

**НАЦІОНАЛЬНА АКАДЕМІЯ ПЕДАГОГІЧНИХ НАУК УКРАЇНИ
ДЕРЖАВНА НАУКОВО-ПЕДАГОГІЧНА БІБЛІОТЕКА УКРАЇНИ
ІМЕНІ В. О. СУХОМЛИНСЬКОГО**



Серія
„На допомогу професійній самоосвіті
працівників освітянських бібліотек”
Випуск 9

ІНДЕКСУВАННЯ ДОКУМЕНТІВ КЛЮЧОВИМИ СЛОВАМИ

МЕТОДИЧНІ РЕКОМЕНДАЦІЇ

Київ 2011

НАЦІОНАЛЬНА АКАДЕМІЯ ПЕДАГОГІЧНИХ НАУК УКРАЇНИ
ДЕРЖАВНА НАУКОВО-ПЕДАГОГІЧНА БІБЛІОТЕКА УКРАЇНИ
ІМЕНІ В. О. СУХОМЛИНСЬКОГО

Серія
*„На допомогу професійній самоосвіті
працівників освітянських бібліотек”*
Випуск 9

ІНДЕКСУВАННЯ ДОКУМЕНТІВ

КЛЮЧОВИМИ СЛОВАМИ

Методичні рекомендації

Київ 2011

УДК 025.43.011(083.13)

ББК 78.372.15я82

I-60

*Затверджено до друку вченою радою
Державної науково-педагогічної бібліотеки України імені В. О. Сухомлинського
(протокол № 3 від 15 березня 2011 р.)*

Укладач *І. Г. Лобановська*, в. о. заступника директора з науково-бібліотечної роботи
ДНПБ України ім. В. О. Сухомлинського

Рецензенти: *Н. Й. Дзюба*, зав. відділу каталогізування НБУ для дітей

Т. Ф. Букишина, наук. співроб. ДНПБ України ім. В. О. Сухомлинського

Відповідальний за випуск *І. Г. Лобановська*, в. о. заступника директора з науково-
бібліотечної роботи ДНПБ України ім. В. О. Сухомлинського

Літературний редактор *І. П. Білоцерківець*, молод. наук. співроб. ДНПБ України ім.
В. О. Сухомлинського

Індексування документів ключовими словами : метод. рек. / НАПН
України, Держ. наук.-пед. б-ка України ім. В. О. Сухомлинського ; [уклад.
І. Г. Лобановська]. – К. Нілан-ЛТД, 2011. – 32 с. – (Серія «На допомогу профе-
сійній самоосвіті працівників освітянських бібліотек» ; вип. 9).

У методичних рекомендаціях докладно відображено способи описування
основного смислового змісту документів за допомогою ключових слів, наведе-
но численні приклади індексування окремих видів документів. Запропонована
методика має універсальний характер і може бути використана як для запов-
нення поля „Ключові слова” електронних каталогів бібліотек незалежно від ви-
користовуваної АБІС („ІРБІС”, „УФД-Бібліотека”, „MARC-SQL” тощо), так і
для визначення ключових слів у процесі оформлення наукових публікацій.

Видання призначено для працівників освітянських бібліотек України всіх
видів, може бути використано в роботі книгозбірень інших систем і відомств, а
також буде корисним для тих, хто займається науковою, педагогічною та нау-
ково-методичною діяльністю.

УДК 025.43.011(083.13)

ББК 78.372.15я82

ЗМІСТ

СПИСОК СКОРОЧЕНЬ	3
ПЕРЕДМОВА	4
ВСТУП	7
КООРДИНАТНЕ ІНДЕКСУВАННЯ ДОКУМЕНТІВ	10
Аналізування й визначення змісту документа як об'єкта індексування	11
Відбирання понять, що характеризують зміст документа	12
Відбирання термінів індексування для позначення понять	13
Формування пошукового образу документа з термінів індексування	14
Відмінок КС	16
Число КС	16
Інверсія в КС	18
Полісемія та омонімія КС	18
Синонімія КС	19
ІНДЕКСУВАННЯ ОКРЕМИХ ВИДІВ ДОКУМЕНТІВ	19
ПРАВИЛА ЗАПОВНЕННЯ ПОЛЯ „КЛЮЧОВІ СЛОВА”	20
Спеціальні правила наведення КС	22
МОЖЛИВІ ПОМИЛКИ ПІД ЧАС ІНДЕКСУВАННЯ ДОКУМЕНТІВ І ЗАПОВНЕННЯ ПОЛЯ „КЛЮЧОВІ СЛОВА”	24
ДОДАТКИ	25
Додаток 1	25
СХЕМА ВИЗНАЧЕННЯ КЛЮЧОВИХ СЛІВ	25
У ПРОЦЕСІ ВІЛЬНОГО КООРДИНАТНОГО ІНДЕКСУВАННЯ ДОКУМЕНТА	25
Додаток 2	25
Додаток 2	26
ПРИБЛИЗНИЙ ПЕРЕЛІК ГАЛУЗЕВИХ РЕЛЯТОРІВ	26
ТЕРМІНИ ТА ВИЗНАЧЕННЯ	27
СПИСОК ЛІТЕРАТУРИ	29

СПИСОК СКОРОЧЕНЬ

АІС	– Автоматизована інформаційна система
АБІС	– Автоматизована бібліотечно-інформаційна система
БД	– база даних
БЗ	– бібліографічний запис
БО	– бібліографічний опис
ГОСТ	– Государственный стандарт
Держспоживстандарт України	– Державний комітет України з питань технічного регулювання та споживчої політики
ДНПБ України В. О. Сухомлинського	ім. – Державна науково-педагогічна бібліотека України імені В. О. Сухомлинського
ДСТУ	– Державний стандарт України
ЕК	– електронний каталог
ІКТ	– інформаційно-комунікаційні технології
ІПМ	– інформаційно-пошукова мова
ІПС	– інформаційно-пошукові системи
КМР	– картотека методичних рішень
КС	– ключові слова
ЛО	– лексична одиниця
ПОД	– пошуковий образ документа
ПОЗ	– пошуковий образ запиту
СИБИД	– Система стандартів по информации, библиотечному и издательскому делу
ПР	– предметна рубрика
ССІБВС	– Система стандартів з інформації, бібліотечної та видавничої справи
УДК	– Універсальна десяткова класифікація
IFLA (International Federation for Library Associations)	– Міжнародна Федерація бібліотечних асоціацій
ISBD (International Standard Bibliographic Description)	– Міжнародний стандартний бібліографічний опис
ISO (International Organization for Standardization)	– Міжнародна організація зі стандартизації
MARC format (Machine Readable Cataloging format)	– MARC формат, формат машиночитної каталогізації
MARC 21 (USMARC, US Machine Readable Cataloging)	– Американський формат машиночитної каталогізації
RUSMARC	– Російський комунікативний формат представлення бібліографічних записів в машиночитній формі
UNIMARC (Universal Machine Readable Cataloging)	– Міжнародний комунікативний формат представлення бібліографічних записів в машиночитній формі

ПЕРЕДМОВА

Останніми роками відбулася і триває глобальна інформатизація таких сфер діяльності, як наука, освіта, культура. Особливо яскраво це проявилось у бібліотечно-інформаційній справі, де за досить короткий час змінилося вже кілька поколінь технічних і програмних засобів. Процес інформатизації охопив усі рівні бібліотек – від національних до шкільних. Упровадження автоматизованих бібліотечно-інформаційних систем (АБІС) стало необхідною реальністю. Книгозбірні почали організовувати електронні каталоги (ЕК) на свої фонди і зіткнулися з проблемою удосконалення наявних інформаційно-пошукових систем (ІПС). Для розв'язання цієї проблеми розробляються лінгвістичні засоби, що забезпечують індексування документів і запитів, здійснення інформаційного пошуку, формування й структурування баз даних (БД).

Серед усіх видів інформаційного пошуку найбільш популярним є тематичний, особливо здійснюваний за допомогою ключових слів (КС). Цей вид пошуку позитивно оцінюють як індексатори, так і користувачі. Нині індексування КС стало обов'язковим елементом технології бібліотечно-інформаційних систем. Водночас цей вид індексування є „ахіллесовою п'ятою” для багатьох вітчизняних бібліотек, оскільки методичних матеріалів з цього питання явно недостатньо, якщо не сказати, що вони практично відсутні. Вибір КС і форма їх представлення не пов'язані жорсткими правилами і здійснюються переважно в довільній формі. Природно, що бібліотеки намагаються знайти вихід із цієї ситуації, розробляючи власні методики з формування КС з урахуванням свого досвіду роботи і на підставі свого бачення проблеми. Виникає запитання, наскільки ці методичні матеріали відповідають загальним тенденціям, потребам фахівців і користувачів бібліотек? Найчастіше недоліки індексування виявляються в процесі використання ЕК, переважно при обслуговуванні користувачів. З'ясовується, що створений ЕК низької якості і мало придатний для тематичного (особливо багатоаспектного) пошуку.

Кількість бібліографічних записів (БЗ) у створюваних бібліотеками електронних каталогах зростає з кожним днем, і варто замислитися, наскільки запланованими будуть ЕК за майбутньої інтеграції у світовий інформаційний простір. Слід відзначити, що чим більше минає часу, тим масштабнішим стає редагування, без якого не обійтися. Про його необхідність свідчить роль суб'єктивних чинників в індексуванні документів КС. Тобто, міра розкриття змісту документів безпосередньо залежить від рівня кваліфікації індексатора, його обізнаності з тематикою галузі, її проблематикою, наявністю або відсутністю нормативно-інструктивних і методичних документів тощо. Зважаючи на це, розроблення загальних методичних рекомендацій для формування КС є надзвичайно актуальним.

