УДК 025.3:004.057.3

Станіслава Клочок,
мол. наук. співробітник Національної бібліотеки України імені В. І. Вернадського

Формування систем метаданих цифрових історико-культурних ресурсів бібліотек

Статтю присвячено питанням інформаційної архітектури електронних історико-культурних ресурсів. Розглянуто проблеми опису і організації полідокументних ресурсів, вибору систем метаданих та інтеграції метаописів. Зроблено огляд основних систем метаданих для бібліотек, архівів і музеїв.

К л ю ч о в і с л о в а: електронні колекції, інформаційна архітектура історико-культурних ресурсів, інтероперабельність колекцій, каталогізація електронних ресурсів, стандарти метаданих.

Підтримка стандартів опису електронних ресурсів відіграє значну роль у їх створенні. Організація ж доступу в інформаційних системах (бібліотечні каталоги, електронні бібліотеки, портали) здійснюється за допомогою метаданих. Аналіз публікацій І. П. Антоненко, М. Бака (M. Baca), О. В. Баркової, Н. В. Браккера, В. Буле, М. Вален (M. Whalen), О. М. Волохіна, М. С. Вудлі (M. S. Woodley), Т. Джил (Tony Gill), А. Дж. Джилліленд (A. J. Gilliland), Л. А. Куйбишева, К. В. Лобузіної, П. М. Марченко, Т. Л. Місхулії, Ю. Г. Селіванової, Н. В. Стрішенець, Ю. Ю. Юмашевої та інших авторів, які досліджують питання вибору стандартів метаданих для опису інформаційних ресурсів, виявив недостатню увагу фахівців до полідокументності бібліотечних фондів та інтеграції метаописів різних інформаційних ресурсів. Крім того, і питання формування метаданих інтегрованих цифрових історико-культурних ресурсів недостатньо (фрагментарно) висвітлені в українських бібліотекознавчих публікаціях. Тому означені питання потребують додаткових досліджень у контексті активізації створення бібліотеками України масштабних цифрових колекцій історико-культурного надбання [2–4; 6; 7; 13; 15].

Проблеми опису та організації ресурсів

Будь-який інформаційний цифровий ресурс нині має свою інформаційну архітектуру, яка є невід'ємним компонентом бібліотечної діяльності. Для полегшення роботи з різними документами, які зберігаються у бібліотечних фондах, фахівцями була вироблена спеціальна інформаційна архітектура. У бібліотечній діяльності сьогодні використовуються складні системи правил, що дають змогу відібрати, оцінити, ідентифікувати, описати, структурувати та організувати колекції документів. Термін «інформаційна архітектура» з'явився нещодавно. Він має декілька варіацій визначень, його дефініція ще не усталилася. Наприклад, Л. Розенфельд і П. Морвіль під інформаційною архітектурою розуміють мистецтво і науку структурування, класифікацію веб-сайтів та інтрамереж, які полегшують користувачам пошук інформації та керування нею. Сюди ж вони включають і структурне проектування інформаційного простору, що сприяє виконанню завдань організації інтуїтивного доступу до змісту. Крім того, автори зазначають: «бібліотекарі давно займаються організацією інформації та доступом до неї, мають підготовку з технологій пошуку, перегляду та індексування» [11]. Отже, сучасна інформаційна архітектура продуктивно використовує досвід бібліотек для організації інформаційних ресурсів різної тематики у всьому світі. Слід зазначити, що цей досвід уможливлює вирішення цілої низки проблем з формування цифрових ресурсів, адже бібліотечні фонди дуже різні – книги, стародруки, рукописи, архівні документи, образотворчі матеріали, музичні твори тощо, і все це потребує опису, структурування.

Проблема полідокументності постає не тільки перед бібліотеками, а й всіма установами культурної спадщини, де також зберігаються різні об'єкти культури, використовуються спеціальні правила і стандарти. У зв'язку з оцифруванням фондів з'являється потреба в описуванні цифрових копій – не лише для обліку, але і для створення цифрових колекцій та надання доступу до них шляхом розміщення цифрових ресурсів в онлайновому доступі. Переведення об'єктів у цифрову форму, створення ретельно продуманої загальної системи метаданих та інформаційної архітектури значно покращують доступ до інформаційних ресурсів, створюють передумови для вільного обміну описами цифрових об'єктів.

Процес вибору стандарту метаданих пов'язаний з вирішенням таких основних проблем:

© С. Клочок, 2015

- організація і структурування інформації;
- описування різних видів документів і творів культури, а також їх цифрових копій;
- врахування багатомовності колекцій та їх представлення у цифровому ресурсі;
- уможливлення обміну описами.

