

УДК 811.161.2'1 (075.8)

DOI <https://doi.org/10.32782/2522-4077-2025-212-9>

КОРПУСНІ ТА ШІ-ТЕХНОЛОГІЇ УПРАВЛІННЯ ТЕРМІНОЛОГІЄЮ В ЛОКАЛІЗАЦІЙНИХ ПРОЄКТАХ

CORPUS AND AI TECHNOLOGIES FOR TERMINOLOGY MANAGEMENT IN LOCALISATION PROJECTS

Бондаренко О. С.,

orcid.org/0000-0002-3148-8753

кандидат філологічних наук,

доцент кафедри перекладу, прикладної та загальної лінгвістики

Центральноукраїнського державного університету імені Володимира Винниченка

Бондаренко К. Л.,

orcid.org/0000-0002-2964-5123

кандидат філологічних наук,

доцент кафедри перекладу, прикладної та загальної лінгвістики

Центральноукраїнського державного університету імені Володимира Винниченка

Статтю присвячено особливостям управління термінологією в локалізаційних проєктах, зокрема, процесу видобування термінів із фахових текстів за допомогою корпусних технологій та технологій, що передбачають використання ШІ. Враховуючи складність і багатшаровість процесу локалізації, ефективність управління термінологією, починаючи з етапу її видобування, безпосередньо впливає на якість надання локалізаційної послуги.

Розглянуто стратегії і тактики видобування термінів, які застосовуються на ринку лінгвістичних послуг. Уточнено термін «видобування термінів», з'ясовано принципи, які використовуються в сучасних інструментах. Проведено експериментальне дослідження видобування термінів інструментами SketchEngine, OpenAir (Plus) та Fodina TermCatch.

Визначення ефективності видобування передбачало створення корпусу фахових текстів галузі автоматизованого програмного забезпечення загальним обсягом 29374 слів (36912 токенів). Результат видобування оцінювався за критеріями відповідності вимогам до термінів та терміносполук (стійкість, точність, нероздільність, дефінітивність).

За допомогою інструмента SketchEngine було видобуто 400 пропозицій, з яких лише 25% можуть уважатися терміносполуками. Використання продукту OpenAir (Plus) на основі ШІ надало 81 пропозицію, з яких відповідними вимогам виявились 38%. Інструментом Fodina TermCatch tool було видобуто 1460 пропозицій, з яких лише 15.93% виявились відповідними. Виявлена варіативність видобутої термінології може бути корисною задля покращення якості вихідного фахового тексту.

Результати статистичного видобування демонструють більшу точність, однак залишають поза увагою нечастотні терміни. Технології ШІ надають більшу кількість кандидатів у терміни, однак потребують фільтрування нерелевантних результатів. Дослідження доводить ефективність гібридного використання корпусних інструментів на основі статистичних принципів та тих, що використовують ШІ. Автоматизація процесу управління термінологією – важливий крок із оптимізації менеджменту локалізаційних проєктів.

Ключові слова: управління термінологією, локалізаційний проєкт, корпусні технології, автоматизоване видобування термінів, термінна база.

The article explores terminology management for localization projects, with a special focus on extraction technologies, emphasizing corpus-based and AI-driven approaches. Given the complexity of localization, the automation of terminology management plays a crucial role in ensuring high-quality linguistic services.

The study examines various terminology extraction strategies applied in Ukraine's language services market. The research objectives include defining the concept of «terminology extraction» systematizing principles of automated extraction tools, and conducting an experimental extraction process using SketchEngine, OpenAir (Plus), Fodina TermCatch.

To assess the effectiveness of these tools, the authors built a specialized corpus, comprising 29,374 words and 36,912 tokens. The extracted terminology was analyzed based on key criteria: stability, precision, indivisibility, and definitional clarity. SketchEngine tool employs statistical keyword analysis, extracting 400 term candidates. However, only 25% (123) were deemed valid terminological units due to issues with phrase segmentation. Utilizing AI-driven models, OpenAir (Plus) tool extracted 81 term candidates, with 38% (31 terms) meeting the established criteria. However, it tended to overgenerate paraphrased variants. Fodina TermCatch tool extracted 1,460 candidates, but only 15.93% (233 terms) were valid. It identified redundant terms variations, which could be useful for refining source content.

