
DOI: <https://doi.org/10.15407/emodel.44.04.105>
УДК 004.91 81'374.2.

О.В. Нестеренко¹, д-р техн. наук, **І.Є. Нетесін**², канд. фіз.-мат. наук

¹ Міжнародний європейський університет,
Україна, 03127, Київ, пр-т Академіка Глушкова, 42
e-mail: oleksandr_nesterenko@ieu.edu.ua;

² Український науковий центр розвитку інформаційних технологій
Україна, 03127, Київ, пр-т Академіка Глушкова, 44
e-mail: inetesin@gmail.com

Експертне формування електронного глосарію

Запропоновано новий методологічний підхід до термінологічної систематики та підтримки понятійного апарату на основі електронного глосарію (ЕГ). Розроблено мета-модель даних і структуру глосарію з використанням понятійного апарату теорії множин. Наведено формалізований опис процедур, а також описано технологію (методику) експертного наповнення і супроводу контенту ЕГ. Розроблено критерії прийняття рішення експертами щодо включення термінів до ЕГ. На відміну від порядку формування термінологічних ресурсів існуючих ЕГ реалізовано алгоритм схвального голосування експертів для оцінки та вибору «найкращих» варіантів визначення термінів ЕГ. Показано можливості веб-інструментарію підтримки технології експертного наповнення і супроводу ЕГ. Подальший розвиток методичного апарату ЕГ пов'язаний з побудовою онтологічної моделі предметної області, яка сприятиме визначенню експертами несуперечливих термінів для кожного поняття та відношення, а також для термінів, пов'язаних між собою.

К л ю ч о в і с л о в а: термінологічний ресурс, глосарій, голосування експертів, веб-інструментарій.

В різних суспільних сферах, зокрема в науці, освіті, адміністративному управлінні особливе значення має підтримка понятійного апарату та приведення до єдиних визначень (тлумачень) термінів, які стосуються предметної діяльності. Цим процесам завжди приділялась певна увага, але наразі, за часів технологічних зрушень і цифрових трансформацій, коли суттєво зростає обсяг потрібної для опрацювання інформації, особливої актуальності набуває формування єдиного інформаційного простору різних напрямів діяльності [1].

Базове ядро такого простору зазвичай містить основні поняття та судження предметної області (ПдО). Це забезпечує подальший суворологічний перехід від абстрактних загальних понять і висловлювань до

© Нестеренко О.В., Нетесін І.Є., 2022

понять і тверджень нижчого рівня універсальності та абстрактності [2]. Основними елементами ядра є словники і глосарії. Словники відіграють важливу роль у покращенні розуміння основних понять.

Глосарій термінів — це понятійно-орієнтована сукупність термінологічних даних (термінологічний ресурс). Однак глосарії мають складний взаємозв'язок, який не можна повністю описати за допомогою стандартних підходів, таких як включення всіх термінів у лінійний алфавітний список. Крім того, що терміни важко піддаються дефінуванню, зазвичай при створенні та використанні глосаріїв виникають й інші проблеми, пов'язані з обсягом словникових статей, забезпеченням швидкості пошуку та відповідності змісту словника даному часу, довговічністю використання та ін.

Щоб подолати ці інформаційні обмеження останнім часом активно створюються електронні глосарії (ЕГ). При створенні таких глосаріїв можуть бути використані різні підходи та методи, найчастіше це прості спостереження, вибірки з текстів та інтерв'ю. У складних проєктах застосовуються семантичні дослідження різних текстових джерел на основі статистичних показників і моделей машинного навчання [3, 4].

Аналіз ЕГ, доступних у відкритих джерелах (деякі наведено у табл. 1), свідчать про те, що їх переважна більшість є або електронними словниками термінів у форматі pdf-файлів з відповідними пошуковими можливостями, або веб-сторінками з упорядкованими за абеткою термінами. Зазвичай інформація щодо розроблення структурної моделі глосарію в описовій частині не наводиться. Не наводиться також й інформація щодо наявності електронних інструментів для підтримки створення і актуалізації термінологічного ресурсу глосарію, про процедури прийняття рішень щодо включення терміну до глосарію, тощо.

