

УДК 004.7:025.2:004.9

Наталія Тарасенко,

мол. наук. співроб. СІАЗ НБУВ

АНАЛІТИЧНІ АСПЕКТИ ЗБОРУ ТА АРХІВУВАННЯ МЕРЕЖЕВИХ РЕСУРСІВ НАУКОВОЮ БІБЛІОТЕКОЮ

Стаття присвячена аналізу теоретичної основи та практичного досвіду сучасних бібліотек щодо відбору, архівування та надання доступу до мережових інформаційних ресурсів. Закцентовано увагу на аналітичних аспектах роботи з ресурсами Інтернет у бібліотеці. Розглянуто досвід зарубіжних країн зі збору та архівування мережових інформаційних ресурсів.

Ключові слова: мережові інформаційні ресурси, бібліотечні фонди, електронний документ, комплектування, архівування, Інтернет, аналітична обробка мережових ресурсів.

Бібліотеки здавна є невід'ємною частиною інформаційного простору і традиційно покликані забезпечувати формування найбільш повного в державі зібрання вітчизняних і зарубіжних видань на всіх видах носіїв інформації. Унаслідок розвитку процесів інформатизації суспільства, комп'ютеризації діяльності соціальних інститутів (органів державної влади, шкіл, вищих навчальних закладів, сфери виробництва та послуг, системи охорони здоров'я тощо) сьогодні дедалі більша частка припадає на інформацію, яка створюється, існує, зберігається та споживається виключно в електронному вигляді. Практично кожен громадянин у своєму повсякденному житті стикається з пошуком, обробкою, копіюванням чи поширенням електронної інформації. Це докорінно змінило інформаційні запити людей, їхні вимоги до бібліотечного обслуговування, що, у свою чергу, привело до змін у читацькому складі, мотивах звернення до бібліотек, читацьких пріоритетах, а це, у кінцевому підсумку, вплинуло на склад інформаційних ресурсів бібліотек, які крім звичних друкованих видань доповнилися електронними ресурсами. Електронна форма інформації надала змогу зберігати її надійніше й більш компактно та поширювати набагато оперативніше й ширше.

Нині значна частина електронних документальних ресурсів існує лише в глобальних комп'ютерних мережах. Згідно із сучасними даними, щодня з'являються десятки мільйонів нових веб-сторінок і зникає близько 7 млн,

що призводить до постійної втрати цінної для суспільства інформації. Водночас гігантські обсяги вже накопиченої інформації, безперервне зростання її кількості, різномірний і роз'єднаний за багатьма ознаками характер збереження й поширення, відсутність уніфікованого доступу до неї створюють істотні і дедалі зростаючі проблеми, що викликають труднощі в її ефективному використанні.

Відтак, завдання відбору мережевих ресурсів, їх нагромадження, довгострокове збереження та організація доступу до їх подальшого використання сьогодні набуває особливої актуальності для інформаційної спільноти взагалі та наукової бібліотеки, як невід'ємної її частини. Зі створенням таких фондів у бібліотеці вирішуються такі завдання:

- забезпечення більш широкого доступу до документів, надання яких читачам у традиційній формі утруднено або обмежено;
- організація фондів плинних мережевих ресурсів, які швидко зникають, їх каталогізація та надання доступу;
- надання користувачам якісно нових можливостей роботи з великими обсягами електронних даних, які доступні лише у фондах конкретної бібліотечної установи.

Суттєвою перевагою бібліотеки в цьому контексті є можливість забезпечення взаємозв'язку мережевих інформаційних ресурсів із традиційними – друкованими ЗМІ, журнальними, книжковими виданнями, що надає можливість формування найбільш повного та всеохоплюючого інформаційного зрізу тієї чи іншої проблеми.

Частково проблему архівування мережевих інформаційних ресурсів сьогодні вирішують пошукові системи та каталоги ресурсів Інтернет. Вони кумулюють та упорядковують інформацію мережевих ресурсів і, у ряді випадків, зберігають ключові фрагменти публікацій. Прикладом такої системи може бути Google, інформаційна база якої нараховує дані про понад 8 млрд документів [17]. Ця система має сервіс, що забезпечує цільове виявлення наукової інформації (Google Scholar) [18]. Кумульований ресурс щодо науково-інформаційних ресурсів має також спеціалізована пошукова система Інтернет Scirus, що надає користувачам не лише бібліографічну, а й реферативну інформацію [19]. Зазначеними системами проіндексовано понад 200 млн статей з наукових журналів, веб-сторінок науково-дослідних установ і університетів тощо. У цих системах відбиваються і закриті (платні) інформаційні ресурси та матеріали, що готуються до друку. Google Scholar крім бібліографічної інформації представляє також індекс цитування публікації в Інтернет. Однак завдання архівування повних текстів публікацій існуючими

пошуковими системами Інтернет не вирішується. Така функція притаманна саме бібліотекам, адже саме вони накопичили величезний досвід роботи з виданнями, документами, рукописами, найтісніше пов'язані з опублікованою інформацією.

