

<https://doi.org/10.15407/bv2019.05.020>

УДК 81.3

Галина Мацюк,

викладач, Тернопільський національний технічний університет ім. Івана Пулюя (м. Тернопіль, Україна)

вул. Руська, 56, м. Тернопіль, 46001, Україна

E-mail: galuna.matsiuk@gmail.com

Удосконалення технологій інформаційного пошуку на основі тезауруса предметної області

У статті аналізуються шляхи удосконалення технологій інформаційного пошуку за допомогою тезауруса предметної області. Запропоновано алгоритм процедури інформаційного пошуку, розкривається технологія формування тезауруса предметної області «Розумне місто». Наголошується, що при формуванні пошукового образу і пошукового запиту повинна використовуватися інформаційно-пошукова мова, яка ґрунтується на спільних термінологічних засадах.

К л ю ч о в і с л о в а: тезаурус, систематизація знань, предметна область «Розумне місто», інформаційний пошук, термінологічний контроль.

Іntenсивний розвиток сучасних інформаційних технологій сприяє формуванню величезних потоків інформації, змінює підходи до проведення пошукових операцій у них, до опрацювання отриманої інформації, вимагає постійного вивчення інформаційних потреб користувачів. Використання інтелектуальних інформаційних систем для реалізації пошукових алгоритмів сприяло підвищенню їх ефективності. Основні напрями розвитку інформаційних технологій пов'язані зі створенням інформаційних систем, заснованих на знаннях відповідних предметних областей.

Аналіз останніх досліджень і публікацій. Створення та застосування можливостей складних інформаційних систем, орієнтованих на знання, розглядаються у дослідженнях цілої низки науковців. Автоматизоване створення тезауруса предметної області на основі синтактико-семантичного аналізу природно-мовних текстів подається у роботах В. Ю. Величко [2]. Формуванням тематично спрямованих пошукових запитів до розподілених інформаційних ресурсів з використанням тезаурусних моделей опікуються О. В. Палагін, В. Ю. Величко, О. Є. Стрижак [12]. У роботах зі створення тезаурусів з різних навчальних дисциплін Т. Б. Андрусенко, О. Є. Стрижака запропонована «тезаурусно-мережева модель предметної області», яка визначається як «модель, побудована на основі опису об'єктів предметної області (ПрО) і тезауруса цієї області», а також включає термінологічну лексико-семантичну систему, програмні засоби реалізації різних режимів роботи, сервісні можливості, що надаються користувачеві. Інфор-

мація про об'єкт ПрО представляється схемою: «поняття – визначення – семантичні відношення даного поняття з іншими»[1]. Методика розроблення термінології інформаційних ресурсів як бази формування тезаурусів подана у працях А. Я. Гладуна, Ю. В. Рогушиної [3]. Засади застосування деяких елементів формування тезауруса у правовій інформатиці, опис окремих підходів, моделей, алгоритмів проаналізовані Д. Ланде [9]. Питання терміносистеми й відбору термінів при укладанні словників досліджують В. А. Широков [14; 15], Б. І. Шуневич [16], І. М. Кульчицький [8]. Водночас, поза увагою дослідників залишаються технології формування тезаурусів у вузько-спеціальних предметних областях.

Мета статті – проаналізувати шляхи удосконалення технології інформаційного пошуку за допомогою розроблення тезауруса предметної області.

Інформаційний пошук: сучасні підходи

Сучасні підходи до проведення інформаційного пошуку ґрунтуються на вивченні інформаційних потреб, на формуванні зручних, ефективних інформаційних запитів. Інформаційний запит відображає інформаційну потребу, сформульовану природною мовою. Перекодування інформаційного запиту інформаційно-пошуковою мовою дає змогу сформувати пошуковий образ запиту. Удосконалені пошукові технології на основі пошукового образу запиту забезпечують вивчення документів, що містять релевантну наукову інформацію. В умовах постійного зростання кількості документів набагато ефективнішим є підхід, котрий ґрунтується на поєднанні двох методів:

1) аналізу контенту документів на основі пошукового образу запиту; 2) порівняння пошукового образу запиту з пошуковим образом документа. Взаємодія цих методів пошуку забезпечує зростання ефективності пошукових процесів, проте тут постає проблема формування коректних пошукових запитів.

