

УДК 681.3.01

ТЕХНОЛОГІЯ ПЕРЕТВОРЕННЯ БІБЛІОГРАФІЧНИХ ОПИСІВ

У статті розглянуто технологію обробки та принципи перетворення бібліографічних описів з ГОСТу 7.1-84 до діючого ДСТУ ГОСТ 7.1:2006. Досліджуються структури бібліографічних описів різних стандартів. Отримана інформація використовується для формування певного набору правил перетворення бібліографічних описів в діючий формат на прикладі автореферату дисертації. Запропоновано застосувати мову XML для розмітки бібліографічного опису

Ключові слова: бібліографічний опис, система підготовки бібліографічних описів, розширювана мова розмітки XML

В статье рассмотрена технология обработки и принципы преобразования библиографических описаний с ГОСТа 7.1-84 в действующий ДСТУ ГОСТ 7.1:2006. Исследуются структуры библиографических описаний разных стандартов. Полученная информация используется для формирования определенного набора правил преобразования библиографических описаний в действующий формат на примере автореферата диссертации. Предложено применить язык XML для разметки библиографического описания

Ключевые слова: библиографическое описание, система подготовки библиографических описаний, расширяемый язык разметки XML

О. В. Касілов

Кандидат технічних наук*

E-mail: o.kasilov@hotmail.com

А. В. Суботіна*

E-mail: subotinaalena@mail.ru

*Кафедра «Системи інформації»

Національний технічний університет

«Харківський політехнічний інститут»

вул. Фрунзе, 21, м. Харків, Україна, 61002

1. Вступ

З 1 липня 2007 року в Україні діє міждержавний стандарт ДСТУ ГОСТ 7.1:2006 «Система стандартів з інформації, бібліотечної та видавничої справи. Бібліографічний запис, бібліографічний опис. Загальні вимоги та правила складання» згідно з Указом Держкомітету України з питань технічного регулювання і споживчої політики від 10.11.2006 р. № 322. Цей стандарт відповідає попередньому ГОСТові 7.1-2003 та є базовим для складання бібліографічного опису всіх видів документів.

Багато видань, оформлених згідно попередніх ГОСТів (ГОСТ 7.1-84, ГОСТ 7.16-79, ГОСТ 7.18-79, ГОСТ 7.34-81, ГОСТ 7.40-82 та інші) необхідно перетворити до діючого стандарту для подальшої обробки.

2. Постановка задачі

Більшість українських бібліотек (наприклад, НБУ ім. В. І. Вернадського, ХГНБ ім. В. Г. Короленко та інші) користуються системою Web-ІРБІС. ІРБІС є інтегрованою бібліотечно-інформаційною відкритою системою. Система автоматизації бібліотек відповідає усім міжнародним вимогам до сучасних бібліотечних систем. У той же час вона підтримує всі традиції бібліотечної справи [1, 2], однак, не дозволяє впровадити ДСТУ ГОСТ 7.1:2006 [3] в повній мірі.

Це обумовлено тим, що в ній відсутні модулі для перетворення бібліографічного запису та опису із старої форми представлення до нової згідно діючого стандарту.

Програмне забезпечення для роботи з бібліографічною інформацією, наприклад, Biblioscape [4, 6], BiblioExpress [4], HotReference [4], Zotero [4], Библио-Мастер [5], EndNote [7] та інші, відносяться до систем управління бібліографічною інформацією. Ці програми дозволяють автоматизувати деякі функції по створенню каталогів публікацій різних видів і пошуку інформації в спеціалізованих БД та інше. Однак і вони не здатні забезпечити користувача інформацією, яка надана в новому форматі згідно діючого стандарту ДСТУ ГОСТ 7.1:2006.

Таким чином, на сьогоднішній день в Україні немає жодної системи підготовки бібліографічних описів, принаймні у вільному доступі. Тому виникає необхідність у розробці саме таких україномовних систем з метою удосконалення процесу перетворення бібліографічних записів та описів із старих стандартів до діючого (ДСТУ ГОСТ 7.1:2006).

3. Основна частина

Згідно до ГОСТу 7.60-2003 «Издания. Основные виды, термины и определения.» [8] існує багато видів видань, для яких потрібне перетворення із старих стандартів до нового діючого (ДСТУ ГОСТ 7.1:2006). Для цього розглянемо процес на прикладі автореферату дисертації та скористаємося ГОСТом 7.1-84 [9].

