



Валентина Білоус,

директор бібліотеки Вінницького державного педагогічного університету імені Михайла Коцюбинського

Лінгвістичне забезпечення — важливий аспект діяльності бібліотеки Вінницького державного педагогічного університету імені Михайла Коцюбинського

У статті висвітлено значення лінгвістичного забезпечення для формування уніфікованого пошукового образу документа та створення електронного каталогу. На досвіді бібліотеки Вінницького державного педагогічного університету імені Михайла Коцюбинського показано методику формування авторитетного файлу "Предметні рубрики", вивчення та впровадження методики координатного індексування у роботу бібліотеки вищого навчального закладу.

Ключові слова: лінгвістичне забезпечення, координатне індексування, інформаційно-пошукові мови, пошуковий образ документа, предметні рубрики, електронний каталог, Вінницький державний педагогічний університет імені Михайла Коцюбинського.

Ефективність роботи бібліотеки вищого навчального закладу визначається повнотою та якістю інформаційного забезпечення навчально-виховного та наукового процесів. У зв'язку з цим головним в інформаційно-бібліографічній діяльності є вивчення методики та винайдення форм роботи щодо опрацювання і пошуку інформації, які б забезпечували інформаційні та пошукові запити користувачів [4, с. 9].

Створення власних бібліографічно-інформаційних ресурсів та електронного каталогу вимагає грамотного лінгвістичного забезпечення (ЛЗ) автоматизованої бібліотечно-інформаційної системи (АБІС), формування пошукового образу документа (ПОД). Зрозуміло, що від грамотності створення ПОД залежить якість електронного каталогу (ЕК) та результативність інформаційного пошуку. Вивчення та впровадження методики ЛЗ передбачають стандартний підхід до формування бібліографічних записів (БЗ), авторитетних файлів, словників, тезауруса та інших допоміжних файлів і баз даних (БД).

Лінгвістичне забезпечення АБІС розуміється як вживана в її технологіях сукупність інформаційних мов, словникових БД, лінгвістичних процесів і засобів ведення та використання словників. Лінгвістичне забезпечення — це комплекс мовних засобів, необхідних для опрацювання документів і запитів, проведення ефективного пошуку в ЕК згідно з тематичними запитами користувачів, здійснення мовної сумісності ЕК бібліотек корпорацій для обміну даними [7, с. 4].

Ефективність використання інформаційних ресурсів бібліотеки багато в чому залежить від інформаційно-пошукових мов (ПМ), які застосовуються для опису і пошуку документів. ПМ — формалізована штучна мова, призначена для індексування документів, інформаційних запитів і опису для наступного зберігання і пошуку (ДСТУ 6096:2009 "Система стандартів з інформації бібліотечної та видавничої справи. Універсальна десятикова класифікація. Структура, ведення та індексування"). Використання уніфікованої ПМ у бібліотеках забезпечує, при якісному індексуванні документів, високу точність пошуку в АБІС.

Лінгвістичне забезпечення можна схематично виразити арифметичною дією: ПМ + Нормативні документи + Методика індексування + Авторитетні дані = Лінгвістичне

забезпечення. Основою ЛЗ для ЕК є: мова бібліографічного опису; мова предметних рубрик; ієрархічні класифікації.

Питання ЛЗ та використання ПМ неодноразово розглядалися бібліотечною спільнотою і лишаються актуальними не лише в Україні. Теоретичні та практичні аспекти ЛЗ для створення уніфікованого пошукового образу та пошукового запиту документа висвітлювали такі науковці, як Т. Вілгжаніна, К. Зайцева, О. Збанацька, О. Лавренцова, Е. Сукіасян, Я. Шрайберг та інші. Існує міжнародний досвід з лінгвістичного забезпечення АБІС, бібліотеки різних країн обирають для створення уніфікованого пошукового образу різноманітні стандартизовані документи, що використовують при індексуванні: словники предметних рубрик (ПР), тезауруси, застосовують різні ПМ.