Якісне індексування досягається, якщо систематизатор прагне до повноти і точності відображення змісту документів КС, відбиваючи всі аспекти предмета, що розглядається в документі, або об'єкта. При цьому повинні в першу чергу враховуватися інформаційні інтереси користувачів системи. Вивчення результатів пошуку при обслуговуванні читачів, коригування й доповнення пошукових образів документів дозволяють зробити ЕК або БД інформативними, з високими пошуковими можливостями. Іншою складовою успіху високої якості

індексування є складання словників КС за єдиною методикою, при цьому враховуються особливості індексованого масиву документів, оскільки він призначається в першу чергу для потенційних споживачів конкретної бібліотеки або певної галузі.

Останнім часом обговорюються проблеми кооперування бібліотечно-інформаційної діяльності, а тому необхідно розробити документ, який би забезпечив єдиний підхід до координатного індексування вхідного потоку документів у всіх освітянських книгозбірнях.

З метою науково-методичного забезпечення бібліотек освітянської галузі у Державній науково-педагогічній бібліотеці України імені В. О. Сухомлинського було підготовлено методичні рекомендації „**Індексування документів ключовими словами**”. Роботу було розпочато у 2005 р. під час виконання наукового дослідження за темою „Створення, становлення і розвиток мережі освітянських бібліотек України (X–XX ст.)” і продовжено в межах НДР „Теоретичні та науково-практичні аспекти створення інтегрованого галузевого інформаційного ресурсу в Державній науково-педагогічній бібліотеці України імені В. О. Сухомлинського” (наук. керівник – канд. іст. наук, старш. наук. співроб., директор ДНПБ України ім. В. О. Сухомлинського П. І. Рогова).

Робота виконувалася з дотриманням вимог чинних стандартів: ГОСТ 7.66–92 (ИСО 5963–85) „Индексирование документов. Общие требования к координатному индексированию”, ДСТУ 2395–2000 (ГОСТ 30671–99)(ISO 5963:1985) „Обстеження документа, встановлення його предмета та відбір термінів індексування. Загальна методика”, ДСТУ 2398–94. „Інформація та документація. Інформаційні мови. Терміни та визначення”, ДСТУ ГОСТ 7.59:2003 „Індексування документів. Загальні вимоги до систематизації та предметизації” та ін.

У процесі підготовки рекомендацій було використано ряд методичних документів, зокрема Російської національної бібліотеки [26], Національної бібліотеки Республіки Казахстан [23], наукових бібліотек Томського [31] та Казанського [22] державних університетів та ін.

Підготовці рекомендацій сприяли численні публікації, присвячені питанням індексування ключовими словами. Було виявлено, що переважно і досить активно в цьому напрямі працюють російські бібліотекознавці, зокрема Д. І. Блюменау [2; 3], Ф. С. Воройський [7–10], М. В. Єрахторін [18], О. М. Зайцева [19], Є. Н. Пименов [28–30], Г. А. Скарук [32], І. Л. Скіпор [33], Е. Р. Сукіасян [36], Г. С. Щербініна [44] та ін. Серед вітчизняних дослідників слід відзначити Т. І. Вилегжаніну [4], Г. В. Власову [5], Н. Й. Дзюбу [13].

Під час роботи над рекомендаціями враховувався досвід роботи провідних бібліотек Росії, України та Білорусі, зокрема Російської національної бібліотеки, бібліотечних корпоративних мереж Москви і Казані, Національної парламентської бібліотеки України, Національної бібліотеки України для дітей та ін., а також власні напрацювання

Головне завдання пропонованого методичного посібника – надати практичну допомогу працівникам освітянських бібліотек в удосконаленні їхніх електронних каталогів, тим самим безпосередньо сприяти підвищенню якості та ефективності обслуговування користувачів.

Метою створення цих рекомендацій є усунення недоліків у використанні методу координатного індексування документів ключовими словами. Цій меті підпорядковані структура, зміст і наповнення методичного посібника, в якому докладно відображено правила описування основного смислового змісту документа за допомогою КС.

Методичні рекомендації ілюстровано великою кількістю прикладів. Приділено увагу досвіду роботи в галузі індексування документів різних видів. Розглянуто основні проблеми, з якими стикаються каталогізатори в роботі, а також запропоновано шляхи їх розв'язання. Запропонована методика має універсальний характер і може бути використана як для заповнення поля „Ключові слова” електронних каталогів, незалежно від використовуваної АБІС („ІРБІС”, „УФД-Бібліотека”, „MARC-SQL” тощо), так і для визначення ключових слів у процесі оформлення наукових публікацій. Пропонований посібник сприятиме поліпшенню якості та оперативності каталогізування документів, ефективному використанню електронних каталогів для задоволення інформаційних потреб користувачів.

Наприкінці посібника наведено основні терміни та визначення, пов'язані з індексуванням документів, із посиланнями на стандарти, з яких вони запозичені. Переклади з російськомовних стандартів подано в авторській редакції. Корисними бібліотечним працівникам будуть і додатки. Подано також список літератури з висвітлених питань.

Видання розраховане на працівників освітянських бібліотек України всіх видів, може бути використано в роботі книгозбірень інших систем і відомств, а також буде корисним для тих, хто займається науковою, педагогічною та науково-методичною діяльністю.

Зауваження та пропозиції просимо надсилати на адресу:

ДНПБ України ім. В. О. Сухомлинського, вул. М. Берлинського, 9, 04060, Київ

Електронна пошта: vnokd@i.ua

ВСТУП

Під індексуванням документів розуміється описування їхнього змісту засобами інформаційно-пошукової мови. Мета індексування – забезпечення високої ефективності пошуку необхідних документів і даних за показниками „точність” і „повнота”. Операції індексування документів належать до категорії інтелектуальних (аналітичних) видів робіт, що вимагають від каталогізаторів високої професійної теоретичної та практичної підготовки [7, с. 42]. Основними видами ІПМ є: класифікаційна, предметизаційна і дескрипторна.

Класифікаційна мова призначена для індексування документів та інформаційних запитів за допомогою понять і кодів будь-якої класифікаційної системи (УДК, ББК, ДКД тощо). В основі класифікаційних мов лежить систематична класифікація понять, тобто класифікація, що відбиває смислові відношення між поняттями. Класифікаційні мови призначені насамперед для формалізування логічних зв'язків слів природної мови. Усі ІПМ класифікаційного типу характеризуються рядом властивостей, що утруднюють їх використання під час інформаційного пошуку, особливо із застосуванням технічних засобів. До таких властивостей можна віднести: попередню координацію (зв'язок) слів і словосполучень у рубриці, складність оновлення й доповнення, практичну неможливість повного і детального розроблення схеми класифікування, трудомісткість використання під час індексування тощо.

Усе це призвело до створення мов вербального типу. Ці мови використовуються для представлення лексичних одиниць слова та вираження природної мови в їхній орфографічній формі. На відміну від класифікаційних мов, які використовуються для систематизації документів, вербальні мови орієнтовані на позначення не класифікаційних галузей, а конкретних об'єктів або предметів змісту документів у пошукових образах документів (ПОД) або в пошукових образах запитів (ПОЗ). До мов вербального типу належать предметизаційна і дескрипторна мови.

Предметизаційна мова призначена для індексування документів за допомогою рубрикаторів (словників предметних рубрик), а в разі їх відсутності – за методикою предметизації. В основу предметизаційної мови покладено перелік предметних рубрик (ПР), який є коротким формулюванням теми документа природною мовою. Предметизаційна мова є штучною системою і має бути побудована стандартизовано та уніфіковано (нормалізовано). Водночас предметні рубрики не призначені для глибокого, вірогідного пошуку за темою. За ПР важко знайти який-небудь аспект теми, неможливо перелічити всі власні назви предметів, про які йдеться в документі, тощо. Використання в ІПС тільки даних предметних рубрик для забезпечення інформаційного пошуку не може повністю задовольнити інформаційні запити користувачів. Трапляються випадки, коли деякі побічні теми документа не описуються ПР. Особливо це стосується наукових збірників, матеріалів конференцій, симпозіумів тощо, в яких інколи містяться статті, доповіді, що лише віддалено стосуються основної тематики, але становлять інтерес для користувачів конкретної бібліотеки. Найімовірніше, такі побічні теми не будуть відображені ні класифікаційними індексами, ні ПР, але можуть бути описані ключовими словами.

Дескрипторна мова слугує для координатного індексування документів і запитів за допомогою тезаурусів (словників дескрипторів) або ключових слів з використанням природної мови. В основу дескрипторних ППМ покладено абетковий перелік лексичних одиниць (ЛО). Головною різницею у використанні предметизаційної та дескрипторної мов є те, що в процесі предметизації індексується предмет (тема), а під час координатного індексування – термін, який цьому предмету відповідає. Комплекс ключових слів є свого роду лексичною моделлю наукового тексту. Функціональна значимість КС визначається тим, що вони є одним з найоптимальніших способів класифікування, зберігання й передавання інформації. Відбиваючи розвиток і термінодинаміку визначеної наукової галузі, комплекс КС є ще й системою відстеження й поширення сучасної термінології.

КС доцільно застосовувати для додаткового розкриття змісту документа на глибшому рівні, тобто використовувати ПР для описування основних предметів документа та їхніх аспектів, а КС – для подальшої деталізації, а також описування побічних тем документа. Тоді в ПОД будуть міститися ПР (нормалізована лексика) і КС у редакції автора. Такий підхід зручний для всіх категорій користувачів: як для тих, хто приблизно знає, що хоче знайти (для цього краще використовувати ПР), так і для тих, кого може зацікавити щось дуже специфічне (пошук відбувається за КС). Особливої цінності список КС може набути на останньому етапі пошуку, коли за запитом знайдено якийсь масив документів (за ПР). Ознайомившись зі списком КС знайдених документів, можливо одразу визначити потрібні документи, навіть не переглядаючи їх. Таким чином, використання КС збільшує пошукові можливості ЕК. Їх слід розглядати як додатковий засіб для збільшення повноти індексування та представлення різних за значимістю тем документа, не описаних іншими ППМ.