Процедуру інформаційного пошуку можна подати у вигляді такого алгоритму:

Крок перший: формулювання інформаційної потреби і запиту.

Крок другий: окреслення інформаційних масивів, у яких здійснюватиметься пошук.

Крок третій: співставлення пошукового образу запиту з пошуковими образами документа та вмістом документа.

Крок четвертий: видобування інформації, що відповідає пошуковому образу запиту.

Крок п'ятий: формування відповіді користувачу.

Такий підхід ефективно спрацьовує, якщо при формуванні пошукового образу і пошукового запиту використовується інформаційно-пошукова мова, що ґрунтується на використанні спільних термінологічних засад. Безперечно, синтаксис і семантика інформаційно-пошукової мови визначаються структурою і змістом документів. Завдання щодо формування пошукового образу документа і пошукового образу запиту спрощується при чіткому представленні знань про конкретну предметну область, з використанням системи, придатної для використання інформаційних технологій. Тезауруси предметних областей і є саме такою системою подання знань.

Формування тезауруса

Термін «тезаурус», що в перекладі з грецької «thesauros» означає «скарбниця», не є новим у науковій літературі, адже його фахівці застосовують у лінгвістиці, педагогіці, психолінгвістиці та в інших наукових галузях протягом тривалого часу [10]. У минулому тезаурусами позначалися переважно словники з максимальною повнотою представлення лексики, з прикладами її вживання в текстах. Нині під тезаурусом розуміється складний тип словника, в якому всі терміни пов'язані між собою семантичними відношеннями, що відображають основні співвідношення понять предметної області знань, котра описується.

Поняття «тезаурус» нині активно використовується в галузях штучного інтелекту та інформаційних технологій. Дедалі більше інтелектуальних завдань, пов'язаних з опрацюванням знань,

наприклад, таких як індексування документів, інформаційний пошук, автоматизоване опрацювання текстів, передбачають використання тезауруса.

Тезаурус представляє повний систематизований набір даних будь-якої предметної області, що допомагає людині або комп'ютеру орієнтуватися в ній. Саме тезаурус більш повно подає термінологічну систему предметної області з урахуванням складності і багатовимірності семантичних взаємозв'язків. Він представляє систему термінів відповідної галузі. За його допомогою можна не лише здійснювати пошук, а й вивчати окремі терміни (поняття), отримувати семантичні зв'язки з іншими термінами (поняттями), усвідомлюючи їхню роль у системі знань певної галузі чи при розв'язанні конкретного завдання. Тезаурус зумовлює оновлення термінології певної області знань та робить доступними наукові джерела інформації, допомагаючи різномовним носіям зрозуміти один одного, сприяти поглибленому вивченню тих мов, якими він представлений.

Оскільки тезаурус передусім репрезентує термінологію певної галузі знань, то зважаючи на специфіку його семантики, завдання упорядкування термінів співвідноситься із завданням систематизації наукових знань. Це дає підстави розуміти статус тезауруса як спосіб представлення наукових знань.

Характерною перевагою тезауруса є його здатність завдяки своїй спеціальній структурі відображати існуючі відношення між включеними в нього терміноодинацями. На відміну від тлумачного словника, тезаурус дає змогу з'ясувати значення одиниці не стільки за допомогою визначення, скільки за допомогою співвіднесення змісту слова з іншими одиницями тезауруса.

Головна відмінність тезауруса від традиційного словника полягає у тому, що він подає не лише компіляцію статей, а й систему термінів-дескрипторів. Ключові слова – еквівалентні і близькі за значенням, а за їх допомогою проводиться пошук та опрацювання інформації. Вони об'єднуються в клас умовної еквівалентності. Кожен такий клас є словниковою одиницею, яка подається у формі окремого слова, словосполучення або коду. Ця словникова одиниця і є дескриптором. Класифіковані за змістом дескриптори утворюють тезаурус.