Основна мета цієї статті — показати значення та важливість ЛЗ у створенні грамотного ПОД, формуванні єдиного інформаційного простору. На досвіді бібліотеки Вінницького державного педагогічного університету (ВДПУ) імені Михайла Коцюбинського показано методику формування авторитетного файлу "Предметні рубрики", вивчення та впровадження методики координатного індексування в роботу бібліотеки вищого навчального закладу.

Одне із пріоритетних завдань бібліотеки ВНЗ — створення електронного каталогу, власних електронних ресурсів. ЛЗ відіграє значну роль у функціонуванні АБІС, здійснюючи індексування документів і запитів, ефективний пошук у ресурсних БД за тематичними запитами, визначаючи її пошукові можливості, лінгвістичну сумісність з іншими системами. Зрозуміло, що створення правильно уніфікованого ПОД зумовлює чітке дотримання правил методики у частині уніфікації окремих елементів ПР, бо саме від цього залежить успішність пошуку інформації, що є важливим для університетської бібліотеки.

Для грамотного створення ПОД та релевантного пошуку документа важливе значення у бібліотеці педуніверситету надається формуванню бібліографічних записів, дотриманню єдиної методики створення БЗ, формалізації та нормалізації подання інформації. Структура бібліографічних записів базується на міжнародному стандарті UNIMARC. Бібліотека здійснює імпорт/експорт даних у комунікативних форматах (UNIMARC/USMARC/RUSMARC/ISO). Широкий вибір екранних форм орієнтований на різні типи бібліогра-

фічного опису та подання різноманітних видів документів. Автоматизована технологія лінгвістичного опрацювання (систематизація, предметизація, координатне індексування) забезпечує можливість навігації за рубрикаторм, алфавітно-предметним показником до класифікаційних систем (АПП), тезаурусом та авторитетним файлом "Предметні заголовки".

Для формування якісних БД ЕК велике значення має вивчення та впровадження лінгвістичного забезпечення існуючої у бібліотеці АБІС ІРБІС-32. Наразі бази даних ЕК книгозбірні становлять понад 200 тис. записів, у тому числі доступні користувачам більше 180 тис. записів.

Іntenсивне впровадження у практику роботи бібліотеки нових технічних і програмних засобів висуває до ЛЗ все більші вимоги. Уніфікація ЛЗ ЕК при формуванні пошукового образу та пошукового запиту документа, що особливо важливо в корпоративній каталогізації, є дуже складним та трудомістким процесом, оптимізувати який можна шляхом співпраці з науковими бібліотеками [5, с. 22].

На корпоративних засадах співробітництва з освітянськими бібліотеками ВДПУ імені Михайла Коцюбинського бере участь у формуванні інтегрованого галузевого інформаційного ресурсу психолого-педагогічної тематики, є партнером у створенні корпоративного каталогу освітянських бібліотек, здійснює обмін між галузевими книгозбірнями БЗ статей з періодичних видань і наукових збірників університету.

Серед проблем, які виникли перед бібліотекою педуніверситету насамперед у процесі корпоративної каталогізації, — ЛЗ для створення уніфікованого пошукового образу та пошукового запиту документа, зокрема: розподіл аналітичних описів за ПР (спрощені розподіли для окремих тем, важливих для ВДПУ, серед яких тема "Педагогічні науки"). Саме ці проблеми призвели до необхідності вивчення питання ЛЗ ЕК та прийняття рішення щодо індексування документів в університетській книгозбірні. Перед її співробітниками гостро постали питання ЛЗ АБІС: в якому професійному напрямі рухатися бібліотеці: у вербальному індексуванні документів обирати координатне індексування чи предметизацію?

Клавіатурний набір ключових слів (КС) працівниками бібліотеки, відсутність методичного керівництва призвели до великої питомої ваги ненормалізованих термінів, дублювання понять, орфографічних та стилістичних помилок. Вихідний масив КС, необхідний для побудови словника, отримували шляхом накопичення ключових слів і словосполучень при вільному індексуванні документів.

Аналізуючи ефективність діючого у бібліотеці з 2005 р. електронного каталогу нових надходжень, одним із пошукових елементів якого є предметні рубрики, її фахівці дійшли висновку, що предметні рубрики, сформульовані без використання нормалізованої лексики, не можуть бути ефективним засобом пошуку.