У зв'язку з необхідністю забезпечити пошук як за широкотематичними запитамі, так і за запитамі, сформульованими з точністю до детальних понять, ефективними для пошуку за широкотематичними запитамі є тільки класифікаційні ППМ, для пошуку за детальними запитамі – тільки вербальні ППМ. Виходячи з цього, у складі лінгвістичного забезпечення ІПС має бути хоча б одна вербальна і одна класифікаційна ППМ.

Плануючи і створюючи лінгвістичне забезпечення, передусім необхідно виходити з різноманітних потреб користувачів. Практика роботи бібліотек доводить, що користувачі частіше зацікавлені вести пошук не за класифікаційними системами, а за запитамі тематичного, предметного характеру. При цьому предметні рубрики менш популярні, ніж ключові слова, тому що в бібліотеках немає єдиної системи предметизації; відсутній універсальний словник предметних рубрик; користувачі не володіють методикою предметизації, що ускладнює формулювання пошукового образу запиту, адекватного пошуковому образу документа; не в усіх ЕК реалізується доступ користувачів до внутрішніх словників предметних рубрик тощо. Тому видається доцільним розширити можливості тематичного пошуку за допомогою використання ключових слів, які для користувачів є найбільш зрозумілим пошуковим засобом, за допомогою якого вдається знайти велику кількість релевантних документів.

За відсутності в Україні єдиного науково-дослідного, науково-методичного та інформаційного центру з розроблення лінгвістичних засобів бі-

біотеки на практиці стикаються з цілою низкою складних проблем, не вирішених на теоретичному рівні. До них, наприклад, належить проблема вибору ІПМ або їхніх елементів. Під час організації електронного каталогу неминуче виникає запитання – якою ІПМ скористатися для розкриття предметного змісту документа: мовою предметних рубрик, мовою ключових слів чи їх поєднанням? Незважаючи на принципові відмінності цих мов, практика їх застосування ускладнюється через невирішені проблеми. Ні в теорії предметизації, ні в теорії координатного індексування немає чітких рекомендацій відносно вибору таких параметрів, як довжина лексичної одиниці (слово або словосполучення), сталість словосполучень, прямий або інверсований порядок наведення лексичних одиниць у словосполученнях тощо. На практиці це призводить до суб'єктивізму, неможливості порівняння результатів семантичного опрацювання документів і запитів, знижує показники повноти і точності інформаційного пошуку.

Проблема нормалізації лексичних одиниць загострюється й через те, що немає практичних засобів контролю лексики, рівною мірою доступних всім бібліотекам. Ефективність реалізації мови предметних рубрик різко падає через відсутність виданих нормалізованих списків ПР та інформаційно-пошукових тезаурусів, які б підтримувалися в робочому стані, були придатні для поповнення й внесення змін, враховували галузеву специфіку.

Отже, вивчаючи теоретичні аспекти та практичне застосування ІПМ у процесі каталогізування, можна зробити висновки, що кожна з них має свої переваги та недоліки. Саме тому проблему підвищення ефективності пошуку доцільно вирішувати з використанням кращих якостей кожної з мов індексування. Паралельне використання кількох ІПМ – необхідна умова якісної систематизації та предметизації документів, що забезпечують і простий, і міжгалузевий, і міждисциплінарний пошук [19]. Значну увагу слід звернути на уніфікування ІПМ, оскільки це зробить процес індексування якіснішим і надасть більші пошукові можливості користувачам. Розглядаючи питання використання ІПМ, завжди слід пам'ятати: хоч би які швидкісні сучасні комп'ютери ми мали, без правильного використання інформаційно-пошукових мов ІПС буде працювати неефективно.

Державна науково-педагогічна бібліотека України імені В. О. Сухомлинського, що за своїми функціями та завданнями є головною бібліотекою освітянської галузі, науково-інформаційним, методичним і координаційним центром, використовує як лінгвістичні засоби кілька ІПМ, що забезпечує швидкий і багатоаспектний доступ користувачів до інформації найбільш відомою йому інформаційно-пошуковою мовою. Для найширшого розкриття змісту документів і надання їх користувачам наукове опрацювання документів здійснюється в рамках моделі „Комплексне індексування” [21].

Галузевий пошук інформації в ЕК здійснюється через формулювання нормалізованих тематичних рубрик, узятих з таблиць УДК, виражених за допомогою класифікаційних індексів. Пошукова система, побудована за галузевим систематичним принципом, дає змогу відповісти на запити, пов'язані з галуззю загалом, її окремими частинами або конкретним питанням, темою, якщо вони цікавлять користувача тільки у зв'язку з даною галуззю.

Для розширення пошукових можливостей ЕК бібліотеки застосовується предметний пошук, який на противагу галузевому є міждисциплінарним, між-галузевим, комплексним. За таким запитом користувачі отримують комплекс усіх документів з предмета, проблеми, теми незалежно від галузей знань, до яких належать усі ці документи.

Для подальшої деталізації ПОД і ПОЗ використовується мова ключових слів. Це обумовлено такими чинниками:

- ✓ універсальним тематичним складом документів, які опрацьовуються в бібліотеці, що робить проблематичним використання жорстко нормованої лексики;

- ✓ семантичними труднощами індексування документів і запитів жорстко нормованими термінами для рівня підготовки користувачів баз даних;

- ✓ відносною простотою використання лексики КС, близької до природної мови;

- ✓ можливістю постійного доповнення й оновлення списку КС, оскільки до абеткового переліку може бути введена будь-яка лексична одиниця, необхідна для індексування.

Це дозволяє опрацьовувати політематичний вхідний документний потік бібліотеки, найбільш повно розкриває семантичний зміст документів. У процесі аналітико-синтетичного опрацювання створюється комплексний пошуковий образ документа.

КООРДИНАТНЕ ІНДЕКСУВАННЯ ДОКУМЕНТІВ

Координатне індексування [coordinate indexing] документів – це інформаційно-пошукова мова вербального (словесного) типу, що передбачає вираження змісту документа множиною ключових слів і/або дескрипторів. У процесі індексування здійснюється інтелектуальний аналіз документа, у результаті якого з нього вибираються поняття, що перетворюються потім на терміни індексування [12].

Індексування змісту документа оцінюється з точки зору як поточних, так і передбачуваних потреб користувачів. Слід намагатися відобразити в ПОД усі найважливіші поняття, що розглядаються в документі. З літературних джерел відомо, що до 51% інформації втрачається через те, що індексатор не ввів у ПОД необхідний термін [33].

У створенні записів для ЕК за відсутності тезаурусів застосовують метод вільного координатного індексування, що базується на уявленні про те, що основний смисловий зміст документа може бути з достатнім ступенем точності й повноти виражений набором КС, які містяться в індексованому тексті.

Вільне координатне індексування означає індексування ключовими словами, вибраними безпосередньо з повного тексту документа й представленими в ПОД у термінології автора без нормалізації, з мінімальним контролем над лексикою і без урахування того, які КС уже використовувалися раніше для індексування таких самих або близьких за змістом документів [12]. На відміну від методики індексування з використанням тезауруса, де ПОД складається на основі КС документа, але представлених у вигляді стійких понять, зв'язаних з тезаурусом, під час вільного індексування словник КС створюється разом з формуванням бази.

Основними особливостями пропонованої моделі смислового опрацювання інформації та формування ПОД є максимальна повнота індексування інформації, можливість описання сфери інтересів користувачів у вигляді переліку основних термінів (ключових слів), які предметно незалежні, можуть бути не регламентовані тезаурусами та іншими нормативними словниками. Крім того, ця технологія становить значний інтерес для бібліотек будь-яких типів, видів і рівня автоматизації, оскільки вона універсальна, орієнтована на користувачів як локальних, так і глобальних телекомунікаційних мереж і покликана сприяти поліпшенню якості смислового пошуку.

Процес індексування містить такі етапи, що здійснюються у зазначеній послідовності:

- ⇒ аналізування й визначення змісту документа як об'єкта індексування;
- ⇒ відбирання понять, що характеризують зміст документа;
- ⇒ відбирання термінів індексування для позначення понять;
- ⇒ формування пошукового образу документа з термінів індексування.

Ці етапи можуть бути об'єднані в складі технологічних процедур за умови належного виконання кожного з них.

Аналізування й визначення змісту документа як об'єкта індексування

Аналізування документа залежить від його фізичної форми. Розрізняють такі два випадки:

- індексування друкованих документів;
- індексування аудіовізуальних творів або фонограм і документів на електронних носіях інформації.

Найповніше розуміння друкованих документів досягається в результаті докладного прочитання всього тексту. У разі необов'язковості або неможливості вичерпного ознайомлення з документом, індексатор має гарантувати повноту використання корисної інформації. Для цього залежно від виду та обсягу документа (книга, журнал, стаття тощо) ретельно вивчають його текстові частини, що є основним джерелом індексування:

- ✓ довідковий апарат документа:
 - основна назва, продовження назви;
 - анотація, реферат;
 - зміст;
 - передмова, висновки тощо
- ✓ вступ;
- ✓ назви частин, глав;
- ✓ перші фрази глав, розділів, параграфів;
- ✓ ілюстрації, схеми, таблиці та підписи до них;
- ✓ слова і групи слів, які в тексті підкреслено або виділено поліграфічними засобами;
- ✓ фрази, що містять вирази на зразок „основні завдання дослідження”, „мета дослідження”, „метод дослідження” тощо.

