

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
НАЦІОНАЛЬНА БІБЛІОТЕКА УКРАЇНИ ІМЕНІ В. І. ВЕРНАДСЬКОГО

МАЦЮК ГАЛИНА РОСТИСЛАВІВНА

УДК 025.4.06:303.436.2:004/001.4

**ТЕЗАУРУС ЯК ІНСТРУМЕНТ ІНФОРМАЦІЙНОГО
ЗАБЕЗПЕЧЕННЯ МІЖДИСЦИПЛІНАРНИХ НАУКОВИХ ДОСЛІДЖЕНЬ**

27.00.03 – книгознавство, бібліотекознавство, бібліографознавство

27 – соціальні комунікації

АВТОРЕФЕРАТ

дисертації на здобуття наукового ступеня кандидата наук із соціальних
комунікацій

Київ – 2020

Дисертацією є кваліфікаційна наукова праця на правах рукопису.

Робота виконана в Тернопільському національному технічному університеті імені Івана Пулюя, Міністерство освіти й науки України

Науковий керівник: доктор наук із соціальних комунікацій, професор
Кунанець Наталія Едуардівна,
Національний університет «Львівська політехніка»,
професор кафедри інформаційних систем та мереж

Офіційні опоненти: доктор наук із соціальних комунікацій, доцент
Збанацька Оксана Миколаївна,
Національна академія керівних кадрів культури і
мистецтв,
професор кафедри культурології та інформаційних
комунікацій

кандидат наук із соціальних комунікацій
Струнгар Артур Валерійович,
Національна бібліотека України
імені В. І. Вернадського,
старший науковий співробітник відділу
політологічного аналізу

Захист відбудеться «18» лютого 2021 р. о 12.00 год. на засіданні спеціалізованої вченої ради Д 26.165.01 із захисту дисертацій на здобуття наукового ступеня доктора (кандидата) наук у Національній бібліотеці України імені В. І. Вернадського за адресою 03039, Київ, Голосіївський просп., 3.

З дисертацією можна ознайомитися в читальному залі бібліотекознавчої літератури Національній бібліотеці України імені В. І. Вернадського за адресою 03039, Київ, Голосіївський просп., 3.

Автореферат розіслано «15» січня 2021 р.

Учений секретар спеціалізованої вченої ради,
кандидат наук із соціальних комунікацій

Л. А. Литвинова

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Обмін інформацією, отримання релевантних даних є важливими аспектами функціонування наукових комунікацій. Потреба якісного пошуку документів та швидкого інформаційного обміну ресурсами генерується в роботі бібліотек. Тому сучасним напрямом діяльності бібліотечних установ є організація ефективних процесів опрацювання та пошуку інформації, що сприяє покращенню процедур виконання інформаційних запитів користувачів. З огляду на це виникає необхідність уніфікувати пошуковий образ документа й пошуковий образ запиту користувачів з метою їхнього подальшого пошуку в автоматизованих бібліотечних інформаційних системах. Цьому слугують штучні інформаційно-пошукові мови, призначені для передачі змісту та формальних ознак документів, а також формування інформаційних запитів. Вагомими серед них є інформаційно-пошукові тезауруси, використання яких сприяє швидкому та результативному пошуку релевантних документів.

Активне впровадження в роботу бібліотек автоматизованих технологій ставить високі вимоги до лінгвістичного забезпечення автоматизованих бібліотечних інформаційних систем, тому воно потребує постійного вдосконалення.

Важливим у розробленні лінгвістичного забезпечення для створення уніфікованого пошукового образу та пошукового запиту документа є спеціально створені тезауруси. Сьогодні існує значна кількість тезаурусів, які використовуються для оброблення та пошуку інформації в базах даних. Серед таких ресурсів загальномовні – WordNet, GermaNet, DanNet, MultiWordNet тощо; спеціалізовані (представляють терміносистему певної галузі науки) – SNOMED CT, MeSH, EUROVOC, AGROVOC тощо. Українські інформаційні ресурси представлені інформаційно-пошуковим тезаурусом Національної бібліотеки України імені Ярослава Мудрого, галузевим тезаурусом лінгвістичного забезпечення електронної бібліотеки ХДУХТ, тезаурусами з комп'ютерної лексикиграфії та лінгвістичних термінів Інституту філології Київського національного університету імені Тараса Шевченка, а також вузькоспеціалізованими галузевими тезаурусами з нанотехнології, Державної прикордонної служби України, будівельної термінології, авіації, надтвердих матеріалів тощо.

Дослідженню проблем тезауруса присвячені роботи таких вітчизняних і зарубіжних науковців як Т. Вилегжаніна, Н. Гендіна, А. Гладун, В. Дубичинський, О. Збанацька, Ю. Караулов, І. Кудашев, Є. Купріянов, К. Лобузін, Н. Лукашевич, Л. Пшенична, Р. Робертс (R. Roberts), Ю. Рогушина, Н. Соколова, В. Шульгіна та ін.

Але, незважаючи на це, слід констатувати, що дослідження проблеми розроблення інформаційно-пошукового тезауруса предметної області «Розумне місто» не проводилися.

З огляду на той факт, що наукові дослідження в інноваційній міждисциплінарній галузі «Розумне місто» (оскільки до досліджень залучені фахівці з інформаційних технологій, соціальних комунікацій та телекомунікацій, економіки та електронного урядування) набирають інтенсиву, для інформаційного забезпечення її дослідників виникає потреба формування єдиної понятійної бази та

відповідної спеціальної термінології, щоб створювати пошукові образи документів, що висвітлюють результати наукових розвідок. Зазначене завдання лежить у площині лінгвістичного забезпечення бібліотечних інформаційно-пошукових систем, узгодження українських термінів з іншомовними при формуванні пошукових образів та пошукових запитів. Такий стан речей генерує потребу розроблення двомовного інформаційно-пошукового тезауруса «розумного міста», який відповідав би вимогам сучасної бібліотечної спільноти щодо формування систематизованої терміносистеми як інструмента створення пошукових образів документів та пошукових запитів користувачів на рівні двох мов.

Отже, актуальність дисертаційної роботи полягає в представленні технології побудови інформаційно-пошукового тезауруса як інструмента інформаційного забезпечення міждисциплінарних наукових досліджень на основі формалізованого подання термінології предметної області «Розумне місто» для підвищення ефективності формування пошукових образів документів та пошукових образів запитів користувачів бібліотек.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконана в Тернопільському національному технічному університеті імені Івана Пулюя в рамках науково-дослідних тем: «Класи соціокомунікаційних технологій в проектах «Розумне місто» (2017-2019 рр., державний реєстраційний №0117U002240), «Класи інформаційних технологій в проектах «Розумне місто» (2017-2019 рр., державний реєстраційний №0117U002241)) та науково-дослідної теми Національного університету «Львівська політехніка» «Науково-освітнє соціокомунікаційне середовище великого міста: моделювання, прототипування, інформаційні технології» (2016-2019 рр., № 0116U006723).

Мета та завдання дослідження. Мета – розроблення концептуальних засад формування двомовного інформаційно-пошукового тезауруса як формалізованого представлення знань предметної області «Розумне місто» та інструмента ефективного пошуку релевантної інформації в автоматизованих бібліотечних інформаційних системах для інформаційного забезпечення наукових досліджень, що проводяться на міждисциплінарних засадах.

