

На наш погляд, найповніше відповідає цим вимогам міжнародно визнана система CDS/ISIS/M, розроблена в ЮНЕСКО і призначена для структурованих нечислових баз даних [3].

Ця система підтримує сітьовий режим, що дуже важливо при організації роботи бібліотек. Система висуває мінімальні вимоги до апаратних засобів - функціонує на комп'ютерах IBM-PC/XT/AT та сумісних з ними і вимагає об'єм пам'яті 512 Кб. Існує версія системи для комп'ютерів серії VAX, що працюють під керуванням операційної системи VAX/VMS. Важливим є й те, що ЮНЕСКО підтримує систему CDS/ISIS, супроводжує та вдосконалює її, готуючи нові версії та додаткові засоби, що розширюють її можливості. У Києві існує регіональний центр ЮНЕСКО з розповсюдження системи CDS/ISIS.

Важливою особливістю системи є її призначення для управління полями змінної довжини, що дає змогу раціонально використовувати дисковий простір комп'ютера і вільно визначати максимальну довжину кожного поля.

Для полегшення і спрощення процесу взаємодії користувача з системою CDS/ISIS/M, італійською асоціацією документальних служб, бібліотек та архівів розроблено систему HEURISKO [4], що розповсюджується ЮНЕСКО і має дружній інтерфейс кінцевого користувача для організації пошуку інформації, перегляду та експорту баз даних, підготовлених у системі CDS/ISIS/M. Система дуже зручна для використання читачами як електронний каталог у бібліотеках. Крім зручного меню пошуку, система HEURISKO дає змогу експортувати зроблений запит

на магнітний носій (дискету) у файл формату ISO 2709, що дає можливість користувачеві поповнити свою власну інформаційно-пошукову систему результатами, одержаними в режимі ретроспективного пошуку (пошук умасиви за певною темою) або в режимі вибіркового розповсюдження інформації (регулярний пошук за заданою темою тільки в масиві поточних надходжень).

Передача по модемному зв'язку файлів імпорту-експорту у форматі стандарту ISO 2709 дасть змогу бібліотекам обмінюватися інформацією, поповнюючи свої бази даних за умов, що ці бази спроектовані в єдиному форматі.

Оскільки кожна бібліотека самостійно вирішує за допомогою яких програмних засобів здійснюватиметься ведення бібліографічних баз даних, відділи автоматизації повинні передусім керуватися вимогами, які відповідають вимогам формату UNIMARC, а не зручністю інтерфейсу або розповсюдженістю тієї чи іншої інформаційно-пошукової системи. При визначенні структури бібліографічної бази даних необхідно забезпечити відповідність позначок полів та кодів підполів вимогам формату UNIMARC. Наприклад, у Державній бібліотеці України для юнацтва експлуатується база даних з аудіовізуальних матеріалів; розроблена в системі CDS/ISIS/M.

Використання в бібліотеках України програмних засобів, що не відповідають сформульованим вимогам, ускладнить інтеграцію електронних каталогів різних бібліотек у єдиний електронний каталог, вимагатиме значних коштів і часу на конвертацію даних, підготовлених для обміну.

1. Сенченко Н.И. Библиотeki и компьютеры / Под ред. В.В.Васильева. - Киев: Наук. думка, 1990. - 216 с., ил.
2. Руководство по UNIMARC: Руководство по применению международного коммуникативного формата UNIMARC / Пер. с англ. авт. коллектива под руководством А.И.Земскова, Я.Л.Шрайберга - М., 1992. - 320 с.: табл.
3. Пакет прикладных программ CDS/ISIS/M: версия 2.3/ Меж-

- дународный центр научно-технической информации. - М., 1991. - Вып. 72. - 256 с.: табл. (Метод. материалы и документация по пакетам прикладных программ).
4. HEURISKO: Дружественный интерфейс конечного пользователя CDS/ISIS / Пер. на рус. язык; Ип-т программных систем АН Украины - Киев, 1993. - 51 с., ил.

Олексій
Бодня

Обробка документальних потоків (на основі розпізнавання бібліографічної інформації)

Втілення в життя Національної програми інформатизації неминуче приведе до суттєвого зростання обсягів електронної інформації. Від джерел інформації до споживача йдуть інформаційні або документальні потоки [1]. Зрозуміло, що інформаційні потоки мають дійти до кожного конкретного споживача, незважаючи на колосальні обсяги загального документообігу.