Індексування тільки за назвою не рекомендовано. У разі індексування за рефератом або анотацією необхідно стежити за адекватністю передачі в них змісту документа. Як назви, так і реферати (анотації) можуть бути невідповід-

ними тексту і в такому разі не можуть вважатися надійним джерелом, необхідним індексаторові.

Для аналізування аудіовізуальних творів, фонограм або документів на електронних носіях інформації додатково потрібні спеціальні технічні пристрої. У разі неможливості обстеження самого документа, індексування здійснюють на основі назви і/або стислого опису. Якщо письмовий опис виявиться невідповідним, неточним або недостатнім, індексатор повинен мати можливість переглянути або прослухати інформацію безпосередньо з носія.

Відбирання понять, що характеризують зміст документа

Прочитавши текст (переглянувши або прослухавши запис), визначають, до якої теми підходить документ і формулюють тезу (проблему). Далі слід написати найбільш значимі слова або поняття з тексту, що несуть основне змістове навантаження, позначають ознаки предмета, стан, процес, дію тощо.

Під час складання первинного списку КС враховується зацікавленість різних користувачів у різних аспектах проблеми: одному буде цікаво дізнатися про нові наукові факти, іншому з'ясувати специфічні терміни тощо. Кількість характеристик і понять, відображених у ПОД, визначає його повноту і є найважливішим показником якості індексування. У ПОД необхідно відобразити всі поняття, що можуть мати цінність для користувачів. Якщо в документі виявлено більше однієї теми зі сфери інтересів користувачів, ці теми мають бути розглянуті окремо. Тематика, що відображається під час індексування, не повинна обмежуватися вузькими рамками безпосередніх інтересів користувачів інформаційно-пошукової системи. У ПОД слід включати також поняття, пов'язані з побічними аспектами документа, наприклад соціальні та економічні проблеми наукових досліджень.

Основним критерієм у відборі понять має бути їхня потенційна цінність для вираження змісту документа або для його пошуку. Необхідно орієнтуватися на типові запити до інформаційно-пошукових систем:

- відбирати поняття, найбільш застосовувані користувачами ІПС;
- уточнювати склад лексики і граматичні правила ІПМ на основі зворотного зв'язку з користувачами.

Число термінів індексування, що приписуються одному документові, визначається не їхньою кількістю, а необхідністю найбільш повного описування всіх важливих аспектів змісту документа. Обсяг документа також може впливати на повноту індексування. Обмеження числа термінів має ґрунтуватися на змістовому відборі найбільш важливих понять. За даними, опублікованими в спеціальній літературі, для кожного документа оптимальним вважається від 8 до 12 КС, які не є синонімами [7, с. 44 ; 33]. У практиці окремої бібліотеки їх може бути як більше, так і менше. Список ключових слів має передусім найбільш повно відображати ті частини змісту документа, які мають найбільший читацький попит. Порядок КС у списку не залежить від їхньої ролі та характеру відношень між ними в змісті кожного конкретного документа.

Не рекомендовано описувати документ як фізичний об'єкт з точки зору його форми, обсягу, матеріалу тощо. Допускається відобразити в ПОД подібну інформацію, якщо це дозволить точніше визначити відповідність документа інформаційним потребам користувачів.

Відбирання термінів індексування для позначення понять

Головний принцип і критерій для відбирання ключових слів (словосполучень) полягає, по-перше, у збереженні їхньої науковості. По-друге, будучи „ключем” у розкритті змісту, вони мають бути нормативними. Нормативність необхідна для досягнення однозначності трактування понять, для збереження системи понять, яка склалася, що забезпечить ясність і доступність використання термінів користувачами. *При вільному координатному індексуванні слова або словосполучення, що визначаються як КС, обов’язково мають бути в змісті документа.*

Виносити як КС слід тільки ті поняття, що безпосередньо пов’язані з дослідженням автора, виражають його основну суть, є найбільш значимими для конкретного дослідження. У список КС нові терміни або поняття, що вводяться автором, уміщують тільки в тому випадку, якщо в роботі обґрунтовано необхідність їх введення в понятійний словник науки і/або вони органічно пов’язані з проблемою дослідження й розкривають її суть. При цьому слід утримуватися від „засмічення” списку КС невластивими науці поняттями й термінами, оскільки це ускладнить пошук користувача. Для цього терміни, відібрані як КС, перевіряють на точність і прийнятність за словниками, енциклопедіями, довідниками, класифікаційними таблицями, інформаційно-пошуковими тезаурусами, термінологічними стандартами та іншими джерелами, визнаними авторитетними в конкретній галузі.

Якщо в документі міститься принципово нове поняття (термін), його також використовують як КС, навіть якщо воно на момент створення БД не описано в жодному довідковому виданні, ані в інших БД. Слід мати на увазі, що комплекс КС не повинен містити інформацію про зміст і значення цих слів. КС – це пошукові слова, їхнє основне призначення – сприяння пошуку бібліографічної інформації про документи, в яких дається опис і докладне тлумачення цих понять та/або явищ. Потрібно зазначити, що слово може стати терміном лише через деякий час після виникнення потреби в ньому. У період невизначеності для позначення одного і того ж предмета різні фахівці можуть вживати різні слова.

У бібліографічній практиці сутність розуміння тексту і формування його пошукового образу полягає в представленні тексту документа короткими та суттєвими формулами, визначенням у кожній формулі центрального за змістом поняття, зв’язуванні понять між собою й створенні в результаті єдиного логічного ланцюга ідей. Отже, у смисловому аспекті комплекс КС документа будується за такою схемою:

Суб’єкт – якісні характеристики – властивості
Дії – способи – технології
Галузі діяльності (застосування)

З отриманих логічних ланцюжків слів або словосполучень виявляють ті, що становлять домінують тексту. При цьому треба враховувати, що більшість сучасних АБІС передбачає можливість пошуку не тільки в окремих полях БЗ, але й за всіма полями одночасно. Тому під час остаточного відбирання КС слід

уникати повного дублювання назви, продовження назви документа і/або предметних рубрик.

Формування пошукового образу документа з термінів індексування

ПОД формується з вибраних термінів індексування за допомогою граматичних засобів ІПМ даної ІПС. КС оформляються на основі визначених правил, які забезпечують однозначність їх розуміння й застосування.

Насамперед з конкретного тексту документа вибирають спеціальні (специфічні) терміни, що використовуються тільки в окремій галузі та є її розпізнавальними знаками. До КС передусім належать терміни, найважливішими властивостями яких є номінативність (здатність називати предмет), однозначність, точність, відсутність експресії. Як КС можуть застосовуватися: повнозначні слова (унітерми), складні та складноскорочені слова, термінологічні словосполучення, фрази і речення (як виняток), аббревіатури, чисельні характеристики, хронологічні дані, власні імена, символічні позначення, формули, що несуть смислове навантаження в тексті документа. Слова і словосполучення, визначені як КС, за формою вираження мають виконувати інформаційну функцію. Найбільш прийнятною є називна форма:

комп'ютери
французький живопис
особистісно орієнтоване навчання
тематико-типологічний план
аудіовізуальні засоби навчання
День перемоги
21 ст.
1990–1995 рр.
Ортеги гіпотеза
США
Ньютона закон тяжіння; $F = \gamma m_1 m_2 / r^2$

Як КС використовуються іменники окремо або в поєднанні з прикметниками або числівниками, власні імена, іменні словосполучення, числівники.

Займенники до складу ключових слів не входять.

Дієслова, прислівники й дієприслівники видозмінюють:

- дієслова записують у вигляді віддієслівних іменників:

перетворюють – перетворення;

- прислівники змінюють на відповідні прикметники:

автоматично – автоматичний;

- дієприслівники перетворюють на відповідні іменники:

навчаючи – навчання

граючи – гра

Не рекомендується включати в ПОД як самостійні КС неінформативні поняття (*теорія, типологія, основи, методологія* тощо), які можуть застосовуватись у сполученні майже з усіма предметами.

Для уточнення таких понять їх включають у ПОД у сполученні з іншими словами, що звужують (конкретизують) їхні значення:

теорія історії
типологія психології
технології навчання
методологія науки

Як КС не рекомендується використовувати службові слова (частки, прийменники, сполучники, вигуки тощо).

Кількість слів у словосполученнях має відповідати кількості слів у прийнятому терміні. Рішення про розділення або збереження словосполучень або складних слів ухвалюють з урахуванням лінгвістичних і прагматичних критеріїв.