The study underscores the hybrid nature of terminology extraction today, combining statistical, AI-based, and hybrid approaches. While statistical methods provide greater precision, they may miss rare terms, whereas AI-driven tools offer high recall but require filtering.

In conclusion, automated terminology extraction enhances efficiency in localization. However, manual validation remains essential for ensuring accuracy. The research confirms that combining multiple extraction tools yields the best results, making terminology extraction – an essential part of terminology management – a crucial step in localization project management.

Key words: terminology management, localisation project, corpus technologies, automated term extraction, termbase.

Постановка проблеми. Локалізація – це «лінгвістична та культурна адаптація цифрового продукту до вимог та особливостей певної локалі, а також послуги та технології, покликані забезпечити багатомовність світового інформаційного потоку» [1, с. 157]. Комплексність процесу адаптації, динамічність ринку локалізації, великі обсяги і водночас висока конкуренція надавачів лінгвістичних послуг роблять автоматизацію етапів локалізації запорукою успішності на ринку лінгвістичних послуг. Зважаючи на те, що вербальна складова продуктів для локалізації, – це першопочатково фахові мови, згодом адаптовані під масового споживача, термінологія є однією з найважливіших елементів таких мов.

Актуальність дослідження полягає в тому, що завчасно (перед власне виконанням перекладацької частини локалізації) підготовлена, унормована та узгоджена з клієнтом термінологія – запорука якісно виконаної адаптації продукту, підтвердженої безпомилковим розумінням текстової складової продукту споживачами [2; 3].

Аналіз останніх досліджень і публікацій. Управління термінологією в виробничих процесах локалізації та перекладу не так часто привертає увагу дослідників, оскільки потребує повноцінного занурення в процес, однак широко обговорюється на ключових галузевих подіях, зокрема, Translating Europe Forum 2024, темою якого була *Translators at the Forefront of AI-driven Data, Terminology and Technology*. Деякі вчені розглядають коректність термінології у контексті забезпечення якості та контролю якості кінцевого продукту або виробничого процесу як такого [4; 5]. Розвиток цифрових інструментів, зокрема, автоматизованого опрацювання корпусів або таких, що залучають системи штучного інтелекту в контексті управління термінологією цікавив європейських учених із початку 2000-них, [6], в Україні публікації почали з'являтися дещо пізніше [7; 8].

Поточне дослідження переважно ґрунтується на доробку вчених, які розглядають специфіку роботи з термінологією крізь призму виробничого процесу надання лінгвістичних послуг [2; 9; 10; 11; 12].

Мета нашої розвідки – виокремлення основних стратегій видобування термінів, що застосовуються на сучасному українському ринку лінгвістичних послуг. Серед **завдань**: уточнення форми і змісту терміносполуки «видобування термінів»; систематизація принципів функціонування інструментів, які забезпечують автоматизоване видобування термінів із фахових текстів; власне видобування термінів за допомогою інструментів OpenAir (версія Plus), Fodina TermCatch, SketchEngine із корпусу фахових текстів галузі автоматизованого перекладу задля подальшої локалізації відповідного ПЗ для української локалі.

Для перевірки ефективності кількох систем автоматизованого видобування термінів на основі посібника Trados Studio Level 1 Trados Studio 2022 було створено корпус загальним

обсягом 29374 слів (36912 токенів). До корпусу було застосовано процедуру видобування термінів інструментом SketchEngine (функція Key words/terminology extraction). Це ж джерело було оброблено інструментами OpenAir (версія Plus) та Fodina TermCarch. Згодом видобутий матеріал (потенційні кандидати в терміни) був переглянутий задля визначення відповідності вимогам до термінів та терміносполук, які мають нероздільними, стійкими, точними та дефінітивними [13]. Результатом перегляду стали уточнені списки термінів та терміносполук, які можуть бути рекомендовані для довідкових матеріалів (термінних баз) для подальшої локалізації.

Виконання завдань передбачало використання низки методів: дистрибутивно-статистичний аналіз корпусу [14] застосування аналітичного методу передбачало перегляд і оцінку видобутих термінів на відповідність вимогам до термінів (стійкість, точність, нероздільність, дефінітивність); елементи порівняльного аналізу було використано для зіставлення результатів автоматизованого видобування термінів різними інструментами.