У бібліотеках серед загальних документальних потоків чільне місце займають документи з бібліографічною інформацією [2]. Вона може розповсюджуватися серед користувачів у форматі

найпоширеніших баз бібліографічних даних, у форматі ISO [3] чи інших міжнародних комунікативних форматах [4,5], а також текстовими файлами. Одним з варіантів текстових файлів буде файл бібліографічних записів з сукупністю описів різних видань (статей) за загальноприйнятими правилами таких описів. Функціонально він може вміщувати в собі якийсь бібліографічний покажчик, каталог бібліотеки чи певну його частину, реферативний журнал, індекс цитування та ін.

Чимало таких документів уже безпосередньо готуються на машинних носіях. За наявності відповідного програмного забезпечення у користувача не викликає жодних проблем і перенос інформації з паперових носіїв у файли EOM з допомогою сканерів.

Та цінність такої інформації істотно зростає, коли вона буде структурована у складі якоїсь бази даних (БД) чи в інший спосіб.

Отже, при обміні інформацією між бібліографічними базами даних у вигляді файлів бібліографічних записів з'являється проблема перетворення інформації з форматів використаних БД у бібліографічні записи й навпаки.

Питання перетворення інформації з баз даних простіше і практично реалізується звичайними засобами будь-якої СУБД (системи управління базами даних) широкого користування.

Що ж до приєднання до БД інформації, яка існує у вигляді бібліографічних записів, то тут питання складніше. Вже є засоби в бібліотечних автоматизованих системах, які дозволяють безпосередньо користувачу, сидючи за ПЕОМ, розбивати бібліографічні записи відповідно до структури бібліотечної БД. Однак при обробці великої кількості інформації така ручна технологія буде неефективною.

Звідси випливає потреба автоматичного розпізнавання інформації у бібліографічних записах з метою її перетворення під формат існуючих БД.

При вирішенні цього завдання виникає потреба у створенні інтелектуальних систем аналізу бібліографічної інформації, які повинні аналізувати текстову інформацію щодо наявності в ній бібліографічної інформації, а також розбивати виділену бібліографічну інформацію по інформаційних полях.

Під інформаційними полями ми розуміємо відокремлені частки інформації, які містять у собі елементи структурованої інформації і відповідають полям певних БД.

Якщо поглянути на проблему ширше, то такі системи можуть застосовуватися для розпізнавання будь-якої слабкоструктурованої інформації.

Демо визначення різних ступенів структурованості інформації.

Під повністю структурованою інформацією нами розуміється механічне об'єднання інформаційних полів з точно відомою структурою об'єднання і визначеними межами між інформаційними полями. А під *неструктурованою інформацією* - звичайний текст на природній мові. Отже, слабкоструктурованою називатимемо інформацію, що має деякі елементи структуризації, та межі між інформаційними полями не є в ній точно визначеними, а загальна схема об'єднання інформаційних полів є багатоваріантною.

У бібліографічному описі маємо типовий приклад слабкості структурованої інформації. Завдання її обробки полягає в тому, що в бібліографічному записі за місцем, розділювачами і формальними ознаками інтелектуальній системі обробки бібліографічної інформації треба відокремити значення полів існуючої БД (інформаційних полів) з щонайменшою кількістю помилок. Для цього необхідно створити спеціалізовані алгоритми і відповідне програмне забезпечення.

Навіть найпростіший інтелектуальний аналіз бібліографічного запису вимагає відокремити авторів книги (можливо кількох), назву, інформацію про видавництво і рік виходу книги з друку в умовах можливої відсутності відповідного значення і практично багатоваріантного використання розділювачів.

Робота зі створення інтелектуальних систем аналізу бібліографічної інформації потребує аналізу всіх можливих варіантів розташування інформаційних полів у бібліографічному записі, варіантів роздільників між

ними і формальних семантичних ознак належності інформації до відповідного інформаційного поля. Відокремивши інформаційне поле, можна перенести його інформацію до поля БД.

Перенесення однозначно структурованої текстової інформації до будь-якої БД звичайно можна здійснювати засобами самої БД.

Приміром, для перетворення повністю структурованої текстової інформації у формат ISO 2709 використовується конвертор FANGORN. Він після конвертування видає файл у стандарті ISO 2709. Надалі такий файл може бути завантажений до будь-якої БД, засоби експорту/імпорту котрої відповідають ISO-стандартам.

Прикладом обміну бібліографічною інформацією, якісний рівень якого вдалося покращити за рахунок впровадження інтелектуальних систем, може бути здійснюваний ЦНБ збір від обласних бібліотек України підготовлених ними рефератів газетних статей з місцевих видань.