Рекомендується зберігати:

– ідіоматичні вислови (семантично неподільні звороти, значення яких не виходять буквально зі значень його компонентів, стійкі вислови з самостійним значенням):

нічні метелики
нові українці
Содом і Гоморра
спальні райони
чорні діри
чума ХХІ ст.
шокова терапія

– словосполучення, окремі компоненти яких мають широке значення:

дерев'яні конструкції
інформаційне забезпечення
педагогічні системи

– словосполучення, що позначають обладнання, машини, матеріали тощо і пишуться через дефіс:

зволожувачі-охолоджувачі
нумератори-реєстратори
цукор-сирець

Якщо словосполучення стає, його використовують як КС, керуючись при цьому такими критеріями:

– при утворенні словосполучення одне зі слів змінює своє значення, внаслідок чого відбувається переосмислення одного з компонентів словосполучення, наприклад:

гаряча лінія
гарячі точки
гарячі новини (*переосмислюється прикметник „гарячий”*);

– словосполучення, в якому при перетворенні прикметника в іменник загубиться його сенс, наприклад:

бібліотечний почерк (*неможна замінити на: бібліотеки; почерк*);

– словосполучення, що є власними іменами або містять їх, наприклад:

Латинська Америка
Нижня Саксонія
Пржевальського кінь
рентгенівські промені

Словосполучення розділяють, якщо:

– розглядається процес і об'єкт, наприклад:

будівництво бібліотек = бібліотеки; будівництво

– розглядається частина і ціле, наприклад:

читальний зал бібліотеки = бібліотеки; читальні зали

– розглядається об’єкт і його призначення, наприклад:

соціальне середовище для соціалізації особистості =
соціальне середовище; соціалізація; особистість

– об’єднуються два іменники, наприклад:

бібліотеки Канади = бібліотеки; Канада

– іменник супроводжується кількома прикметниками, наприклад:

обласні наукові медичні бібліотеки =
обласні бібліотеки; медичні бібліотеки; наукові бібліотеки

Терміни-речення залишають, не поділяючи їх на словосполучення, тільки в тому випадку, якщо такий поділ призводить до втрати смислу вихідного речення. Наприклад, такими КС можуть бути різноманітні команди, лозунги, заклики.

Після визначення й відбору КС необхідно дотримуватися їхньої одноманітності або нормалізування. КС записують в єдиній граматичній формі (відмінок, число тощо). Форму представлення КС уточнюють за словниками ключових слів (тезаурусами), термінологічними, тлумачними словниками тощо.

Слід зауважити, що формулювання КС в ЕК багато в чому залежить від пошукових можливостей ІПС. У переважній більшості ЕК можливо створити для кожного запису фактично необмежену кількість точок доступу, причому як пошукові можна вибирати будь-які поля та підполя БЗ документа, видобувати окремі слова або навіть цілі фрази. В ЕК реалізовано різноманітні види пошуку: простий, розширений, за словниками тощо, розрахованими як на непідготовленого, так і на досвідченого користувача. Застосування логічних операторів „і”, „не”, „або” дозволяє: виділяти записи, в яких містяться кілька заданих термінів одночасно; здійснювати пошук за словосполученнями, скомбінованими з уні-терм, за скороченими або точно заданими термінами. У такому випадку можна було б не враховувати відмінок і число іменників, але ми вважаємо за потрібне використовувати „м’яку” нормалізацію для точнішого створення ПОД.

Відмінок КС

Ключові слова формулюють, як правило, в називному відмінку. У словосполученнях допускається застосування називного відмінка в поєднанні з іншими відмінками, частіше родовим і місцевим. При цьому значуще слово наводять у називному відмінку. У разі доцільності допустиме також використання прийменників:

бази даних
боротьба за владу
боротьба з тероризмом
капіталістичний спосіб виробництва
роль особистості в історії
ракоподібні
церква і держава

Число КС

Вибір форми однини або множини в нормалізуванні КС визначається як мовними чинниками, так і специфічними особливостями терміносистеми. Загальне правило використання однини або множини таке:

✓ форма множини використовується для обчислювальних іменників, про які можна спитати „скільки?“:

вчителі
документи
школярі

✓ форма однини використовується для необчислювальних іменників:

інформація
історія
мислення

В однині використовують:

➤ терміни, що позначають процеси:

державне регулювання
інформаційний супровід
методичне забезпечення
навчання

➤ назви наукових дисциплін:

бібліотекознавство
педагогіка

➤ назви державних або громадських організацій, установ, партій, товариств, об'єднань, спілок тощо:

Кабінет Міністрів України
Інститут педагогіки
Спілка письменників України
ЮНЕСКО
УБА

➤ назви подій:

Велика Вітчизняна війна

Якщо існує змістова різниця в застосуванні термінів в однині або множині, використовують обидві форми. У таких випадках рекомендується до КС додавати слова пояснення – релятори. Вони записуються після КС у круглих дужках і не мають самостійного значення:

пісок (матеріал)	–	піски (простір, зайнятий піском)
шум (будь-який)	–	шуми (сердечні)
глибина (вимір)	–	глибини (глибокі місця)
груша (дерево)	–	груші (плоди)
дух (запах)	–	духи (пахуча рідина)

Використовувати одне і те ж КС одночасно в однині та множині в одній БД неприпустимо. Якщо той самий термін набуває різного значення (або відтінків значення) в однині та множині, записують ці терміни як в однині, так і в множині, а до КС в однині додають релятори:

роман (жанр)	романи
Земля (планета)	землі
приказка (жанр)	приказки

Інколи один і той же термін в одних галузях використовується тільки в однині, а в інших – тільки в множині:

вода

але

підземні води
стічні води

У практиці індексування трапляються терміни, що можуть використовуватись у формі множини лише в поетичній, спеціальній, технічній мові. У цьому випадку рішення про вибір форми числа КС ухвалюється на підставі аналізу словників, довідників, а також застосування цього терміна в професійній лексиці:

ризик
економічний ризик
страховий ризик
управління ризиками

ринок
зовнішній ринок
національні ринки
ринки золота
ринки збуту
світовий ринок
сировинні ринки
цільові ринки

Інверсія в КС

Зважаючи на те, що КС формуються на основі природної мови, перевага надається прямому порядку слів. Інверсію (перестановку) застосовують в окремих випадках. Наприклад, інверсуються власні імена, що входять до складу КС. На перше місце виноситься власне ім'я в родовому відмінку:

Допплера ефект
Лаптевих море
Макаренка система
Менделєєва таблиця
Ньютона закон
Патона міст

Полісемія та омонімія КС

Під час визначення КС велике значення має вибір повної або короткої форми назви предмета. Уникненню полісемії та омонімії сприяє переважне використання повної форми назви предмета замість короткої (аббревіатури):

ДПА – Державна податкова адміністрація
довідково-пошуковий апарат

Виняток становлять аббревіатури, що майже повністю витіснили коротку форму, частіше використовуються не тільки в терміносистемі, але й у звичайній мові та не потребують розкриття:

США
НАТО
ООН
ЮНЕСКО

але

вищі навчальні заклади (а не ВНЗ)

Якщо в документі використано аббревіатуру, рекомендується включати в КС як аббревіатуру, так і повну форму словосполучення:

ІПС; інформаційно-пошукові системи

Використання як стійких, так і нестійких словосполучень, а також додавання до загальнонаукових термінів інших слів, які конкретизують їхнє значення, також допомагають уникнути полісемії та омонімії:

економічні проблеми
екологічні проблеми

бібліотечні мережі
електричні мережі

страхові агенти
атмосферні агенти

акції компаній
політичні акції

В окремих випадках, коли неможливо додати інші пояснювальні слова (іменники або прикметники), використовуються релятори, які вказують на дисципліну, що вивчає цей предмет:

губки (зоол.)
лігатура (металург.)
лігатура (мед.)
лігатура (муз.)
Відродження (іст. епоха)
радикали (політ.)
радикали (матем.)

Синонімія КС

Синоніми КС включаються в ПОД якщо: терміни ще не є сталими, існують паралельно, є поширеними:

мемуари
спогади

етологія
поведінка тварин

довкілля
навколишнє середовище

густосологія
естетичне виховання

андрагогіка
освіта дорослих

Відродження
Ренесанс

Це ж стосується й різних форм написання одного і того ж слова:

неперервна освіта
безперервна освіта

ІНДЕКСУВАННЯ ОКРЕМИХ ВИДІВ ДОКУМЕНТІВ

Індексуванню ключовими словами підлягають усі документи (їхні складові частини), крім творів художньої літератури, біографічних (біобібліографічних) видань, присвячених окремим особам (персоналіям), а також енциклопедій, словників, довідників і бібліографічних посібників універсального змісту („Українська радянська енциклопедія”, „УСЕ. Український словник енциклопедичний”, „Літопис книг” тощо).

Під час індексування багатотомних, серіальних і продовжуваних видань КС визначають для видання загалом і, у разі необхідності, для кожного тому (випуску, частини) окремо.

Для індексування збірників, матеріалів наукових заходів (з'їздів, симпозіумів, конференцій, семінарів тощо) КС визначають як на збірники в цілому, так і на окремі статті.

Під час описування підручників і навчальних посібників обов'язково зазначають клас:

англійська мова; 5 клас

Для індексування галузевих (тематичних) енциклопедій, словників, довідників, бібліографічних посібників визначають узагальнені КС, які виражають тему або галузь знань:

„Педагогічний словник” – педагогіка; терміни; словники

Під час індексування телефонних або адресних довідників як КС записують також вид документа і назву міста:

„До послуг киян” – служба побуту; телефонні довідники; Київ

Індексування документів за формальними ознаками (форма, вид, жанр, читацьке призначення тощо) найперше здійснюють за допомогою формальних підзаголовків предметних рубрик з відображенням їх у відповідному підполі. У полі *Ключові слова* зазначають переважно формальні ознаки документів, відмінні від них або детальніші.

При аналітичному опрацюванні періодичних видань (газет, журналів, бюлетенів, календарів) здійснюється індексування окремих статей.

Індексування електронних ресурсів відбувається за загальними правилами.

ПРАВИЛА ЗАПОВНЕННЯ ПОЛЯ „КЛЮЧОВІ СЛОВА”

Однією з основних вимог до електронного каталогу є забезпечення ним широких пошукових можливостей, у тому числі тематичних. У комунікативних форматах MARC, які використовуються в ЕК, це досягається завдяки використанню інформаційно-пошукових елементів у полях тематичного пошуку, наприклад:

- для індексів класифікації:
 - *Індекси УДК*
 - *Індекси ББК*
- для предметних рубрик:
 - *Персоналія*
 - *Організація/захід як предметна рубрика*
 - *Основна предметна рубрика*
 - *Географічна назва як предметна рубрика*
- для ключових слів:
 - *Ключові слова*

Загальні правила наведення КС

Ключові слова записують у поле *Ключові слова*.