Під час створення метаописів цифрових копій історико-культурних ресурсів виникає проблема з описом сурогатів * цих ресурсів. Тут важливо розрізняти ресурс і його сурогат. Метадані сурогату не можуть збігатися з описом вихідного ресурсу. Наприклад, оригінал художнього твору може бути поза сферою дії авторського права, тоді як авторські права на фотографію цього твору охороняються. Ресурси можуть бути пов'язані один з одним різними відносинами (створення, оприлюднення (публікація), права на використання). Таким чином, крім описових метаданих у цифрового ресурсу виникає ще цілий спектр додаткових елементів опису, які поділяються на такі групи:

- адміністративні метадані – управління та адміністрування ресурсів (придбання, права, місце зберігання);
- описові метадані – опис або ідентифікація ресурсів (каталожні описи);
- метадані збереження – управління процесами збереження (стан, міграція даних);
- технічні метадані – функціональні особливості системи (формати, паролі, водяні знаки, обсяг ресурсів);
- метадані використання – рівень і тип використання ресурсів (правила доступу, користувач та інформація про нього) [9].

Для вирішення завдань ідентифікації та збирання додаткових метаданих цифрових ресурсів у 2001 р. Федерація цифрових бібліотек (Digital Library Federation) за підтримки Бібліотеки Конгресу США розробила стандарт метаданих METS (Metadata Encoding and Transmission Standard – Стандарт кодування і передавання метаданих). Він використовується для кодування описових, адміністративних і структурних метаданих про об'єкти, що зберігаються у цифровій бібліотеці, опису цифрових сурогатів на рівні одиниці опису та для пошуку цих об'єктів на колекційному або груповому рівнях [18].

Важливим компонентом стратегій збереження цифрової спадщини є метадані, що містять інформацію, необхідну для документування процесу зберігання цифрових ресурсів. Метадані збереження містять відомості про формат, структуру і використання цифрового ресурсу, історію всіх операцій, пов'язаних з об'єктом збереження, у т. ч. про будь-які зміни, автентичність, технічну історію, історію зберігання, відповідальність, права, пов'язані зі збереженням тощо. Для розроблення метаданих збереження у 2003 р. OCLC (Online Computer Library Center – Онлайновий комп'ютерний бібліотечний центр) і RLG (Research Libraries Group – Товариство дослідницьких бібліотек) було створене наукове товариство PREMIS (Preservation Metadata: Implementation Strategies – Збереження метаданих: стратегії реалізації). До нього ввійшли PREMIS-представники наукових універсальних та національних бібліотек, музеїв, архівів, урядових і комерційних організацій з шести країн світу. За допомогою цього формату можуть описуватися всі атрибути, обставини та події, що супроводжують цифровий об'єкт: сканування, цифрова обробка, публікація, місце зберігання, права доступу до нього [20]. Рекомендації цього стандарту були враховані НБУВ під час розроблення бази даних обліку робіт з оцифрування документів на традиційних носіях.

За останнє десятиліття з'явилося безліч ініціатив, які продукували нові стандарти на основі вже існуючих. Систем метаданих як універсальних, так і орієнтованих на окремі види електронних колекцій, наприклад, у сфері науки, культури та освіти, налічується вже більше 50 (див. рис. 1) [1]. Таке розмаїття стандартів породжує цілу низку проблем: інколи метадані, які мають

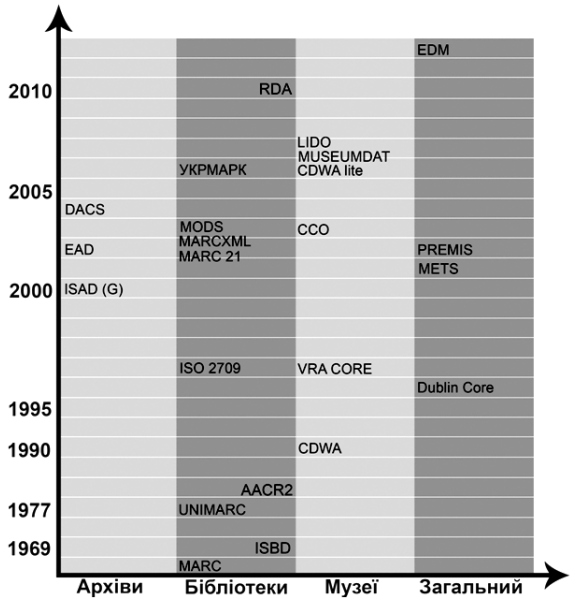

Рис. 1. Збільшення кількості стандартів метаданих об'єктів культури

* Сурогат – це подання ресурсу у відмінній від оригіналу формі. Наприклад, цифрове зображення або фотографія твору мистецтва, подання деякої території у вигляді віртуальної реальності або факсиміле деякого об'єкта.

різний зміст, отримують у різних системах однакові найменування; спостерігається неузгодженість метаданих у межах однієї колекції; фіксується занадто багато зайвої інформації; відсутня ключова контекстна інформація; не завжди має місце відповідність технічним стандартам.

Іноді дуже важко обрати єдиний стандарт опису, тому що не існує такого мірила, яке б давало змогу описувати всі види об'єктів культури та їх цифрові копії. В останні роки багато дискутують щодо установ культурної спадщини, зокрема бібліотек, архівів і музеїв. У структурі ІФЛА з'явилася навіть асоціація цих установ, яка отримала назву LAM (Libraries, Archives and Museums – Бібліотека – архів – музей) [21]. Кожна з цих структур продовжує працювати незалежно від інших, водночас всі вони мають спільні цілі, ресурси і користувачів, а отже, повинні об'єднуватися для реалізації спільних завдань, у т. ч. у сфері метаданих. Найбільше змістове навантаження у будь-якій інформаційній архітектурі несуть описові метадані, тому їм необхідно приділяти особливу увагу.