Інформаційно-пошуковий тезаурус робить можливим вхід через будь-яку характеристику терміна. Останній тут не обмежений тільки своєю словниковою статтею, як це має місце у паперовому слов-

нику. Електронне подання тезауруса уможливило через систему гіперпосилань встановлення логічних та асоціативних зв'язків між термінами.

Електронні тезауруси, що ґрунтуються на великих можливостях сучасної комп'ютерної техніки для зберігання інформації, здатні містити значно більше інформації, ніж окремих паперовий словник специфічного спрямування. Йдеться про можливість реалізації критеріїв комплексності і вичерпності при подаванні та зберіганні інформації, яка стосується кожного окремого поняття.

Тезауруси набувають дедалі більшого значення в процесі структурування та моделювання знань, систематизації та впорядкування наукової термінології, формування термінологічних банків даних та їх використання. У сучасній науці навіть побутує «тезаурусний метод», який широко застосовується у теорії та практиці інформаційного пошуку. Завдяки йому основний акцент робиться на змістовий аналіз призначених для відбору термінів із текстів, які є важливими для запиту користувача.

Тезаурус предметної області – система знань, подана у вигляді набору ключових термінів або дескрипторів (концептів) цієї області, пов'язаних між собою певними семантичними відношеннями, що відображають основні співвідношення понять предметної області. Основне завдання тезауруса – підвищення ефективності пошуку необхідної інформації: на основі семантичних відношень тезауруса відбувається розширення запиту, навігація по зв'язках допомагає чіткіше сформулювати сам запит. Отже, тезаурус – це спосіб систематизації знань певної предметної області, інструмент ефективного інформаційного пошуку.

Тезаурус застосовується як інструмент термінологічного контролю під час аналізу, опрацювання документів, інформаційних запитів, а також у процесі пошуку інформації. Функціональна роль тезауруса в інформаційно-пошукових системах висуває високі вимоги до якості підготовки тезауруса. Адже саме від ступеня досконалості останнього великою мірою залежить ефективність пошуку. Створення тезауруса – це один із засобів моделювання предметної області у формі множини термінів та множини відношень між цими термінами.

Отже, тезаурус сприяє пошуку та відбору необхідної інформації. Проте, структура, принципи організації, котрі забезпечують достатнє наповнення термінами, глибину розкриття семантичних взаємозв'язків, дають змогу тезаурусу моделювати термінологічну систему предметної області

найбільш повно і системно. Практична цінність тезауруса предметної області визначається можливістю використовувати його одночасно для аналізу і конструювання лексичної системи, класифікації і зберігання термінологічних даних, опрацювання інформації в пошукових системах для піднесення ефективності проблемно орієнтованого пошуку в Web [11].

Моделювання предметної області – доволі непросте завдання, оскільки потрібно поєднати індивідуальні концептуальні моделі різних фахівців для побудови єдиної моделі, яка не викликає розбіжностей. Модель предметної області вибудовується сукупністю понять, відношень та взаємозв'язків між ними. Саме такою моделлю предметної області «Розумне місто» має стати створений нами тезаурус.

Розроблення тезауруса є складним поетапним процесом. На основі аналізу робіт [4–7; 16] ми консолідували підходи до технології формування тезауруса і сформулювали алгоритм, якого дотримувалися, працюючи над створенням тезауруса предметної області «Розумне місто»:

перший етап – аналіз наявних словників, дотичних до предметної області «Розумне місто», формування вимог до визначення термінів тезауруса та його структури;

другий етап – укладання реєстрової частини (відбір лексикографічних одиниць зі словників, наукових статей, матеріалів конференцій, стандартів тощо і їх переклад українською мовою), визначення та уніфікування відношень, притаманних термінам у текстах статей;

третій етап – побудова словникових статей, упорядкування тезауруса відповідно до вибудованої структури, розроблення семантичних зв'язків;

четвертий етап – уточнення і редагування отриманого матеріалу, узгодження з користувачем.

Складним, багатоопераційним є процес побудови тезауруса. Кожен етап роботи тут пов'язаний з аналізом багатьох варіантів. Для того, щоб отримати новий інтелектуальний продукт, необхідно забезпечити певні умови, а саме:

- чітке розуміння мети та постановки завдання;
- достатність інформаційної бази;
- повний опис об'єктів предметної області;
- злагоджену співпрацю групи дослідників.

Науково-дослідницька робота з укладання тезауруса розпочинається з відбору лексичного матеріалу. Певні труднощі на цьому етапі пов'язані насамперед з необхідністю розмежувати термінологічну і загальноживану лексику. Далі, по мірі

накопичення термінологічного матеріалу, що видобувається з текстів, визначаються основні термінологічні поля галузі, яка досліджується.

При побудові інформаційно-пошукового тезауруса першим завданням є відбір термінів для включення до нього. Існує декілька джерел для відбору термінів при розробленні тезауруса. Насамперед повинні бути вивчені тезауруси близьких предметних областей, котрі містять значну кількість термінів, які можуть увійти до структури нового тезауруса. Терміни – кандидати на внесення до тезауруса – можуть пропонуватися експертами предметної області. Ще одним джерелом отримання термінологічних одиниць є наукові тексти.

На першому етапі формування тезауруса предметної області «Розумне місто» головну роль відіграли: британський стандарт PAS 180:2014 Smart cities – Vocabulary [13] як першоджерело термінів та знань про «розумне місто»; стандарти для забезпечення технологій «розумного міста», розроблені Британським інститутом стандартів (BSI) та Міжнародною організацією зі стандартизації ISO; науково-технічна література українською, російською, англійською мовами. Оскільки перевагами названих джерел є надійність, високий рівень загальноприйнятої інформації, зрозуміла мова подання, доступність, висока насиченість термінологічною лексиною маємо підстави для формування коректного тезауруса «розумного міста», який дасть змогу структурувати і накопичувати інформацію, релевантну для цієї галузі знань, вибудовувати розумовий пошук даних в інформаційних сховищах і мережах. Крім того, він сприятиме піднесенню рівня професійної підготовки фахівців не тільки в галузі «розумного міста», а й інформаційних технологій взагалі.

Опрацювання статей – також досить копітка робота, вона потребує надзвичайно багато часу. Проте, слід зазначити, що наукові статті, матеріали конференцій на відміну від словників, довідників, енциклопедій та інших, дають можливість виявляти конкретні терміни, які перебувають в ужитку, їх взаємозв'язки, бачити реальну картину їхнього функціонування у текстах.

Отже, вивчати терміни треба не тільки за допомогою словників, довідників, а й різної науково-технічної літератури, яка засвідчує факт їхнього реального функціонування.

Наступним етапом роботи над тезаурусом «розумного міста» є формування переліку ключових понять, що містяться у вищезазначених наукових джерелах. З цього переліку відбиралися поняття,

кандидати в терміносистему тезауруса, і подавалися як складники системи, також репрезентувалася системність галузі, що досліджується через термінологічні концепти. Відбір термінів здійснювався за такими критеріями: інформаційність, частотність, відповідність темі, актуальність, практична значущість.

При наповненні словника ставиться завдання з максимальною повнотою подати термінологію, яка використовується для опису понять предметної області, що досліджується.

При цьому враховувалося, що терміни, які формують концепти предметної області «Розумне місто», не ізольовані, вони встановлюють тісні зв'язки не тільки із загальнонауковими, а й галузевими поняттєвими категоріями. Тому термінологічні концепти розподіляються на загальнонаукові, міжгалузеві та вузькоспеціальні. Безперечно, провести чітке розмежування між ними достатньо складно, оскільки вони є дифузними завдяки постійній взаємодії та перехідності.

Робоча структура тезауруса предметної області «Розумне місто» формується з низки тематичних комплексів: загальні поняття, системи «розумного міста», процеси управління ресурсами, технології, інфраструктури, транспорт, мобільність, енергетика, телекомунікації, фінанси і економіка, місцеві та центральні органи влади, спільноти, освіта, охорона здоров'я, державні і приватні моделі надання послуг, система безпеки, навколишнє середовище.