Для досягнення поставленої мети передбачено вирішення таких **завдань**:

- виявити та проаналізувати джерельну базу дослідження щодо формування пошукових образів та запитів на документи, що містять результати міждисциплінарних досліджень;

- обґрунтувати вибір моделі представлення знань предметної області «Розумне місто» з метою підвищення ефективності інформаційного пошуку в автоматизованих бібліотечних інформаційних системах для забезпечення наукових досліджень;

- простежити історію розвитку технологій формування тезаурусів та проаналізувати роль тематичного інформаційно-пошукового тезауруса в процесі систематизації термінології певної галузі та можливості його застосування, як засобу підвищення ефективності пошуку в автоматизованих бібліотечних інформаційно-пошукових системах;

- удосконалити в сучасних бібліотеках технології побудови інформаційно-пошукового тезауруса, що описує категоріально-понятійний апарат певної предметної області та взаємозв'язки між концептами;

- розробити двомовний інформаційно-пошуковий тезаурус «розумного міста» для формування уніфікованих пошукових образів документів та пошукових запитів користувачів з метою їхнього подальшого пошуку в автоматизованих бібліотечно-інформаційних системах.

Об'єкт дослідження – процеси формування пошукових образів документів та пошукових запитів користувачів у бібліотеках, створених на основі двомовного подання термінів предметної області «Розумне місто».

Предмет дослідження – структуроване подання знань предметної області «Розумне місто» у формі двомовного інформаційно-пошукового тезауруса для проведення ефективної процедури пошуку в бібліотеках.

Методи дослідження. Для досягнення мети та розв'язання завдань, поставлених у дисертаційній роботі, використано методи загальнонаукові, інформатики, математичні, соціокомунікаційний та історико-хронологічний.

У процесі дослідження були поєднані загальнонаукові методи та прийоми дослідження: джерелознавчого пошуку й аналізу профільних публікацій за темою дослідження для визначення стану розробленості наукової проблеми; дедуктивний та індуктивний методи для побудови тезауруса на основі переходу від ширших термінів до вузких та навпаки; метод зіставлення для виявлення відмінних рис мовної форми вираження та змістового наповнення термінів, їхніх структурних компонентів та елементів в українській та англійській мовах; термінологічний аналіз для уточнення поняттєвого апарату дослідження та термінів галузі «розумне місто» з метою їх унормування та упорядкування.

Методи інформатики (метод інформаційного підходу, консолідації інформації, метод онтології, тезаурусний підхід) було використано для виявлення особливостей та розроблення інформаційно-пошукового тезауруса; математичний інструмент системного підходу до складних проблем прийняття рішень – метод аналізу ієрархій – для вибору способу представлення знань предметної області «Розумне місто».

Застосований соціокомунікаційний метод було використано для обґрунтування сутності тезаурусів, які реалізуються в цілеспрямованих комунікаційних зв'язках між суб'єктами та об'єктами системи, суб'єктами системи та зовнішнім середовищем. Історично-хронологічний метод дозволив дослідити виникнення та розвиток тезаурусів у хронологічній послідовності.

Наукова новизна дослідження полягає в тому, що дисертація є одним з першим досліджень, в якому проведена робота над створенням формалізованого представлення знань інноваційної предметної області «Розумне місто» у формі двомовного інформаційно-пошукового тезауруса, що відображено у таких положеннях:

уперше:

- застосовано метод аналізу ієрархії для вибору способу представлення знань предметної області «Розумне місто» при формуванні інформаційних ресурсів бібліотек;

- на основі аналізу наукових публікацій *консолідовано* підходи до формування двомовного тезауруса, який сприяє ефективному формуванню метаданих інформаційних ресурсів бібліотек, та сформовано алгоритм його розроблення;

- *розроблено* двомовний інформаційно-пошуковий тезаурус «розумного міста» на основі формалізованого представлення знань предметної області, що забезпечує інформаційний пошук релевантних документів у бібліотеках, сприяє підвищенню ефективності інформаційного забезпечення наукових досліджень, слугує самостійним термінологічним довідником, який допомагає орієнтуватися в даній галузі;

удосконалено:

- визначення поняття «інформаційно-пошуковий тезаурус предметної області» для використання в бібліотечних інформаційних системах;

доповнено:

- класифікацію типів семантичних відношень з урахуванням міждисциплінарних тенденцій сучасного наукового знання (виділено ряд проблемно-специфічних відношень характерних для предметної області «Розумне місто»), що сприяло підвищенню інформаційного пошуку в бібліотеках;

набули подальшого розвитку:

- технологія поповнення бібліотечними працівниками тезауруса «розумного міста»: наповнення новими лексичними одиницями для двомовного відображення найновіших термінів галузі, відслідковування міжпредметних зв'язків;

- методи розроблення електронної версії двомовного тезауруса предметної області «Розумне місто» для бібліотечних інформаційних систем.

Практичне значення дослідження. Практичне значення дисертаційної роботи полягає в розробленні двомовного інформаційно-пошукового тезауруса предметної області «Розумне місто» як інструмента пошуку релевантних документів для інформаційного забезпечення міждисциплінарних наукових досліджень.

Сьогодні для опису будь-якої предметної області та створення інформаційних систем у бібліотеках чимале значення має розроблення нових алгоритмів та методик формування тезаурусів на основі онтологій. Для адекватного відображення моделей подання знань, здатних до взаємодії, у роботі запропонована методика, яка використовує проблемно-специфічні відношення для реалізації пошукових запитів та дозволяє формувати якісні знання про досліджувану галузь, що сприяє підвищенню ефективності інформаційного забезпечення.

Результати дослідження впроваджені та використовуються в роботі науково-технічної бібліотеки Тернопільського національного технічного університету імені Івана Пулюя, Державного архіву Тернопільської області, а також у навчальному процесі кафедри комп'ютерних наук Тернопільського національного технічного університету імені Івана Пулюя. Водночас результати дослідження можна використовувати в дослідницькій, навчальній та практично-фаховій діяльності у різних суміжних галузях наук. Вони можуть слугувати науково-методологічним базисом для розширення та поглиблення навчальних курсів із інформаційного

забезпечення наукових досліджень та інформаційного менеджменту. Їхнє активне впровадження сприятиме розвитку бібліотечних та інформаційних установ, де аналіз документів здійснюється з використанням штучних мов. У бібліотеках та архівах тезаурус можна використовувати як інструмент формування уніфікованих пошукових образів та запитів документів для пошуку релевантних документів, формування авторитетних файлів предметних рубрик, якісного індексування документів для забезпечення адекватності пошуку в автоматизованих бібліотечно-інформаційних системах.

Впровадження матеріалів дослідження підтверджено відповідними актами.

Особистий внесок здобувача. Дисертаційне дослідження виконано самостійно, викладені наукові результати, висновки й рекомендації належать дисертанту. Внесок здобувача в працях, які виконані в співавторстві, конкретизовано в списку публікацій.

Апробація результатів дисертації. Основні теоретичні положення, висновки й результати дисертаційного дослідження було викладено в опублікованих матеріалах та збірниках праць, обговорено та апробовано на міжнародних наукових та науково-практичних конференціях, серед яких: VI Міжнародна науково-практична конференція «Математика. Інформаційні технології. Освіта» (Луцьк-Світязь, 5-7 червня 2017 р.); XV Международная научно-практическая конференция «Современные информационные технологии в экономике и управлении предприятиями, программами и проектами» (Харьков, 29 сентября 2017 г.); V Науково-технічна конференція «Інформаційні моделі, системи та технології» (Тернопіль, 1-2 лютого 2018 р.); VII Міжнародна науково-практична конференція «Математика. Інформаційні технології. Освіта» (Луцьк-Світязь, 3-5 червня 2018 р.); 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT) (Lviv, September 11-14, 2018); Міжнародна наукова конференція «Бібліотека. Наука. Комунікація: 100-річчя Національної бібліотеки України імені В. І. Вернадського» (Київ, 6-8 листопада 2018 р.); VI Науково-технічна конференція «Інформаційні моделі, системи та технології» (Тернопіль, 12-13 грудня 2018 р.); The 3rd International Conference on Computational Linguistics and Intelligent Systems (COLINS-2019) (Kharkiv, April 18-19, 2019); The VIII International Scientific Conference «Mathematics. Information Technology. Education», MoMLeT&DS'2019 (Shatsk, June 2-4, 2019); 2019 IEEE 14th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT) (Lviv, September 17-20, 2019); X Львівський міжнародний бібліотечний форум «Нова бібліотека – відповідальність кожного» (Львів, 18-21 вересня 2019 р.); Міжнародна наукова конференція «Бібліотека. Наука. Комунікація: актуальні тенденції у цифрову епоху» (Київ, 8-10 жовтня 2019 р.); VII Науково-технічна конференція «Інформаційні моделі, системи та технології» (Тернопіль, 12-13 грудня 2019 р.).