КС записують з дотриманням правил сучасної орфографії.

КС записують мовою каталогізації, прийнятою в бібліотеці, переважно – українською. Допускається паралельне наведення ключових слів іншими мовами. У разі необхідності зазначення терміна мовою оригіналу документа, КС забезпечуються помітками про приналежність до тієї або іншої природної мови. Для цього використовують коди мов MARC-форматів, які додаються до терміна в круглих дужках як релятори. Їх записують малими літерами латинської абетки, наприклад:

англійська – (eng)
німецька – (ger)
французька – (fra)

Назви фірм, торговельних марок, моделей виробів, марок матеріалів тощо записують у формі, використаній у тексті документа, і в загальноприйнятій транскрипції:

Microsoft (eng); IBM; Panasonic (eng); BMW; Coca-Cola (eng); Кока-Кола;
Internet (eng); Інтернет

Назви фірм мають бути, за можливості, короткими. Якщо в документах широко вживається скорочена назва або абревіатура, то їх слід використовувати. Наприклад:

фірма Bosch (eng) – замість Robert Bosch
фірма GM – замість General Motors Co

Скорочення типу *Co, Ltd, GmbH, KK* тощо наводити не слід. Замість них використовують слово *фірма*.

У разі вибору як КС запозичених термінів, на які в українській мові транскрипція ще не закріпилася („дистрибутори” і „дистриб’ютори”, „франчайзинг” і „франшизинг” тощо), рекомендується використовувати тільки одну з форм на основі ухвалення спільного з іншими індексаторами рішення та із занесенням його до картотеки методичних рішень (КМР).

Якщо термін, визначений як КС, у тексті документа відсутній і введений на розсуд індексатора (синонім), він має бути позначений. Для цього до терміна без пробілів додається астериск (*). Наприклад:

Ренесанс
Відродження*

КС записуються малими літерами, крім власних імен. Для іноземних мов застосування великих літер обумовлено правилами граматики відповідної мови.

У написанні ключових слів жодні скорочення не допускаються, крім хронологічних ознак (р., pp., ст., до н.е. тощо). Якщо наявна абревіатура поняття, її вносять окремим КС:

агропромисловий комплекс; АПК
інформаційно-пошукові мови; ІПМ

Релятори записують з малої літери в круглих дужках після КС через один пробіл. Скорочення реляторів допускається застосовувати відповідно до положень стандартів.

Власні назви наводять без лапок, з додаванням у разі необхідності реляторів:

Ранок (журнал)
Тарас Шевченко (теплохід)
Просвіта (товариство)
Волга (автомоб.)
Руслан (літак)
Динамо (футбол. клуб)
Освіта (видавництво)

Лапки в ключових словах використовують у виняткових випадках (використання терміна в переносному значенні, в авторській інтерпретації тощо):

„мозковий штурм”
„п’ята колона”

Порядок наведення КС довільний, оскільки на пошук в ЕК не впливає.

Спеціальні правила наведення КС

✦ Національно-визвольні рухи записують як назву руху і назву нації, народності:

національно-визвольний рух; українці; гайдамаки; Коліївщина

✦ Офіційно прийняті й стійкі поняття, такі, як світські та релігійні свята, терміни народного календаря записують повним словосполученням без лапок:

День незалежності
1 Травня; Першотравень
1 вересня; День знань
Новий рік
Тетянин день
Івана Купала
Покров Пресвятої Богородиці

✦ У разі існування кількох варіантів назв перевага надається коротшому:

Різдво Христове – Різдво

але

Різдво Пресвятої Богородиці

✦ Назви окремих технічних засобів записують із конкретизацією назви або номера моделі. При цьому ключовим словом описується тип виробу:

автомобілі; ВАЗ-2106; Запорожець (автомоб.)
комп’ютери; Intel (eng)

✦ Для програмного забезпечення зазначають версію продукту:

електронні таблиці; Excel 5.0;
операційні системи; Windows 2000;
програми векторної графіки; CorelDRAW 8

✦ Назви державних установ і організацій записують повністю і в загальноприйнятому скороченні:

Верховна Рада України; ВР України
Національна академія педагогічних наук України; НАПН України

✦ Назви війн, боїв і битв записують як стійкі словосполучення:

Перша світова війна
Велика Вітчизняна війна
Столітня війна

битва за Дніпро
Куликовська битва
бої за Київ
бій під Крутами

✦ Поняття, пов'язані з мистецтвом, культурою, народною творчістю нації або народності, записують як іменник із прикметником:

французький живопис
українські пісні; народні пісні
білоруська вишивка

✦ Поняття, пов'язані з територіальною ознакою, записують як окремі ключові слова:

США; література
Україна; архітектура

✦ Народні промисли описують нормативною лексикою як стійки словосполучення:

Димківська іграшка
Петриківський розпис

✦ Символічні позначення застосовують у прийнятій формі:

C4-фотосинтез

✦ Чисельні характеристики (діапазон температур, тиску тощо) наводять зі словесним формулюванням:

діапазон температур 80–90°C

✦ Для описування дат і часових інтервалів дотримуються таких правил:
– стійкі назви ер, періодів, епох, віків записують як одну лексичну одиницю:

Палеозойська ера
Юрський період
епоха бронзи
Срібний вік
бронзовий вік
середні віки

– для описування сторіччя використовують арабські цифри і скорочення слова „сторіччя” з пробілом:

19 ст.
21 ст.

– для описування року використовують арабські цифри і скорочення слова „рік” з пробілом:

2005 р.

– роки „від” і „до” записують через тире без пробілів:

1941–1945 рр.

– десятиліття записують із нарощуванням закінчення:

90-ті рр.
30–40-ві рр.

Окрім цих правил, індексатори можуть, залежно від особливостей індексованого потоку документів і запитів споживачів, виробити власні. Їх слід зафіксувати і дотримуватися при створенні ЕК і БД.

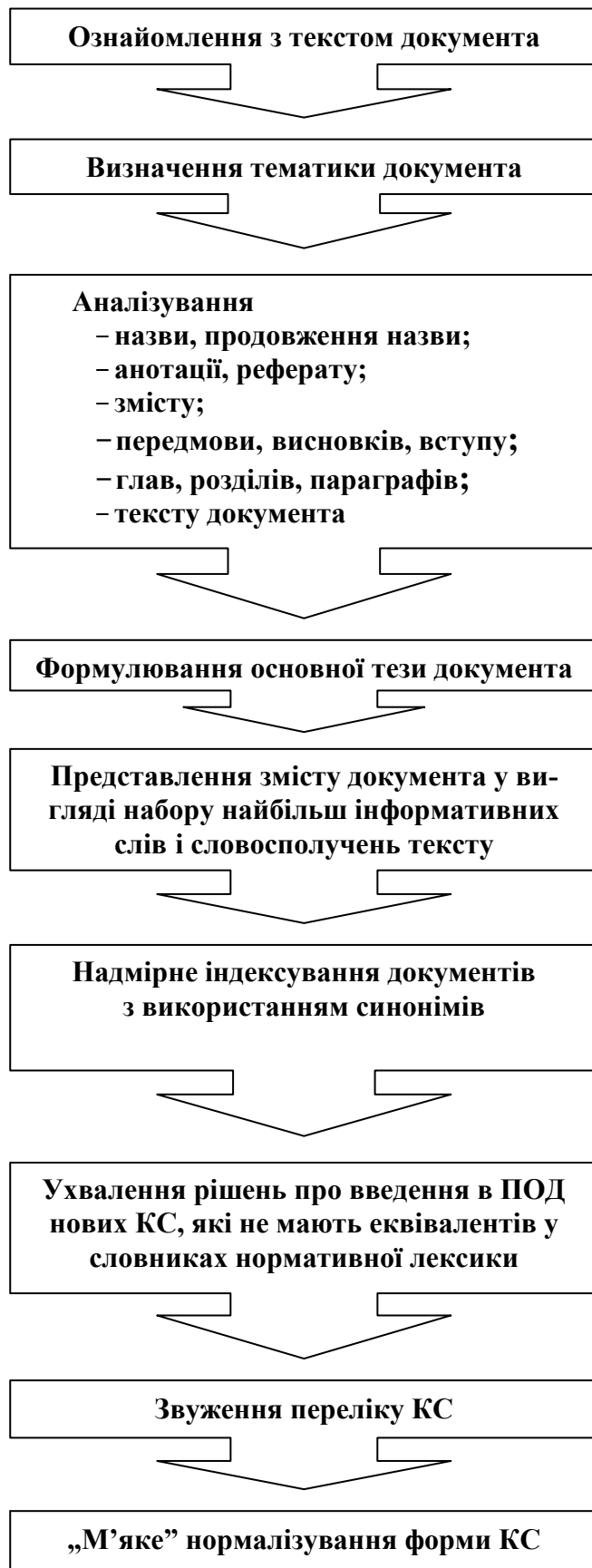
МОЖЛИВІ ПОМИЛКИ ПІД ЧАС ІНДЕКСУВАННЯ ДОКУМЕНТІВ І ЗАПОВНЕННЯ ПОЛЯ „КЛЮЧОВІ СЛОВА”

У процесі індексування й виділення ключових слів трапляються помилки, яких слід уникати. Серед найпоширеніших помилок можна відзначити такі:

- * неповне відображення змісту документа в ПОД;
- * смислові помилки індексування;
- * включення в ПОД ключових слів, які відповідають малоінформативним аспектам змісту документа (коли об’єкт тільки згадується);
- * використання випадкових, маловживаних у тексті термінів;
- * неправильна побудова ключових слів, зокрема використання невиправдано довгих словосполучень;
- * винесення як КС понять, які не розкривають суті питання (наприклад, термін „активізація” може бути віднесений як до розумової діяльності, так і до роботи в колективі тощо);
- * винесення як КС широкого поняття, що виходить за межі дослідження, у той час як у документі висвітлено тільки один аспект, що його стосується (наприклад, як КС виділено: *Верховна Рада України*, у той час як у роботі йдеться про законодавчу діяльність Верховної Ради України в галузі освіти). Користувач одержує неточну інформацію, оскільки за ключовим словом розраховує на ширші відомості, більший обсяг інформації.
- * випадкові орфографічні помилки (описки);
- * технічні помилки: відсутність або наявність зайвих пробілів, запис КС не в те поле тощо.