У ЦНБ створено БД «Реферати газетних статей», яка підтримується засобами СУБД CDS/ISIS, містить у собі інформацію з понад 200 газет України та СНД і призначена для задоволення потреб читачів і абонентів бібліотеки в кумульованій інформації з питань внутрішньої, зовнішньої, економічної та соціальної політики.

Інформація для БД по центральних і київських газетних виданнях готується у самій ЦНБ. Однак в Україні існує багато місцевих видань, які не завжди своєчасно надходять до Бібліотеки. І тому реферати їх статей зручніше готувати в обласних бібліотеках.

В Україні не створено типову автоматизовану систему для обласної бібліотеки і можна припустити, що в різних областях будуть різні автоматизовані бібліотечні системи. Слід також врахувати, що більшість бібліотек роблять на шляху автоматизації тільки перші кроки.

Отже, найдієвішою є методика, за якою інформація з регіональних бібліотек передається у вигляді текстових файлів, що містять у собі бібліографічні записи, кожен з яких складається з аналітичного опису статті та її реферату.

Файли, що містять бібліографічні записи з рефератами газетних статей, електронною поштою надходять до ЦНБ. Тут для їх обробки застосовується створений у ній програмно-технологічний комплекс розпізнавання бібліографічної інформації.

У ньому використовуються методи синтаксичного аналізу для розпізнавання типу бібліографічного опису і з'ясування його структури, семантичного аналізу для відокремлення інформаційних полів. З цією ж метою застосовуються і елементи морфологічного аналізу. Після розпізнавання інформація конвертується у міжнародний формат ISO 2709, що відповідає внутрішній структурі БД «Реферати газетних статей».

Надалі інформація з цього файлу приєднується до БД засобами імпорту даних системи CDS/ISIS.

Інтелектуальні системи аналізу бібліографічної інформації знайдуть широке застосування і в інших бібліографічних та інформаційних процесах. Наявно можна побачити переваги їх застосування у процесах комплектування, каталогізації та інформаційного аналізу. Найбільшого ефекту, на нашу думку, можна досягти при використанні таких систем для переносу інформації з традиційних карткових каталогів до електронних.

Отже, існування великих обсягів слабкоструктура-

ної бібліографічної інформації на машинних носіях потребує проведення робіт з їх структуризації. Для здійснення таких робіт з метою кумуляції інформації у бібліографічних БД доцільно використовувати інтелектуальні системи розпізнавання бібліографічної інформації.

Дослідна експлуатація однієї з таких систем (програмно-технологічного комплексу розпізнавання бібліографічної інформації в текстових файлах з рефератами газетних статей) у ЦНБ довела її ефективність.

1. Першиков В.И., Савинков В.М. Толковый словарь по информатике. - М.: Финансы и статистика, 1991. - 543 с.
2. Информатика / Под ред К.В.Тараканова. - М.: Книга, 1986. - 304 с.
3. ISO 2709 - 1981: Документация - Формат для обмена библиографической информацией на магнитной ленте.
4. Информационная технология. Комплекс стандартов на автоматизированные системы. - М.: Изд-во стандартов, 1991. - 127 с.
5. Руководство по применению международного коммуникативного формата UNIMARC. - М.: ГПНТБ России, 1992. - 319 с.

РЕЦЕНЗІЇ

Марина Дмитрієнко

Дослідження історії інтелігенції

(Рубльов О.С., Черченко Ю.А. Сталінщина і доля західноукраїнської інтелігенції (20 - 50-ті роки ХХ ст.) - К.: Наук. думка, 1994. - 350 с.).

Серед українських етнічних земель західний регіон посідає унікальне місце в національній історії. Знайомство західноукраїнців з «владою пролетаріату» розпочалося в 20-ті роки через культурні та наукові зв'язки з УСРР, сумнозвісного 1939, після приєднання регіону до СРСР воно вже набуло майже канонічних для сталінізму «взаємовідносин». Історія західноукраїнської інтелігенції, зокрема її взаємин з радянською владою за доби сталінізму, й стала предметом наукового дослідження, результати якого репрезентовано в обговорюваній книзі.