Співробітники ВДПУ імені Михайла Коцюбинського відмовились від хаотичного накопичення ключових слів та прийняли рішення про перехід до методики предметизації. Провідні спеціалісти книгозбірні дослідили стан та розробили концепцію ЛЗ, визначили основи, що становлять інтегроване ядро інформаційно-пошукової системи. Рішення про застосування таблиць ББК при індексуванні документів було продиктовано традиціями ведення паперових каталогів.

Для лінгвістичного забезпечення АБІС, покращення доступу до інформаційних ресурсів, багатоаспектного розкриття їхнього змісту, підвищення якості бібліографічного та аналітичного опису документів в ЕК було створено робочу групу з формування, редагування та доповнення авторитетного файла "Предметні рубрики", розроблено "Методику

створення предметних рубрик", "Методичні рекомендації з використання, редагування та поповнення авторитетного файла "Предметні рубрики" засобами АРМ "Каталогізатор" та "Методику індексування документів ключовими словами у бібліотеці ВДПУ", які затвердила методична рада бібліотеки. Створенню та впровадженню цих методик передували аналіз вітчизняного та зарубіжного досвіду лінгвістичного забезпечення АБІС.

Члени методичної ради зауважили, що індексування документів лише ключовими словами небажане та недостатнє для створення пошукового запиту та образу документа, тому паралельно потрібно створювати ПР. Використання вербальних інформаційно-пошукових мов є найактуальнішим для створення ПОД. До цього виду мов належить предметизація та мова координатного індексування (КІ).

Бібліотека розпочала роботу над створенням контрольованого списку предметних рубрик, виділенням дескрипторів і створенням асоціативних зв'язків, що є основою для формування власного нормативного лінгвістичного забезпечення авторитетного файла "Предметні рубрики" з метою усунення недоліків: багатозначності, синонімії, невизначеності використовуваних термінів, застосування понять різного ступеня узагальнення, родових та видових, загальних та конкретних ПМ при найменуванні класів, підкласів, уживання неуніфікованих, багатозначних термінів — для дотримання однозначності, точності, чіткості при інтерпретації результатів. Цей тип індексування і є координатним індексуванням.

Головний бібліотекар відділу комплектування та наукової обробки літератури І. Зарванська приділяє значну увагу питанню створення пошукового образу документа та розглядає координатне індексування як базовий засіб індексування та пошуку документів в електронному середовищі. Мета індексування ключовими словами — забезпечити найповніше та найточніше розкриття усього змісту документа для цілкового задоволення інформаційних потреб користувачів бібліотеки. Саме завдяки координатному індексуванню, тобто використанню ключових слів та дескрипторів або ключових слів чи дескрипторів, створюється грамотний пошуковий образ документа. Спеціаліст бібліотеки вважає, що ключові слова "відчиняють двері" електронного каталогу та надають доступ користувачеві до інформації [3, с. 27]. Безпосередньо І. Зарванською розроблено методику індексування документів ключовими словами у ВДПУ імені Михайла Коцюбинського.

Фахівці деяких бібліотек координатне індексування замінюють використанням ключових слів, але КІ передбачає насамперед наявність тезаурусів. Спеціалісти бібліотеки педінституту зіткнулися з проблемою відсутності необхідних предметних рубрик, потребою їхньої уніфікації для аналітичного опрацювання журнальних статей. Виникла необхідність самостійно формувати ПР для створення бібліографічних записів в електронному каталозі.

Роботу з уніфікації лінгвістичних засобів фахівці бібліотеки вишу розпочали з вивчення методик створення ПР, існуючого світового та вітчизняного досвіду складання різних за тематикою тезаурусів. Робоча група з розроблення методик формування, редагування, поповнення та використання авторитетного файла "Предметні рубрики" прийняла за основу існуючий в АБІС "ІРБІС" авторитетний файл "Предметні заголовки" (словникову БД), який постійно оновлюється, підтримується в робочому стані та використовується під час пошуку.