Будь-який словник КС з часом потребує редагування, тому що „обростає” великою кількістю синонімів, омонімів, порушень правил при складанні КС. Щоб уникнути його глобального редагування, необхідно вже при складанні продумати способи усунення явищ, що заважають пошуку інформації.

**СХЕМА ВИЗНАЧЕННЯ КЛЮЧОВИХ СЛІВ
У ПРОЦЕСІ ВЛЬНОГО КООРДИНАТНОГО ІНДЕКСУВАННЯ ДОКУМЕНТА**



ПРИБЛИЗНИЙ ПЕРЕЛІК ГАЛУЗЕВИХ РЕЛЯТОРІВ

авіа. – авіаційний термін	комп'ют. – комп'ютерний термін
автомоб. – автомобільний термін	лінгв. – лінгвістичний термін
аеродин. – аеродинамічний термін	літ. – літературний термін
анат. – анатомічний термін	лог. – термін логіки
антр. – антропологічний термін	мат. – математичний термін
археол. – археологічний термін	мед. – медичний термін
архіт. – архітектурний термін	мет. – металургійний термін
астр. – астрономічний термін	мех. – термін механіки
біб. – бібліотечний термін	мист. – термін мистецтва
біол. – біологічний термін	мінер. – мінералогічний термін
бот. – ботанічний термін	мкб. – мікробіологічний термін
буд. – будівельний термін	мор. – морський термін
бухг. – бухгалтерський термін	муз. – музичний термін
вет. – ветеринарний термін	опт. – термін оптики
військ. – військовий термін (від „військо”)	пед. – педагогічний термін
воєн. – воєнний термін (від „війна”)	політ. – політичний термін
геогр. – географічний термін	психіатр. – психіатричний термін
геод. – геодезичний термін	психол. – психологічний термін
геол. – геологічний термін	радіоелектрон. – термін радіоелектроніки
геом. – геометричний термін	радіотехн. – радіотехнічний термін
гідродин. – гідродинамічний термін	реліг. – релігійний термін
гідрол. – гідрологічний термін	с.-г. – сільськогосподарський термін
гірн. – термін гірничої справи	соціол. – соціологічний термін
граф. – графічний термін	спорт. – спортивний термін
дет. маш. – термін деталей машин	стат. – статистичний термін
дипл. – дипломатичний термін	театр. – театральний термін
дор. буд. – термін дорожнього будівництва	техн. – технічний термін
екон. – економічний термін	турист. – туристичний термін
електр. – електротехнічний термін	фарм. – фармацевтичний термін
естет. – естетичний термін	фіз. – фізичний термін
етич. – етичний термін	фізіол. – фізіологічний термін
етногр. – етнографічний термін	філол. – філологічний термін
залізн. – залізничний термін	філос. – філософський термін
зоол. – зоологічний термін	фін. – фінансовий термін
інф. – термін інформатики	фолькл. – фольклорний термін
іст. – історичний термін	хім. – хімічний термін
кібернет. – кібернетичний термін	шах. – шаховий термін
	юр. – юридичний термін

ТЕРМІНИ ТА ВИЗНАЧЕННЯ

Індексування – процес вираження змісту документа та/або запиту інформаційно-пошуковою мовою (ГОСТ 7.27-80).

Інформаційно-пошукова мова (ІПМ) – мова, призначена для вираження змісту документів або запитів або описування фактів з метою подальшого пошуку (ГОСТ 7.27-80).

Пошуковий образ документа (ПОД) – текст, який виражає інформаційно-пошуковою мовою основний зміст документа й використовується для інформаційного пошуку (ГОСТ 7.27-80).

Інформаційно-пошуковий масив – упорядкована сукупність документів, фактів або відомостей про них, призначена для інформаційного пошуку (ГОСТ 7.27-80).

Інформаційно-пошукова система (ІПС) – призначена для інформаційного пошуку сукупність інформаційно-пошукового масиву, інформаційно-пошукової мови, правил її застосування, критерію видачі й технічних засобів (ГОСТ 7.27-80).

Координатне індексування (КІ) – індексування за допомогою перелічення ключових слів або дескрипторів (ГОСТ 7.27-80); індексування, мета якого полягає у всебічному відображенні змісту документа або запиту за допомогою включення в пошуковий образ усіх необхідних для цього термінів індексування (ГОСТ 7.66-92).

Контрольоване індексування – індексування, при якому передбачена заміна інформативних слів тексту дескрипторами, зазначеними у визначеному інформаційно-пошуковому тезаурусі або іншому словнику індексування (ГОСТ 7.66-92).

Вільне індексування – індексування, технологія якого не передбачає заміни інформативних слів тексту відповідно до рекомендацій спеціального словника індексування (ГОСТ 7.66-92).

Лексична одиниця (ЛО) ІПМ – послідовність символів, слово, словосполучення, фрагмент слова або умовне позначення, які розглядаються в даній ІПМ як елементарна одиниця, що використовується для представлення в пошукових образах документів або запитів визначеного поняття, об'єкта або значення параметра.

Ключове слово (КС) – слово або словосполучення з тексту документа або запита, що несе в даному тексті суттєве змістове навантаження з точки зору інформаційного пошуку (ГОСТ 7.27-80).

Дескриптор – словникова одиниця інформаційно-пошукової мови, виражена словом, словосполученням або кодом, яка є ім'ям класу умовної еквівалентності, до якого включені еквівалентні й близькі за значенням ключові слова (ГОСТ 7.27-80).

Інформативне слово – слово або словосполучення в тексті документа або запиту, яке несе в ньому суттєве змістове навантаження (ГОСТ 7.66-92).

Специфічний термін – інформативне слово, що найбільше відображає зміст документа, використання якого відрізняє цей документ від інших тематично близьких документів (ГОСТ 7.66-92).

Повнота індексування – ступінь відображення в пошуковому образі змісту документа (або запити), що визначається як відношення числа специфічних термінів і фактографічних відомостей, які включаються в пошуковий образ, до числа таких термінів і відомостей, які є в тексті документа або запити (ГОСТ 7.66-92).

СПИСОК ЛІТЕРАТУРИ

1. Алешин, Л. И. Поиск документов: анализ и новые возможности / Л. И. Алешин // Библиография. – 2002. – № 4. – С. 14–22.
2. Блюменау, Д. И. Индикаторный метод компьютерного свертывания в процессе обучения аналитико-синтетической переработке информации / Д. И. Блюменау, Л. Н. Афанасова // Науч. и техн. б-ки. – 2001. – № 12. – С. 29–41.
3. Блюменау, Д. И. Информационный анализ/синтез для формирования вторичного потока документов : учеб.-практ. пособие / Блюменау Д. И. – СПб. : Профессия, 2002. – С. 46–56.
4. Вилегжаніна, Т. І. Інформаційно-пошуковий тезаурус / Т. Вилегжаніна, О. Збанацькі ; Нац. парлам. б-ка України. – Донецьк : Альфа-Прес, 2004. – 1104 с.
5. Власова, Г. В. Індексуння як процес аналітико-синтетичної переробки інформації : навч. посіб. / Г. В. Власова. – К., 2006. – С. 172.
6. Возможности тематического поиска в электронном каталоге ГПНТБ СО РАН / Л. А. Жарикова, Г. А. Скарук, А. В. Бобко, Г. И. Сергеева // Науч. и техн. б-ки. – 2002. – № 8. – С. 30–38.
7. Воройский, Ф. С. Индексирование документов в АИБС / Ф. Воройский // Библиотека. – 1996. – № 9. – С. 42–44.
8. Воройский, Ф. С. Информационные ресурсы АРБИКОН – пора подумать об их качестве / Воройский Ф. С., Острая С. А. // Науч. и техн. б-ки. – 2005. – № 2. – С. 76–83.
9. Воройский, Ф. С. Некоторые пути повышения качества поисковых характеристик электронных каталогов / Ф. С. Воройский // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества : материалы 3-й Междунар. конф. „Крым – 96”. – 1996. – Т. 1. – С. 286–289.
10. Воройский, Ф. С. Основные принципы обеспечения информационного поиска в корпоративных электронных каталогах [Электронный ресурс] / Ф. С. Воройский // Информационные технологии, компьютерные системы и издательская продукция для библиотек : Восьмая Междунар. конф. и Выставка LIBCOM–2004. 15–19 нояб. 2004 г. – Режим доступа: <http://www.gpntb.ru/libcom4>. – Назва з екрану.
11. ГОСТ 7.25–2001. Тезаурус информационно-поисковый одноязычный. Правила разработки, структура, состав и форма представления. – Изд. офиц. – Минск, 2001. – 16 с. – (Система стандартов по информации, библиотечному и издательскому делу).
12. ГОСТ 7.66–92 (ИСО 5963–85). Индексирование документов. Общие требования к координатному индексированию. – Изд. офиц. – М. : Госстандарт России, 1992. – 16 с. – (Система стандартов по информации, библиотечному и издательскому делу).
13. Дзюба, Н. Й. Інформаційно-пошукові мови: проблеми теорії та практики / Н. Й. Дзюба // Бібліотекознавство. Документознавство. Інформологія. – 2005. – № 1. – С. 50–57.
14. ДСТУ 2392–94. Базові поняття. Терміни та визначення. – Вид. офіц. – К. : Держстандарт України, 1994. – 53 с. – (Інформація та документація).