Бібліографічний опис включає до свого складу бібліографічні відомості про документи, які повинні бути наведені за певними правилами. При цьому встановлюється наповнення та порядок проходження областей та елементів призначених для ідентифікації

і загальної характеристики документа. Розглянемо наступний приклад.

За діючим стандартом ДСТУ ГОСТ 7.1:2006 *заголовок бібліографічного опису має вигляд* (з врахуванням вживання великої та малої літер):

Основна назва [*Загальне позначення матеріалу*] = Паралельна назва : відомості, які відносяться до назви/відомості про Авторство чи Відповідальність ; про інших Осіб. – Відомості про повторність видання/Відповідальність за видання. – *Зона специфічних відомостей*. – Місце видання : Вид-во, рік. – Фізична (кількісна) характеристика. – (Серія і підсерія; №, т.). – *Примітки (додаткова інформація від бібліографа, напр.: системні вимоги до електрон. ресурсів)*. – ISBN.

У бібліографічному описі використовують такі умовні розділові знаки: [. –] крапка і тире; [.] крапка; [,] кома; [:] двокрапка; [;] крапка з комою; [/] коса риска; [//] дві косі риски; () круглі дужки; [] квадратні дужки; [+] плюс; [=] знак рівності.

Пунктуація у бібліографічному описі виконує функції звичайних розділових знаків та умовних розділових знаків (за термінологією нового стандарту – знаки приписної пунктуації). Вони ставляться між зонами бібліографічного опису та їх елементами і виконують розпізнавальну функцію елементів та зон.

Розглянемо бібліографічний опис згідно до ГОСТу 7.1-84 на прикладі опису автореферату дисертації, який буде мати вигляд:

Мацюк Г. П. Кодифікація української мови в галицьких граматах першої половини XIX ст.: Автореф. дис. д-ра філол. наук: 10.02.01 / Ін-т мовознав. ім. О. О. Потебні НАН України. К., 2002. 32 с.

За діючим стандартом ДСТУ ГОСТ 7.1:2006 бібліографічний опис автореферату дисертації буде мати вигляд:

Мацюк Г. П. Кодифікація української мови в галицьких граматах першої половини XIX ст. : автореф. дис. на здобуття наук. ступеня д-ра філол. наук : спец. 10.02.01 “Українська мова” / Мацюк Галина Петрівна; Ін-т мовознав. ім. О. О. Потебні НАН України. – К., 2002. – 32 с.

Підкреслення в обох випадках вказують на невідповідності у бібліографічних описах.

Зробимо більш детальний аналіз бібліографічних описів та позначимо невідповідності на прикладах.

Перша відмінність полягає в тому, що в попередньому ГОСТі відсутній пробіл перед «:» після назви. Крім того, Автореф. пишеться з великої літери, а зараз обмежуємося прописною, і не вказуємо «на здобуття наукового ступеня».

Наступна відмінність – перед «:» з номером спеціальності відсутній пробіл у попередньому описі. Далі в новому ДСТУ у “ ” вказуємо назву спеціальності, а у попередньому увага на цьому зовсім не акцентується. Після цього в новому стандарті треба враховувати відомості про відповідальність, а тільки потім пишемо пробіл, двокрапка та знов пробіл.

Знаки пунктуації перед містом видавництва – у попередньому стандарті являє собою наступну відмінність це крапка та пробіл, а далі йде власне назва міста (якщо є скорочення – вказується крапка). У новому стандарті ставиться спочатку крапка, пробіл, тире та знову пробіл. Після року в попередньому стандарті вказують крапку та пробіл, а в новому до цього

додається тире та пробіл. Бібліографічний опис року ідентичний.

Згідно до приведених вище відмінностей маємо справу з багатьма проблемами, які можна вирішити такими шляхами:

1. Відповідно до наказу Міністерства освіти і науки, молоді та спорту України від 14.09.2011 №1057 «Про затвердження Переліку наукових спеціальностей» порівнювати дійсність номерів спеціальностей та доповнювати їх відповідними назвами.

2. Отримати доступ до Всеукраїнської бази даних авторів дисертацій (та авторефератів) з метою визначення повного ім'я та по батькові автора.

3. Підключити користувача до інших БД.