У книгозбірні ВДПУ імені Михайла Коцюбинського при індексуванні документів та вихідним продуктом при поповненні списку ПР виступають такі пошукові засоби:

— ПМ класифікаційного типу (Універсальна десятикова класифікація (УДК); Бібліотечно-бібліографічна класифікація (ББК); Бібліотечна класифікація для обласних бібліотек; електронна версія Тезауруса Національної парламентської бібліотеки України тощо);

— вербальні мови — дескрипторна мова з контрольованою лексикою та мова ПР (енциклопедії та енциклопедичні словники; тлумачні та термінологічні словники; довідники та словники-довідники тощо).

Вибір нової вербальної мови зумовлений перспективами використання авторитетного файлу "Предметні рубрики" в обмінно-інформаційних корпоративних процесах та створення грамотного ПОД. У якості мов індексування вербального типу виступають мови предметних рубрик і ключових термінів. Бібліотека формує власні авторитетні файли ПР та визначає ключові слова. При прийнятті рішення щодо методики використання мови предметних рубрик в автоматизованому режимі було враховано, що при формуванні БД ЕК стали очевидними обмеження традиційної мови предметних рубрик (МПР).

Аналіз документів та визначення їхнього предмета здійснюється індексаторами бібліотеки за допомогою уніфікованої пошукової мови. Мова ПР — один із різновидів вербальних ПМ з нормалізованою лексикою, що застосовується у бібліотеці для організації якісного пошуку інформації в БД ЕК. Повноцінне та якісне управління ЕК неможливе без авторитетних файлів, які використовуються у бібліографічних записках. Вони об'єднують під одним терміном усі можливі варіанти певного поняття. Це допомагає уніфікувати формулювання авторських та тематичних заголовків та уникнути дублювання.

Серед вимог до створення списку предметних рубрик — повно і точно передавати зміст, а при необхідності форму і призначення документа, що відображається у певній інформаційно-пошуковій системі (ІПС); забезпечувати однозначність тлумачення термінів індексування; допускати багатаспектне індексування та внесення змін (доповнень і виправлень); забезпечувати простоту та зручність індексування, інформаційного пошуку та ведення певної ІПС; відображати сучасний стан терміносистеми у цій галузі знань [2].

При виборі формулювання ПР допускається інверсія для введення у початкову позицію найважливішого терміна, що несе основне смислове навантаження. Мова ПР має задовольняти потреби користувачів, відображати сучасну термінологію та сучасний стан лексики природної мови. Лексичними одиницями (ЛО) мови списку ПР виступають заголовки і підзаголовки. Заголовки є основними ЛО, підзаголовки — допоміжними ЛО, тобто використовуються тільки разом із заголовком. При індексуванні та формуванні ПР особливо важливо дотримуватися основної вимоги методики предметизації, правила адекватного відображення змісту документа та привласнювати йому адекватні рубрики, виходячи з конкретного змісту.

При формуванні авторитетних файлів "Предметні рубрики" використовується ієрархія предметних зв'язків: від загального терміна — до конкретного предмета, при яких одна ПР представляє клас (категорію) або ціле, а інша — підклас цього класу (відносини "рід—вид") або частину цього цілого (відносини "частина—ціле").

При створенні складної ПР застосовується категоріальний метод. Завдання індексатора — вибрати потрібний для певної теми ряд категорій і розробити правила поєднання підзаголовків з відповідними цим категоріям заголовка-

ми. Необхідно відзначити, що грамотно складена ПР — це рубрика, побудована згідно з правилами категоріального аналізу, який передбачає поєднання кожного підзаголовка з певними категоріями предметів.

Перший елемент предметної рубрики — заголовок (визначає розміщення предметної рубрики у списку ПР, дає змогу зібрати в ньому подібні чи близькі за тематикою ПР). Рештою елементів виступають підзаголовки. Кожен конкретизує поняття, виражене попередніми елементами.

Приклади предметних рубрик:

- математика — урок;
- фізика — міжпредметні зв'язки;
- українська мова — методи навчання;
- педагогіка — вчитель;
- політика — внутрішня — Україна;
- історія України — козаччина.