Стандарти описових метаданих

FRBR (Functional Requirements for Bibliographic Records) – Функціональні вимоги до бібліографічних описів) (IFLA, 2008) описують головну функцію метаданих як задоволення чотирьох основоположних потреб користувачів інформаційних систем: знайти, ідентифікувати, відібрати та отримати інформаційні об'єкти [18]. В епоху комп'ютеризації на основі бібліографічних стандартів були розроблені спеціальні, машиночитані, стандарти. Це дало змогу створити електронні бази описів бібліотечних документів, забезпечити пошук, управління, багатоцільове використання кожного опису, уможливило обмін описами з іншими організаціями. Сьогодні метадані є тим елементом інформаційної архітектури, на якому тримається будь-який цифровий ресурс. Комплекс стандартів, розроблених у рамках ініціатив системи бібліотека – архів – музей (LAM), можна формально розділити на три основні групи: стандарти структури даних, стандарти змісту, стандарти представлення у машиночитаній формі (див. рис. 2).

Розглянемо більш детально процес становлення та розвитку цих систем стандартів метаданих у різних установах документної спадщини.

Бібліотечні стандарти. У свій час розвиток бібліотечних метаданих мав на меті, насамперед, забезпечення інтелектуального та фізичного доступу до бібліотечних фондів. Стандартом структури бібліотечних даних сьогодні є зведене видання ISBD

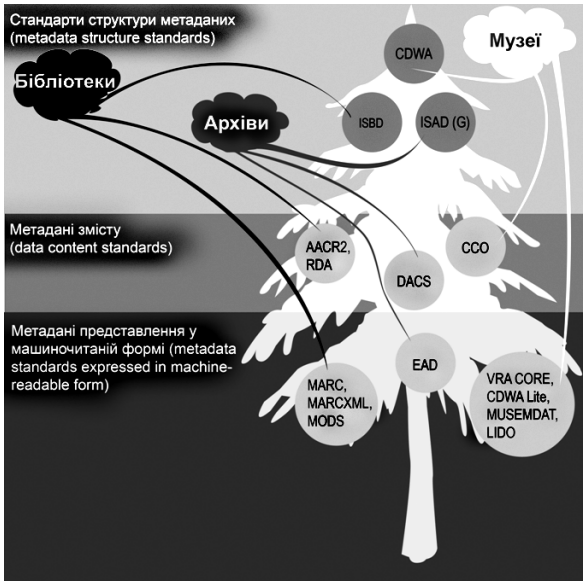

Рис. 2. Стандарти метаданих
у бібліотечній, архівній та музейній практиці

(International Standard for Bibliographic Description – Міжнародний стандарт бібліографічного опису) від 2007 р. Воно замінило більш ранні окремі ISBD (стародруків, картографічних матеріалів, періодики, електронних ресурсів, некнижних матеріалів і нотних записів). Еволюція ISBD пов'язана з набуттям ним у 1960 р. статусу стандарту Міжнародної організації зі стандартизації ISO 2709. З того часу усі стандарти бібліографічного опису, які затверджуються й використовуються у країнах, що співпрацюють у інформаційній галузі, узгоджуються з ISBD, зберігаючи при цьому певні національні відмінності. Бібліотечні метадані включають у себе бібліографічний опис, індекси, предметні рубрики, анотації, примітки тощо, створені відповідно до правил каталогізації (стандартів змісту даних), таких як AACR2 та їх комп'ютерної реалізації – стандартів машиночитаної каталогізації, таких як MARC формат.

Реалізація ідей FRBR та необхідність опису цифрових ресурсів спричинили перегляд AACR (Anglo-American Cataloguing Rules – Збереження метаданих: стратегії реалізації) та появу більш сучасного їх варіанта (розробленого у 2010 р.) – RDA (Resource Description and Access – Опис ресурсів та доступ), призначеного для каталогізації ресурсів у цифрову епоху, яка, у свою чергу, диктує необхідність опису ресурсів, представлених у базах даних. При цьому робиться наголос на універсальності RDA та незалежності від формату і відображення на екрані. Стандарт RDA призначений для полегшення пошуку, ідентифікації, відбору та отримання необхідної інформації користувачами [5].

Формати MARC широко застосовуються для ство-

рення електронних каталогів. Програма MARC1 була розроблена Бібліотекою Конгресу США у 1965–1966 рр. з метою отримання даних каталогізації у машиночитаній формі. У 1977 р. Міжнародна федерація бібліотечних асоціацій випустила видання «Універсальний формат MARC» (Universal MARC Format, UNIMARC). Цей формат повинен був стати посередником між будь-якими національними версіями форматів MARC і, тим самим, забезпечувати конвертування даних з національного формату у UNIMARC, а з нього – в інший національний формат. 1999 р. у результаті узгодження та наступного злиття бібліографічних форматів США і Канади (USMARC і CANMARC) було оголошено про створення на їх основі нового формату («Формату XXI століття») – MARC21, що включає у себе формати: бібліографічних даних, авторитетних даних, даних про фонди, класифікаційних даних, суспільної інформації. Сьогодні формат MARC21 використовується у США, Канаді, Австралії, Новій Зеландії, в університетських бібліотеках Великої Британії, Франції, Угорщини, Данії, Іспанії, Швеції, Фінляндії, Італії.