Публікації. Основні положення та результати дисертаційного дослідження викладено у 23 наукових публікаціях, серед яких 4 статті у вітчизняних наукових фахових журналах, 5 статей у наукових періодичних виданнях інших держав, 13 публікацій у збірниках матеріалів міжнародних конференцій; 1 стаття, що додатково

відображає наукові результати дослідження. П'ять із вказаних наукових праць опубліковано в журналах та збірниках, що індексуються в міжнародних наукометричних базах: Scopus, Web of Science.

Структура та обсяг дисертації. Дисертація складається з титульного аркуша, анотації українською та англійською мовами, списку публікацій здобувача за темою дисертації, змісту, переліку умовних скорочень, вступу, 4 розділів, висновків у кінці розділів та загальних висновків, списку використаних джерел та додатків. Загальний обсяг дисертації – 292 сторінки: основна частина – 212 сторінок (8,9 авт. арк.) у тому числі 22 рисунки та 6 таблиць, список використаних джерел – 253 позиції на 27 сторінках, додатки – 26 сторінок.

ОСНОВНИЙ ЗМІСТ ДИСЕРТАЦІЇ

У **вступі** обґрунтовано актуальність теми дисертаційної роботи; окреслено мету, завдання, предмет та об'єкт дослідження; наведено дані про зв'язок роботи з науковими програмами, планами, темами; охарактеризовано методи дослідження; визначено наукову новизну, практичне значення отриманих результатів; наведено відомості про апробації результатів дослідження та публікації автора.

У першому розділі **«Інформаційно-пошуковий тезаурус як об'єкт бібліотекознавчих досліджень»**, який складається з трьох підрозділів, на підставі аналізу інформаційних ресурсів за тематикою дисертаційної роботи визначено її теоретичну базу та обґрунтовано методологічні підходи й методи дослідження.

В умовах розвитку сучасного інформаційного суспільства працівникам бібліотек для того, щоб орієнтуватися в накопичених інформаційних ресурсах і потужному потоці інформації, наявної в глобальній мережі «Інтернет», здійснювати ефективний пошук даних, релевантних запитам користувачів, організувати їхнє ефективне інформаційне забезпечення, необхідні допоміжні засоби, якими є спеціальні методи опрацювання документів, серед яких вагома роль відводиться інформаційно-пошуковим тезаурусам.

У розділі проаналізовано особливості інформаційно-пошукових тезаурусів, які використовуються у закордонних та вітчизняних бібліотеках, як інструментів ефективного формування метаданих, створення пошукових образів та пошукових запитів. Досліджено основні види, змістовне наповнення та кількісні показники цього різновиду інформаційних ресурсів бібліотек.

Теоретичну базу дослідження становили інформаційні ресурси: 1) база даних англійських лексичних одиниць – комп'ютерний тезаурус WordNet (створений у Принстонському університеті, США), а також ряд тезаурусів аналогічної тематики іншими мовами (GermaNet, DanNet, MultiWordNet тощо), в основу розроблення яких лягла концепція WordNet-у, що свідчить про ефективність інформаційних ресурсів такої структури; 2) комп'ютерні тезауруси, що використовуються для індексування, каталогізації документів і пошуку інформації в галузевих інформаційних ресурсах бібліотек з проблем біології, медицини, охорони здоров'я та суміжних наук – SNOMED CT, MeSH, Web-IPBIC; освіти, культури, природничих, соціальних та гуманітарних наук, комунікації та інформації – багатомовний тезаурус ЮНЕСКО;

галузей діяльності інститутів Європейського Союзу – тезаурус EUROVOC; сільського господарства – тезаурус AGROVOC та інші.

Певний внесок у цю галузь зроблено й в Україні. Розроблено тезаурус з комп'ютерної лексикографії, тезаурус лінгвістичних термінів (Інститут філології Київського національного університету імені Тараса Шевченка), інформаційно-пошуковий тезаурус Національної бібліотеки України імені Ярослава Мудрого, галузевий тезаурус лінгвістичного забезпечення електронної бібліотеки Харківського державного університету харчування та торгівлі та інші.

Проведений аналіз наукових праць з питання створення галузевих комп'ютерних тезаурусів засвідчує наявність вагомих напрацювань із розроблення вузькоспеціалізованих галузевих тезаурусів з проблем нанотехнології, будівельної, авіаційної галузей, для прикордонної служби України тощо.

Дослідження інформаційно-пошукових тезаурусів, принципів їхньої побудови й функціонування в різних галузях дає підстави зробити висновок, що специфіка кожної предметної області знаходить відображення в структурі її тезауруса та підходах до його формування.

Розглянуто особливості міждисциплінарних досліджень і необхідність їхніх відображення у тезаурусі. Встановлено, що суть міждисциплінарного підходу у сфері наукових досліджень – це збагачення суміжних наук на основі запозичення методологічного інструментарію, об'єднання зусиль для пояснення нових понять та з'ясування тенденцій суспільного розвитку. Таким чином, міждисциплінарність – це синергія різних наук, що передбачає розвиток інтеграційних процесів, зростаючу взаємодію методів, інструментарію задля отримання нового наукового знання (йдеться про утворення нових міждисциплінарних об'єктів, що формують інноваційну галузь «Розумне місто»). Тому міждисциплінарне дослідження предметної області «Розумне місто» зводиться до виявлення нових відношень між поняттями вихідних дисциплін, установлення нової системи законів, що їх пов'язують, та синтезі прагматики вирішення нових завдань.

Такими чином, розроблення інформаційних тезаурусів для ефективного формування пошукових образів документів у бібліотеках через постійне зростання обсягів різногалузевої інформації набуває все більшої актуальності.

Представлено широкий спектр наукових методів та підходів, використаних в дисертаційному дослідженні; обґрунтовано доцільність застосування кожного з них. Глибина вивчення наявного стану дослідження проблеми та широкий спектр використаних методів, принципів та підходів забезпечили об'єктивність, наукову обґрунтованість, системність дисертаційного дослідження.

У другому розділі **«Вибір моделі представлення знань предметної області «Розумне місто» для підвищення ефективності інформаційного пошуку в бібліотеках»**, який складається з трьох підрозділів, проаналізовано інструменти та засоби вирішення завдання побудови інформаційно-пошукового тезауруса для підвищення ефективності процедур формування метаданих та пошуку в автоматизованих бібліотечних інформаційних системах з метою налагодження ефективної системи інформаційного забезпечення міждисциплінарних наукових досліджень.

Відзначено, що основу будь-якої предметної області складають сукупність знань та даних про об'єкти, процеси та явища, які підлягають проектуванню та зберіганню в базах даних. Моделлю представлення знань предметної області є її формальний опис. Оскільки кожній предметній області притаманні поняття й зв'язки між ними, термінологія, закони, які пов'язують між собою об'єкти даної предметної області, процеси й події, тому для вирішення таких завдань необхідно використовувати моделі, які містять основні характеристики певної предметної області.

Для вибору оптимальної моделі представлення знань предметної області використано метод аналізу ієрархій. Для вибору найкращого способу представлення знань була задана множина альтернатив для досягнення поставленої мети: логічні моделі, продукційні моделі, фреймові моделі, семантичні мережі та онтології. Проведений аналіз дозволив аргументовано визначити онтологічний підхід як найкращу модель для формування пошукових образів у інформаційних системах різної складності, у тому числі й бібліотечних, та подати його у формі тезауруса, як окремого випадку онтологічного подання знань.

Основне завдання онтологічного підходу полягає в тому, щоб полегшити процедуру пошуку інформації через систематизацію знань, створення єдиної ієрархії понять, уніфікацію термінів та правил їхньої інтерпретації. Такий підхід формалізує процеси інформаційного забезпечення наукових досліджень, що проводяться на міждисциплінарних засадах.

Продемонстровано алгоритм розроблення онтологічної моделі двомовного подання знань на прикладі предметної області «Розумне місто», яка ґрунтується на формалізованому описі об'єктів цієї галузі англійською та українською мовами. Запропонована модель сприяє чіткому визначенню термінів предметної області, уніфікації процедур їх тлумачення та перекладу. Сукупність термінів предметної області утворює її термінологічну систему, на основі якої відбувається формування тезауруса, що продемонстровано на прикладі предметної області «Розумне місто».

У третьому розділі «**Тезаурус як засіб підвищення ефективності сучасних інформаційно-пошукових систем бібліотек**», який включає сім підрозділів, досліджується тезаурус у аспекті історії його формування. Історія виникнення тезаурусів сягає коренями праць мислителів стародавніх часів, серед яких «Ономастикон» Юлія Поллукса (II ст. н. е.), «Амаракоша» Амара Сіна (кінець II – початок III ст. н. е.). Відзначено, що поняття «тезаурус» як науковий термін було вперше використано у XIII ст. флорентійським ученим Брунетто Латіні в заголовку його праці – систематизованої енциклопедії «Книга про скарб» («Livres dou trésor»), що повністю збігалось із семантикою слова «thesauros», тобто «скарб», «багатство». Сучасний етап історії ідеографічних праць відкривається роботою П. М. Роже «Тезаурус англійських слів і виразів» (1852 р.), що вважається першою науково обґрунтованою спробою логічного впорядкування лексики, його схема класифікації була адаптована для різних мов. Відомими ресурсами є «Аналогічний словник французької мови» П. Буассьєра (1862 р.), автор якого відмовився від створення логіко-філософської схеми класифікації понятійного змісту лексики та скористався алфавітно-змістовим розташування слів, «Німецький словниковий склад за

предметними групами» Ф. Дорнзайфа (1959 р.), «Ідеологічний словник іспанської мови» Х. Касареса (1959 р.). Акцентовано увагу на тому, що перші тезауруси були пов'язані з фундаментальними проблемами пізнання, відображеннями картини світу та закономірностями його сприйняття й не враховували особливостей інформаційної діяльності. Як інструмент інформаційного пошуку тезаурус почав розглядатися у ХХ ст.

Дисертаційна праця містить аналіз наукової дискусії, пов'язаної з поданням терміна «тезаурус» у різних дослідженнях, та обґрунтування трактування терміна «інформаційно-пошуковий тезаурус предметної області» як сукупності структурованих знань, поданих у вигляді набору дескрипторів та аскрипторів, які пов'язані між собою певними відношеннями, що відображають семантичні зв'язки між поняттями певної предметної області. Наявність зв'язків між термінами створює передумови для використання тезаурусів у бібліотеках у процесах індексування документів та інформаційного пошуку.

Розглянуто структуру інформаційно-пошукового тезауруса як засобу систематизації наукових понять певної предметної області, яка включає множину термінів та множину відношень між ними, і створюється для підвищення якості інформаційного пошуку в інформаційних ресурсах бібліотек. Головна особливість інформаційно-пошукового тезауруса полягає в тому, що він подає не лише компіляцію термінологічних статей, а й систему дескрипторів та аскрипторів, уніфікованих мовних виразів, які подаються засобами природної мови, відображають предмет документа й використовуються для отримання контрольованого доступу до його змісту. Дескриптор відповідає групі ключових слів природної мови, відібраних з тексту документа досліджуваної предметної області. Кожне поняття виражається лише одним дескриптором і кожен дескриптор виражає лише одне поняття або групу споріднених понять. Класифіковані за змістом дескриптори та аскриптори утворюють інформаційно-пошуковий тезаурус певної предметної області.

Характерною особливістю тезауруса є те, що статті в ньому побудовані таким чином, що відображають семантичні відтінки близьких за значенням слів. Якщо традиційний тлумачний словник дає тільки визначення слова та приклади його використання в реченнях, то статті тезауруса розкривають значення слова як через визначення, так і через зіставлення з іншими словами й понятійними групами, що сприяє суттєвому спрощенню пошукових процедур у інформаційних ресурсах бібліотек та підвищенню рівня релевантності інформаційного забезпечення наукових досліджень, що відбуваються на міждисциплінарних засадах.

Відзначено, що алгоритм розроблення тезауруса передбачає дотримання низки правил щодо представлення понять у описовій формі, подано формальні та семантичні вимоги до формування дефініції термінів інформаційно-пошукового тезауруса в бібліотеках.

Продемонстровано основні процедури та принципи розроблення тезаурусів на прикладі побудови двомовного українсько-англійського інформаційно-пошукового тезауруса «розумного міста». Запропонована концепція розроблення двомовного

тезауруса ґрунтується на консолідованому алгоритмі, що враховує найвагоміші результати досліджень вітчизняних та зарубіжних науковців.

Таким чином, вважаємо, що розроблення двомовного інформаційно-пошукового тезауруса потребує реалізації таких етапів: 1) дослідження потреб користувача, визначення меж предметної області, підбір джерельної бази, формування вимог до визначення термінів тезауруса та його структури; 2) формування реєстрової частини: виділення лексикографічних одиниць зі словників, наукових статей, матеріалів конференцій, стандартів тощо і їх переклад відповідно українською чи англійською мовами; складання списку ключових слів; побудова класів умовної еквівалентності (для усунення синонімії, полісемії і переходу до дескрипторної інформаційно-пошукової мови, тобто приведення лексичних одиниць до заданої стандартної форми); 3) побудова двомовних словникових статей (формулювання дефініцій термінів); розроблення набору семантичних зв'язків та їх актуалізація (установлення відношень між термінами); упорядкування тезауруса відповідно до розробленої структури; 4) уточнення та редагування отриманого матеріалу, узгодження з користувачем.

Розглянуто особливості отримання інформації для побудови інформаційно-пошукових тезаурусів для використання в бібліотеках, а саме лінгвістичні та гносеологічні аспекти спілкування з експертом, що сприяють отриманню інформації за допомогою комунікативних та текстологічних методів. Основним при цьому є розуміння змісту документа, що дозволяє здійснити компресію тексту у вигляді ключових слів, які передають зміст тексту, та слугує зручною методичною основою для проведення текстологічних процедур отримання інформації.

Розглянуто тезаурус як засіб лінгвістичного забезпечення бібліотечних та архівних інформаційних систем. Саме лінгвістичні засоби є інтерфейсом між природною мовою та інформаційно-пошуковою системою. Від лінгвістичних засобів тематичного пошуку залежить ефективність використання інформаційних ресурсів бібліотек та архівів у контексті інформаційного супроводу науково-дослідної роботи користувачів. Створення інформаційних ресурсів бібліотек генерує потребу використання коректного лінгвістичного забезпечення для автоматизованих бібліотечних інформаційних систем, що сприятиме ефективному формуванню метаданих для пошукового образу документа. Тезауруси слугують одним із засобів лінгвістичного забезпечення в процесах індексування, формування предметних рубрик для електронних каталогів бібліотек. Саме від їхньої якості залежить ефективність інформаційного пошуку, а в подальшому – і релевантність інформаційного забезпечення міждисциплінарних наукових досліджень.

У четвертому розділі **«Двомовний тезаурус предметної області «Розумне місто» як інструмент вирішення інформаційно-пошукових задач у бібліотеках»**, який складається з п'яти підрозділів, подано хід роботи над розробленням двомовного тезауруса на прикладі міждисциплінарної предметної області «Розумне місто».

Для формування терміносистеми предметної області «Розумне місто» було використано автентичні фахові тексти, перевага яких обумовлена їхньою особливістю – наявністю термінологічної лексики, що розкриває наукові поняття

досліджуваної галузі та представляє їх однозначне пояснення. Не менш вагомим джерелом формування термінологічного фонду досліджуваної галузі є електронні каталоги наукових бібліотек та реферативні бази даних. Проаналізовані джерела дали змогу виявити терміни, які знаходяться в науковому обігу та встановити зв'язки між ними. Такий підхід сприяє ефективній процедурі формування тезауруса, оскільки перевага таких джерел – це надійність, високий рівень загальноприйнятої інформації, зрозуміла мова подання, доступність (можливість отримати вільний доступ до даних джерел), висока насиченість термінологічною лексикою.

Відбір термінів здійснювався на основі таких критеріїв як інформаційність, відповідність темі, актуальність, практична значущість. Серед термінів тезауруса є загальнонаукові, міжгалузеві та вузькоспеціальні терміни, бо досліджувана галузь розвивається в тісному взаємозв'язку з іншими сферами, таким чином, носить міждисциплінарний характер.

Доведено, що специфіка структури інформаційно-пошукового тезауруса міждисциплінарної предметної області «Розумне місто» визначається набором наукових напрямів, які її формують. Саме для галузі «Розумне місто» системоформуючим науковим напрямом є «цифрові технології» («інформаційні технології», «комунікаційні технології»). Виходячи з цього, ми вважаємо, що терміни цього напрямку є ключовими, а терміни, які використовуються в інших наукових дисциплінах (економіка, медицина, освіта, охорона навколишнього середовища, управління та ін.) – допоміжними. Міждисциплінарне дослідження предметної області «Розумне місто» та формалізоване представлення знань даної галузі у формі двомовного інформаційно-пошукового тезауруса для бібліотек полягає у виявленні понять, якими оперують у вихідних дисциплінах, встановленні нових відношень між цими поняттями та нової системи законів, що їх пов'язують. Міждисциплінарний обмін знаннями в тезаурусі «розумного міста» представляється як проблема інтеграційних відношень на таких рівнях: логічному (рід-вид), системно-структурному (частина-ціле), еквівалентності (синоніми), асоціативному.

Сформовано структурні класифікації термінів предметної області «Розумне місто». У ході аналізу термінологічного матеріалу було з'ясовано, що термінологія цієї галузі складається з лексичних одиниць двох типів: термінів-слів та термінів-словосполучень, останні, у свою чергу, поділяються на двокомпонентні та багатокомпонентні. У терміносистемі «розумного міста» переважають термінологічні словосполучення, які виражають більш складні поняття в порівнянні з термінами-словами.

Двомовний інформаційно-пошуковий тезаурус «розумного міста» включає 649 словникових статей (з них 560 дескрипторів і 89 аскриптори).

Робоча структура тезауруса «розумного міста» формується із низки тематичних комплексів: загальні поняття, інформаційні системи «розумного міста», управління ресурсами, технології, інфраструктури, мобільність, енергетика, телекомунікації, транспорт, освіта, охорона здоров'я, навколишнє середовище та ін. Таке тезаурусне міждисциплінарне моделювання знання дає можливість науковцям виходити за межі термінологічної системи однієї наукової галузі знань у порівнянні

з іншою. Таким чином відбувається обмін інформацією між науковцями різних галузей досліджень.

При побудові інформаційно-пошукового тезауруса розкрито типи відношень між поняттями, які виражені термінами. Для цього була сформована класифікація відношень між термінами «розумного міста», яка включає два типи: традиційні (ієрархічні («рід-вид», «частина-ціле», «екземпляр-клас») й еквівалентні) та проблемно-специфічні (асоціативні) відношення («виробляє», «використовується (в)», «визначає», потребує, «є джерелом утворення», «зберігає», «є напрямком», «оцінює» та ін.).

Подано структуру двомовного інформаційно-пошукового тезауруса «розумного міста», який складається зі вступу, алфавітного лексико-семантичного покажчика термінів та пермутаційного покажчика. В основній частині тезауруса, алфавітному лексико-семантичному покажчику, представлено дескриптори й аскриптори зі словниковими статтями. Він використовується для пошуку відповідного дескриптора та його парадигматичних зв'язків, забезпечуючи можливість адекватного формулювання запиту, а при незадовільних результатах пошуку – зміни стратегії пошуку, зокрема використовуючи дескриптори, що входять у словникову статтю. У розділі визначено за яких умов термін отримує статус дескриптора чи аскриптора. Подано структуру словникової статті дескриптора та аскриптора, відношення, які позначаються відповідними посиланнями. Допоміжним покажчиком до алфавітного лексико-семантичного покажчика в тезаурусі «розумного міста» є пермутаційний покажчик, який містить дескриптори й аскриптори без словникових статей. Його особливістю є формування словникових гнізд, у які включені всі значущі слова, що входять до складу термінів тезауруса. Таким чином тезаурус «розумного міста» забезпечує такі можливості пошуку термінів при формулюванні запиту: а) пошук словникових статей за заголовними дескрипторами; б) пошук дескрипторів і аскрипторів у словниковому гнізді.

Доведено, що інформаційно-пошуковий тезаурус «розумного міста» виступає інструментом інформаційного забезпечення міждисциплінарних наукових досліджень. Він забезпечує однакове семантичне навантаження певного терміна для різних галузей, сприяючи передусім однозначному розумінню його в колі науковців цих сфер науки. Одночасно тезаурусне представлення знань галузі «Розумне місто» дозволяє працівникам бібліотеки сформувати однозначне трактування певного терміна при створенні пошукового образу документа.

Окреслено основні завдання та функції двомовного інформаційно-пошукового тезауруса «розумного міста».

Представлено алгоритм інформаційного пошуку релевантної інформації за допомогою тезауруса «розумного міста».

ВИСНОВКИ

У процесі роботи над дисертаційним дослідженням досягнуто його мету, виконано всі поставлені завдання та отримано такі основні науково-практичні результати:

1. Проведений аналіз наукових джерел за темою дослідження щодо формування пошукових образів та запитів документів, що містять результати міждисциплінарних досліджень показав, що незважаючи на значну кількість досліджень щодо побудови й функціонування інформаційно-пошукових тезаурусів та наявність значної кількості робіт, пов'язаних з предметною областю «Розумне місто», відсутній тезаурус, який нормалізував би лексику досліджуваної галузі й був би інструментом для індексації та пошуку документів у автоматизованих бібліотечно-інформаційних системах.

2. Обґрунтовано вибір моделі представлення знань предметної області «Розумне місто» для підвищення ефективності інформаційного пошуку в бібліотеках за допомогою використання методу аналізу ієрархій. Для парного порівняння було запропоновано такі альтернативи: логічні моделі, продукційні моделі, фреймові моделі, семантичні мережі та онтології.

Застосування методу аналізу ієрархій дозволило на основі науково-обґрунтованих підстав обрати найзручніший та найінформативніший спосіб подання концептів – онтологічний підхід як найкращий інструмент для формування інформаційно-пошукового тезауруса, що сприяє ефективній процедурі пошукових образів документів у автоматизованих бібліотечних інформаційних системах.

3. Простежено розвиток тезаурусів у процесі історії їхнього розвитку. Акцентовано увагу на тому, що історія виникнення тезаурусів ґрунтується на працях мислителів стародавніх часів, їхня побудова пов'язана з фундаментальними проблемами пізнання, відображення картини світу та закономірностями їхнього сприйняття. У інформаційному аспекті тезаурус почав розглядатися в ХХ ст.

Проаналізовано роль тематичних інформаційно-пошукових тезаурусів у процесах систематизації термінології предметних областей та можливість їхнього застосування як засобу підвищення ефективності пошуку в сучасних автоматизованих бібліотечних інформаційних системах. Використання інформаційно-пошукових тезаурусів сприяє підвищенню ефективності змістового опрацювання документів, зокрема при розробленні уніфікованого підходу до формування пошукового образу документа й пошукового образу запиту користувачів бібліотек для реалізації процедур пошуку в автоматизованих бібліотечних інформаційних системах для інформаційного забезпечення міждисциплінарних наукових досліджень. Використовуючи тезаурус, бібліотеки можуть формувати авторитетні файли, предметні рубрики або визначати ключові слова, індексувати документи й запити засобами дескрипторної мови, замінюючи ключові слова відповідними дескрипторами. При якісному індексуванні документів використання уніфікованої інформаційно-пошукової мови в бібліотеках та архівах дозволяє досягти високої точності пошуку та оперативно реагувати на появу нових термінів. Інформаційно-пошуковий тезаурус сприяє уніфікації довідкового апарату бібліотек, оскільки чітко фіксує місце терміна в понятійній системі певної предметної області, відображає його структурні зв'язки.

4. Запропоновано засадничий підхід до формування двомовного тезауруса предметної області, на прикладі предметної області «Розумне місто», який є одним із важливих практичних інструментів інформаційної діяльності бібліотек.

Сформовано алгоритм його розроблення на основі консолідації принципів побудови тезаурусів, який складається з чотирьох етапів: 1) визначення меж предметної області, пошук джерел та накопичення галузевої інформації; 2) формування реєстрової частини тезауруса; 3) побудова словникових статей; 4) уточнення та редагування отриманого тезауруса. Запропонований підхід сприяє вирішенню актуальних завдань у досліджуваній галузі в контексті інформаційного забезпечення міждисциплінарних наукових досліджень, автоматизації процесів розроблення двомовних пошукових образів; створення алгоритмів формування пошукового запиту у вузькоспеціальній предметній області.

5. Побудова тезауруса передбачала реалізацію низки послідовних кроків. Було виокремлено терміни з досліджуваних джерел на основі критеріїв: інформативність, частотність, відповідність темі, актуальність, практична значущість, що описують поняття предметної області. З'ясовано, що концепти предметної області «Розумне місто» складається з двох типів: термінів-слів (однокомпонентних) та термінів-словосполучень (багатокомпонентних). Переважають термінологічні словосполучення, які виражають більш складні поняття. Система термінів охоплює поняття, які належать до таких категорій: загальні поняття, інформаційні системи «розумного міста», управління ресурсами, технології, інфраструктури, мобільність, енергетика, телекомунікації, транспорт, фінанси й економіка, місцеві та центральні органи влади, спільноти, освіта, охорона здоров'я, системи безпеки, навколишнє середовище.

Простежено систему відношень між термінами «розумного міста». Класифікація зв'язків містить два типи: традиційні (ієрархічні й еквівалентні) та проблемно-специфічні (неієрархічні або асоціативні).

Розроблено двомовний інформаційно-пошуковий тезаурус «розумного міста», який максимально наповнений «робочими» термінами, що відповідають існуючому документному потокові, і складається зі вступу, алфавітного лексико-семантичного покажчика термінів та пермутаційного покажчика. В алфавітному лексико-семантичному покажчику термінів представлено всі дескриптори та аскриптори зі словниковими статтями. Він використовується для пошуку дескриптора та його парадигматичних зв'язків і забезпечує можливість адекватного формулювання запитів. Допоміжним покажчиком до алфавітного лексико-семантичного є пермутаційний покажчик термінів, який містить дескриптори та аскриптори без словникових статей і використовується для пошуку за окремими словами, що входять до складу термінів-словосполучень, у відповідних словникових гніздах.

Інформаційно-пошуковий тезаурус «розумного міста» дозволяє формувати уніфіковані пошукові образи документів та запитів з метою подальшого пошуку в автоматизованих бібліотечно-інформаційних системах; індексувати документи; формувати авторитетні файли, які сприяють забезпеченню якості опрацювання документів; уніфікувати введення інформації, що дозволяє отримати одноманітність вводу та скорочення чи уникнення помилок; здійснювати релевантний пошук у електронних каталогах, базах даних бібліотек; організувати обмін авторитетними записами між бібліотеками; заощаджувати час на опрацювання документів.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Статті в наукових фахових виданнях України:

1. Мацюк Г. Представлення знань у вузькоспеціальних предметних областях за допомогою тезауруса / Г. Мацюк, Н. Кунанець // Бібліотечний вісник. – 2019. – № 1. – С. 19–23.

Розкрито роль галузевого інформаційно-пошукового тезауруса в процесах систематизації термінології предметної області.

2. Мацюк Г. Алгоритм формування двомовного тезауруса інформаційного масиву предметної області «Розумне місто» / Н. Кунанець, Г. Мацюк // Бібліотекознавство. Документознавство. Інформологія. – 2019. – № 1. – С. 110–117.

Проаналізовано технологічні засади створення двомовного тезауруса предметної області «Розумне місто».

3. Мацюк Г. Інформаційно-пошукові тезауруси: світовий та вітчизняний досвід формування / Г. Мацюк // Бібліотекознавство. Документознавство. Інформологія. – 2019. – № 2. – С. 106–115.

4. Мацюк Г. Р. Удосконалення технологій інформаційного пошуку на основі тезауруса предметної області / Г. Р. Мацюк // Бібліотечний вісник. – 2019. – № 5. – С. 20–25.

Статті в наукових періодичних виданнях інших держав:

5. Мацюк Г. Р. Підготовка масиву термінів для побудови двомовного тезауруса предметної області «Розумне місто» / Г. Р. Мацюк // Social and Human Sciences. PolishUkrainian scientific journal. – 2018. – № 3 (19). – URL : https://sp-sciences.io.ua/s2634980/matsiuk_halyna_2018_.stages_of_preparation_of_terms_array_for_creation_of_bilingual_thesaurus_of_the_smart_city_data_domain_social_and_human_sciences_polish-ukrainian_scientific_journal_03_19.

6. Matsuik H. Selective Dissemination of Information – Technology of Information Support of Scientific Research / A. Rzhеuskyi, H. Matsuik, N. Veretennikova, R. Vaskiv // Advances in Intelligent Systems and Computing III. – 2018. – Vol. 871. – P. 235–245. (Індексовано в Scopus).

Подано технологію вибіркового поширення інформації.

7. Matsuik H. Formation of Search Queries Based on Thesaurus of Narrowly Specialized Subject Areas / H. Matsuik, A. Rzhеuskyi, N. Kunanets, N. Veretennikova // Econtechmod. An international quarterly journal. – 2019. – Vol. 08. – No. 3. – P. 11 – 14.

Проаналізовано роль галузевого інформаційно-пошукового тезауруса для підвищення ефективності інформаційного забезпечення наукових досліджень, що здійснюється з використанням технологій вибіркового розповсюдження інформації.

8. Halyna Matsuik Open online training courses for engineering purpose / Ihor Baran, Nataliya Kunanets, Halyna Matsuik, Mykola Mytnyk, Khristina Shunevich, Yuriy Skorenky and Volodymur Jaskilka // Workshop Proceedings of the VIII International Scientific Conference «Mathematics. Information Technology. Education». MoMLeT&DS'2019 (Ukraine, Shatsk, June 2-4, 2019). – Shatsk, 2019. – Vol. 2386. – Pp. 331-339. – ISSN 1613-0073. (Індексовано в Scopus).

Запропоновано використовувати тезаурус при створенні онлайн-курсів дистанційного навчання інженерного спрямування.

9. Matsuik H. Use of the Smart City Ontology for Relevant Information Retrieval / Nataliya Kunanets, Halyna Matsuik // Proceedings of the 3rd International Conference on Computational Linguistics and Intelligent Systems (COLINS-2019) (Ukraine, Kharkiv, April 18-19, 2019). – Kharkiv, 2019. – Pp. 322–333. – ISSN 1613-0073. (*Індексовано в Scopus*).

Подано алгоритм побудови онтології предметної області «Розумне місто».

Наукові праці, які засвідчують апробацію матеріалів дисертації:

10. Мацюк Г. Р. Онтологічне дослідження предметної області «Розумне місто» / Н. Е. Кунанець, В. В. Пасічник, Г. Р. Мацюк, О. В. Мацюк, О. П. Левченко // Математика. Інформаційні технології. Освіта : VI Міжнар. наук.-практ. конф. (Україна, Луцьк-Світязь, 5-7 червня 2017 р.). – Луцьк-Світязь, 2017. – С. 48–52.

Подано онтологічне дослідження предметної області «Розумне місто як основи для побудови бази знань про «інтелектуальне» місто, запропоновано класифікаційне дерево онтології «Розумне місто» на етапі представлення першого рівня.

11. Мацюк Г. Р. Двомовний тезаурус як соціокомунікаційна складова проекту «Розумне місто» / Г. Р. Мацюк, Н. Е. Кунанець // Современные информационные технологии в экономике и управлении предприятиями, программами и проектами: материалы XV Международной научно-практической конференции (Украина, Харьков, 29 сентября 2017 г.) – Харьков : Национальный аэрокосмический университет им. Н. Е. Жуковского «Харьковский авиационный институт», 2017. – С. 85–87.

Обґрунтовано актуальність створення комплексного двомовного тезауруса як соціокомунікаційної складової проекту «Розумне місто».

12. Мацюк Г. Р. Вербальне моделювання предметної області засобами побудови тезауруса / Н. Е. Кунанець, Г. Р. Мацюк // Матеріали V науково-технічної конференції «Інформаційні моделі, системи та технології» Тернопільського національного технічного університету імені Івана Пулюя (Україна, Тернопіль, 1-2 лютого 2018 р.). – Тернопіль, 2018. – С. 36.

Подано особливості вербального моделювання предметної області засобами побудови тезауруса, консолідовано підходи до технологій формування тезауруса.

13. Мацюк Г. Метадані інформаційних ресурсів у контексті інтеграції різнотипних даних у соціокомунікаційних інформаційних системах / Г. Липак, О. Липак, Н. Кунанець, Г. Мацюк // Матеріали V Міжнародної науково-технічної конференції «Інформаційні моделі, системи та технології» Тернопільського національного технічного університету імені Івана Пулюя (Україна, Тернопіль, 1–2 лютого 2018 р.). – Тернопіль, 2018. – С. 38.

Здійснено порівняльний аналіз стандартів опису метаданих інформаційних ресурсів різних за профілем установ.

14. Мацюк Г. Р. Застосування методу аналізу ієрархій для вибору моделі представлення предметної області «Розумне місто» / Н. Е. Кунанець,

А. В. Ржеуський, Г. Р. Мацюк // Математика. Інформаційні технології. Освіта : тези доп. VII Міжнар. наук.-практ. конф. (Україна, Луцьк – Світязь, 3-5 червня 2018 р.). – Луцьк – Світязь, 2018. – С. 68 – 70.

Розглянуто вибір моделі представлення знань предметної області «Розумне місто» у рамках створення інформаційно-пошукового тезауруса.

15. Matsiuk Halyna Application of Saaty Method while Choosing Thesaurus View Model of the “Smart city” Subject Domain for the Improvement of Information Retrieval Efficiency / Kunanets Nataliia, Matsiuk Halyna // 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT) (Ukraine, Lviv, September 11-14, 2018). – Lviv, 2018. – P. 21–25. (Індексовано в Scopus, Web of Science).

Подано методу вибору моделі представлення концептів для побудови тезауруса предметної області «Розумне місто» методом Сааті.

16. Мацюк Г. Р. Формування тезауруса для виконання інформаційного пошуку у вузькоспеціальних предметних областях / Н. Е. Кунанець, Г. Р. Мацюк // Бібліотека. Наука. Комунікація: 100-річчя Національної бібліотеки України імені В. І. Вернадського: матеріали міжнар. наук. конф. (Україна, Київ, 6-8 листопада 2018 р.) / Нац. б-ка України ім. В. І. Вернадського. – Київ, 2018. – С. 366–369.

Розкрито поняття тезауруса як інструмента інформаційного пошуку ресурсів у автоматизованих інформаційно-бібліотечних системах для створення пошукових образів для вузькоспеціальних предметних областей.

17. Мацюк Г. Стандарти, що регламентують формат подання інформаційно-пошукового тезауруса / Г. Мацюк, Н. Кунанець // Інформаційні моделі, системи та технології: матеріали VI науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя (Україна, Тернопіль, 12-13 грудня 2018 р.). – Тернопіль, 2018. – С. 39.

Проаналізовано міжнародні та вітчизняні стандарти, що регламентують побудову інформаційно-пошукових тезаурусів.

18. Мацюк Г. Побудова онтології як інтероперабельної моделі подання знань у процесі наукового дослідження предметної області «Розумне місто» / Г. Мацюк, Н. Кунанець // Інформаційні моделі, системи та технології: матеріали VI науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя (Україна, Тернопіль, 12-13 грудня 2018 р.). – Тернопіль, 2018. – С. 40.

Обґрунтовано обрання онтологічного підходу для методологічного забезпечення застосування системного підходу до аналізу та вивчення предметної області.

19. Halyna Matsiuk Application of the “Smart City” data domain thesaurus as the tool for representing knowledge while improving the problem-oriented Web search effectiveness / Halyna Matsiuk, Nataliya Kunanets // 2019 IEEE 14th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT) (Ukraine, September 17-20, 2019). – Lviv, 2019. – P. 31–34. (Індексовано в Scopus, Web of Science).

Розкрито використання тезауруса предметної області «Розумне місто» як інструмента термінологічного контролю в процесі аналізу, опрацювання документів та пошуку інформації.

20. Мацюк Г. Р. Тезаурус як засіб семантизації понять предметної області / Г. Р. Мацюк, Н. Е. Кунанець // Бібліотека. Наука. Комунікація: актуальні тенденції у цифрову епоху: матеріали міжнар. наук. конф. (Україна, Київ, 8-10 жовтня 2019 р.) / Нац. б-ка України ім. В. І. Вернадського. – Київ, 2019. – С. 94–98.

Розкрито поняття тезауруса як засобу семантизації понять предметної області.

21. Мацюк Г. Р. Лексикографічна систематизація галузевої термінології за допомогою тезауруса / Н. Е. Кунанець, Г. Р. Мацюк // X Львівський міжнародний бібліотечний форум «Нова бібліотека – відповідальність кожного» : збірник матеріалів. – Київ : УБА, 2019. – С. 64–66.

Розкрито роль галузевого інформаційно-пошукового тезауруса як лексикографічного систематизатора термінології певної галузі.

22. Мацюк Г. Р. Лінгвістичний аспект спілкування з експертом при формуванні пошукового тезауруса / Г. Р. Мацюк, Н. Е. Кунанець // Інформаційні моделі, системи та технології: матеріали VII науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя (Україна, Тернопіль, 11-12 грудня 2019 р.). – Тернопіль, 2019. – С. 70.

Розглянуто особливості лінгвістичного аспекту спілкування з експертом під час отримання інформації для побудови інформаційно-пошукового тезауруса для використання в бібліотеці.