15. ДСТУ 2394–94. Комплектування фонду, бібліографічний опис, аналіз документів. Терміни та визначення. – Вид. офіц. – К. : Держстандарт України, [1995]. – 88 с. – (Інформація та документація).
16. ДСТУ 2395–2000 (ГОСТ 30671–99) (ISO 5963:1985) Обстеження документа, встановлення його предмета та відбір термінів індексування. Загальна методика. – Вид. офіц. – К. : Держстандарт України, 2001. – 8 с. – (Інформація та документація).
17. ДСТУ 2398–94. Інформаційні мови. Терміни та визначення. – Вид. офіц. – К. : Держстандарт України, 1994. – 45 с. – (Інформація та документація).
18. Ерахторин, М. В. Индексирование ключевыми словами – основа информационного обслуживания читателей / Ерахторин М. В. // Б-ки учеб. заведений. – 2010. – № 36. – С.24–30 ; № 37. – С. 14–45. – (Справочный аппарат).
19. Зайцева, Е. М. Лингвистическое обеспечение автоматизированных библиотечно-информационных систем: современные требования и направления развития / Е. М. Зайцева // Науч. и техн. б-ки. – 2000. – № 3. – С.54–57.
20. Козлова, Л. М. Методика формирования ключевых слов: взгляд на проблему [Електронний ресурс] / Л. М. Козлова // Менеджмент библиотек учреждений образования : 4-я Междунар. конф. 27–29 мая 2003 г. – Витебск : ВГУ им. П. М. Машерова, 2003. – Режим доступа: <http://www.lib.vsu.by/ru/file/index>. – Назва з екрану.
21. Лобановська, І. Г. Лінгвістичне забезпечення електронного каталогу Державної науково-педагогічної бібліотеки України імені В. О. Сухомлинського: комплексний підхід / І. Г. Лобановська // Програмне забезпечення у сфері освіти і науки : матеріали конф., 12–13 трав. 2010 р. / М-во освіти і науки України, НАПН України ... Держ. наук.-пед. б-ка України ім. В. О. Сухомлинського [та ін.]. – К., 2010. – С. 63–65.
22. Методические рекомендации по координатному индексированию документов / Сост. Г. А. Аухадиева, И. К. Фомичева. – Казань, 1998. – 8 с.
23. Методические рекомендации по составлению ключевых слов Национальной библиотеки Республики Казахстан [Електронний ресурс] / Нац. б-ка Республики Казахстан. – Караганда. – Режим доступа: <http://www.library.karaganda.kz:8001>. – Назва з екрану.
24. Методические требования к выделению ключевых слов / К. З. Асатурова ; НИИ общей педагогики АПН СССР. – М., 1975.
25. Михайлов, А. И. Основы информатики / А. И. Михайлов, А. И Черный, Р. С. Гиляревский. – 2-е, перераб. и доп. изд. – М. : Наука, 1968. – 756 с.
26. Никольцева, Н. П. О возможности содержательного индексирования документов ключевыми словами (по материалам эксперимента, проведенного в РНБ на базе комплекса „Охрана окружающей среды”) [Електронний ресурс] / Н.П. Никольцева, О.А. Седышева ; Рос. Нац. б-ка. – Режим доступа: <http://www.nlr.ru/cat/predmet/article.htm>. – Назва з екрану.
27. Оранская, Л. И. Некоторые особенности использования дескрипторного поискового языка в библиографической информационно-поисковой системе универсальной библиотеки / Л. И. Оранская // Науч. и техн. б-ки. – 1997. – № 9. – С. 13–23.

28. Пименов, Е. Н. Нормативность и некоторые проблемы разработки тезаурусов и других лингвистических средств ИПС / Е. Н. Пименов // НТИ. Сер. 1. Организация и методика информ. работы. – 2000. – № 5. – С. 7–16.
29. Пименов, Е. Н. О тезаурусе замолвите слово : [Лингвистическое обеспечение] / Е. Пименов // Библиотека. – 2001. – № 7. – С. 42–45.
30. Пименов, Е. Н. О факторах, влияющих на индексирование: индексирование и предметная область / Е. Н. Пименов // НТИ. Сер. 1. Организация и методика информ. работы. – 2000. – № 2. – С. 15–23.
31. Правила составления и написания ключевых слов для электронного каталога НБ ТГУ [Электронный ресурс] / сост.: Г. А. Байтингер ; при участии О. А. Дубовицкой. – Томск, 2001. – Режим доступа: <http://www.lib.tsu.ru>. – Назва з екрану.
32. Скарук, Г. А. Тематический поиск в электронном каталоге: проблемы лингвистического обеспечения / Г. А. Скарук // Библиотековедение. – 2001. – № 3. – С. 48–57.
33. Скипор, И. Л. Координатное индексирование в информационно-библиотечных системах и сетях : Результаты исследования инструктивно-методической документации [Электронный ресурс] / Скипор И. Л., Сбитнева Е. А. // Деятельность б-ки и развитие информ. культуры общ-ва : материалы 10-й Междунар. юбилей. конф. „Библиотечное дело–2005” (Москва, 20–22 апр. 2005) / Мос. гос. ун-т культ. и иск. – М., 2005. – Режим доступа: <http://libconfs.narod.ru/2005>. – Назва з екрану.
34. Смагина, Н. Н. Лингвистическое обеспечение электронного каталога: Консультация специалиста / Смагина Н. Н. // Вузовские б-ки Алтай. края: Сборник. – Барнаул, 2003. – Вып.3. – С. 40–42.
35. Сукиасян, Э. Р. Координатное индексирование : выбор терминов индексирования и формирование поискового образа документа / Э. Р. Сукиасян // Библиотека. – 2005. – № 3. – С. 42.
36. Сукиасян, Э. Р. Школа индексирования : практ. пособие / Э. Р. Сукиасян. – М. : Либерия, 2005. – 143 с. - (Библиотекарь и время. XXI в. ; вып. 39).
37. Сукиасян, Э. Р. Электронный каталог и читатель / Э. Р. Сукиасян // Науч. и техн. б-ки. – 2000. – № 9. – С. 79–85.
38. Типовая технологическая инструкция по индексированию документов в узлах КСПБМ // Создание корпоративной сети публичных библиотек Москвы (КСПБМ): Раб. проект. Т. 1. Порядок формирования, хранения и использования машиночитаемых ресурсов общего пользования КСПБМ. – М., 2001. – С. 55–57.
39. Універсальна десяткова класифікація (УДК): у 2 кн. : [пер. з англ.] / [голов. ред. М. І. Сенченко; UDC Consortium; Кн. палата України]. – К. : Кн. палата України, 2000. – Кн. 1–2.
40. Ханжин, А. Г. Разработка методики координатного индексирования документов / А. Г. Ханжин // НТИ. Сер. 2. Информ. процессы и системы. – 1995. – № 8. – С. 22–26; № 9. – С. 14–19.
41. Шутько, Е. В. Аналитическая обработка текстовой информации и смысловой поиск в электронных базах данных: проблемы противоречий / Е. В. Шутько // Университетская библиотека: проблемы и решения: Пятая Всерос. науч.-практич. конф. – Екатеринбург, 2003. – С. 62–68.

42. Шутько, Е. В. О некоторых проблемах лингвистического обеспечения создания баз данных и поиска в электронном каталоге [Электронный ресурс] / Е. В. Шутько // Университетская библиотека на рубеже тысячелетий: выбор пути: Третья Всерос. науч.-практич. конф. – Екатеринбург, 2001. – Режим доступа: <http://consensus./conferences/unilib2000>. – Назва з екрану.

43. Щербинина, Г. С. Координатное индексирование в свете интеграционных процессов / Г. С. Щербинина // Науч. и техн. б-ки. – 2002. – № 10. – С. 72 – 76.

44. Щербинина, Г. С. Философия координатного индексирования / Г. С. Щербинина // Науч. и техн. б-ки. – 2000. – № 9. – С. 67–78.

45. Якушева, Ф. Л. Ключевые слова в электронной библиографической записи: проблемы и решения [Электронный ресурс] / Якушева Ф. Л. // Университетская библиотека на рубеже тысячелетий: выбор пути: Вторая Всероссийская науч.-практич. конф. 23–25 нояб. 1999 г. – Режим доступа: <http://lib/usu/ru>. – Назва з екрану.

ДЛЯ ПОТАТОК

ДЛЯ ПОДАТОК

НАУКОВО-ВИРОБНИЧЕ ВИДАННЯ

ІНДЕКСУВАННЯ ДОКУМЕНТІВ

КЛЮЧОВИМИ СЛОВАМИ

Методичні рекомендації

*Затверджено до друку вченою радою
Державної науково-педагогічної бібліотеки України імені В. О. Сухомлинського
(протокол № 3 від 15 березня 2011 р.)*

Укладач Лобановська Інна Георгіївна

Літературний редактор Білоцерківець Ірина Петрівна

Відповідальний за випуск Лобановська Інна Георгіївна

Оригінал-макет

Підписано до друку

Формат 60x84/8. Папір офсетний.

Гарнітура Times New Roman.

Ум. друк. арк.

Наклад 300 прим.

Зам. №