Виклад основного матеріалу. Виконання термінологічної роботи не завжди покладається безпосередньо на лінгвістів-локалізаторів, оскільки великі гравці, як правило, надають готові термінні глосарії, над якими працює команда термінологів та галузеві експерти. У випадку, якщо компанія випускає декілька релізів, продукт набуває нових функцій, супроводжується застосунками тощо, термінологія зазнає впливу і цих процесів: може фіксуватися варіативність, деякі терміни виходять з ужитку тощо [15]. Варіативність термінів в індустрії лінгвістичних послуг, на відміну від усталених фахових спільнот, розглядається як нормальний стан речей, оскільки сама локалізаційна діяльність максимально адаптована до процесу виробництва і поширення цифрових продуктів. У професійних застосунках (наприклад, Multiterm), призначених як для управління термінологією, так і для використання готових термінних глосаріїв, наприклад, передбачені випадні опції для «term candidates» та «validated terms».

Процес ручного добирання термінів із фахових текстів, їхнє стандартизування і гармонізування налічує не одне століття. Цілком очевидно, що розвиток технологій не міг не вплинути і на цей процес. Вчені і практики погоджуються, що процес упорядкування термінології все ще здійснюється переважно здійснюється, однак на етапі добору того, що згодом стає термінними базами, глосаріями, фаховими словниками тощо нехтування цифровими інструментами не раціональне.

Процес автоматизованого видобування термінології розпочато західними дослідниками, тому початковою терміносполукою на позначення процесу слід уважати «term extraction». Незважаючи на те, що одним із вимог до терміну однозначна відповідність поняттю, дотичні публікації містять синоніми «terminology mining» [16], «glossary extraction» [17], «terminology acquisition» [18]. Українські позначення поняття так само варіативні: «видобування термінів» [7], «добування термінів» [8], «вилучення термінів» (у значенні пошуку в тексті), «екстрагування термінів» [19] тощо.

Враховуючи те, що за міжнародними вимогами термін має бути: відповідним позначеному ним поняттю (transparent); системним (consistent); раціонально стислим (linguistic economy); дериваційно здатним (derivability and compoundability); переважно рідною мовою (preference for native language) тощо [20], не вважаємо термін *екстрагування* оптимальним, бо він не є питомим, а відтак достатньо дериваційно здатним (наприклад, утворювати дієслово на позначення процесу або іменник на позначення особи, яка здійснює цю операцію). «Вилучення» не передає специфіки і призначення процесу, бо «вилучено» терміни не назавжди, а щоб згодом «повернути» їх у фаховий текст у вигляді відповідника мовою локалі. Зважаючи на словотвірні особливості вихідного (англійського) терміну (ex+traction), вважаємо термін видобування придатним для опису процесу, оскільки він є відповідним (навіть на рівні афіксу) позначуваному поняттю, питомим, системним; раціонально стислим і дериваційно здатним.

Видобування термінів зазнає стрімких змін: від одномовного статистичного автоматизованого видобування термінології, двомовного автоматизованого видобування до використання ШІ. Можемо констатувати, що є свідками гібридного підходу до цього процесу, коли одночасно можуть використовуватися кілька підходів (спочатку одномовне статистичне (для виявлення найпоширеніших, а відтак значущих для проєкту понять, виражених термінами), згодом – двомовне (для пошуку потенційних відповідників) тощо.

Одномовне автоматизоване видобування (automatic term extraction) – це процес видобування термінів із фахових текстів однієї мови. Від початку цей добір ґрунтувався на припущенні про частотність уживання термінів у фахових текстах. Однак проста вибірка keywords (ключових слів) із фахових текстів надавала хибні результати, оскільки до ключових слів за такою логікою входили артиклі, модальні дієслова, одиниці на кшталт *people, user, client* (для ПЗ) тощо. В подальшому гіпотеза трансформувалась у припущення про перспективність зіставлення статистичної значущості словосполучень у фокусному корпусі (фахових текстів), із тим, наскільки характерними такі сполуки є для загального корпусу національної мови, що обирається як референтний. Насиченість (у порівнянні з загальним корпусом) корпусу фахових текстів термінами, на думку розробників, є ознакою того, що потрібні одиниці будуть сприйматися програмою як типові для фахових, нетипові для загальних, а відтак, з високим ступенем вірогідності, термінні одиниці. Такий підхід активно застосовується SketchEngine – програмним забезпеченням для створення та обробки корпусного матеріалу. Автоматизоване видобування термінів може здійснюватися із застосуванням ШІ (великих мовних моделей) алгоритм дії яких не настільки прозорий, а результат – не настільки передбачуваний, хоча має свої переваги.