Означена структура тезауруса, як і визначення, відкрита для змін і поповнень у майбутньому, вона досить широка, дає змогу охопити перший етап відбору понять.

Розроблення тезауруса предметної області «Розумне місто» передбачає опис слів і словосполучень, що зустрічаються в текстах даної предметної області. Тезаурус містить не тільки терміни, які характеризують важливі поняття в текстах цієї області, а й охоплює широке коло більш специфічних термінів, виявлення яких у конкретному тексті зробить пошук ефективнішим.

У процесі дослідження було розроблено класифікацію відношень між термінами предметної області «Розумне місто», проте набір цих семантичних зв'язків між поняттями нашого тезауруса не є обмеженим.

Основні типи понятійних зв'язків утворюють ієрархічні відношення «рід – вид» та «частина – ціле».

Родо-видові відношення охоплюють велику групу термінів, утворюючи мікросистеми: мінімаль-

ну кількість термінологічних одиниць – гіпонімів, які об'єднані родовим терміном – гіперонімом. Родо-видовий зв'язок встановлюється між двома термінами, якщо поняття, що знаходиться нижче, входить у значення поняття, що стоїть вище. Саме класифікація термінів, вибудована на відношеннях «рід–вид», дає можливість об'єктивно виокремити гіпероніми (головні поняття) в предметній області «Розумне місто» та класифікувати навколо них гіпоніми, створюючи загальну класифікаційну схему понять.

Семантичний простір термінів «Розумного міста» – складна ієрархічна структура, елементи якої пов'язані гіперо-гіпонімними відношеннями. Тезаурусна модель предметної області «Розумне місто» постає у вигляді схеми з вершиною «розумне місто», яка розщеплюється на шість головних тематичних термінополів: «розумне життя», «розумні люди», «розумна економіка», «розумне довкілля», «розумна мобільність», «розумне управління». Відповідні терміни є гіпонімами по відношенню до гіпероніма «розумне місто».

До традиційних відношень належать і відношення еквівалентності, які встановлюються між термінами, що позначають одне і те ж поняття (подібні терміни перебувають у відношенні синонімії). Такі терміни належать до однієї категорії термінів, мають тотожне значення і вони взаємозамінні у всіх контекстах. Наприклад: інтелектуальне місто (intelligent city) – розумне місто (smart city), стійкі джерела енергії (sustainable energy sources) – відновлювальні джерела енергії (renewable energy sources) та ін.

Для більш повного розкриття специфіки предметної області, що досліджується, разом із визначенням традиційних відношень між термінами нами було виділено низку особливих проблемно-специфічних відношень, характерних для області «Розумне місто», які дають змогу зробити опис кожного терміна тезауруса більш повним. Детальна класифікація проблемно-специфічних зв'язків уможливує більш точне вираховування функції семантичної близькості двох понять, що може істотно підвищити ефективність автоматичних систем, побудованих тезаурусним методом.

Висновок. Розроблений тезаурус предметної області «Розумне місто» містить великий масив термінів, виразів, які використовуються при дослідженні інтелектуальних систем міста, систем комунікації, слугують термінологічним інструмен-

том проведення ефективних інформаційно-пошукових операцій.

Концепти тезауруса становлять основу для формування пошукового образу документа і пошукового образу запиту. Процеси їх побудови спрощуються завдяки чіткому представленню знань про конкретну предметну область. Тезаурус можна розглядати як модель термінологічної системи, формальну систему знань предметної області «Розумне місто».