Як справедливо зазначають автори, можливість повноцінного дослідження історії інтелігенції в тоталітарну добу з'явилася тільки в останні роки, коли було скасовано цензурні обмеження на друковану продукцію та відкрилися спецфонди архівів. Головні міркування, якими дослідники керувалися в цій праці, доволі повно викладено в наступній цитаті: «Поділяючи думку, що на сучасному етапі розвитку вітчизняної історіографії новітньої історії України першорядне значення має, насамперед, навіть не розвиток теорії (так званої історіософії), а звичайне буденне накопичення фактів, автори прагнули розглянути життєвий і творчий шлях не лише осіб, так би мовити, «першого ряду» (до чого традиційно тяжіє історіографія), а й «рядових» інтелігентів, уродженців західноукраїнських земель, без з'ясування якого неможливе об'єктивне відтворення складної і суперечливої картини взаємодії Західної і Східної України в 1920-х - на початку 1950-х років» (с.12). Монографія є колективним витвором істориків Олександра Рубльова та Юрія Черченка, відомих за численними публікаціями, що стосуються періоду сталінізму на західноукраїнських землях. Чотири розділи праці: «І. Участь західноукраїнської інтелігенції у здійсненні політики українізації в УСРР (1923 -

початок 1930-х років)», «ІІ. Наукові та культурні зв'язки інтелігенції УСРР та Західної України у 20 - 30-х роках», «ІІІ. «Похід на галичан» та масові репресії щодо української інтелігенції у 30-х роках», «ІV. Ставлення громадськості Західної України до утвердження тоталітаризму в УСРР (кінець 20-х - 30-ті роки)», - написані О.Рубльовим; два заключні: «У. Політика нової влади щодо західноукраїнської інтелігенції після воз'єднання (вересень 1939 - червень 1941 р.)» і «УІ. Сталінщина і західноукраїнська інтелігенція (40-ві - початок 50-х років)», - Ю.Черченком.

У книзі розглядається вельми широкий спектр проблем, пов'язаних з історією західноукраїнської інтелігенції; повернення в радянську Україну, співробітництво з радянською владою, участь у процесі «коренізації», наукові та культурні зв'язки, ідеологічні впливи, соціально-психологічне сприйняття сталінського режиму, особливості національної самосвідомості та вже традиційна за останні роки проблема долі інтелігенції у 30 - 50-ті роки. Монографія серед низки інших праць виділяється насиченою джерельною базою. В ній широко використано недоступні раніше для істориків архівні матеріали, зокрема архіви Служби безпеки України. Залучено також епістолярну спадщину, спогади, мемуари визначних науковців, партійних керівників, діячів культури, освіти, зрештою, звичайних сучасників тоталітарного режиму. В процесі висвітлення канви історичних подій у монографії враховано також фактичний матеріал та дослідницькі підходи, подані в студіях зарубіжних українознавців Джона Армстронга, Ярослава Білінського, Богдана Кравченка, Джеймса Мейса, Олега Геруса, Петра Мірчука та ін.

Послідовно досліджено в книзі основні етапи радянської політики щодо західноукраїнської інтелігенції, її склад та чисельність у різні періоди, роль та функції в державному будівництві й народному господарстві. Простежено життєві шляхи та

долі окремих осіб з перспективи названої проблематики. Джерельну вартість праці підвищує ряд документальних додатків, матеріали архівів СБУ, зокрема «Заява письменника Я.О.Галана прокурору УСРР з проханням переглянути справу його дружини Г.І.Галан (Геник)», «Лист дирекції Інституту суспільних наук Львівської філії АН УСРР № 424-а від 24 серпня 1956 р. до ЦК Компартії України та прокуратури УСРР з проханням з'ясувати долю та дати оцінку діяльності відомого діяча галицької соціал-демократії С.Г.Вітика, який був репресований у 1930-х роках в УСРР» тощо. Ці додатки виконують й ілюстративну функцію, подаючи під різними кутами зору тогочасні історичні реалії.

Документальний матеріал викладено логічно й послідовно, і всю книгу написано гарною літературною мовою, що робить її приступною широкому загалу читачів, які мають добру нагоду отримати відносно повне уявлення про ряд визначних особистостей, діяльність та доля яких і досі залишається маловідомою для сучасників, а також про загальні контури цієї трагічної епохи. Обґрунтовані та закономірні висновки авторів дослідження пояснюють використання механізму перманентних репресій, морального цькування як провідного засобу політики радянської щодо західноукраїнського населення для його інтегрування в тоталітарну систему.

Зважаючи на нерівномірність розробки дослідницьких проблем з історії тоталітарної доби в Україні та значну наукову цінність вміщених у монографії матеріалів, вона певною мірою компенсує існуючий брак подібної літератури і добре прислужиться численним дослідникам української культури.