Для вирішення завдань інтеграції даних, зокрема даних у форматі MARC21 з іншими форматами, фахівці Бібліотеки Конгресу США разом із заінтересованими експертами розробили схему метаданих MODS (Metadata Object Description Standard – Стандарт метаданих опису об'єктів). Формат підтримується у вигляді XML-схеми. MODS-формат розширений порівняно з DC (Dublin Core) і спрощений порівняно з MARC. Використовується бібліотеками США, Канади, Великої Британії, Австралії, Німеччини для ефективного обміну даними.

Слід згадати також розроблення національного формату обміну даними УКРМАРК, який має загальнодержавне значення. Він був створений у рамках міжвідомчої групи розробників (2003–2006 рр.), яка об'єднала фахівців Національної бібліотеки України імені В. І. Вернадського (Національна академія наук України), Національної парламентської бібліотеки України (Міністерство культури і туризму України), Наукової бібліотеки ім. М. Максимовича Київського національного університету імені Тараса Шевченка (Міністерство освіти і науки України) в рамках договору про інформаційну співпрацю за підтримки Міжнародного фонду «Відродження». Результати співпраці провідних українських книгозбірень представлені на порталі НБУВ у вигляді електронної версії формату УКРМАРК [14]. Останній, зберігаючи структуру UNIMARC, має низку особливостей, які випливають з його первісної орієнтації на підтримку не лише комунікативних, а й каталогізаційних функцій. Він призначений забезпечувати комп'ютерну каталогізацію документів, сприяти обміну бібліографічними даними, розробленню сумісних внутрішніх форматів інформаційних систем бібліотек та інших бібліографічних установ України.

Бібліографічні метадані створювалися систематично і за єдиними стандартами. З 1960 р. існує можливість обміну цими описами, вони використовуються бібліотеками і користувачами в автоматизованих системах, таких як OPAC (Online Public Access Catalogue – Онлайновий каталог загального доступу) або у комерційних базах даних. Сьогодні бібліографічні метадані створюються не тільки бібліотечними працівниками, а й автоматизованим способом за допомогою таких засобів, як конвертація метаданих, збір метаданих та веб-індексування. Автоматизація формування метаданих набуває дедалі більшого поширення з розвитком Resource Description Framework (RDF) і Семантичного Вебу (Semantic Web). Інтеграції ресурсів також сприяють протоколи OAI (Open Archives Initiative – Ініціатива «Відкриті архіви») і Z39.50, призначені для збору інформації, що зберігається у базах даних або сховищах, яку неможливо зібрати за допомогою стандартних протоколів http і html. Вони підтримують різний синтаксис описів, але при цьому обов'язково використовують стандарти метаданих. Міжнародні проекти, які розробляють різні сервіси збирання та інтеграції даних, для агрегації метаописів використовують саме ці протоколи.

Архівні стандарти. Важливою складовою створення архівних метаданих є традиційне зосередження на обставинах створення документа. З'ясування і збереження контексту дають змогу виявити і зберегти докази існування, достовірності описів і артефактів; сприяють аутентифікації цих об'єктів, допомагають дослідникам їх аналізувати та інтерпретувати. Такий підхід до опису документів визначає також особливості системи архівних метаданих. Архівні і рукописні метадані (Archival and manuscript metadata), більш відомі як archival description (архівний опис), включають у себе каталог описів, описи окремих документів, допоміжні пошукові елементи. Під час розроблення архівних стандартів структури даних упродовж останніх трьох десятиліть враховувався досвід створення бібліографічних стандартів [8]. Вони включають MARC AMC (MARC Archival and Manuscripts Control) формат, опублікований Бібліотекою Конгресу у 1984 р. (тепер інтегрований у MARC21 для бібліографічного опису); ISAD (G) (General International Standard Archival Description, Second Edition – Загальний міжнародний стандарт

архівного опису), друге видання якого підготовлене Міжнародною радою архівів у 2000 р.; стандарт змісту даних, що описує архіви: DACS (Describing Archives: A Content Standard – Опис архівних матеріалів: стандарт змісту), який вперше був опублікований у 2004 р.; стандарт представлення архівних даних у машиночитаній формі EAD (Encoded Archival Description – Кодований архівний опис), ухвалений Товариством американських архівістів (Society of American Archivists) у 2002 р. Він сумісний зі стандартом MARC21. Значний позитивний досвід упровадження і апробації EAD накопичений архівами США, Франції, Німеччини, Нідерландів, Великої Британії. В українській архівній галузі також існують напрацювання щодо створення власної дескриптивної системи. На основі міжнародного стандарту архівного описування (загального) ISAD(G) розроблено стандарт ДСТУ 4331:2004: «Правила описування архівних документів (ПАО)» [10].

