] // Wikipedia: The Free Encyclopedia. – Access Mode : URL : [en.wikipedia.org/w/index.php?title=PubMed\\_Central&ol-](http://en.wikipedia.org/w/index.php?title=PubMed_Central&ol-)  
[did=321121698](http://en.wikipedia.org/w/index.php?title=PubMed_Central&ol-did=321121698). – Title from Screen. – Date of Access: 28 October 2009.
22. PubMedCentral hits one million article mark: Press Release of U.S. National Library of Medicine [Virtual Resource]. – 2007. – 1 June. – Access Mode : URL : [www.nlm.nih.gov/news/pmcemillion.html](http://www.nlm.nih.gov/news/pmcemillion.html). – Title from Screen. – Date of Access: 10 November 2009.
23. Reich, V. Follow the Money! / Victoria Reich // Serials Review. – 2006. – Vol. 32. – Is. 2. – P. 68–69. – Access Mode : URL : [dx.doi.org/10.1016/j.serrev.2006.03.008](http://dx.doi.org/10.1016/j.serrev.2006.03.008).
24. Reich, V., Rosenthal, D. Distributed Digital Preservation: Lots of Copies Keep Stuff Safe [Virtual Resource] / Victoria Reich, David S.H. Rosenthal // Proceedings of the Indo-US Workshop on International Trends in Digital Preservation (Pune, India, 24–25 March 2009). – 2009. – 5 p. – Access Mode : URL : [www.lockss.org/locksswiki/files/ReichIndiaFinal.pdf](http://www.lockss.org/locksswiki/files/ReichIndiaFinal.pdf).
25. Reich, V., Rosenthal, D. LOCKSS: Lots Of Copies Keep Stuff Safe / Vicky Reich, David S. H. Rosenthal // The New Review of Academic Librarianship. – 2000. – Vol. 6. – P. 155–161.
26. Rosenthal, D. et al. Requirements for Digital Preservation Systems: A Bottom-Up Approach [Virtual Resource] / David S. H. Rosenthal, Thomas S. Robertson, Tom Lipkis, Vicky Reich, Seth Morabito // D-Lib Magazine. – 2005. – Vol. 11. – N. 11. – Access Mode : URL : [dx.doi.org/10.1045/november2005-rosenthal](http://dx.doi.org/10.1045/november2005-rosenthal).
27. Rosenthal, D., Reich, V. LOCKSS, A permanent web publishing and access system: Brief introduction and status report / David S. H. Rosenthal, Vicky Reich // Serials: The Journal for the Serials Community. – 2001. – Vol. 14. – N. 3. – P. 239–244.
28. Roussopoulos, M., Bungale, P. Stealth modification versus nuisance attacks in the LOCKSS peer-to-peer digital preservation system / Mema Roussopoulos, Prashanth Bungale // Peer-to-Peer Networking and Applications. – 2009. – Online First. – P. 1–12. – Access Mode : URL : [dx.doi.org/10.1007/s12083-009-0055-5](http://dx.doi.org/10.1007/s12083-009-0055-5).
29. Souto, P. N. E-publishing development and changes in the scholarly communication system / Patricia Nascimento Souto // Ciencia da Informacao. – 2007. – Vol. 36. – N. 1. – P. 158–166.
30. Symmetry, Integrability and Geometry: Methods and Applications (SIGMA) [Virtual Resource]. – Access Mode : URL : [www.emis.de/journals/SIGMA/about.html](http://www.emis.de/journals/SIGMA/about.html). – Title from Screen. – Date of Access: 19 November 2009.
31. Tansley, R. et al. The DSpace institutional digital repository system: current functionality / Robert Tansley, Mick Bass, David Stuve, Margret Branschofsky, Daniel Chudnov, Greg McClellan, MacKenzie Smith // Proceedings of the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries (Houston, Texas, 27–31 May 2003). – 2003. – P. 87–97.
32. Top Reasons to Use DSpace [Virtual Resource] // DSPACE. – Access Mode : URL : [www.dspace.org/why-use-dspace/why-use/](http://www.dspace.org/why-use-dspace/why-use/). – Title from Screen. – Date of Access: