**Гуренко Євгенія Юріївна**,

ORCID <https://orcid.org/0000-0001-9697-6507>,

молодша наукова співробітниця,

відділ комплексного опрацювання документів,

Національна бібліотека України імені В. І. Вернадського,

Київ, Україна

е-mail: genadinike@gmail.com

**Проблеми організації тематичного пошуку в електронному каталозі Національної бібліотеки України імені В. І. Вернадського**

Розглянуто проблеми організації тематичного пошуку в електронних каталогах наукових бібліотек на прикладі інформаційно-пошукової системи НБУВ. Проаналізовано можливості та доцільність застосування різних пошукових мов для пошуку документів в ЕК НБУВ. Приділено увагу перспективам покращення пошукових алгоритмів з використанням семантичних технологій.

*Ключові слова*: електронний каталог, тематичний пошук, інформаційно-пошукові мови, пошукова система НБУВ

 У сучасній науковій бібліотеці електронний каталог (ЕК) став основним інструментом доступу до інформаційних ресурсів. Для пошуку документів в ЕК користувачі самостійно обирають інформаційно-пошукову мову (ІПМ) залежно від структури та функціональних можливостей електронної бібліотеки. Це можуть бути як класифікаційні мови, так і вербальні (предметизаційні, дескрипторні, мова ключових слів).

 В Національній бібліотеці України імені В. І. Вернадського (НБУВ) пошук здійснюється за класифікаційними індексами УДК і Рубрикатора НБУВ, а також природною мовою або за ключовими словами. Пошук за індексами УДК є найскладнішим для користувачів, які не мають достатньої підготовки. Запам’ятати велику кількість кодів неможливо; крім того, не всі терміни, що найчастіше зустрічаються у запитах, можна знайти в алфавитно-предметних покажчиках до таблиць УДК. Наприклад, поняття «урбоекологія», «біоекономіка» відсутні – індекси, що відповідають цим поняттям, у таблицях УДК не представлені. Скласти індекс для пошуку може бути непросто навіть фахівцю. Користувач може знайти конкретний документ або відповідну йому рубрику, якщо точно знає індекс УДК документа чи рубрики, якій цей індекс відповідає. Наприклад, Бібліотека Конгресу США для організації своєї великої колекції не використовує ані УДК, ані десяткову класифікацію Дьюї, а лише власну систему класифікації LCC. Втім, у Бібліотеці Конгресу можна знайти документи за індексами УДК, якщо вони були проіндексовані іншими бібліотеками чи організаціями та надійшли до її фондів. Таким чином, для міжнародного обміну інформацією можливість пошуку за індексами УДК залишається важливою.

 Практика показує, що більшість читацьких запитів мають тематичний характер, тобто користувачі найчастіше віддають перевагу пошуку за ключовими словами через його простоту і схожість із пошуком в Інтернеті. Для більш точного результату тематичний навігатор НБУВ, який є складовою пошукової системи бібліотеки, пропонує обрати галузь знань, а потім вводити ключові слова – тобто використовувати комбінацію ІПМ класифікаційного і вербального типу.

 Фахівці стверджують, що предметизаційна та дескрипторна ІПМ мають низку переваг над пошуковою мовою ненормованих ключових слів,  оскільки принцип координатного індексування та тематичний контроль дозволяють найповніше розкрити зміст документа і забезпечують найбільш точний результат читацького запиту. У пошуковій системі НБУВ пошук за предметними рубриками або за допомогою дескрипторів наразі не реалізовано, проте в процесі розробки тематичного навігатора ведеться робота над тим, щоб заголовок кожної рубрики найбільш повно розкривав ії зміст і був закінченою думкою. Наприклад, користувач хоче знайти інформацію щодо підприємництва в умовах війни в Україні. Ймовірно, він введе три ключові слова: «підприємництво» («бізнес»), «війна», «Україна» або словосполучення з цих слів – і швидко знайде рубрику: «Підприємництво (організація бізнесу) під час війни в Україні», навіть без зазначення галузі знань. Типовий синонім ключового слова вказаний в дужках у заголовку рубрики, що доповнює її зміст. Фактично, цей приклад демонструє застосування координатного індексування за допомогою слів і словосполучень природної мови, тобто ми маємо справу з дескрипторною ІПМ без фіксованого словника і без жорсткої парагматики.

 Організація тематичного пошуку в ЕК наукових бібліотек стикається з низкою лінгвістичних проблем, повʼязаних із багатозначністю слів, синонімією, а також необхідністю враховувати семантичні звʼязки між поняттями. Крім того, складність становлять великі обсяги даних, різноманітність форматів документів (текстові файли, PDF, мультимедійні ресурси тощо), нерідко застарілий інтерфейс, необхідність забезпечення точного пошуку, оптимізації алгоритмів для швидкої обробки великих обсягів даних.