Музейні стандарти. Консенсус і співпраця у галузі метаданих значно повільніше досягалися у музейній справі, де переваги стандартизації опису, спільної каталогізації та обміну описовими даними були менш очевидні до відносно недавнього часу. З кінця 1990-х рр. такі інструменти, як стандарт CDWA (Categories for the Description of Works of Art – Категорії описів творів мистецтва), CCO (Cataloging Cultural Objects – Каталогізація об'єктів мистецтва і культури), Museumdat (Музейні дані), LIDO (Lightweight Information Describing Objects – Полегшений стандарт інформації для опису об'єктів) і CDWA Lite XML (полегшений CDWA) почали розглядатися і реалізовуватися музеями [9].

Основу музейного опису становить стандарт CDWA, розроблений Інститутом Гетті спільно з Американською асоціацією мистецтв. Розроблення CDWA розпочалося у 90-ті і тривало майже два роки. У проекті брали участь фахівці у галузі мистецтва, інформаційних технологій, бібліотечної та архівної справи. Правила CDWA описують загальні, концептуальні, принципи опису об'єктів мистецтва та зображень, безвідносно до конкретної реалізації музейних баз даних, вони лише задають орієнтири для розроблення мистецьких інформаційних систем. У CDWA подано схематичне представлення вимог і принципів, які обґрунтовані дисципліною «Історія мистецтва». Категорії CDWA розташовуються таким чином, щоб розкриватися від опису твору мистецтва до опису, аналізу текстових / візуальних джерел інформації про нього. У певному сенсі це розгортання інформації від об'єктивних даних до суб'єктивних інтерпретацій.

Категорії стали основою розроблення першого стандарту, спеціально призначеного для каталогізації музейних об'єктів та предметів мистецтва – CCO (Cataloging Cultural Objects – Каталогізація культурних об'єктів). CCO орієнтований на якісну каталогізацію. Твір культури, мистецтва як результат творчої діяльності людини описується за такими критеріями: що це, хто це зробив, де це зроблено, як це зроблено, який матеріал використаний і т. п. Під копією розуміється візуальне відтворення твору. Унікальність правил стандарту полягає також у тому, що вони утримують рівновагу між каталогізацією об'єктів культури та їх візуальними сурогатами, де копія об'єкта культури зазвичай представлена у вигляді фотографії або образу об'єкта, якщо йдеться про цифровий формат. Для музейних стандартів характерно включення контрольованих точок доступу до основних правил опису предмета мистецтва.

Для машиночитаного представлення опису музейних об'єктів використовується XML-схема VRA Core (Visual Resoures Associations Core – Базова схема Асоціації візуальних ресурсів). Вона повністю узгоджена з CCO, що гарантує найкраще представлення інформації про музейний об'єкт. Схема визначає елементи метаданих, котрі використовуються для опису творів мистецтва та їх копій. Набір складників VRA містить дві групи елементів, категорії опису об'єкта (Work Description Categories) і категорії опису візуального документа (Visual Document Description Categories).

Одним з найбільш значних досягнень CDWA стало створення у 2006 – 2007 рр. схеми XML під назвою MUSEUMDAT, яка перевершує схему CDWA Lite. Вона включає більше зібрань, що належать до історії природи та історії культури, і представляє елементи відповідно до моделі CIDOC. Формат призначено для агрегування (збору) даних (XML-схема), оптимізованого пошуку і публікації на музейних порталах, а також для автоматичного збору ключових даних.

XML-схеми CDWA Lite і MUSEUMDAT виявилися досить вдалими, вони утворили модель, що визначає мінімальний набір інформації, необхідної для полегшення розкриття ресурсів, а також уможливили рішення, яке дає змогу установам безперешкодно брати участь у постачанні своїх колекцій до зведених каталогів та електронних сховищ. Всі ці починання засвідчують, що при зростаючих очікуваннях універсальних засобів пошуку кожен повинен мати можливість зробити свої колекції доступними онлайн. Для цього можна успішно використовувати існуючі стандарти. CDWA Lite була першовідкривачем на цьому шляху, а MUSEUMDAT розширила

сферу свого застосування до такої міри, щоб охопити більше колекцій і стати гнучкішою, враховувати описові концепти і зв'язки, котрі використовуються для опису документів культурної спадщини. Схема LIDO (Lightweight Information Describing Objects – Полегшений стандарт інформації для опису об'єктів) – це спільна ініціатива творців форматів CDWA Lite і MUSEUMDAT, яка дала змогу об'єднати ці формати і забезпечити сумісність. LIDO добре зарекомендував себе у проекті ATHENA і дозволив агрегувати не тільки музейні, а й бібліотечні, архівні інформаційні ресурси. Він широко використовується в проектах Europeana.

Бібліотечні, архівні і музейні метадані спочатку призначалися для локальних цілей і зберігалися в окремих сховищах до кінця 1990-х рр., але тепер вони поширюються в Інтернеті через різні ресурси, зокрема OCLC, WorldCAT, World Digital Library (Світова цифрова бібліотека), Europeana, MINERVA, Мультикультурна Канада (Multicultural Canada), Архіви США (Archives USA), Пам'ять Америки Бібліотеки Конгресу США (Library of Congress's American Memory Project).