Збір ресурсів Інтернету для цілей довготривалого збереження й надання доступу до них може також здійснюватися автоматично за допомогою програм-роботів або шляхом вибіркового відбору, глибокого збору та архівування сайтів.

Результатом автоматичного збору мережевих ресурсів, або веб-харвестінгу (Web-harvesting – від англ. harvest – збирати врожай), є всі матеріали певного сегмента мережі в момент збирання даних. Веб-харвестінг здійснюють програми-роботи, або веб-кроулери (Web-crawler – від англ. to crawl – повзти), засновані на тих самих принципах, що й пошукові машини. На початку процесу виконується ручне налаштування параметрів збору інформації, що визначає, з яких доменів збираються матеріали для зберігання (наприклад, національний домен або матеріали з певної тематики). Після закінчення роботи кроулерів необхідні верифікація та архівування зібраної інформації, що вимагає участі людини.

Як правило, процедура веб-харвестінгу виконується регулярно через певні досить великі проміжки часу (наприклад, раз на півроку). Зміни, що відбулися в мережі за цей період, не архівуються і повністю втрачаються.

Якість і повнота результатів веб-харвестінга залежить від використовуваних роботів, які постійно вдосконалюються. Але поки що кроулери можуть збирати тільки поверхневу інформацію, а глибинні матеріали залишаються неохопленими. Результат роботи кроулерів – статичні подання інтернет-сторінок, як правило, першого й другого рівня.

Тим не менш, у результаті веб-харвестінга утворюються величезні обсяги інформації для зберігання. Ця інформація не може бути каталогізована звичайним способом, тому для автоматичного анотування та структурування розробляються й використовуються спеціальні програми, засновані на методах семантичного вебу. Недоліком харвестінга є дублювання, оскільки архівуються дзеркала сайтів, тобто один і той же матеріал збирається кілька разів.

Цей метод недостатньо ефективний для збору та збереження таких інтернет-ресурсів, як газети, потокові відео- та аудіоресурси, результати роботи веб-камер, інтерактивні документи, цифрові матеріали різних типів, що зберігаються в базах даних.

Водночас важко дати таке визначення мережевої публікації, щоб

воно давало змогу знаходити всі матеріали, які потребують збереження. Мережеві публікації не є лінійними й закінченими, містять у собі ресурси різних форматів (текст, статичні зображення, відео- та аудіофрагменти, потокові відео та аудіо), мають складну структуру. Часто складно встановити автора й видавця матеріалу. Дата публікації також є розмитим поняттям з огляду на динамічність і постійне оновлення сайтів.

Потребують особливого підходу також інтернет-ресурси з коротким життєвим циклом. Необхідний вибірковий тематичний відбір із глибоким багаторівневим збором і архівуванням таких матеріалів, який реалізується на основі закону про обов'язковий примірник або на основі договорів з видавцями і дає більш якісний результат на невеликому сегменті мережі. Співпраця з видавцями дає можливість каталогізувати зібрані ресурси. Зазвичай використовується поєднання обох методів – повний автоматичний збір мережевих матеріалів якихось сегментів мережі через певні періоди часу і глибоке вибіркове тематичне архівування найбільш цінних ресурсів Інтернету.

Для реалізації проектів з веб-архівування необхідно дотримання як мінімум двох умов: достатнє довгострокове фінансування і адекватна законодавча база, переважно у формі законодавства про обов'язковий примірник документів. Утім, швидкий розвиток технологій породжує нові види мережевих видань, і законодавство не завжди може відповісти цим змінам. Якщо орієнтуватися на існуючі сьогодні технології і типи мережевих публікацій, законодавство доведеться переглядати занадто часто. Якщо ж закон має більш загальний характер, орієнтується на можливі майбутні технології, які важко передбачити, то неможливо точно визначити всі необхідні для збирання й збереження інформації процедури, забезпечити контроль виконання, передбачити штрафні санкції.

Очевидно, що неможливо зобов'язати всіх, хто публікує свої матеріали в Інтернеті, передавати їх на зберігання до бібліотек, тому веб-харвестінг є єдиним варіантом, що забезпечує повноту охоплення певного сегмента мережі (наприклад, національного домену). А от визначення найбільш цінних для держави й суспільства ресурсів Інтернет та збір їх або на основі закону про обов'язковий примірник, або на основі договорів з видавцями, цілком можливий і практикується в багатьох країнах.