Список бібліографічних посилань

1. Андрусенко Т., Стрижак А. Управление знаниями в учебном процессе на основе тезаурусов. *E-Learning World*, 2007. № 1. С. 56–62. URL : <http://www.elw.ru/magazine/22/200/>
2. Величко В., Волошин П., Свитла С. Автоматизированное создание тезауруса терминов предметной области для локальных поисковых систем. «*Knowledge – Dialogue – Solution*» *International Book Series «INFORMATION SCIENCE & COMPUTING»*. Sofia, 2009. № 15. P. 24–31.
3. Гладун А. Я., Рогушина Ю. В. Методики розробки термінології предметної області як базису для формування онтологій та тезаурусів. *Проблеми інформатизації та управління*. 2009. № 3 (27). С. 26–34.
4. Гладун А. Я., Рогушина Ю. В. Основы методологии формирования тезаурусов с использованием онтологического та мереологического анализа. *Искусственный интеллект*. 2008. № 5. С. 112–124.
5. Городецкий Б. Ю. Проблемы и методы современной лексикографии. *Новое в зарубежной лингвистике*. 1983. Вып. XIV. С. 5–22.
6. Дубичинский В. В. *Теоретическая и практическая лексикография*. Wien; Charkov. 1998. 160 с.
7. *IDEF5 – Ontology Description Capture Method*. URL : www.idef.com/IDEF5.html
8. Кульчицький І. М. Відображення структурних відношень реєстру термінологічних словників у базах даних. *Вісн. Нац. ун-ту «Львів. Політехніка»*. 2010. № 689. С. 271–280.
9. Ланде Д. В. *Елементи комп'ютерної лінгвістики в правовій інформатиці*. К. : НДІП НАПрН України, 2014. 168 с.
10. Лукашевич Н. В. *Тезаурусы в задачах информационно-го поиска*. М. : Изд-во МГУ, 2011. 495 с.
11. Мацюк Г., Кунаець Н. Представлення знань у вузько-спеціальних предметних областях за допомогою тезауруса. *Бібліотечний вісник*. 2019. № 1. С. 19–23.
12. Мінцер О., Палагін О., Величко В., Стрижак О., Тахере Г. Інструменти підтримки процесів аналітичної діяльності експерта при тематичному дослідженні інформаційних ресурсів і джерел. *Медична інформатика та інженерія*. 2011. № 2. С. 12–23.
13. *PAS 180 Smart city technology*. URL : <http://www.bsi-group.com/smart-cities/Smart-Cities-Standards-and-Publication/PAS-180-smart-cities-terminology/>
14. Широков В. А. *Елементи лексикографії*. Київ : Довіра, 2005. 304 с.
15. Широков В. А. *Інформаційна теорія лексикографічних систем*. Київ : Довіра, 1998. 331 с.
16. Шуневич Б. І. Сучасні способи відбору термінів та укладання перекладних словників нових терміносистем. *Вісник Житомирського держ. ун-ту ім. Івана Франка*. 2008. Вип. 38. С. 90–93.