На рис. 1 наведено алгоритм роботи системи підготовки бібліографічного опису, де описано порядок обробки інформації. На початку роботи програми користувач вводить вхідний текст у попередньому ГОСТі 7.1-84, та після цього виконується перетворення до нового ДСТУ ГОСТ 7.1:2006. Далі відбувається перевірка на необхідність отримання додаткової інформації з допоміжних баз даних. Якщо є така необхідність, робимо пошук певних баз даних, підключаємось до них та доповнюємо недостатню інформацію. Для остаточної обробки знайденої інформації можливо буде потрібна допомога експерта. Якщо перетворення виконано без перешкод, тоді ці дії не виконуються. Таким чином, по закінченню роботи програми усі результати подаються користувачеві.

Розглянувши структуру бібліографічного опису, сформулюємо набір правил для перетворення вхідних даних (бібліографічний опис) у вихідні дані (бібліографічний опис ДСТУ ГОСТ 7.1:2006 в XML запису) [10].

$$PR_n(T_n^j) = P_n, \quad j = 1, M,$$

де PR_n – програма, що здійснює перетворення; T_n^j – складова частина бібліографічного опису; R_n – результат перетворення; j – номер складової частини бібліографічного опису.

Розмітка документа (бібліографічного опису) має на меті: виділення смислових частин (логічних елементів) документа і зв'язків між ними (структурна розмітка). У подальшому визначаються дії, які мають бути здійснені з цими елементами. Для цього використовують знаки приписної пунктуації (розділові знаки між зонами бібліографічного опису та їх елементами).

Розглянемо докладніше.

Згідно попереднього стандарту (ГОСТ 7.1-84):

Мацюк Г. П. Кодифікація української мови в галицьких граматах першої половини XIX ст.: Автореф. дис. д-ра філол. наук: 10.02.01 / Ін-т мовознав. ім. О. О. Потебні НАН України. К., 2002. 32 с.

Сформулюємо правила обробки бібліографічного опису за ГОСТом 7.1-84.

П1: Прізвище (*пробіл*)

П2: Ім'я (Символ у верхньому регістрі + .) (*пробіл*)

П3: По батькові (Символ у верхньому регістрі + .) (*пробіл*)

П4: Назва (: + *пробіл*)

П5: Автореф. дис. *** (: + *пробіл*)

П6: Код спеціальності (*пробіл* / *пробіл*)

П7: Видавництво (. + *пробіл*)

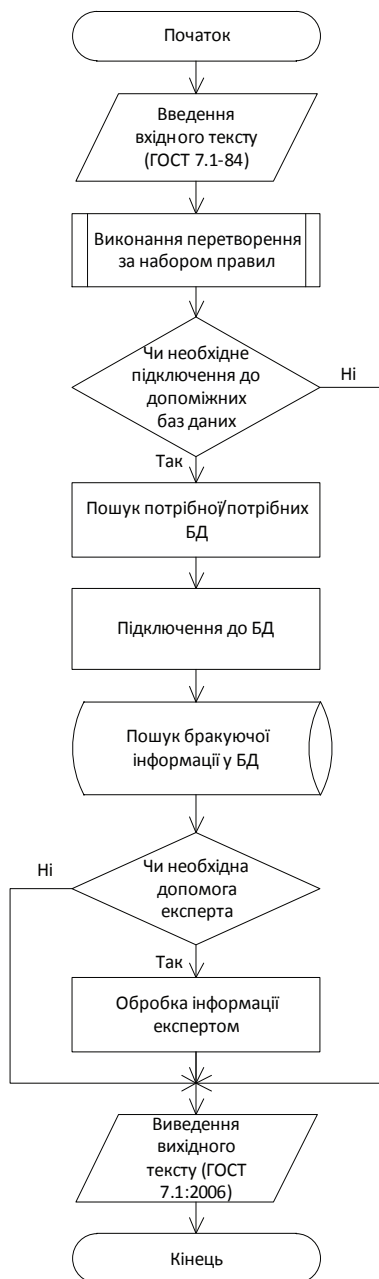


Рис. 1. Алгоритм роботи системи підготовки бібліографічного опису

П8: Місто (, + пробіл), якщо назва міста скорочена, тоді (. + , + пробіл)

П9: Рік видання (4 цифри + . + пробіл)

П10: Кількість сторінок (цифри + пробіл + с.)

Відповідно до стандарту ДСТУ ГОСТ 7.1:2006:

Мацюк Г. П. Кодифікація української мови в галицьких граматиках першої половини XIX ст. : автореф. дис. на здобуття наук. ступеня д-ра філол. наук : спец. 10.02.01 "Українська мова" / Мацюк Галина Петрівна ; Ін-т мовознав. ім. О. О. Потебні НАН України. – К., 2002. – 32 с.