Деякі країни, у яких архівування Інтернету не закріплено законодавчо, проте, реалізують проекти з веб-харвестінгу й глибокого тематичного архівування на основі добровільного надання видавцями матеріалів на зберігання. Можливості доступу до заархівованих матеріалів визна-

чаються законами про охорону прав на інтелектуальну власність, тому в більшості країн доступ до них обмежений або навіть закритий.

Формування фондів електронних, у тому числі й мережевих, інформаційних ресурсів бібліотек постійно перебувають у полі зору вітчизняних бібліотекознавців. У дослідженнях переважає проблематика, пов'язана з вивченням термінології та типології електронних документів, технологічного процесу електронного фондоутворення, визначення місця електронного документа в бібліотечних фондах.

Важливими в розробці теоретичних основ розвитку бібліотек в електронному середовищі та аналізу вже набутого досвіду зі створення й розвитку електронних бібліотек у світі та Україні є наукові напрацювання таких дослідників, як Я. Шрайберг, Т. Бахтуріна, Т. Кузмішина, І. Моргенштерн, Ю. Столяров, Е. Сукіасян та ін. Визначенню сучасних тенденцій та перспектив розвитку електронних інформаційних ресурсів бібліотек присвячені роботи Г. Євстігнєєвої, А. Земскова, Б. Ленга, М. Лінквіста, О. Онищенко та ін.

Особливу групу становлять публікації дослідників, що займалися розробками науково-організаційних та правових засад створення збірань електронних документів, серед них: Е. Азгальдов, О. Бойкова, С. Васильченко, К. Генєєва, С. Китс, В. Наумов, С. Пушкова, М. Слободяник, А. Стрельцов, І. Тюніна, П. Хугенхольтц та ін.

Вивченню проблем класифікації електронних видань присвячені публікації А. Антопольського, Т. Майстрович, Т. Бахтуріної, В. Вуль, Ю. Столярова, Г. Швецової-Водки. Можливості електронних бібліотек досліджуються А. Антопольським, К. Вігурським, А. Віслим, А. Земськовим, Ю. Нохрїним, А. Федоровим, Я. Шрайбергом та ін.

Питанням аналізу та використання електронних ресурсів присвячені праці П. Бакута, Б. Шумилова, А. Гиржона, В. Широкова, В. Голуба, В. Губарець, В. Степанова, А. Соляник, Г. Шемаєвої, Т. Ярошенко.

Помітний внесок у вивчення наукових засад формування електронних бібліотечно-інформаційних ресурсів поряд з іншими вченими вносить група дослідників Національної бібліотеки України імені В. І. Вернадського. Зокрема, у працях В. Горового, Л. Костенка, Г. Ковальчук, Т. Павлуші, І. Павлуші, Т. Добко, О. Воскобойнікової-Гузевої, Н. Стрішенець, Н. Самохіної вивчаються та узагальнюються проблеми системного підходу до формування фондів електронних бібліотек, принципи відбору електронних ресурсів, тенденції та особливості розвитку бібліотек в інформаційному суспільстві. При вивченні проблеми комплектування сучасних бібліотек мережевими ресурсами на окрему

увагу заслуговують дослідження в цій галузі В. Копанєвої. У дисертаційній роботі «Формування фонду мережевих ресурсів у науковій бібліотеці (90-ті роки ХХ ст. – поч. ХХІ ст.)» (2008) [6] і монографії «Бібліотека як центр збереження ресурсів Інтернету» (2009) [7] вона обґрунтувала необхідність формування в інтранет-середовищах вітчизняних бібліотек нових інформаційних складових – фондів мережевих ресурсів, створила концепцію кооперативного архівування науково-інформаційних і суспільно значущих ресурсів Інтернету провідними галузевими й регіональними бібліотеками України, дослідила тенденції розвитку правового підґрунтя архівування мережевих ресурсів – перехід від заборонної парадигми класичного авторського права до дозвільної системи поширення знань. В. Копанєва також розробила методологічні засади інформаційної технології опрацювання інтранет-ресурсів у бібліотеках, що базуються на використанні Дублінського ядра метаданих і повнотекстовому індексуванню.

Важливим аспектом наукових праць з досліджуваної проблематики є визначення принципів відбору мережевих інформаційних ресурсів, які підлягають архівуванню, виокремлення видів документів, якими комплектується ресурс (текстові, графічні тощо), їхньої тематики (універсальна, тематична, проблемно-орієнтована тощо), прийнятих обмежень (за хронологією, мовою, призначенням та іншими ознаками).

Загальноприйнятною є точка зору, що об'єктами комплектування не можуть бути ресурси, що містять визначену чинним законодавством шкідливу інформацію, а саме: заклики до насильства, розпалювання расової і міжнаціональної ворожнечі тощо. Інтернет-ресурси будь-якого формату можуть бути джерелом поповнення фонду бібліотеки, якщо вони відповідають тематичним, видовим критеріям відбору й бібліотека має програми-клієнти для їх використання. Обмеження за способом подання документів мають значення лише тоді, коли документ не можна конвертувати у формати, прийняті в певній бібліотеці [14].