Метадані та інтеграція ресурсів

У книзі М. Бака «Вступ у метадані 3.0» (принцип 6, розділ «Практичні принципи створення метаданих та технічне обслуговування») [2] наголошується , що немає єдиного стандарту метаданих, який був би достатнім для опису всіх типів колекцій і матеріалів; вибір найбільш відповідного пакета стандартів метаданих та інструментів, створення прозорої, узгодженої схеми метаданих відповідно до цих стандартів не тільки дасть змогу добре описати конкретні колекційні матеріали, а й забезпечить сумісність метаданих з різними стандартами установ культурної спадщини.

На сьогодні цілком очевидним є те, що застосування загальноприйнятих стандартів та забезпечення їх інтероперабельності – єдиний спосіб інтеграції інформаційної співпраці установ культурної спадщини. Загальновизнаним стандартом, який є мінімально необхідним для опису цифрових об'єктів виступає стандарт метаданих Дублінського ядра. Він містить елементи опису, спільні для всіх об'єктів комплексу документної спадщини бібліотека–архів–музей. Дублінське ядро характеризується простотою створення і підтримки, зрозумілою семантикою; набір основних елементів тут компактний, але дає змогу задавати практично всі необхідні атрибути. Структурні блоки метаданих Dublin Core (DC), розроблення якого розпочалось у 1995 р., включають у себе елементи бібліографічного опису, бібліотечних контрольованих словників і класифікаційних схем. Стандарт був затверджений у США ANSI (American National Standards Institute – Американський інститут національних стандартів) як стандарт Z39.85. У 2003 р. стандарт ISO 15836 затвердив набір метаданих DC в якості міжнародного стандарту [17]. Важливу роль у концепції метаданих DC відіграють контрольовані словники – значення для відповідного поля метаописів обираються з фіксованої множини слів або кодів, яка обмежена набором ретельно підібраних пошукових термінів. Завдяки універсальності формату атрибути Дублінського ядра так чи інакше входять практично до усіх більш складних форматів метаданих.

Метадані, що використовуються фахівцями установ культурної спадщини, часто регулюються стандартами, створеними різними співтовариствами і установами з метою забезпечення якості, узгодженості та сумісності. Для вирішення проблем інтеграції цих систем метаданих у рамках проекту Europeana було розроблено власні стандарти метаданих Europeana Semantic Elements (ESE) і Europeana Data Model (EDM), а також сервіс агрегування і конвертування MINT (Metadata interoperability Service – Сервіс сумісності метаданих). ESE і EDM містять елементи Дублінського ядра, а також елементи, необхідні для задоволення потреб проекту Europeana. В останньому зібрано більше 20 млн текстів, зображень, відео та звукозаписів. Всі описи цих об'єктів, з якими працюють сервіси Europeana, зберігаються у форматах LIDO, ESE і EDM. Частиною моделі EDM є система Linked Open Data (LOD) – гіпертекстова система, призначена для публікації структурованих даних. Вона зв'язує метадані між собою і уможливлює масштабування, представлення різних аспектів змісту, виявлення і створення перехресних посилань між відповідними ресурсами [16; 22].

MINT дає змогу збирати метадані з кількох джерел, конвертувати імпортовані описи у проміжні формати метаданих, розподіляючи необхідні елементи метаданих однієї схеми в іншу, потім конвертувати їх в остаточний формат ESE і зберігати метадані у репозиторії. MINT використовує протокол OAI-PMH і конвертує описи у ESE відповідно до вимог проекту.

Фонд НБУВ є унікальним зібранням джерел інформації. Він містить книги, періодичні видання, карти, ноти, образотворчі матеріали, рукописи, стародруки тощо. Однією з важливих проблем інтеграції полідокументних ресурсів фонду НБУВ є забезпечення сумісності опису архівних та бібліотечних документів. Для стандартизованого обміну за-

писами між архівними каталогами призначений формат метаданих для архівного опису EAD, котрий забезпечує функції, аналогічні MARC-формату стосовно бібліографічних записів. Формат передбачає забезпечення сумісності у кодованих даних з MARC21. На рівні окремих описів забезпечена сумісність основних точок доступу. Для здійснення міжнародної інтеграції створених електронних ресурсів книжкових та архівних фондів формуються відповідні конвертери. Конвертування даних у формат MARC або DC дає змогу скористатися вже розробленими протоколами передавання даних.

Крім завдань повноцінного опису полідокументного фонду НБУВ, особливої актуальності набуває адаптація прийнятих рішень до міжнародних стандартів, що зумовлено участю Бібліотеки у світових цифрових проектах. Саме тому, для представлення метаданих у Всесвітній цифровій бібліотеці (World Digital Library – WDL) і у Бібліотечній асоціації Євразії (БАЄ) нашими фахівцями були розроблені спеціальні робочі аркуші. Основу стандарту описових метаданих становить DC-формат. Для цих проектів мета-описи створювалися на трьох мовах – українській, російській та англійській, за необхідності використовувалися Virtual International Authority File (VIAF), списки MARC List for Relators, ISO 639-3 (трибуквені коди мов), Getty's Thesaurus of Geographic Names (TGN), Library of Congress Subject Headings (LCSH), Десяткова класифікація Дьюї (Dewey classification number, DCD).