Найпопулярнішою сферою застосування ЕГ вважається сфера освіти, зокрема у галузі викладання іноземних мов і культур [5], а також у мульти-лінгвальних дослідженнях [6]. Однак наразі ця технологія використовується більш широко — від класичних застосувань у лексикографії [7] до сільськогосподарських підприємств [8].

Аналіз публікацій свідчить, що поширенню ЕГ сприяє напрямок їх створення у вигляді веб-додатку з інтерактивним дизайном, щоб забезпечити автоматизоване наповнення глосарію та пошукові механізми для зручності у використанні. Однак у численних ПдО при створенні глосаріїв існує чимало проблем, пов'язаних з низкою факторів, які подолати поки ще не вдається навіть із застосуванням сучасних математичних методів та засобів штучного інтелекту.

Для багатьох ПдО існує міждисциплінарність у визначенні термінів, а також декілька їх визначень у міжнародних і вітчизняних стандартах та інших нормативних документах. При впровадженні термінології в прак-

тичну діяльність необхідно враховувати особливості організаційної структури відомства, специфічні організаційні процеси та процедури. Зрештою, необхідною є чітка регламентація визначень пов'язаних термінів, таких, як юридичні, фінансові, безпекові, лінгвістичні (професійний жаргон) та ін. Тому одним із шляхів при створенні глосаріїв вбачається залучення експертів з даної ПдО.

Експертні методи демонструють певну ефективність в задачах оптимізації та прийняття рішень різної інформаційної складності, зокрема для оцінювання і вибору альтернатив [9]. При створенні глосарію завдання полягає ще й у тому, щоб з застосуванням інформаційних технологій за-

Таблиця 1. Характеристики ЕГ

Назва ЕГ	Загальні відомості	Характеристика
Інформаційно-довідковий ресурс стандартизованої військової термінології http://www.mildic.com	Предметна область: військова термінологія <i>Мови глосарію:</i> українська, англійська	Методика формування: проекти термінів проходять лінгвістичну експертизу, реєструється в органі військової стандартизації, змістовне наповнення є предметом обговорення на науково-практичних конференціях, семінарах та інших наукових заходах
Глосарій термінів архівної справи та діловодства Товариства американських архіваріусів (A Glossary of Archival and Records Terminology)	Предметна область: архівна справа та діловодство у США та Канаді <i>Мови глосарію:</i> англійська	Технологія: веб-ресурс, текстовий формат Наявність моделі: структуру глосарію та його статей наведено в [4] Технологія: pdf-файл, текстовий формат
Глосарій онлайн-бібліотеки товариства «Вартова башта» https://wol.jw.org/uk/wol/publication/r15/lp-k/nwtstg	Предметна область: біблійні і небіблійні терміни <i>Мови глосарію:</i> українська та ін.	Технологія: веб-ресурс, пошук за абеткою
Klasyfikacja zaopatrzenia NATO (NSC – NATO Supply Classification) https://wcnjnk.wp.mil.pl/u/43H2_-_2021-12-01.pdf	Предметна область: військова термінологія <i>Мови глосарію:</i> польська	Технологія: pdf-файл, текстовий формат

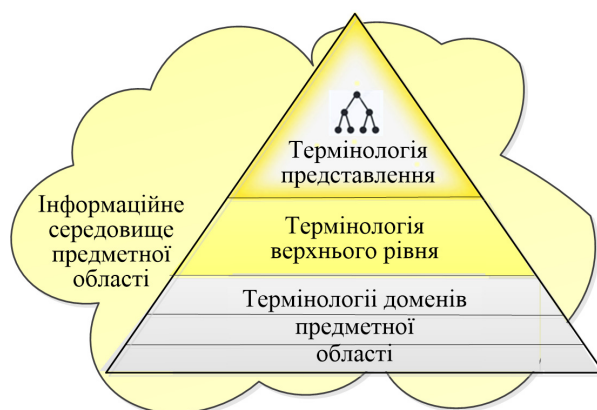


Рис. 1. Концептуальна модель термінологічного ресурсу ЕГ

безпечити ефективну співпрацю експертів, зокрема з урахуванням очевидної парадигми, а саме чим більше експертів беруть участь в опрацюванні інформаційних матеріалів, тим вищий рівень результатів можна очікувати.

