

Електронна наукова періодика відкритого доступу: семантичні веб-технології для бібліотек

Наявність величезних обсягів слабоструктурованої інформації у сучасному веб-середовищі унеможливила задоволення інформаційних потреб користувачів за допомогою простого навігаційного пошуку. У багатьох випадках засоби статичного перегляду веб-сторінок у край неефективні, або взагалі непридатні для вирішення завдань пошуку наукової інформації. Зі зміною поглядів на організацію інформації та пошукових систем глобальної мережі з'явилися нові технологічні моделі проектування і створення веб-ресурсів. Сьогодні ми можемо спостерігати розвиток технології Веб 3.0, яка ґрунтується на ідеях семантичного вебу. Останні зводяться до того, що додавання певної структури даних (метаданих) неструктурованому контенту глобальної мережі робить його більш зрозумілим пошуковим роботам, що, у свою чергу, дає можливість перетворити ресурси Інтернету в аналог бази даних. Ці процеси у середовищі веб-технологій спричинили посилення уваги до бібліотечних технологій, бо саме в них накопичений великий досвід упорядкування знань.

Питання про те, що може привнести технологічна модель Веб 3.0 у бібліотечну діяльність і як виглядатиме Бібліотека 3.0 або Семантична Бібліотека порушувались ще у 2007 р. на Генеральній конференції IFLA (Південна Африка)¹. За узагальненим прогнозом, основною місією бібліотек у найближчому майбутньому буде упорядкування неорганізованих електронних ресурсів у систематичні категорії. Для цього використовуватимуться семантичні веб-технології, онтології та метадані. Великого значення набудуть електронні каталоги бібліотек як сховища метаданих документів різної природи. Набудуть чинності стандарти бібліотечного упорядкування на все веб-середовище; спостерігатиметься подальша інтеграція електронних каталогів бібліотек з наданням доступу до електронних інформаційних ресурсів; створюватимуться бібліотечні портали, що надаватимуть доступ до всіх бібліотечних ресурсів у режимі «єдиного вікна». Запровадження цих технологій передбачається приблизно у 2016 р.². Зважаючи на широкі можливості, які відкривають технології семантичного вебу для бібліотек, на 77-му Конгресі ІФЛА в Пуерто-Рико (2011 р.) було утворено спеціальну групу за

¹ *Saw G., Todd H.* Library 3.0: where art our skills? // World library and information congress : 73rd IFLA general conference and council (19–23 August 2007, Durban, South Africa). – 2007. – Р. 1–15.

² *Лобузін К. В.* Бібліотека 3.0: знання, сховища даних, експерти // Бібліотекознавство. Документознавство. Інформологія. – 2012. – № 1. – С. 26–35; *Лобузін К. В.* Технології організації знанневих ресурсів у бібліотечно-інформаційній діяльності: монографія; відп. ред. О. С. Онищенко; НАН України, Нац. б-ка України ім. В. І. Вернадського. – К., 2012. – 252 с.

інтересами – Semantic Web Special Interest Group (SWSIG), орієнтовану на вивчення цього напрямку (<http://www.ifla.org/about-swsig>).

Сьогодні процеси інтеграції бібліотечних та веб-технологій особливо інтенсивно розвиваються в галузі організації доступу до наукової інформації. Набуття популярності у всьому світі ідеї відкритого доступу до наукової інформації, спричинилося до накопичення величезних обсягів наукових та освітніх ресурсів в електронному вигляді. Загальнодоступним інтегратором цих ресурсів стала глобальна пошукова система Google Академія (Google Scholar), призначена для виокремлення з глобального інформаційного потоку джерел наукової інформації. Останнім часом відбулись істотні зміни в наукометричних та бібліометричних можливостях системи. Зокрема, було удосконалено алгоритми екстрагування метаданих веб-сторінок та автоматичного виявлення цитування наукових публікацій, забезпечено можливість створення наукометричних профілів вчених та наукових фахових видань, що дає змогу авторам стежити за цитуванням своїх статей. Ці технологічні інновації в сервісних можливостях інформаційної системи Google Академія привернули увагу користувачів наукових інтернет-комунікацій. Якщо говорити про Україну, то тут слід відзначити розроблення інформаційного сервісу «Бібліометрика української науки», зорієнтованого на показники індексів цитування Google Академія³.

Важливим показником ступеня інтеграції наукового е-ресурсу (бази даних, бази знань або інформаційного порталу) в наукову комунікацію суспільства є OPENNESS (Відкритість). Цей показник входить до оцінок рейтингу Webometrics Ranking of World Universities, системи, яка визначає світовий рейтинг університетських сайтів відповідно до Берлінських принципів, визначених ЮНЕСКО (<http://www.webometrics.info>). Поняття відкритості передбачає глобальні зусилля зі створення цінних файлів (PDF, DOC, PPT), опублікованих та проіндексованих академічною пошуковою системою Google Академія. Під цінними файлами (Rich files) слід розуміти файли у форматах, які використовують автори для представлення та поширення власних робіт. Враховуються файли MS Word (doc), PowerPoint (ppt), Adobe Acrobat (pdf), а також деякі інші. Автори рейтингу вважають, що велика кількість таких документів засвідчує: на сайті зберігаються не лише адміністративна інформація, а й наукові матеріали.

Незмінною складовою міжнародних систем наукової інформації, які представляють в організованому інтерфейсі наукові публікації, стали також персональні профілі вчених. Вони уможливають введення та ідентифікацію персональної інформації за принципами інтерактивного обміну даними Веб 2.0.

³ Бібліометрика української науки: інформаційно-аналітична система / Л. Костенко, О. Жабін, О. Кузнецов, Є. Кухарчук, Т. Симоненко // Бібл. вісн. – 2014. – № 4. – С. 8–12.