Складовою архіву насамперед мають стати колекції наукових і суспільно значущих документів національної складової Інтернету, тобто слід забезпечити повноту відображення в архіві публікацій, створених на території держави. Нормативною базою для архівування друкованих та інших видань є законодавчі акти щодо обов'язкового примірника документів, однак сьогодні не лише в Україні, а й в інших державах мережеві публікації не підпадають під сферу дії цих нормативних актів [16].

Варто зазначити, що архівування ресурсів глобальних інформаційних мереж не має належної нормативно-правової основи через невизначен-

ність у міжнародному й національному законодавствах прав бібліотек щодо їх функціонування в електронному середовищі. Певною мірою розроблено питання, пов'язані з дотриманням авторських прав під час архівування електронних ресурсів узагалі й мережових як їхньої складової частини. Серед українських бібліотекознавців значну увагу цим питанням приділяла В. Копанєва. Вона констатує, що, оскільки мережева інформація належить до «відкритого контенту», у разі, якщо вона є у відкритому доступі, її захищено однією з публічних ліцензій (*GNU Open Public Licence*) або ліцензіями *Creative Commons*, що вирішує питання її поширення й використання [9].

Архів має включати також колекції зарубіжних мережових науково-інформаційних ресурсів. При визначенні тематичного складу зарубіжних мережових ресурсів, що мають архівуватися, слід виходити зі стратегічних та середньострокових пріоритетних напрямів інноваційної діяльності в Україні, що затверджені на законодавчому рівні.

Мають підлягати архівуванню також колекції інтернет-публікацій про Україну, а також зібрання мережових документів українською мовою та публікації українців та українських інституцій в інших державах.

Разом з тим, як слушно зазначає В. Копанєва, технологія архівування науково-інформаційних ресурсів глобальних комп'ютерних мереж повинна бути спрямована на досягнення максимальної ресурсоощадності й давати змогу уникнути дублювання під час обробки документів [8]. На думку науковця, для досягнення цієї мети за умов обмежених матеріально-технічних і кадрових ресурсів та наявності розвинутих телекомунікаційних мереж до архівування доцільно залучити всі зацікавлені інституції – суб'єкти вітчизняної системи документальних наукових комунікацій (провідні галузеві й регіональні бібліотеки, інформаційні центри, наукові установи й навчальні заклади тощо). Виходячи з цього, В. Копанєва робить висновок, що основною концептуальною засадою побудови цілісної системи архівування науково-інформаційних ресурсів в Україні слід обрати поєднання принципів децентралізованого збору й первинної обробки галузевих і регіональних фрагментів інформаційних мереж із централізованим формуванням зведеного довідково-пошукового апарату для підтримки багатоаспектного використання інформаційних ресурсів розподіленого архіву [там само].

Треба зазначити, що електронні колекції або веб-архіви інтернет-ресурсів нині створено в багатьох бібліотеках України. Насамперед заслуговують на увагу фонди мережових ресурсів Національної бібліотеки України імені В. І. Вернадського. Загальний архів науково-

інформаційних ресурсів Інтернету Національна бібліотека України імені В. І. Вернадського формує з 2005 р. Інтранет, екстранет-середовище НБУВ (<http://www.nbuv.gov.ua/home>) має універсальний у видовому й тематичному аспектах електронний фонд понад 700 тис. публікацій. Цей фонд містить універсальні, галузеві й тематичні енциклопедії, довідники й підручники; колекції наукової, науково-популярної і художньої літератури; наукові журнали; музейні колекції; мультимедійні матеріали; зібрання матеріалів з бібліотечної справи й науково-інформаційної діяльності тощо. Переважна більшість складових архіву мережових ресурсів у фондах інших бібліотек України відсутня, інформаційні ресурси архіву надаються користувачам в інтранет-середовищі бібліотеки та її філіях.

Пошуковий апарат бібліотеки має у своєму складі систему бібліотечних каталогів і картотек і фонд довідково-бібліографічних видань обсягом 200 тис. примірників. Цей фонд містить документи нормативного характеру (закони, укази, постанови тощо), енциклопедії, тлумачні словники, довідники, бібліографічні посібники. Систему бібліотечних каталогів і картотек утворюють генеральний алфавітний каталог, читачкї алфавітний і систематичний каталоги і понад 30 каталогів і картотек підрозділів бібліотеки.