Концептуальна модель термінологічного ресурсу ЕГ. Основним принципом формування глосарію є термінологічне відтворення даних з використанням термінів, якими позначене одне поняття, що надає можливість визначати синоніми і еквіваленти перекладу на іншу (наприклад, англійську) мову. Водночас, можливе застосування і лексичного подання даних при необхідності задати різні сенси або значення слова [10]. До глосарію включаються поняття, які узагальнюють множину предметів, об'єктів, сутностей, абстракцій з загальними ознаками, та концепти (екземпляри) — назви конкретних об'єктів. Отже, інформаційна картина ПдО зазвичай є значною за обсягом системою взаємодії понять і концептів, заданих за допомогою бінарних відношень.

Така система зазвичай має ієрархічну структуру шарів ранжування (рис. 1). Метамоделю термінологічної системи починається від абстракцій загальних понять і відношень між ними (термінологія представлення, або вища термінологія (Upper Terminology)). Формування термінології верхнього рівня є обов'язковою частиною загального алгоритму розробки глосарію. Наступний шар (Core Theories, термінологія ядра) відтворює загальні факти про простір і зв'язки, універсальні поняття і відношення, які є важливими для загальноприйнятих суджень у даній ПдО. Цей термінологічний шар орієнтований на створення моделей предметних областей у вигляді «моделей світів» — завершених, взаємопов'язаних і взаємообумовлених систем термінологічних знань.

Термінологія окремих частин ПдО (Domain-Specific Theories) описує специфіку ПдО і надає терміни, що стосуються окремих екземплярів (individuals) ПдО. Отже, у предметно-орієнтованих глосаріях на нижчому рівні використовуються конкретні, специфічні для даної ПдО, терміни і відношення, тобто термінологічна база цього рівня формується через конкретизацію понять, визначених у термінологіях ядра.

Глосарій складається з термінологічних записів, створених з елементів — категорій даних. Термінологічний запис — це упорядкована сукупність термінологічних даних, які стосуються одного поняття, екземпляра. До обов'язкових категорій термінологічного запису належать: терм україномовний; англomовний відповідник терму; визначення терму (тлумачення); джерело походження. Можливі включення й додаткових категорій, таких як аббревіатура (скорочення терму), синоніми та ін.

Терміни (або терми), які вибираються з множини існуючих і мають потрапити до глосарію, повинні відповідати області спеціалізації глосарію, тобто конкретній ПдО. Фактично тут постає задача оптимізації:

$$f(t) \rightarrow opt, t \in T, \quad (1)$$

де $f(t)$ — цільова функція; T — допустима множина термінів; t — термін як допустима точка задачі. Задача оптимізації формулюється як пошук оптимуму opt цільової функції в сенсі «найкращої» відповідності термінів області спеціалізації глосарію. Для забезпечення відповідності термінів області спеціалізації визначаються певні обмеження щодо включення термінів у вигляді множини термінів (множина O), які не повинні включатися.

На практиці поширені випадки, коли один і той же термін, вибраний з різних джерел, може мати різні визначення — стандартизовані і такі, що історично сформувалися в обігу в даній ПдО. Для запобігання дублювання до глосарію по кожному терму має потрапити лише один термінологічний запис, тому задача (1) перетворюється у задачу багатокритеріальної оптимізації. На основі гіпотетико-дедуктивної методології пропонується застосування трьох цільових функцій, які утворюють опис критерію задовільності для вибору остаточного визначення:

$$opt_{\vec{t}} \{f_1(\vec{t}), f_2(\vec{t}), f_3(\vec{t})\}, \vec{t} \in T, \quad (2)$$

де $f_1(\vec{t})$ — визначається безпосередньо сферою застосування глосарію і описується в правилах формування глосарію з використанням загальних понять термінології представлення та ядра; $f_2(\vec{t})$ — з декількох визначень терміну, які узгоджуються з практикою їх використання, переважає

визначення, яке є наближеним до міжнародних стандартів (ISO/IEC); $f_3(\vec{t})$ — з декількох визначень терміну, які можна застосувати, кращим є більш просте визначення.

Вочевидь, в цьому переліку одні цільові функції важливіші за інші, тоді критерій оптимальності можна визначити за лексикографічним порядком, який вимагає ранжування критеріїв у тому сенсі, що оптимізація за наступним критерієм можлива лише тоді, коли досягнуто оптимум для попередніх критеріїв. У даному випадку це означає, що перший критерій має найбільший пріоритет, і лише у випадку існування декількох розв'язків за цим критерієм здійснюється пошук розв'язків за другим, а потім і за третім критерієм.

Існування такої ієрархії (лексикографічного порядку) дозволяє розв'язувати задачу послідовно, крок за кроком оптимізуючи за кожним наступним критерієм та використовуючи оптимальні значення попередніх критеріїв як обмеження.

Множина термінів, які розташовані в області спеціалізації глосарію та можуть бути включеними до глосарію за першим критерієм (T_1), є власною підмножиною множини I термінів, що існують в інформаційному просторі даної ПдО ($T_1 \subset I$). Множину T термінів з їх остаточними визначеннями з урахуванням другого та третього критеріїв (T_2, T_3), і обмежень O можна подати у графічному вигляді (рис. 2), а також у вигляді формули

$$T = (\overline{T_2 \cup T_3} \cup O) \setminus T_1. \quad (3)$$

Оцінювання ефективності наведеного підходу до вибору термінів залежить від того, про що свідчать отримані дані за межами формальної задачі моделювання. Це зумовлює проблему вибору методу, за яким має здійснюватися оцінювання. Вочевидь, таким методом може бути розвідувальний аналіз даних як один з підходів до аналізу наборів даних для узагальнювання їхніх основних характеристик, зокрема з використанням статистичної моделі та за допомогою візуалізації. Його застосування можливе на існуючому наборі даних, тобто після формування глосарію (розгляд цього підходу заслуговує окремої публікації).

Метод експертного оцінювання і вибору визначення терміну. Як було зазначено, для розв'язання задачі формування глосарію пропонується надати перевагу експертному методу. Найбільш розповсюдженим способом прийняття колективного рішення, зокрема в експертних групах, які створюються для вирішення деяких спеціалізованих питань, є голосування. В літературі описано й в світовій практиці використовується багато десятків процедур голосування, які істотно відрізняються одна від одної.

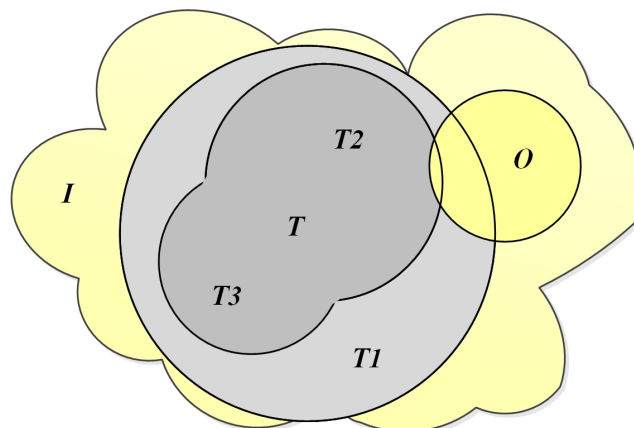


Рис. 2. Концептуальна модель належності терміну до області спеціалізації глосарію

Правильний вибір процедури голосування сприяє отриманню узгодженого рішення, однак виділити з існуючих кращу процедуру, що має переваги у порівнянні з іншими, або розділити всі ситуації, що виникають, на типові і для кожної з них вказати кращу процедуру, на жаль, неможливо. Процедури голосування, навіть якщо вони зовні здаються простими, насправді є складними і витонченими способами прийняття рішення та узгодження інтересів. Кожна з таких процедур має свої переваги і недоліки, є «зручною» або «незручною» в конкретних умовах. Крім того, критерії, за якими потрібно оцінювати і порівнювати процедури, є неочевидними і продовжують бути предметом дискусій і досліджень.