Публікації, що додатково відображають наукові результати дисертації:

23. Мацюк Г. Р. Тезаурус предметної області «Розумне місто» / Н. Е. Кунанець, Г. Р. Мацюк // Вісник Національного університету «Львівська політехніка» «Інформаційні системи та мережі». – 2017. – № 872. – С. 179–189.

Окреслено особливості формування тезауруса предметної області «Розумне місто», подано приклад побудови фрагменту онтології з предметної області «Розумне місто».

АНОТАЦІЯ

Мацюк Г. Р. Тезаурус як інструмент інформаційного забезпечення міждисциплінарних наукових досліджень. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня кандидата наук із соціальних комунікацій за спеціальністю 27.00.03 – книгознавство, бібліотекознавство, бібліографознавство. – Тернопільський національний технічний університет імені Івана Пулюя; Національна бібліотека України імені В. І. Вернадського. Київ, 2020.

У дисертації здійснено огляд основних видів та змістове наповнення інформаційно-пошукових тезаурусів, які використовуються в зарубіжних та вітчизняних бібліотеках, як інформаційний ресурс структури певної предметної області та інструмент формування пошукових образів документів і пошукових запитів у бібліотеках.

Застосовано метод аналізу ієрархій для вибору моделі представлення знань предметної області «Розумне місто» при формуванні інформаційних ресурсів бібліотек.

Розглянуто еволюцію тезаурусів у процесі історії їхнього розвитку. Проаналізовано роль тематичних інформаційно-пошукових тезаурусів у процесах систематизації термінології предметних областей та можливість їхнього застосування як засобу підвищення ефективності пошуку в сучасних автоматизованих бібліотечних інформаційних системах.

На основі аналізу наукових публікацій консолідовано підходи до формування двомовного інформаційно-пошукового тезауруса, який сприяє ефективному формуванню метаданих інформаційних ресурсів бібліотек, та сформовано алгоритм його побудови.

Розроблено двомовний інформаційно-пошуковий тезаурус «розумного міста», який можна використовувати для формування уніфікованого пошукового образу документа й пошукового образу запиту користувачів бібліотеки, індексування документів, здійснення релевантного пошуку в електронних каталогах, базах даних бібліотек.

Ключові слова: інформаційно-пошуковий тезаурус, дескриптор, аскриптор, термін, онтологія, предметна область, інформаційно-пошукова система, інформаційний пошук, пошуковий образ, бібліотеки, автоматизована бібліотечна інформаційна система.

АННОТАЦІЯ

Мацюк Г. Р. Тезаурус как инструмент информационного обеспечения междисциплинарных научных исследований. – Квалификационный научный труд на правах рукописи.

Диссертация на соискание ученой степени кандидата наук по социальным коммуникациям по специальности 27.00.03 - книговедение, библиотекведение, библиографоведение. – Тернопольский национальный технический университет имени Ивана Пулюя; Национальная библиотека Украины имени В. И. Вернадского. Киев, 2020.

В диссертации сделан обзор основных видов и содержательное наполнение информационно-поисковых тезаурусов, которые используются в зарубежных и отечественных библиотеках, как информационный ресурс структуры определенной предметной области и инструмент формирования поисковых образов документов и запросов в библиотеках.

Применен метод анализа иерархий для выбора модели представления знаний предметной области «Умный город» при формировании информационных ресурсов библиотек. Рассмотрена эволюция тезаурусов в аспекте истории их развития. Проанализирована роль отраслевых информационно-поисковых тезаурусов в процессах систематизации терминологии предметных областей и возможность их применения в качестве средства повышения эффективности поиска в современных автоматизированных информационных библиотечных системах.

На основе анализа научных публикаций консолидировано подходы к формированию двуязычного информационно-поискового тезауруса, который способствует эффективному формированию метаданных информационных ресурсов библиотек, и сформирован алгоритм его разработки.

Разработан двуязычный информационно-поисковый тезаурус «умного города», который можно использовать для формирования унифицированного поискового образа документа и поискового образа запроса пользователей библиотек, индексирования документов, осуществления релевантного поиска в электронных каталогах, базах данных библиотек.

Ключевые слова: информационно-поисковый тезаурус, дескриптор, аскриптор, термин, онтология, предметная область, информационно-поисковая система, информационный поиск, поисковый образ, библиотеки, автоматизированная библиотечная информационная система.

SUMMARY

Matsiuk H. R. Thesaurus as the instrument for information support of interdisciplinary scientific research. – Qualification scientific work on the rights of manuscripts.

Thesis for a Candidate Degree in Social Communications (PhD) by specialty 27.00.03 – book science, library science, bibliography. – Ternopil Ivan Puluj National Technical University, V. I. Vernadsky National Library of Ukraine. Kyiv, 2020.

The overview of the main types and content of information retrieval thesauruses used in foreign and national libraries, as the information resource of the particular domain-specific structure and the tool for search document images and library retrieval requests formation is carried out in this thesis.

The hierarchy analysis method is used in order to select the “Smart City” domain knowledge presentation model while developing the libraries information resources. This makes it possible to select the most convenient and informative way of presenting the concepts of the given domain for search images formation in automated information library systems on the scientifically proved basis, and presenting it in the form of thesaurus using ontological approach.

The thesauri evolution in their history development aspect is considered.

The definition “domain information retrieval thesaurus” is clarified as the set of structured knowledge, presented in the form of descriptors set interrelated by certain relationships reflecting semantic connections between the particular domain concepts. The thesaurus makes it possible to present the concept meaning not only by defining it, but by correlating it to other concepts or groups of concepts as well. This correlation is usually determined by establishing paradigmatic connections between concepts. The availability of such links creates the preconditions for the thesauruses use in document indexing and information retrieval in library systems.

The role of specific purposes information retrieval thesauruses in the processes of domain terminology systematization and the possibility of their use as the means of increasing search efficiency in modern automated information library systems are analyzed.

On the basis of the scientific publications analysis, approaches to the formation bilingual information retrieval thesaurus development are consolidated contributing to the efficient formation of libraries information resources metadata, and the algorithm for its development is constructed.

The “Smart City” bilingual information retrieval thesaurus consisting of three parts: introduction, alphabetic lexical-semantic index of terms and permuted index is developed. The alphabetic lexical semantic index of terms listing all descriptors and non-descriptors with lexical entry is used in order to search the descriptor and its paradigmatic relationships and provide high-quality request formulation. The permutation term index containing descriptors and non-descriptors without lexical entry is used to search single keywords for the descriptors in appropriate group of words with the same root presented in one dictionary entry.

The system of relations between the thesaurus terms containing two types of relations: traditional - hierarchical and equivalent and problem-specific (non-hierarchical or associative) is developed. Traditional hierarchical relationships show the logical descriptors subordination and are established if the relationship is true regardless of context. The equivalence relation is established between descriptor and non-descriptor, i. e., between the terms-synonymous, if replacing one of them with another does not change the content of the text relevant to the information search. Associative links use makes it possible to generate additional descriptors useful for indexing or searching.

The “Smart City” bilingual information retrieval thesaurus used to form the unified document search image and users request search image; document indexing; performing the relevant search in electronic catalogs and libraries databases is developed.

Key words: information retrieval thesaurus, descriptor, non-descriptor, term, ontology, domain, information retrieval system, information request, information search, search image, libraries, automated library information system.

Підписано до друку 12.01.2021. Формат 60×90, 1/16.
Друк лазерний. Папір офсетний. Гарнітура TimesNewRoman.
Умовно-друк. арк. 0,9. Наклад – 100 прим.
Замовлення № 12012021

Друк ФОП Паляниця В. А.
Свідоцтво ДК №4870 від 20.03.2015 р.
м. Тернопіль, вул. Б. Хмельницького, 9а, оф.38.
тел. (0352) 528-777.