Збиранням й архівуванням мережових інформаційних ресурсів у НБУВ опікується відділ формування й обліку електронних інформаційних ресурсів (ВФЕІР), створений у 2010 р. з метою впровадження нової форми інформаційної діяльності – обслуговування читачів повнотекстовими мережовими ресурсами Інтернету. Основними функціями відділу є формування та опрацювання зібрання інформаційних ресурсів українського сегмента Інтернету й актуальних для країни зарубіжних мережових матеріалів у рамках вирішення завдань програми «Бібліотека-XXI» [14]. Пріоритетну тематику архівування електронних інформаційних ресурсів у НБУВ визначають Закон України «Про пріоритетні напрями інноваційної діяльності в Україні», Постанова Кабінету Міністрів України «Про затвердження Інструкції переліку пріоритетних тематичних напрямів наукових досліджень і науково-технічних розробок на період до 2015 р.» від 07.09.2011 р. та постанова президії НАН України від 21.09.2011 р. щодо організації виконання згаданої Постанови Кабінету Міністрів України.

Електронні видання пріоритетної тематики становлять 94 % від загальної кількості всієї наукової електронної бібліотеки (ЕБ) НБУВ:

- повнотекстові електронні видання науково-технічної спрямованості – 14% від загальної кількості електронних інформаційних ресурсів;

- електронні видання природничої тематики (у тому числі фізико-математичні науки, хімічні науки) – 13 %;
- інформаційні й комунікаційні технології – 8 %;
- енергетика – 6 % від загальної кількості всього зібрання;
- природокористування – 6 %;
- науки про життя (у тому числі ресурси медичної тематики) і розвиток біотехнологій – 9 %;
- нові речовини й матеріали, нанотехнології і технології наноматеріалів – 7 %;
- суспільні й гуманітарні науки (економіка, мовознавство, філософія, психологія, історія) – 22 %;
- юридична література – 9 % від загальної і1082 кількості ресурсів наукової ЕБ.

На сьогодні загальна кількість опрацьованих за допомогою програмного середовища ІРБІС-64 видань у науковій ЕБ НБУВ перевищує 35 тис. бібліографічних описів електронних документів. Якщо видання пріоритетної тематики становлять 94 % від загальної кількості ресурсів ЕБ, решта повнотекстових мережевих ресурсів охоплює загальнонаукове, міждисциплінарне знання, військову справу тощо. Також широко представлено в науковій ЕБ довідкову літературу: словники, енциклопедії, довідники з усіх галузей знань.

До електронних матеріалів, які архівуються, належать:

- наукові, науково-популярні видання: аналітичні доповіді, монографії, колективні видання, популярні брошури;
- навчально-методичні видання: підручники, методичні посібники;
- довідкові видання: словники, енциклопедії, довідники;
- інші видання: збірки статей, атласи, альбоми, життєписи, нормативно-правові акти, закони, кодекси тощо.

Поповнення наукової ЕБ новими матеріалами відбувається за рахунок відбору інформаційних ресурсів із джерел відкритого доступу. Співробітники ВФЕІР проводять постійний моніторинг ресурсів мережі Інтернет, переглядають електронні бібліотеки, архіви, сайти, портали установ, організацій з метою знаходження і подальшого архівування необхідних ресурсів, які відповідають певним умовам. Наукові співробітники відділу проводять також роботу з відстеження і першочергового архівування ресурсів, які в подальшому можуть зникнути з мережі Інтернет або не існуватимуть у вільному доступі.

Опрацювання архівованих електронних ресурсів відбувається у ВФЕІР за допомогою системи автоматизації бібліотек ІРБІС-64, роботу

з якою підтримує відділ програмно-технологічного забезпечення комп'ютерних мереж НБУВ. Для продукування вихідної форми опрацьованого видання відомості про кожний електронний документ заносяться співробітниками ВФЕР у 24 відповідні поля. Вихідна форма електронного документа містить бібліографічний опис, анотацію, тип ресурсу, розділ знань, ключові слова. Також наводяться відомості про формат завантаженого документа, його обсяг, подається гіперпосилання на повний текст ресурсу.

При створенні архівів мережевих ресурсів важливу роль відіграють правові питання, адже багато ЕБ публікують твори без попередньої згоди авторів. І хоча на першу вимогу авторів такі тексти знімаються зі сайту, де-юре має місце порушення їхніх прав. Виходом із цієї ситуації стало рішення про розміщення електронних інформаційних ресурсів у локальній мережі НБУВ (лише для перегляду в читальних залах), що не суперечить Закону України «Про авторське право і суміжні права» від 23.12.1993 р. № 3792-ХІІ, ст. 22 якого закріплює це як один з випадків, на який не поширюється виключне право власника авторських прав забороняти використання його творів. Ідеться про відтворення одного примірника твору бібліотеками та архівами, якщо їхня діяльність не спрямована на отримання прибутку і відповідний ресурс представлено для ознайомлення. Частиною повнотекстових електронних інформаційних ресурсів наукової ЕБ НБУВ також можуть користуватися віддалені читачі (електронні версії видань, що передані авторами та видавництвами на договірній основі).