Експертний вибір через голосування (процедура голосування) має наступні особливості:

- 1) будь-яким чином формується набір варіантів (альтернатив), щодо яких має бути прийняте рішення;
- 2) кожен учасник процедури формує свою думку про ці варіанти і відображає її згідно з інструкцією;
- 3) відповідно до деякого формального регламенту ця інформація обробляється і визначається колективне рішення.

Численні процедури голосування відрізняються одна від одної значенням, яке вкладається в кожен з цих трьох пунктів. У процедурах голосування можливі принаймні три варіанти інструкцій для учасників:

- 1) провести дихотомію всіх альтернатив, тобто висловитись «згоден» або «не згоден»;
- 2) поділити всі альтернативи на три групи: «згоден», «не згоден», «утримуюся»;

3) впорядкувати всі альтернативи, наприклад розташувати їх у порядку переваги без права (або з правом) вказувати однаково кращі варіанти, розбити всі варіанти на більшу кількість типів та ін.

Зазвичай результатом колективного вибору повинен бути один варіант. Однак практично кожна процедура голосування може привести до вибору декількох (кращих згідно з цією процедурою) варіантів. В цьому випадку проводиться додаткове голосування або приймається окреме рішення, наприклад в результаті обговорення або уточнення (змін) правил голосування.

З численного переліку існуючих процедур на практиці найчастіше використовуються лише кілька стандартних, а саме метод відповідної більшості, метод схвального голосування, метод Борда, метод Кондорсе, метод аналізу ієрархій, які детально описані в літературі, наприклад [11]. Аналізуючи ці методи, можна зробити висновок про неоднозначність результатів голосування (вибору) за різними методами. Розглянемо такий приклад.

Нехай є група з 21 експерта, які повинні зробити вибір з чотирьох альтернатив: *A*, *B*, *C*, *D*. Результати ранжування від «найкращої» до «найгіршої» альтернативи наведено в табл. 2.

За методом відповідної більшості перемагає альтернатива *A*, тобто вісім голосів, попри те, що для явної більшості експертів (13 голосів) *A* є найгіршим варіантом. За методом Борда перемагає альтернатива *B* (сім експертів по три бали, дев'ять по два бали та п'ять експертів по одному балу, що дорівнює 44 балам). За методом Кондорсе перемагає альтернатива *C*, тому що ця альтернатива перемагає усі інші (у кожному рядку вона стоїть перед однією з множини альтернатив).

В результаті проведеного аналізу оцінки та вибору пропозицій термінологічних записів для включення у глосарій основним обрано метод модифікованого схвального голосування (рис. 3). За цим методом кожний експерт має право не лише підтримати одну або декілька пропозицій та утриматись від рішення, але й відхилити будь-яку кількість альтернатив (голосування проти). Кожен варіант голосування оцінюється за баль-

Таблиця 2. Приклад голосування експертів

Кількість експертів	Результати ранжування			
3	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
5	<i>A</i>	<i>C</i>	<i>B</i>	<i>D</i>
7	<i>B</i>	<i>D</i>	<i>C</i>	<i>A</i>
6	<i>C</i>	<i>B</i>	<i>D</i>	<i>A</i>