 Багатозначність слів і синонімія стають проблемою у випадках, коли один і той самий термін має різне значення в різних контекстах, а різні терміни можуть позначати одне поняття. Наприклад, в біологічній  таксономії один і той самий біологічний таксон (вид, рід тощо) може мати кілька назв українською і латинською мовами, одна з яких є дійсною, а інші – синонімами. Це ускладнює пошук інформації, оскільки система може некоректно інтерпретувати запит або не знаходити релевантні документи, що містять синоніми. Для вирішення цієї проблеми назви-синоніми вводяться в поле «Допоміжні заголовки» тематичного навігатора та беруть участь у пошуку за ключовими словами.

 Розробка і застосування таблиць відповідностей між рубриками УДК та Рубрикатора НБУВ також позитивно вплинула на функціонування ЕК НБУВ. Звʼязки між рубриками двох класифікаційних систем реалізовані у тематичному навігаторі НБУВ, що підвищує точність і повноту результатів пошуку, забезпечуючи всебічне використання інформаційного потенціалу бібліотеки.

 Серед ключових напрямів покращення тематичного пошуку в ЕК НБУВ в перспективі можна зазначити:

1. Інтеграція семантичного пошуку – створення тематичних словників, охоплення тем яких визначається документним потоком НБУВ, використання онтологій та штучного інтелекту для автоматичного розпізнавання змісту документів. Це дозволить формувати індекси на основі змісту, а не лише формальних ознак.
2. Розширення таблиць відповідностей – охоплення нових галузей, зокрема військової справи, цифрової безпеки, соціальних наук, міждисциплінарних досліджень.
3. Удосконалення тематичного навігатора – деталізація рубрик, покращення інтерфейсу, додавання функцій фільтрації та комбінування індексів.
4. Створення авторитетного файлу тематичних рубрик – уніфікація термінології, що використовується у Рубрикаторі, ключових словах та описах документів, із включенням відповідностей до УДК. Такий ресурс дозволить забезпечити цілісність пошуку, уникнути дублювання термінів і створити основу для автоматизованої класифікації.
5. Розробка внутрішньої платформи для обміну класифікаційними рішеннями – бази кейсів, що охоплюють нові теми, приклади систематизації складних документів з коментарями фахівців. Це дозволить не лише зберігати методичний досвід, але й виробити спільні практики всередині бібліотекитауникнути розбіжностей при опрацюванні документів.

 У статті «Поняття предметних рубрик в американському бібліотекознавстві» Н. Стрішенець зазначає, що у минулому рубрика зʼявлялася у Списку предметних рубрик Бібліотеки Конгресу лише після того, як вона була присвоєна конкретному виданню, що надійшло до Бібліотеки. Список містив лише рубрики, які охоплювали коло видань, що потрапили до фондів, і використання його іншими бібліотеками могло бути обмеженим. У НБУВ склалася протилежна ситуація: останнім часом у Рубрикаторі зʼявилося багато рубрик, які не містять жодного документа. Йдеться не про сполучні рубрики, необхідні для цілісної структури Рубрикатора, а про специфічні, що не мають фактичного наповнення. Наприклад, рубрика розділу етнографії «Рогата худоба народів України», де під виразом «народи України» маються на увазі етнічні групи країни, навряд чи коли-небудь наповниться хоча б одним окремим документом. Надмірне створення рубрик без урахування фактичного документного потоку бібліотеки не сприяє покращенню інформаційного-пошукової системи  НБУВ. Це створює ілюзію повноти класифікації, але фактично ускладнює навігацію, збільшує кількість «порожніх» результатів і знижує ефективність пошуку.

 Таким чином, організація тематичного пошуку в ЕК НБУВ є динамічним процесом, що потребує постійної методичної підтримки, технологічного розвитку та аналізу класифікаційних практик. Розробка таблиць відповідностей, уніфікація термінології та впровадження пошукових алгоритмів з використанням семантичних технологій – це ключові інструменти створення ефективного знаннєвого інтерфейсу для українського наукового середовища.

**Yevheniia Hurenko**,

ORCID <https://orcid.org/0000-0001-9697-6507>,

Junior Researcher,

Department of integrated processing of documents,

V. I. Vernadskyi National Library of Ukraine,

Kyiv, Ukraine

е-mail: genadinike@gmail.com

**Problems of organizing thematic search in the electronic catalog of Vernadskyi National Library of Ukraine**

The article considers the problems of organizingsubject search in electronic catalogs of scientific libraries using the example of the information search system of V. I. Vernadskyi National Library of Ukraine (VNLU). It analyzes the possibilities and feasibility of using different search languages for document retrieval in VNLU. Attention is paid to the prospects for improving search algorithms through the use of semantic technologies.

*Keywords*: electronic catalog, subject search, query languages, VNLU search engine