

РЕЦЕНЗІЯ

офіційного рецензента, провідного наукового співробітника Інституту кібернетики ім. В.М. Глушкова НАН України, доктора технічних наук,
професора Кургаєва Олександра Пилиповича
на дисертаційну роботу Василенка Олега Олеговича
«Розроблення засобів опрацювання авторитетних записів для автоматизації зведеного бібліографічного каталогу», подану на здобуття наукового ступеня доктора філософії за спеціальністю 123 «Комп'ютерна інженерія»

Актуальність теми дослідження. Сучасне інформаційне суспільство вимагає оперативного й достовірного доступу до авторитетних бібліографічних даних, однак в Україні досі немає національного авторитетного файлу, а бібліотечні системи використовують різноманітні формати та протоколи. Запропоновані в дисертації методи інтеграції авторитетних записів через ILS-DI, OAI-PMH та REST-API прямо відповідають на потребу уніфікованого технологічного рішення для мережі Коха-бібліотек і, таким чином, є своєчасними та суспільно значущими.

Зв'язок роботи з науковими програмами та практичними завданнями. Дослідження проводилося в руслі «Стратегії розвитку бібліотечної справи на період до 2025 року “Якісні зміни бібліотек для забезпечення сталого розвитку України”», схваленої постановою Кабінету Міністрів України від 23 березня 2016 р. № 219-р.

Наукова обґрунтованість результатів. У роботі наведено формальні описи алгоритмів, схеми обробних процедур і результати експериментів на реальних даних українських бібліотек, ISNI та Wikidata. Висновки логічно випливають із отриманих даних.

Структура та зміст дисертації. Дисертація складається зі вступу, чотирьох розділів, висновків, додатків, 42 рисунків і 15 таблиць.

- Розділ 1 – аналітичний огляд форматів MARC, MODS, Dublin Core та протоколів Z39.50, SRU/SRW, OAI-PMH, ILS-DI, Koha API; обґрунтування вибору MARCXML + OAI-PMH/ILS-DI для гарвестингу.
- Розділ 2 – розроблення XSLT-конвертера MARC21 → UNIMARC, методи семантичного збагачення записів даними ISNI/Wikidata.
- Розділ 3 – методи очищення, нечіткого порівняння, GPU-прискореної кластеризації та об'єднання авторитетних записів.
- Розділ 4 – архітектура й функціонал прототипу зведеного каталогу «КОБЗА»; результати дослідної експлуатації.

Матеріал викладено логічно, структура відповідає вимогам МОН України.

Розглянувши звіт подібності за результатами перевірки дисертаційної роботи на текстові співпадіння, можна зробити висновок, що дисертаційна робота Василенка Олега Олеговича є результатом самостійних досліджень здобувача і не містить елементів фальсифікації, компіляції, фабрикації, плагіату та запозичень. Використані ідеї, результати і тексти інших авторів мають належні посилання на відповідне джерело

Практична значимість. Створено працездатний прототип, впроваджений в бібліотеці Українського гуманітарного інституту. Інструментарій конвертації та алгоритми верифікації авторитетних записів дозволяють істотно скоротити ручну працю при супроводі каталогів.

Наукова новизна:

1. Вперше запропоновано методику отримання авторитетних записів через протоколи ILS-DI, OAI-PMH та API, що усуває потребу у серверних протоколах Z39.50 та SRU/SRW і забезпечує корпоративну взаємодію бібліотек на платформі Koha.
2. Вперше розроблено XSL-трансформацію авторитетних записів про авторів з MARC21 у UNIMARC, яка не вводить додаткових полів, відсутніх у стандарті UNIMARC, є платформонезалежною й дозволяє швидко імпортувати записи в АБІС Koha.
3. Удосконалено методику верифікації авторитетних записів шляхом паралельного зіставлення з національними авторитетними базами, міжнародною базою імен ISNI та семантичною базою знань Wikidata, з урахуванням мовної форми заголовка та різної послідовності компонентів особового імені.
4. Вперше створено комплексний алгоритм нечіткого зв'язку (fuzzy matching) для кластеризації та об'єднання авторитетних записів за відсутності національного нормативного файлу; алгоритм враховує українську специфіку (омогліфи, ініціали) і поєднує дані бібліотек з інформацією міжнародних ресурсів.

Повнота апробації результатів. Матеріали дисертації оприлюднено у трьох статтях, надрукованих у фахових виданнях України (категорії «Б»), а також ще в чотирьох інших наукових публікаціях. Окремі положення роботи доповідалися на шести міжнародних наукових конференціях. Таким чином, кількісні та якісні показники публікацій повністю відповідають вимогам МОН України, установленим для здобувачів ступеня доктора філософії.

Зауваження та рекомендації.

1. Сторінки, що передують змісту не містять пагінації.
2. В назвах підрозділів 3.5 і 3.6 дублюється слово «об'єднання».
3. Доцільно підписати осі на Рисунку 2.1, що містить результати тестування ефективності специфікації конвертування MARC21-UNIMARC: X – кількість підполів, Y – номер поля.
4. До переліку умовних позначень має сенс додати LRM та BIBFRAME, що зустрічаються в 2 розділі.
5. Відкриту семантичну базу знань Wikidata слід використовувати обережно, бо як видно з рисунку 3.7, альтернативні назви можуть містити не ті імена, під якими автори публікувались, а ті, якими їх називали інші. В таких випадках варто проводити модерацію за участі спеціаліста-людини.
6. Оскільки дані віртуального міжнародного авторитетного файлу VIAF є одним з джерел даних для міжнародного стандартного ідентифікатора імені ISNI, збереження окремо ідентифікатора VIAF в авторитетних файлах може бути надлишковим.

Наведені зауваження не знижують наукової та практичної цінності роботи.

Висновок

Дисертаційна робота Василенка О. О. є завершеним науковим дослідженням, що містить вперше одержані результати, важливі для розвитку автоматизованих систем авторитетного контролю та формування зведених каталогів. За актуальністю, науковою новизною, теоретичною і практичною значущістю вона відповідає вимогам «Порядку присудження наукових ступенів» (постанова КМУ № 167 від 24.03.2022 р.).

Вважаю, що здобувач заслуговує присудження ступеня доктора філософії за спеціальністю 123 «Комп'ютерна інженерія».

Провідний науковий співробітник
Інституту кібернетики
ім. В.М. Глушкова НАН України
доктора технічних наук, професор



 Кургаєв О.П.