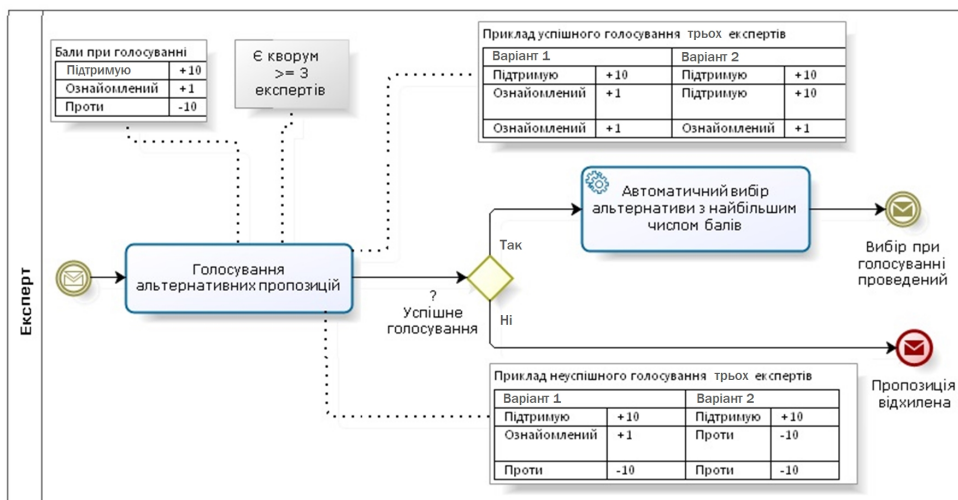


Рис. 3. Формалізований опис процедури схвального голосування

ною шкалою. Підрахунок результатів виконується за умови, що в процедурі оцінки брали участь не менше трьох експертів. У разі необхідності врахування кваліфікації експертів в залежності від сфер їх компетентності та досвіду роботи в даній предметній ПдО можливе застосування відповідних вагових коефіцієнтів.

Цей метод дозволяє експертам приймати рішення, найближче до консенсусного, ніж інші методи. Додатково пропонується метод відповідної більшості або метод аналізу ієрархій.

Технологія експертного наповнення ЕГ. Спочатку формується експертна група, до складу якої доцільно залучати представників різних рівнів діяльності, намагаючись як найширше охопити усю ПдО. На одного з експертів групи покладають функції відповідального експерта. Діяльність експертів з застосуванням комп'ютерного інструментарію може відбуватись у будь-який зручний для них час (24/7), регламентується лише тривалість роботи (наприклад, місяць або ін.). Склад експертної групи упродовж визначеного строку може доповнюватися (змінюватися), чим керує відповідальний експерт.

Затверджені правила формування термінологічного ресурсу ЕГ визначають основні функції експертів та додаткові функції відповідального експерта. Згідно з ними у будь-який час експерт може включити свою пропозицію до таблиці термінів та має можливість запропонувати обговорення терміну на внутрішньому форумі експертної групи з метою отримати рекомендації з удосконалення своєї пропозиції. Під пропозицією розуміється термін з його визначенням, який експерт вважає за доцільне