У процесі зіставлення фокусного корпусу галузі автоматизованого перекладу з референтним корпусом English Web 2021 (enTenTen21) програмою SketchEngine (функція Key words/terminology extraction) було видобуто 400 словосполучень, які можуть розглядатися як кандидати в терміни (рис. 1).

Term	Term	Term	Term	Term
1 trados studio ***	11 character format ***	21 following step ***	31 placeholder tag ***	41 segment in the document ***
2 following screenshot ***	12 current segment ***	22 sample file ***	32 concordance search ***	42 right-hand side ***
3 fuzzy match ***	13 language pair ***	23 tm file ***	33 segment number ***	43 sdi xiff ***
4 file for translation ***	14 cloud capability ***	24 project file ***	34 target document ***	44 bilingual file ***
5 target segment ***	15 repealed segment ***	25 editor view ***	35 exact match ***	45 document preview ***
6 source segment ***	16 mouse pointer ***	26 preview window ***	36 project template ***	46 studio for the first time ***
7 project with multiple files ***	17 trados studio editor ***	27 translation engine ***	37 first segment ***	47 target file ***
8 multiple file for translation ***	18 match value ***	28 project creation ***	38 pressing ctrl ***	48 preview tab ***
9 multiple file ***	19 studio editor ***	29 new file ***	39 following segment ***	49 single file ***
10 target language ***	20 machine translation ***	30 target cell ***	40 side-by-side editor ***	50 cloud project ***

Рис. 1. Візуалізація результатів видобування термінів програмою SketchEngine (функція Key words/terminology extraction) (фрагмент)

За нашими підрахунками лише 25% (123) видобутих словосполук можуть уважатися термінними. Аналіз оточення одного з ключових для CAT-інструментів терміну *segment* доводить, що система видобування розглядає модифікатори іменника такими, які слід розглядати частиною терміносполуки, що не завжди так. Сполуки *target segment* та *source segment* у контексті використання CAT-інструментів термінними вважати можна, а *current segment* – ні, оскільки

остання не є ані стійкою, ані нероздільною, ані точною, що доводять інші результати видобування програмою: *same segment*, *remaining segment*, *considered segment* ітп.