Разом з тим наукового опрацювання потребують аналітичні аспекти архівування мережевих ресурсів, врахування і подальша розробка яких дасть змогу вирішити проблеми щодо більш ефективного відбору та архівування мережевих документів до бібліотечних фондів.

Оскільки аналітична обробка мережевих ресурсів нині широко затребувана в різних галузях діяльності людини, бібліотека, з-поміж іншого, має зосереджуватись як на наданні доступу до аналітичної інформації, представленої в мережі Інтернет, так і на виробленні власних продуктів аналітичного спрямування. Таким чином бібліотека надає не лише технічні можливості доступу до первинної, необробленої інформації, а й створює якісно інший рівень доступу – інтелектуальний, який надає структуровану, систематизовану інформацію, зріз з певної проблеми, зроблений на основі аналізу великих масивів первинних джерел. Це значно скорочує час пошуку та обробки необхідної інформації,

підвищує ефективність її використання, водночас потребуючи значних інтелектуальних та наукових технологій.

Особливого значення набуває визначення переліку та змісту критеріїв, на підставі яких може здійснюватися відбір та аналітична обробка мережевих інформаційних ресурсів, адже їх обсяги на кілька порядків перевищують традиційний документальний потік бібліотеки. Це унеможливає їх обробку методами, що використовуються під час опрацювання друкованих матеріалів, призводить до дублювання робіт, надмірності щодо збору первинної інформації.

Насамперед слід виходити з вивчення ймовірних інформаційних потреб користувачів аналітичної інформації, а також збереження інформації, яка б ґрунтувалася на аналізі нагромаджених фактів з історії життя країни та різних верств суспільства, надавала можливість ретроспективного моніторингу подій, супроводжуваного співмірною їм аналітикою фактів і прогнозів.

З огляду на це необхідно здійснювати порівняльний аналіз мережевих ресурсів щодо міри презентативності інформації з теми на різних сайтах, що впливатиме на їх вибір як джерела для відбору. Рівень прогнозованої інформативності сайту формально залежить від рівня кумуляції інформації аналітичного спрямування, офіційної інформації, неофіційної (опозиційної або незалежної) точки зору.

З іншого боку, міра кумуляції інформації з теми має узгоджуватися з такою змістовною характеристикою, як повнота інформації, ступінь достовірності інформації, яка має визначатися на рівні відбору. Зазначений критерій може мати різні ступені: від сприйняття інформації як недостовірної до наявності її офіційного підтвердження. Сайт державного органу або центральної установи, який містить інформацію, підтверджує її офіційність, має більшу міру вірогідності, що оцінюється позитивно. З іншого боку, повідомлення, розміщене на сайті пересічної юридичної або фізичної особи, може мати значно меншу тривалість існування, але й більшу міру унікальності порівняно з попередніми видами джерел, а значить бути більш цінним.

Слід також врахувати, що на різних сайтах можуть розміщуватися різні версії фактів, їхні коментарі, аналітичні матеріали, що зумовлює необхідність копіювання ресурсів, які містять різні (іноді полярно протилежні) оцінки фактів. З цієї точки зору найбільшу цінність для відбору представляють сайти аналітики новин, на яких часто спостерігається дублювання інформації, що вже зафіксована (ефект поглинання інформації). Разом з тим аналітико-синтетичні операції щодо вилучення дублюючих

частин сайтів можуть займати багато часу. Тобто копіювання сайтів без їх аналітичного опрацювання буде охоплювати значно менше часу, але у цьому разі є більша ймовірність зберігання значних обсягів нерелевантної інформації.

Особливу увагу слід приділити так званим «чуткам» або «пліткам». Ця інформація зникає швидше, але саме вона є атрибутом соціально-комунікаційного тла епохи і її наявність не дає можливості представити історію за офіційними виданнями як суцільну пряму лінію. Важливим є збереження вихідного повідомлення, яке стає першоджерелом для створення аналітичних мережевих ресурсів.

Успішним досвідом у цьому напрямі є діяльність Служби інформаційно-аналітичного забезпечення органів державної влади (СІАЗ), що працює в структурі Національної бібліотеки України імені В. І. Вернадського. Однією з важливих складових джерельної бази інформаційно-аналітичних продуктів СІАЗ є мережеві ресурси – інтернет-видання, електронні версії друкованих ЗМІ, спеціальні інтернет-проекти, сайти владних структур, наукових установ, громадських і політичних організацій, галузеві електронні ресурси. Моніторинг та відбір інформації здійснюється за визначеними напрямками, серед яких: розвиток науки в Україні, проблеми інформатизації, українсько-російські відносини, регіональні конфлікти, розвиток соціальних мереж тощо.