References

1. Andrusenko, T. & Strizhak, A. (2007). Upravlenie znaniiami v uchebnom protsesse na osnove tezaurosov [Knowledge management in the learning process based on thesauri]. *E-Learning World*, 1, 56-62. Retrieved from <http://www.elw.ru/magazine/22/200/> [In Russian].
2. Velichko, V., Voloshin, P. & Svitla, S. (2009). Avtomatizirovannoe sozdanie tezaurosa terminov predmetnoi oblasti dlia lokalnykh poiskovykh sistem [Automated creation of a domain term thesaurus for local search engines]. «*Knowledge – Dialogue – Solution*» *International Book Series «INFORMATION, SCIENCE & COMPUTING»*, 15, 24-31. Sofia, Bulgaria. [In Russian].
3. Hladun, A. Ya. & Rohushyna, Yu. V. (2009). Metodyky rozrobky terminolohii predmetnoi oblasti yak bazysu dlia formuvannia ontolohii ta tezaurosov [Methods for developing domain terminology as a basis for the formation of ontologies and thesauri]. *Problemy informatyzatsii ta upravlinnia*, 3 (27), 26-34. [In Ukrainian].
4. Hladun, A. Ya. & Rohushyna, Yu. V. (2008). Osnovy metodolohii formuvannia tezaurosov z vykorystanniam ontolohichnoho ta mereolohichnoho analizu [Bases of methodology of formation of thesauruses with use ontologic and the mereologic analysis]. *Iskusstvennyi intelekt*, 5, 112-124. [In Ukrainian].
5. Gorodetskii, B. Yu. (1983). Problemy i metody sovremennoi leksikografii [Problems and methods of modern lexicography]. *Novoe v zarubezhnoi lingvistike*, XIV, 5-22. [In Russian].
6. Dubichinskii, V. V. (1998). *Teoreticheskaia i prakticheskaia leksikografiia* [Theoretical and practical lexicography]. Wien; Charkov. [In Russian].
7. *IDEF5 – Ontology Description Capture Method*. Retrieved from www.idef.com/IDEF5.html
8. Kulchytskyi, I. M. (2010). Vidobrazhennia strukturnykh vidnoshen reiestru terminolohichnykh slovnykiv u bazakh danykh [Display of structural relations of the register of terminological dictionaries in databases]. *Visn. Nats. un-tu «Lviv. Politehnika»*, 689, 271-280. [In Ukrainian].
9. Lande, D. V. (2014). *Elementy kompiuternoï linhvistyky v pravovii informatytsi* [Elements of computational linguistics in legal informatics]. Kyiv, Ukraine: NDIIP NAPrN Ukrainy. [In Ukrainian].
10. Lukashevich, N. V. (2011). *Tezaury v zadachakh informacijnogo poiska* [Thesauruses in information search tasks]. Moscow, Russia: Izd-vo MGU. [In Russian].
11. Matsiuk, H. & Kunanets, N. (2019). Predstavlennia znan u vuzkospetsialnykh predmetnykh oblastiakh za dopomohou tezaurosa [Presentation of knowledge in the specialized subject areas with the help of a thesaurus]. *Bibliotechnyi Visnyk*, 1, 19-23. [In Ukrainian].
12. Mintser, O., Palahin, O., Velychko, V., Stryzhak, O. & Takhere, H. (2011). Instrumenty pidtrymky protsesiv analitychnoi diialnosti eksperta pry tematychnomu doslidzhenni informatsiinykh resursiv i dzherel [Tools to support the processes of analytical activity of the expert in case study of information resources and sources]. *Medychna informatyka ta inzheneriia*, 2, 12-23. [In Ukrainian].
13. *PAS 180 Smart city technology*. Retrieved from <http://www.bsigroup.com/smart-cities/Smart-Cities-Standards-and-Publication/PAS-180-smart-cities-terminology/>
14. Shyrovkov, V. A. (2005). *Elementy leksykohrafii* [Elements of lexicography]. Kyiv, Ukraine: Dovira. [In Ukrainian].
15. Shyrovkov, V. A. (1998). *Informatsiina teoriia leksykohrafichnykh system* [Information theory of lexicographic systems]. Kyiv, Ukraine: Dovira. [In Ukrainian].
16. Shunevych, B. I. (2008). Suchasni sposoby vidboru terminiv ta ukladannia perekladnykh slovnykiv novykh terminosystem [Modern ways of selecting terms and compiling new vocabularies for new terms]. *Visnyk Zhytomyrskoho derzh. un-tu im. Ivana Franka*, 38, 90-93. [In Ukrainian].

Halyna Matsiuk,

Lecture, Ternopil Ivan Puluj National Technical University (Ternopil, Ukraine)

IMPROVEMENT OF INFORMATION RETRIEVAL TECHNOLOGIES ON THE BASIS DATA DOMAIN THEZAURUS

The ways of information retrieval technologies improvement due to the data domain thesaurus are analyzed in this paper. The algorithm of the information retrieval procedure and the algorithm of the technology of the «Smart City» data domain thesaurus formation are presented. Considerable attention is paid to the fact that while creating the search image and search request, the information retrieval language based on the common terminology application should be used.

K e y w o r d s: thesaurus, data domain, «Smart City», information search, term.

Ternopil Ivan Puluj National Technical University,
56, Ruska st., Ternopil, 46001, Ukraine
E-mail: galuna.matsiuk@gmail.com

Стаття надійшла до редакції 7.08.2019 р.