Працівники відділу комп'ютеризації та фахівці бібліотеки, які безпосередньо здійснюють індексування документів, створили у зарезервованих полях: поле 965 Тезаурус — 3629 записів, поле 606 авторитетний файл "Предметний заголовок" — 2661 предметна рубрика, поле 600 авторитетний файл "Індивідуальні автори" — 5791 запис, поле 601 авторитетний файл "Колективні автори" — 1937 записів. Систематично доповнюються меню-довідники матеріалів, відомостей до назви, розроблено "розділ знань" тощо. Сформовано словники до предметного заголовка та підзаголовка: географічний — 545 назв, хронологічний — 123, формальний — 128.

Авторитетний файл "Предметні рубрики" використовується при формуванні мови пошукових предметних рубрик для переважної більшості БД ЕК бібліотеки педуніверситету. Географічні та хронологічні підзаголовки можуть приєднуватися до будь-якої рубрики для відображення географічних або хронологічних аспектів. При наявності географічних даних у джерелі каталогізатор використовує меню-довідник географічних назв, сформований бібліотекою. Формальні та хронологічні підзаголовки розташовуються в кінці предметної рубрики.

Застосування методик лінгвістичного забезпечення — складний логічний процес, який потребує не лише відповідної фахової підготовки, а й "лексичної інтуїції", особливо "відчуття слова" [5, с. 22]. Метою створення списку ПР є алгоритмізація дій індексатора при відборі, нормалізації та систематизації лексики для формування ПР. Індексатори бібліотеки на практиці самі вирішують проблему, що виражається, зокрема, у розробленні списків підзаголовків, необхідних для відображення певних галузей знань.

Призначення впровадженій у діяльність університетської бібліотеки методики полягає у зниженні суб'єктивності дій каталогізатора при відборі та нормалізації лексики й оформленні її у вигляді предметних рубрик. Ця методика дає можливість уніфікувати мову ПР та уникнути перевантаження ПОД зайвими термінами.

Незважаючи на те, що бібліотека активно використовує предметні рубрики у практиці індексування, самі індексатори висловлюються на користь мови ключових термінів. Також важливим, на їхню думку, є створення та надання користувачам своєї бібліотеки необхідної та релевантної інформації, що можливо лише за умови наближення інформаційно-пошукових запитів до природної мови, взаємодії систематичного і предметного принципів розкриття інформації. Від якості індексації документів залежить ефективність тематичного пошуку в ЕК. Користувачі також зацікавлені оперувати ключовими термінами, які є для них найзрозумілішим пошуковим засобом. Бібліотечна практика показує, що під час пошуку за допомогою ключових слів

можна знайти більшу кількість релевантних документів, ніж при зверненні до предметних рубрик.

Незалежно від того, чи йде пошук за ключовими словами, чи за ПР, найкращі результати забезпечує набір лексичних одиниць, що становлять тему запиту. У разі пошуку за ключовими словами грамотно складена ПР дає найменшу кількість шуму. При цьому через відсутність єдиної системи предметизації та універсального словника ПР мова предметних рубрик є менш ефективним засобом тематичного пошуку. Користувач більше зацікавлений у застосуванні інформаційних мов вербального типу, ніж інформаційних мов класифікаційного типу. Релевантний пошук інформації можливий при активному використанні різних засобів ЛЗ ЕК [6, с. 50], класифікаційні мови забезпечують пошук за запитами галузевого характеру, універсальний і формалізований, а вербальні мови — за запитами предметного характеру, що частіше застосовується користувачами.

Виходячи з викладеного, закономірно, що правильний вибір і якість застосування у бібліотеці лінгвістичних засобів сприяють високому рівню обслуговування. Тому доцільніше використовувати в АБІС у якості ППМ вербального типу мову ключових термінів, а ПОД вести на мові ключових термінів або транслювати з мови ПР на мову ключових термінів.

Зазначена система орієнтована на адміністраторів словникових баз даних та індексаторів, які підтримують її в актуальному стані та здійснюють творчу роботу з удосконалення лінгвістичних процесорів і засобів автоматизованого ведення та використання словникових БД. Адміністрування БД авторитетного файлу "Предметні рубрики" здійснює завідувач відділу комп'ютеризації та експлуатаційно-технічного обслуговування бібліотечних процесів С. Наугольних, яка брала безпосередню участь у розробленні цих методик.