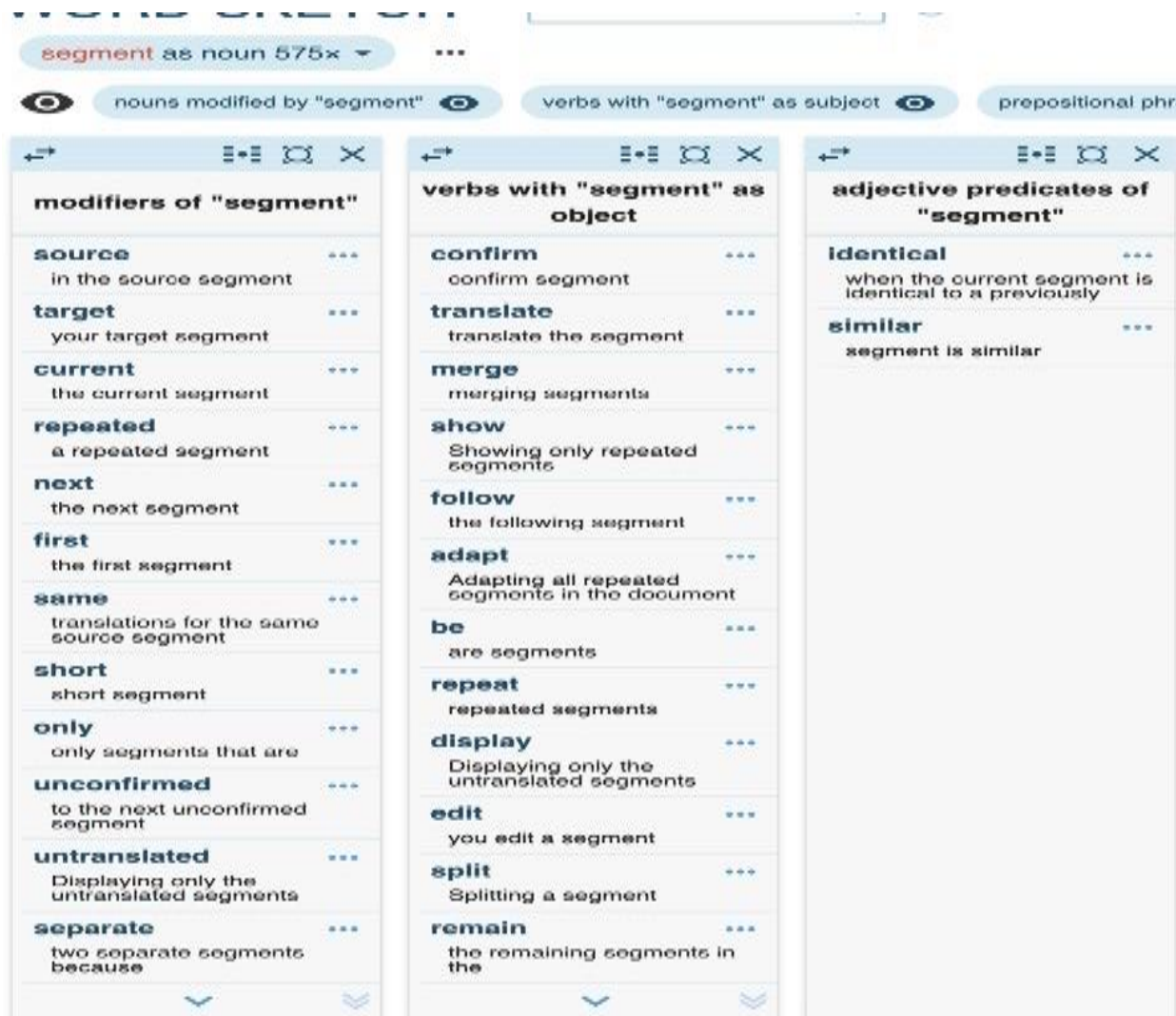


Рис. 2. Візуалізація сполучення терміну *segment* у корпусі галузі автоматизованого перекладу програмою SketchEngine (фрагмент)

Автоматизоване видобування термінів може здійснюватися із застосуванням ШІ (великих мовних моделей) алгоритм дії яких не настільки прозорий, а результат – не настільки передбачуваний, хоча вбачається гравцями ринку доволі значущим [21; 22; 23].

В результаті опрацювання корпусу посібника Trados Studio Level 1 Trados Studio 2022 інструментом OpenAir (версія Plus) було запропоновано 81 сполук, 31 з яких (38%) можна вважати термінними. При цьому, причини того, що решта пропонованих словосполук не є термінними, – інші, ніж в аналізованих вище результатах автоматизованого видобування на основі виключно статистичних даних. У випадку з результатом видобування інструментом, який використовує для генерування великі мовні моделі, обсяг запропонованого може досягатися за рахунок варіацій на кшталт *matched segments/segments matches*, які навіть за формою не можуть уважатися відповідними вимогам до термінів. Доволі частотними є пропозиції уважати терміносполуками терміни з дескрипторами, як-от *Translation Toolbar*, *Editor Ribbon*, *Preview Button*, *Confirmation Buttons*. Останнє може бути корисним для локалізації програмного забезпечення, максимально орієнтованого на кінцевого споживача, оскільки, згідно з Настановами зі стилю провідних компаній, дескриптори рекомендується додавати навіть якщо це суперечить вимогам лаконіч-

ності, як-от: *Report tool bar* ‘the tool bar at the top of the Report section, used to create and edit reports’ (панель “Звіт” і панель інструментів для звітів (рекомендований)); *Business Contact Manager tool bar* ‘the tool bar at the top of an item where you can view existing links or create new links to Account, Business Contact, Opportunity, or Business Project records)’ (панель Диспетчера контактів і панель інструментів Диспетчера бізнес-контактів (рекомендований)) [9].