Самостійне формування спільнотою вчених наукових онлайн-мереж призвело до появи концепції та відповідного терміна наука 2.0 (Research 2.0, Science 2.0) або е-наука (e-Science)⁴. У науковому інтернет-просторі персональний профіль вченого відіграє роль семантичної ланки між публікаціями та їх автором. Персональна інформація – необхідний атрибут метаданих, які повідомляються зовнішнім бібліографічним системам під час розміщення наукової публікації. Від даних персонального профілю слід відштовхуватися у процесі пошуку в Інтернеті робіт певного автора, підрахунку числа цитувань, дослідження наукових зв'язків. У такий спосіб персональна сторінка перетворюється на ключовий елемент інтернет-інфраструктури науки⁵. Технологічні засоби системи Google Академія уможливають також автоматизоване збирання бібліографічних даних та обчислення наукометричних показників, пов'язаних з персональним профілем вченого. Створивши свій персональний профіль, автор може дізнатися, хто посилається на його публікації, переглянути діаграму цитування та обчислити індекси цитування. Наявна також функція автоматичного оновлення списку статей та попереднього перегляду нових надходжень або додавання матеріалів власноруч.

Розуміння важливості відкритості результатів наукової роботи для визначення наукового рейтингу призвело до масового створення електронних інституційних репозиторіїв, які інтегрують та надають у відкритому доступі здобутки науково-дослідних та навчальних установ. Стандартні платформи інституційних репозиторіїв мають готові рішення для постачання метаданих до систем-агрегаторів через протокол OAI-PMH. Цей протокол також дає змогу інтегрувати бібліографічні дані в інформаційні сервіси Google Академія. Правильне постачання якісних метаданих значно підвищує цитованість наукових робіт та покращує якість визначення індексів цитування.

З розвитком інтеграційних процесів у сучасних наукових комунікаціях наукові бібліотеки опиняються перед певною технологічною ділемою: переходити на нову технологічну платформу електронних репозиторіїв⁶, або адаптувати вже існуючі бібліотечні рішення до сучасних вимог представлення

⁴ Журавлєва Е. Ю. Научно-исследовательская инфраструктура Интернет // *Вопр. философии.* – 2010. – № 8. – С. 155–166; Components of a Research 2.0 Infrastructure / T. D. Ullmann, F. Wild, P. Scott, E. Duval, B. Vandeputte, G. Parra, W. Reinhardt, N. Heinze, P. Kraker, A. Fessl // *Sustaining TEL : From Innovation to Learning and Practice.* 5th European Conference on Technology Enhanced Learning, EC-TEL 2010, Barcelona, Spain, September 28 – October 1, 2010. Proceedings – Berlin, Heidelberg : SpringerVerlag Heidelberg, 2010. – P. 590–595.

⁵ Полилова Т. А. Персональные веб-страницы в научном сообществе // *Труды Международной конференции «Научный сервис в сети Интернет: эксафлопное будущее»* (19–21 сентября 2011 г., г. Новороссийск). – Новороссийск, 2011. – С. 476–479. – Режим доступа : <http://agora.guru.ru/abrau2011/pdf/476.pdf>.

⁶ Костенко Л., Симоненко Т. Стратегії технологічної модернізації наукових бібліотек // *Бібл. вісн.* – 2012. – № 6. – С. 53–55.

метаданих цитування. Кожне з цих рішень має як свої переваги, так і недоліки. Суттєвим недоліком першого рішення є необхідність повної перебудови всіх бібліотечних процесів, і, як наслідок, виокремлення репозиторія, що створюється із єдиної системи опрацювання та представлення електронних ресурсів бібліотеки. Друге рішення передбачає значне програмно-технологічне доопрацювання веб-інтерфейсу традиційних бібліотечних електронних каталогів, записи яких представлені у спеціалізованих бібліотечних форматах. Однак, як показав досвід, Національної бібліотеки Франції (НБФ) (The Bibliothèque nationale de France), введення до стандартних слабоструктурованих, з точки зору пошукових робіт, бібліографічних описів упорядкованої системи метаданих вирішує проблеми підвищення доступності бібліографічної інформації в інтернет-середовищі. Результати дослідження Н. Буле з впровадження різних схем метаданих паралельно з традиційним представленням бібліографічних описів представлені на інформаційному порталі «Data.bnf.fr» НБФ (<http://data.bnf.fr/>). Реалізовані в межах проекту технології семантичного вебу дають змогу зв'язувати дані з різних баз даних, каталогів і репозиторіїв, не змінюючи їх індивідуальної внутрішньої структури. Згідно з принципами відкритості, головною вимогою до організації даних є їх доступність і можливість багаторазового використання для різних інформаційних завдань. Реалізація цих принципів стосовно бібліотечних каталогів дає змогу розв'язати суперечність між наявними величезними масивами бібліотечних даних та незручністю їх використання у веб-середовищі. Як зазначають автори проекту, «Data.bnf.fr» є відкритим проектом даних на основі семантичних веб-стандартів та інструментів. Він дає змогу інтегрувати в єдиній точці доступу інформацію різних каталогів НБФ, цифрової бібліотеки Галіка (Galica, <http://gallica.bnf.fr>) та інших веб-ресурсів ⁷... *(Лобузін К. Електронна наукова періодика відкритого доступу: семантичні веб-технології для бібліотек // Бібліотечний вісник. – 2015. – № 3. – С. 18–20).*

⁷ Буле В. Информационная среда, ориентированная на потребности пользователя: будущее за структурированными данными // Электронная библиотека. Вып. 4 : Научные и организационно-технологические основы интеграции цифровых информационных ресурсов: сб. науч. тр. – СПб, 2014. – С. 179–214.