Технологія відбору мережевих ресурсів передбачає переважно моніторинг працівниками служби визначеного переліку веб-сайтів, які надають оптимальну щодо достовірності, актуальності, повноти, обґрунтованості та різносторонності інформацію.

У практиці роботи СІАЗ із мережевою інформацією є також досвід використання автоматизованої інформаційно-моніторингової системи відбору Web-Observer [20]. Функціонування системи полягає у відборі необхідної новинної інформації з бази публікацій на основі критеріїв відбору, що їх обумовлює користувач системи, а також подальшої обробки отриманої інформації. Запити на вибірку повідомлень з бази публікацій формуються на основі параметрів, що попередньо задані користувачем. Зокрема, серед критеріїв фільтрації документів можна виділити такі:

- дата документа;
- часовий проміжок, за який відбуватиметься добірка публікацій (година, доба, тиждень, місяць, три місяці, півроку);
- ключові слова – слово або набір слів, що повинні бути присутні в публікації;

- логіка – логічні присутності ключових слів у повідомленнях, що відбираються;

- мова статті – фільтрація статей визначеною мовою або кількома мовами, наприклад, лише українська, або українська та російська;

- статус – параметр, що регулює відбір серед оброблених або необроблених повідомлень;

- кількість – параметр, що визначає кількість одночасно виведених заголовків публікацій, що відповідають запиту;

- вибір джерел – джерела згруповані за належністю до країни, категорії або рубрики. Можна вибирати як окремі джерела, так і групи джерел.

Комбінуючи фільтри, що описані вище, можна досягти максимальної релевантності знайдених матеріалів. Час пошуку залежить від складності сформованого запиту та тривалості часового проміжку.

Утім, використання автоматизованої інформаційно-моніторингової системи відбору Web-Observer у процесі практичної роботи СІАЗ було визнано не достатньо ефективним, оскільки під задані параметри пошуку часто потрапляла інформація, яка лише дотично відповідала певній проблематиці, або неповна інформація, суттєві елементи якої містяться на більш глибоких рівнях сайтів.

У процесі створення інформаційно-аналітичних продуктів СІАЗ тісно співпрацює з іншими підрозділами НБУВ, використовуючи фонди бібліотеки для створення синтезованих науково обґрунтованих інформаційних продуктів, які мають максимально задовольняти інформаційні потреби користувачів.

Процес інформаційного виробництва СІАЗ складається з кількох етапів:

- пошуку, відбору, структурування мережевої інформації;

- аналізу інформації;

- обробки та створення на базі первинних ресурсів певного виду інформаційного продукту.

Важливим напрямом у роботі з розширення інформаційної бази служби є вивчення регіональних джерел електронної інформації. У результаті такої роботи створено електронну карту регіонів. Веб-ресурси досліджуються за певними критеріями (суспільна значущість, періодичність оновлення мережних ЗМІ; якість підтримки електронного видання; якість оперативної аналітичної інформації; можливості пошукового апарату сайтів, роботи архіву, умови використання матеріалів тощо).

Для вдосконалення роботи з розширення пошуку необхідної інформації запроваджено використання «електронних візиток», які

складаються кожним співробітником відділу технології електронної обробки інформації, поповнюються ним у процесі роботи. Вони сприяють збереженню та накопиченню знайдених веб-ресурсів, чіткому й оперативному перерозподілу шляхів пошуку потрібної інформації. На сьогодні вивчено і введено в робочі процеси близько 200 сайтів центральних українських та російських інтернет-видань, 227 – регіональних.

Під час пошуку, технічного опрацювання та структурування інформації, що є основою роботи служби, враховуються такі критерії відбору матеріалів інтернет-видань, як:

- жанрові особливості інформаційно-аналітичного продукту;
- необхідний рівень деталізації матеріалу;
- потреби замовників дослідження, характер практичного використання.

Сучасні технології надають можливість використовувати дистантну форму надання інформаційно-аналітичної продукції замовниками. Відповідні інформаційно-аналітичні матеріали в електронному вигляді надсилаються на e-mail понад 900 адресатів у вищих органах центральної влади, в обласних центрах, містах та районах України. СІАЗ широко використовує і таку форму дистантного обслуговування читачів, як надання доступу до інформаційно-аналітичної продукції через веб-портал НБУВ у рубриці «Аналітичні матеріали СІАЗ», яка регулярно поповнюється.