В результаті опрацювання корпусу посібника Trados Studio Level 1 Trados Studio 2022 інструментом Fodina TermCarch, який, за словами розробників, визначає та видобуває термінологію, використовуючи III та великі мовні моделі [24] видобуто 1460 одиниць та сполук, які програма пропонує вважати термінними. З них лише 233 (15,93%) можуть уважатися такими, що відповідають вимогам до термінів. Хибно позитивними вважаємо однослівні пропозиції на кшталт *tip, chapter, opportunity, possibility, people, task* тощо які, хоча й є статистично значущими, однак не можуть уважатися такими, що описують поняття, ключові для даного фрагменту дійсності. До пропозицій потрапляють словосполучення, що позначають одне поняття, однак у вихідному тексті помилково номіновані по-різному: *terminology database / termbase, drop-down list dropdown list; upper right corner / upper-right corner; white-space character / whitespace character; upper-case letter / uppercase letter; top right corner / top-right corner*. Виявлення такої варіативності може бути корисним для покращення якості вихідного продукту ще на етапі інтернаціоналізації, який передувє процесу власне локалізації – адаптації для конкретних локалей.

Зауважимо, що незважаючи на різні підходи до видобування термінів, всі три інструменти видобули достатню кількість ідентичних словосполучень, значна частина яких відповідає вимогам до терміносполук. Це такі одиниці, як *Translation Memory, User Interface, Source Segment, Target Segment, Project Settings, Editor View, Ribbon Tab, Sample File, Fuzzy Match, Exact Match, QuickPlace List, Auto-Propagation, Document Layout, Translation Results, Target Language, Source Language, Translation Suggestions, Exact Repetition, Translation Unit* тощо. У відсотковому співвідношенні ця кількість становить 10% від загальної кількості сполук, видобутих статистичним інструментом SketchEngine, 38% від таких, що було видобуто інструментом OpenAir, та 4,1% здобутку інструментом Fodina TermCarch. Зауважимо, що зазначені відсотки не є універсальним показником ефективності інструментів, оскільки: 1) SketchEngine «вважає» термінними тільки сполуки слів; 2) OpenAir, вірогідніше за все «не помічає» важливі для ПЗ, але не настільки частотні терміни, а у відповідь на інструкції (промптинг) щодо збільшення кількості реагує повторами; 3) Fodina TermCarch містить найбільшу кількість видобутих варіантів, частина з яких є нерелевантними, однак загальний обсяг дозволяє вибрати, в тому числі, й однослівні терміни, чого не пропонував SketchEngine. Великі обсяги видобутих кандидатів у терміни, з одного боку, є вадою, оскільки потребують зусиль лінгвіста із «очищення» бази, але з іншого – очищення потребує меншої кількості зусиль і часу, аніж пошук термінів, яких бракує.

Звісно процес удосконалення термінології для локалізаційних проєктів відрізняється від класичної термінографійної роботи: те, що часто йдеться про фахові тексти, призначені не для фахівців (наприклад, користувачів ПЗ), безумовно, впливає на суворість дотримання вимог до термінів чи терміносполук. Обмежений час релізу продуктів (у тому числі, й локалізованих версій), впливає, в тому числі, й на час, який можна приділяти управлінню термінологією.

Висновки. Станом на початок 2025 року можемо констатувати, що автоматизоване видобування термінології є достатньо ефективним етапом, який зменшує кількість часу та зусиль у процесі виконання локалізаційних проєктів. Виключно статистичний метод виявляє значну кількість результатів-потенційних кандидатів у терміни, однак потребує додаткових навичок із налаштування фільтрів для хибно позитивних результатів. Використання OpenAir наразі виглядає ефективнішим, однак кількість видобутих кандидатів у терміни – менша, а половина запропонованого – перефразування найбільш очевидних варіантів. Використання Fodina TermCarch надає найбільшу кількість кандидатів у терміни, однак потребує зусиль із фільтрування зайвого. Зіставлення релевантних результатів трьох систем автоматизованого видобу-