Аналізуючи стан ЛЗ, переконуємось, що ця важлива проблема може бути вирішена лише спільними зусиллями. Нині лінгвістичне забезпечення електронного каталогу ускладнюють невідповідність програмних бібліотечних платформ форматам наведення даних, застосування різних класифікаційних таблиць, заповнення різної кількості полів бібліографічного запису, відсутність єдиного тезауруса, слабе матеріально-технічне забезпечення бібліотеки.

Потрібно і надалі вивчати та узагальнювати досвід різних бібліотек з питань індексування і предметизації. Формуванню єдиного освітнього інформаційного простору, Зведеного українського каталогу та уніфікованого лінгвістичного забезпечення, підвищенню рівня індексування документів суттєво сприятиме кооперативна діяльність бібліотек, у рамках чого планується розпочати роботу над створенням унормованого предметного списку для формування національного авторитетного файлу "Предметні рубрики" з педагогіки та психології. Для якісної корпоративної каталогізації необхідно розробити спільну методику із систематизації, предметизації, координатного індексування та створення уніфікованого лінгвістичного продукту "Авторитетні файли" — "Коллективні автори", "Індивідуальні автори", "Предметні рубрики".

зації, координатного індексування та створення уніфікованого лінгвістичного продукту "Авторитетні файли" — "Коллективні автори", "Індивідуальні автори", "Предметні рубрики".

Список використаної літератури

1. Дрозд О. Корпоративная каталогизация и формирование общего электронного ресурса / О. Дрозд, В. Мыхалик // Библиотечный вестник. — 2011. — № 5. — С. 26—29.
2. Зайцева Е. М. Лингвистическое обеспечение автоматизированных библиотечных систем: современные требования и направления развития / Е. М. Зайцева // Научные и технические библиотеки. — 2004. — № 7. — С. 15—27.
3. Зарванська І. В. Координатне (посткоординатне) індексування в електронному просторі як ефективний засіб створення пошукового образу документа / Ірина Зарванська // Вісник Книжкової палати — 2012. — № 1. — С. 27—31.
4. Збанацька О. Трансформація вербальних інформаційно-пошукових мов в інформаційно-бібліографічній діяльності / О. Збанацька // Бібліотечна планета. — 2008. — № 4. — С. 9—13.
5. Кузьмінч Т. Змістовне опрацювання документного потоку в умовах корпоративної взаємодії бібліотек Білорусі / Т. Кузьмінч, С. Пугачова // Бібліотечний вестник. — 2008. — № 2. — С. 21—23.
6. Рыбальченко Е. Н. Использование средств лингвистического обеспечения электронного каталога для релевантного поиска информации и выполнения различных запросов / Е. Н. Рыбальченко // Научные и технические библиотеки. — 2011. — № 7. — С. 46—51.
7. Татарчук Л. Лінгвістичне забезпечення формування єдиного інформаційного простору в мережі сільськогосподарських бібліотек / Л. Татарчук // Бібліотечний вестник — 2011. — № 3. — С. 3—8.
8. Шемакин Ю. И. Начала компьютерной лингвистики : учеб. пособие / Ю. И. Шемакин. — М. : Изд-во МГОУ, 1992. — 81 с.
9. AGROVOC. — Mode of access: <http://aims.fao.org/website/AGROVOC-Thesaurus/sub>. — Title from screen.

В статті освітлено значення лінгвістичного забезпечення для формування уніфікованого образу документа та створення електронного каталогу. На основі бібліотеки Вінницького державного педагогічного університету імені Михайла Коцюбинського розкрито методику формування авторитетного файлу "Предметні рубрики", вивчення і впровадження методики координатного індексування в роботу бібліотеки вищого навчального закладу.

The article highlights the meaning of linguistic software for formation of a unified search image of a document and for creation of an electronic catalog.

The formation methods of an authority file "Subject headings" and the implementation of coordinate indexing methods into the work of an institution of higher education library are displayed on the experience of the Vinnitsia State Pedagogical University named after Mikhailo Kotsyubynskiy library.

Надійшла в редакцію 29 серпня 2012 року