Отже, наведений аналіз дає підстави стверджувати, що на сьогодні українськими бібліотекознавцями здійснено низку досліджень щодо теоретичних та практичних аспектів архівування бібліотеками мережевої інформації, а також накопичено певний практичний досвід, який, однак, не має належного наукового узагальнення в контексті вивчення аналітичних аспектів створення архівів мережевих ресурсів. Це створює необхідність розв'язання низки питань щодо подальшого вдосконалення діяльності бібліотек з організації комплектування її фондів інформаційними ресурсами мережі Інтернет, забезпечення їх надійного зберігання та ефективного використання в інтересах суспільства.

Список використаних джерел

1. *Горовий В. М.* Соціальні інформаційні комунікації, їх наповнення та ресурс : монографія / В. М. Горовий. – К., 2010. – 356 с.
2. *Горовий В. М.* Бібліотеки оновлюваного суспільства: деякі тенденції

та особливості розвитку / В. М. Горовий. – Бібл. планета. – 2006. – № 3. – С. 27–30.

3. *Льганаєва В. О.* Бібліотека в сучасному інформаційному середовищі: Основні стратегії розвитку / В. О. Льганаєва // Вісн. Харк. держ. акад. культури. Бібліотекознавство. Документознавство. Інформатика : зб. наук. пр. – Х., 1999. – Вип. 1. – С. 30–37.

4. *Казаков В. Г.* Электронный документ как объект библиотечного дела: постановка проблемы [Электронный ресурс] / В. Г. Казаков, Т. В. Майстрович // Электронные библиотеки : рос. науч. электрон. журн. – 2003. – Т. 6. – Вып. 4. – Режим доступа: <http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2003/part4/КМ>. – Загл. с экрана.

5. *Ковальчук Г.* Використання електронних ресурсів в обслуговуванні користувачів у публічних бібліотеках / Г. Ковальчук, Н. Кучерява // Бібл. форум України. – 2010. – №1. – С. 13–16.

6. *Копанєва В. О.* Формування фонду мережевих ресурсів у науковій бібліотеці (90-ті роки ХХ ст. – поч. ХХІ ст.) : автореф. дис. ... канд. іст. наук : 27.00.03 / В. О. Копанєва / НАН України, Нац. б-ка України ім. В. І. Вернадського. – К., 2008. – 21 с.

7. *Копанєва В. О.* Бібліотека як центр збереження інформаційних ресурсів Інтернету : [монографія] / В. О. Копанєва ; НАН України, Нац. б-ка України ім. В. І. Вернадського. – К., 2009. – 198 с.

8. *Копанєва В. О.* Архівування науково-інформаційних ресурсів Інтернет: основні концептуальні положення / В. О. Копанєва // Бібл. вісн. – 2005. – № 2. – С. 14–19.

9. *Копанєва В. О.* Інтернет і авторське право / В. О. Копанєва // Документознавство. Бібліотекознавство. Інформаційна діяльність: Проблеми науки, освіти, практики : зб. матеріалів VII Міжнар. наук.-практ. конф., Київ, 25–27 трав. 2010 р. – К., 2010. – С. 171–173.

10. *Костенко Л. Й.* Розвиток інформаційних технологій та Інтернету в бібліотеках / Л. Й. Костенко // Бібл. вісн. – 2002. – № 1. – С. 16–18.

11. *Майстрович Т. В.* Электронные библиотеки: принципы создания : науч.-метод. пособие / Т. В. Майстрович, А. Б. Антопольский. – М. : Либерия-Бибинформ, 2007. – 283 с.

12. *Майстрович Т. В.* Электронный документ как компонент библиотечного фонда : автореф. дис. ... д-ра пед. наук / Т. В. Майстрович. – М., 2005.

13. *Майстрович Т. В.* Электронный документ в библиотеке : науч.-метод. пособие / Т. В. Майстрович. – М. : Либерия-Бибинформ, 2007. – 139 с.

14. *Самохіна Н. Ф.* Архівування інформаційних ресурсів у системі формування електронної бібліотеки / Н. Ф. Самохіна // Бібл. вісн. – 2012. – № 5. – С. 27–31.
15. *Соляник А. А.* Система документопостачання бібліотечних фондів: закономірності розвитку : монографія / А. А. Соляник. – Х. : ХДАК, 2005. – 230 с.
16. *Чепуренко Я. О.* Державне забезпечення системи обов'язкового примірника творів друку України: етапи розвитку / Я. О. Чепуренко // Нац. б-ка України ім. В. І. Вернадського. – К., 1998. – 69 с.
17. Google [Електронний ресурс]. – Режим доступу: <http://www.google.com.ua/>. – Назва з екрана.
18. Google Scholar [Електронний ресурс]. – Режим доступу: <http://scholar.google.com/>. – Назва з екрана.
19. Scirus [Електронний ресурс]. – Режим доступу: <http://www.scirus.com/>. – Назва з екрана.
20. Web-Observer [Електронний ресурс]. – Режим доступу: <http://www.webobserver.info/>. – Назва з екрана.