Висновки

У галузі опису цифрових історико-культурних об'єктів існує обмежена кількість ключових стандартів, які широко використовуються у світі. Незважаючи на те, що вони є досить популярними у великих міжнародних проектах, основною проблемою залишається забезпечення інтероперабельності та досягнення сумісності. Стандарт Дублінського ядра активно використовується сьогодні як «мінімальний стандарт» метаданих для пошуку ресурсів і як основа багатьох стандартів метаданих, на яку нарощуються більш специфічні елементи. Пріоритетним завданням залишається розроблення механізмів перетворення метаданих із специфічних, для певного сектора культури, стандартів у метадані стандарту Дублінського ядра. Однак, такий «мінімальний» підхід до метаданих призводить до втрат інформації, що міститься в елементах, які не охоплені DC-схемою, і відновити її після такої конвертації неможливо.

Сьогодні, в умовах інтенсивного формування інформаційних ресурсів електронної культури, цифрові бібліотеки, онлайн-архіви і віртуальні музеї активно працюють над створенням інтегрованого історико-культурного середовища, яке б давало змогу кожному ознайомитися з електронними колекціями незалежно від їх місцезнаходження. У цьому контексті зрозуміло, що майбутнє метаданих полягає у здатності успішно інтегрувати дані в інші формати, схеми і стандарти у бібліотеках, архівах і музеях з метою створення функціональної сумісності технологій і реалізації співпраці різних культурних установ.

Список використаних джерел

1. *Антопольский А. Б.* Электронные библиотеки: принципы создания : учебное пособие [Электронный ресурс] / А. Б. Антопольский, Т. В. Майстрович. – Москва : ЛИБЕРЕЯ-БИБИНФОРМ, 2007. – 288. – Mode of access : http://www.big-lib.com/book/71_Elektronnie_biblioteki_principi_sozdaniya

2. *Бака М.* Вступление в метаданные 3.0 [Електронный ресурс] / Мурта Бака (Murtha Baca), Институт Гетти (США) ; Т. Джил (Tony Gill), Глобальная Научная библиотека (США) ; А. Дж. Джиллиленд (Anne J. Gilliland), Университет Калифорнии (США) ; М. Вален (Maureen Whalen) Институт Гетти (США) ; М. С. Вудли (Mary S. Woodley)), Университет Калифорнии (США) // Introduction to Metadata 3.0 : Second Edition // Getty Publications, 2008, Los Angeles, CA, 2008. – Mode of access : http://d2aohiyo3d3idm.cloudfront.net/publications/virtuallibrary/0892368969.pdf

3. *Бака М.* Совместное использование стандартов и специальных знаний в начале XXI века : на пути к созданию единой модели метаданных для разных сообществ [Електронний ресурс] / Мурта Бака (Murtha Baca), Институт Гетти (США) ; Элизабет О'Киф (Elizabeth O'Keefe), Библиотека и Музей Моргана (США) // World library and information congress: 74th IFLA general conference and council. 10 – 14 Aug. 2008, Quebec, Canada. – Quebec, 2008. – Mode of access : http://www.archive.ifla.org/IV/ifla74/papers/156-Baca_OKeefe-trans-ru.pdf.

4. *Баркова О. В.* Стандарты цифровых коллекций: проблемы нормативной базы, особенности национальной стандартизации [Электронный ресурс] // Международная научно-практическая конференция «Оценка и контроль качества сканирования бумажных документов», 19–29 апреля 2011 г., Королев. – Режим доступа : http://www.aiteh.ru/uploads/seminars/barkova.doc

5. *Бахтурина Т. А.* Новый интеллектуальный продукт в сфере каталогизации: Правила RDA «Описание ресурса и доступ» / Т. А. Бахтурина // Научные и технические библиотеки. – 2011. – № 7. – С. 35–45.

6. *Волохін О. М.* Каталогізація цифрових ресурсів Інтернет : Дублінське ядро метаданих : посіб. / О. М. Волохін – Кіровоград, 2003. – 70 с.

7. *Лобузіна К.* Технології організації знаннєвих ресурсів

у бібліотечно-інформаційній діяльності : монографія / К. Лобузіна; відп. ред. О. С. Онищенко ; НАН України, Нац. б-ка України ім. В. І. Вернадського. – К., 2012. – 252 с.

8. *Марченко П.* Стандарти кодованих архівного опису та контексту (проблема впровадження) / Петро Марченко // Архіви України. – 2009. – № 1/2. – С. 79–85.