Аналогічно формулюють правила обробки бібліографічного опису за ДСТУ 7.1:2006.

П1: Прізвище (пробіл)

П2: Ім'я (Символ у верхньому регістрі + .) (пробіл)

П3: По батькові (Символ у верхньому регістрі + .) (пробіл)

П4: Назва (пробіл + : + пробіл)

П5: автореф. дис. *** (пробіл + : + пробіл)

П6: спец. (пробіл)

П7: код спец. (пробіл)

П8: Назва спеціальності у "" (пробіл + / + пробіл)

П9: ПІБ повністю (пробіл + ; + пробіл)

П10: Видавництво (. + пробіл + – + пробіл)

П11: Місто (, + пробіл), якщо назва міста скорочена, тоді (. + , + пробіл)

П12: Рік видання (4 цифри + . + пробіл + – + пробіл)

П13: Кількість сторінок (цифри + пробіл + с.)

4. Висновки

В Україні є необхідність у розробці україномовних систем підготовки бібліографічних записів та описів з метою удосконалення процесу перетворення бібліографічних записів та описів з попереднього виду стандартів до діючого, що дозволяють інтегрувати їх до національних бібліотечних систем та web-ресурсів.

Дослідження структури даних бібліографічних описів згідно з ГОСТ 7.1-84 та ДСТУ ГОСТ 7.1:2006 дозволило розробити принципи та алгоритм перетворення бібліографічних описів до діючого стандарту. Запропоновано застосування мови XML для розмітки бібліографічного опису.

Література

1. Матрікс об'єднання. Автоматизована бібліотечно-інформаційна система ІРБІС [Електронний ресурс] / Режим доступу : <http://matriks-pres.com.ua/index.php/sistemi-avtomatizatsiji/sistemi-avtomatizatsiji>.
2. Асоціація ЕБНИТ. Система ІРБІС. Основні характеристики [Електронний ресурс] / Режим доступу: <http://www.elnit.org/sistema-irbis/osnovnie-characteristiki.html>.
3. ДСТУ ГОСТ 7.1:2006. Бібліографічний запис, бібліографічний опис. Загальні вимоги та правила складання : метод. рекомендації з впровадження [Текст] / уклали: Галевич О. К., Штогрин І. М. – Львів, 2008. – 20 с.
4. Веб-сайт Асоціації LRC [Електронний ресурс] / [Автори асоціації]. Программы для работы с библиографической информацией. - Электрон. дан. - Режим доступу : <http://lrcnetwork.org/tt8r-> Загл. з титул. екрану (прогл. 24 квітня 2013 року).
5. Електронний портал Kazus.ru [Електронний ресурс] / [Автори сайту]. БиблиоМастер v2.0. - Электрон. дан. - Режим доступу : http://kazus.ru/programs/viewdownloaddetails/kz_0/lid_12663.html. - Загл. з титул. екрану (прогл. 24 квітня 2013 року).

6. Biblioscape 9 [Електронний ресурс] / [Автори сайту]. Research Information Manager. - Електрон. дан. - Режим доступу : <http://www.biblioscape.com/features.html#section=page-references>. - Загл. з титул. екрану (прогл. 24 квітня 2013 року).
7. Thomson Reuters. EndNote [Електронний ресурс] / [Автори сайту]. EndNote.Collect.Collaborate.Create.From Anywhere. X6. - Електрон. дан. - Режим доступу : <http://endnote.com>. - Загл. з титул. екрану (прогл. 24 квітня 2013 року).
8. ГОСТ 7.60-2003. СИБИД. Издания. Основные виды, термины и определения [Текст] [Електронний ресурс] / Межгосударственный Совет по стандартизации, метрологии и сертификации. – Мінськ : ИПК Издательство стандартов, 2004. – 41 с. – Режим доступу : <http://standartgost.ru/%D0%93%D0%9E%D0%A1%D0%A2%207.60-2003>.
9. ГОСТ 7.1-84. СИБИД. Издания. Библиографическое описание документа. Общие требования и правила составления [Текст] [Електронний ресурс]. – Москва : ИПК Издательство стандартов, 1984. – 50 с. – Режим доступу : <http://gost.libt.ru/gost-7-1-84.html>.
10. Касілов, О. В. Розробка засобів автоматизації створення бібліографічних описів [Текст] / О. В. Касілов // Східно-Європейський журнал передових технологій. Прикладні інформаційні технології. – Харків.:– 2012. – Т. 7. – № 57. – С. 18–20.