вання дозволило виявити спільні варіанти, які, як правило, відповідні вимогам до термінів та терміносполук, відтак можуть рекомендуватися для термінологічних глосаріїв, придатних для подальшого використання автоматизованими системами перекладу в процесі локалізації.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Schäler R. Translators and localization: Education and training in the context of the Global Initiative for Local Computing (GILC). *The Interpreter and Translator Trainer*. 2007. P. 119–135.
2. Warburton K. Managing terminology in commercial environments. *Handbook of Terminology*. 2015. № 1. P. 359–391.
3. Kageura K., Marshman E. Terminology extraction and management. *The Routledge Handbook of Translation and Technology*. Routledge, 2019. P. 61–77.
4. Popiolek M. Terminology management within a translation quality assurance process. *Handbook of Terminology*. 2015. № 1. P. 341–359.
5. Seljan S. Quality Assurance (QA) of Terminology in a Translation Quality Management System (QMS) in the business environment. *European Parliament: Translation Services in the Digital World*. 2018. P. 92–145.
6. Fulford H. Exploring terms and their linguistic environment in text: A domain-independent approach to automated term extraction. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*. 2001. № 7(2). P. 259–279.
7. Борисова Н. В., Решетило С. С. Автоматизоване видобування термінологічних одиниць з науково-технічних текстів. *Матеріали III Всеукраїнської науково-практичної конференції «Інтелектуальні системи та прикладна лінгвістика»*. Харків, 2014. С. 43–45.
8. Бармак В. Виявлення використань графової бази даних GraphDB за допомогою автоматичного добування термінів. *Запоріжжя: ЗНУ*, 2023. 50с.
9. Бондаренко К. Л., Булгакова К. В. Розвиток української ІТ-термінології в процесі локалізації (на прикладі поштових клієнтів Microsoft Office Outlook та Google Gmail). *Наукові записки. Серія: Філологічні науки*. 2024. Вип. 210. С. 58–65.
10. Бодюл Д. Структурні особливості термінів ІТ на прикладі вебпереглядачів Google Chrome і Microsoft Edge. *Студентський науковий вісник*. 2024. Вип. 31. Кропивницький : РВВ ЦДУ ім. В. Винниченка. С. 231–234.
11. Folaron D. Technology, technical translation and localization. *The Routledge Handbook of Translation and Technology*. 2019. P. 203–219.
12. Bowker L. Terminology management. *The Bloomsbury Companion to Language Industry Studies*. 2020. С. 261–284.
13. Туркевич О. Семантична структура терміносполук методики викладання української мови як іноземної. *Вісник Нац. ун-ту «Львівська політехніка»*. Серія «Проблеми української термінології». 2011. № 709. С. 121–124.
14. Bilyk K. M., Makhachashvili R. K. The analysis of the corpus of texts «Covid-19» using the software «Sketch Engine». *Scientific Trends in Basic and Applied Materials Research of the International Scientific and Practical Conference*. October 30, 2020. Strasbourg : ЛОГОΣ. 2020. Vol. 3. P. 59–65.
15. Бондаренко К. Л., Булгакова К. В. Варіативність термінології програмного забезпечення в контексті локалізації. *Наукові записки. Серія: Філологічні науки*. 2024. Вип. 211. С. 14–19.
16. Kovačević A., Konjović Z., Milosavljević B., Nenadić G. Mining methodologies from NLP publications: A case study in automatic terminology recognition. *Computer Speech & Language*. 2012. № 26(2). P. 105–126.
17. Arora C., Sabetzadeh M., Briand L., Zimmer F. Automated extraction and clustering of requirements glossary terms. *IEEE Transactions on Software Engineering*. 2016. № 43(10). С. 918–945.
18. Dagan I., Church K. Termight: Coordinating humans and machines in bilingual terminology acquisition. *Machine Translation*. 1997. № 12. С. 89–107.
19. Амеліна С. М., Тарасенко Р. О. Вивчення технологій створення електронних термінологічних баз у процесі професійної підготовки перекладачів. *Інформаційні технології і засоби навчання*. 2017. № 60(4). С. 105–115.
20. ISO 704:2009. Terminology work – Principles and methods. Geneva : ISO, 2009.
21. Russo M. ASK THE EXPERTS' 2 – Introduction to prompting. URL: <https://www.youtube.com/watch?v=8bLVcmEX55I>. (дата звернення: 29 січня 2025 р.)
22. Bhaskar B., Braşoveanu C., Rausch I. Contributing to better multilingual AI in Europe. URL: <https://www.youtube.com/watch?v=yOjdEQntOhs&t=8s>. (дата звернення: 29 січня 2025 р.)
23. Goldsmith J. AI-powered terminology extraction: A hands-on guide for translators. URL: <https://www.youtube.com/watch?v=5Y5PhzyeMGI>. (дата звернення: 29 січня 2025 р.)
24. Kaleidoscope. URL: <https://kaleidoscope.at/en/products/fodina-termcatch>.