9. Оцифровка: ландшафт стандартов для европейских музеев, архивов, библиотек : публикация Рабочей группы «Исследование стандартов и подготовка рекомендаций» проекта ATHENA [Электронный ресурс] / Gordon McKenna, Chris De Loof ; перевод на рус. яз. Н. Браккер, Л. Куйбышев. – 2010. – Режим доступа : http://www.minervaplus.ru/publish/standards_landscape.pdf

10. Правила описування архівних документів (ISAD (G):1999, NEQ) : ДСТУ 4331:2004 / розроб. Н. Христова [та ін.]. – Офіц. вид. – Чинний від 01.07.2005. – К. : Держспоживстандарт України, 2005. – IV, 16 с. – (Національний стандарт України).

11. *Розенфельд Л.* Информационная архитектура в Интернете, 2-е издание / Л. Розенфельд, П. Морвиль; пер. с англ. С. Маккавеева. – СПб : Символ-Плюс, 2005. – 544 с., ил. – Mode of access : http://www.novsu.ru/file/1035686

12. *Селиванова Ю.* Международные стандарты метаданных для описания библиотечных, архивных материалов и музейных объектов / Ю. Селиванова, Т. Масхулия // Бібліотечний вісник. – 2012. – № 4. – С. 18–29.

13. *Стрішенець Н. В.* Метадані у сучасному бібліотекознавстві. Метадані – нове чи старе поняття? / Н. В. Стрішенець // Бібліотекознавство. Документознавство. Інформологія . – 2010. – № 2. – С. 4–11.

14. УКРМАРК : нац. формат представлення бібліогр. даних : (проект) [Електронний ресурс]. – К., 2004. – Режим доступу : http://nbuv.gov.ua/node/1811

15. *Юмашева Ю. Ю.* Документы архивного фонда и их представление в архивных, библиотечных и музейных автоматизированных системах: проблемы описания и создания электронных копий // Информационные технологии и письменное наследие : материалы IV междунар. науч. конф. El'Manuscript–2012, Петрозаводск, 3–8 сент. – Петрозаводск ; Ижевск, 2012. – С. 302–306.

16. *Doerr M.* The Europeana Data Model (EDM) [Electronic resource] / M. Doerr // World Library and Information Congress : 76th IFLA General Conference and Assembly: «Information Technology, Cataloguing, Classification and Indexing with Knowledge Management». – Gothenburg, 2010. – Mode of access : http://conference.ifla.org/past-wlic/2010/149-doerr-en.pdf

17. Dublin Core Metadata Initiative (DCMI) [Electronic resource]. – Mode of access: http://dublincore.org.

18. Functional Requirements for Bibliographic Records (FRBR) [Electronic resource] // IFLA. – 1999. – Mode of access : http://www.ifla.org/publications/ functional-requirements-for-bibliographic-records.

19. *McDonough J. P.* METS: standardized encoding for digital library objects / J. P. McDonough // International journal on digital libraries. – 2006. – Т. 6. – № 2. – С. 148–158.

20. PREMIS Data Dictionary for Preservation Metadata [Electronic resource] : Final Report of the PREMIS Working Group. – 2005. – Mode of access : www.loc.gov/standards/premis/v2/premis-2-0.pdf.

21. Public Libraries, Archives and Museums : Trends in Collaboration and Cooperation [Electronic resource] / A. Yarrow, B. Clubb, J. Draper. – Hague : IFLA, 2008. – (IFLA Public Libraries Section. Professional reports, #108). – Mode of access : http://www.ifla.org/en/publications/ifla-professional-reports-108.

22. *Simou N.* AthenaPlus D 3.1 The MINT ingestion platform [Electronic resource] / Nikolaos Simou (NTUA), Eleni Tsalapati. – Roma, 2013. – 55 с. – (Access to cultural heritage networks for Europeana). – Mode of access : http://www.athenaplus.eu/getFile.php?id=271.

Стаття надійшла до редакції 17.02.2015.

UDC 025.3:004.057.3

Stanislava Klochok,
junior scientific researcher, the Vernadsky National Library of Ukraine

FORMATION OF METADATA SYSTEMS FOR DIGITAL CULTURAL AND HISTORICAL RESOURCES OF LIBRARIES

The issues information architecture of cultural heritage resources are devoted. The question of the problems of description multi-document resources, choice of metadata standards for integration and meta descriptions are detail. The main metadata standards for libraries, archives and museums are overview.

K e y w o r d s: digital collections, information architecture of cultural heritage resources, interoperability of collections, cataloging of electronic resource, metadata standards.

УДК 025.3:004.057.3

Станислава Клочок,
мл. науч. сотрудник Национальной библиотеки Украины имени В. И. Вернадского

ФОРМИРОВАНИЕ СИСТЕМ МЕТАДАННЫХ ЦИФРОВЫХ ИСТОРИКО-КУЛЬТУРНЫХ РЕСУРСОВ БИБЛИОТЕК

Статья посвящена вопросам информационной архитектуры электронных историко-культурных ресурсов. Рассмотрены проблемы описания и организации полидокументных ресурсов, выбора систем метаданных и интеграции метаописаний. Сделан обзор основных систем метаданных для библиотек, архивов и музеев.

К л ю ч е в ы е с л о в а: электронные коллекции, информационная архитектура историко-культурных ресурсов, интероперабельность коллекций, каталогизация электронных ресурсов, стандарты метаданных.