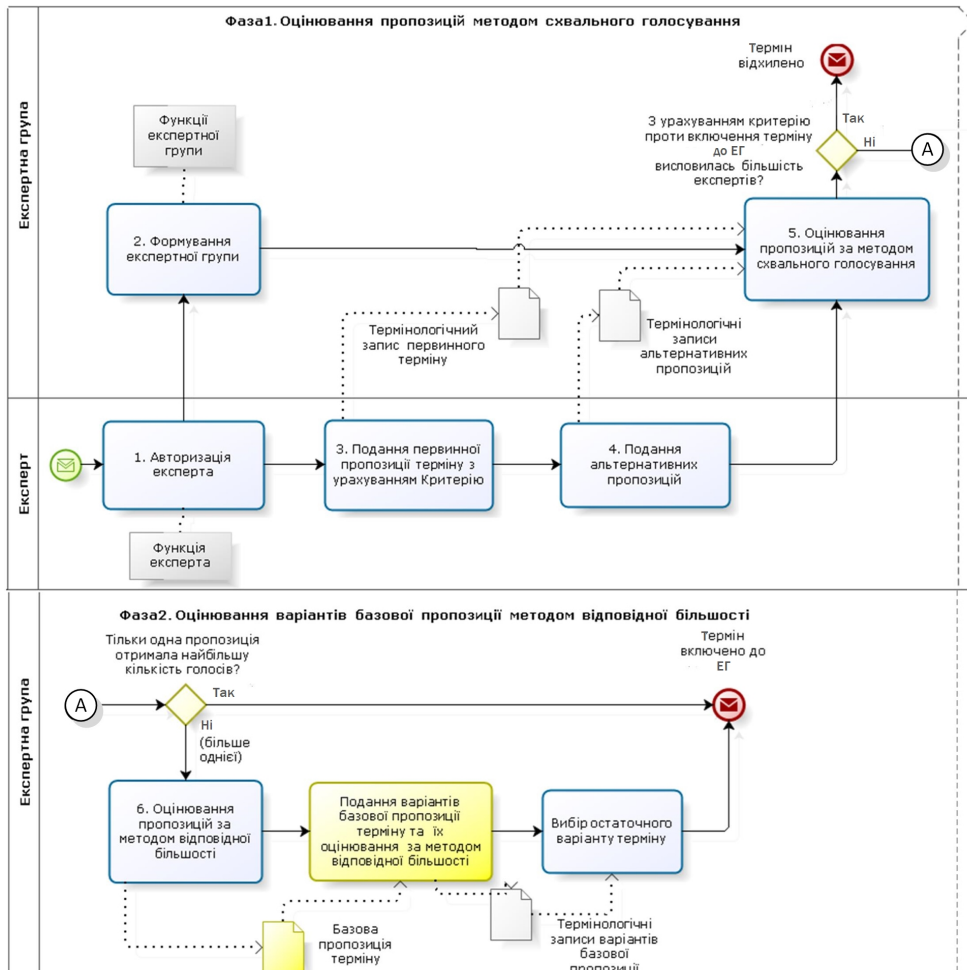


Рис. 4. Формалізований опис технологічного процесу наповнення та актуалізації контенту ЕГ (верхній рівень): 1—6 — етапи експертного наповнення

включити до глосарію. Водночас, кожен експерт у будь-який час може зняти свою пропозицію з розгляду, а також підключитися до оцінювання інших пропозицій на будь-якій фазі.

Таким чином, технологія експертного наповнення ЕГ складається з декількох етапів (рис. 4):

- 1 — подання первинної пропозиції щодо включення нового терміну до ЕГ;
- 2 — подання альтернативних пропозицій;
- 3 — оцінювання пропозицій;
- 4 — перша та друга фази вибору базової пропозиції терміну;

5 — подання варіантів базової пропозиції;

6 — визначення остаточного варіанту терміну.

Після завершення етапів 2 і 3 для кожної пропозиції здійснюється підрахунок поставлених експертами оцінок. На першій фазі методом схвального голосування вибраною вважається пропозиція терміну, яка набрала найбільшу різницю між сумарними кількостями оцінок «підтримати» та «проти» (див. рис. 3). Якщо на першій фазі однакову найбільшу кількість оцінок «підтримати» отримали більше, ніж одна пропозиція, здійснюється друга фаза вибору — експерти оцінюють вибрані на першій фазі пропозиції за методом відповідної більшості. Вибраною на другій фазі вважається пропозиція терміну, яка є найбільш прийнятною для найбільшої кількості експертів (набрала найбільшу кількість голосів «підтримати»).

Вибрана пропозиція отримує статус базової. Якщо більшість експертів висловились проти всіх пропозицій, даний термін не включається до ЕГ. Експерти, для яких базова пропозиція не була найбільш прийнятною, мають можливість подати нові варіанти цієї пропозиції з урахуванням коригувань, які, на їх думку, покращують формулювання обраної базової пропозиції. Після цього ще раз повторюється цикл оцінки та вибору, як під час другої фази.

У разі необхідності остаточне рішення щодо вибору терміну для включення або не включення до ЕГ приймається відповідальним експертом за результатами обговорення на спеціально створеному форумі.

Веб-інструментарій ЕГ. Веб-інструментарій підтримки технології експертного супроводу ЕГ реалізовано у вигляді веб-ресурсу в режимі віддаленого доступу. Веб-інструментарій підтримує рольовий механізм доступу до контенту кінцевими користувачами, експертами та персоналом, що здійснює адміністрування ЕГ. Особливістю інструментарію є шаблони для подання експертами первинних та альтернативних пропозицій термінів, їх редагування та видалення. Шаблони надають можливість подавати термінологічний запис з повним або скороченим набором категорій у зручній для користувача формі. Веб-інструментарій реалізовано програмними засобами, що вільно розповсюджуються.

Для роботи експертів інтерфейс інструментарію ЕГ складається з функціонального, інструментального та інформаційного блоків меню. Функціональний блок (рис. 5, а) дозволяє додавати рядки в таблиці термінів для внесення нового термінологічного запису, створювати дублі вже існуючого термінологічного запису, редагувати та видаляти термінологічні записи, а також здійснювати пошук інформації будь-якої категорії.

Інструментальний блок (рис. 5, б) надає широкі можливості з налаштування інформаційної таблиці та форми редагування, визначення сфери

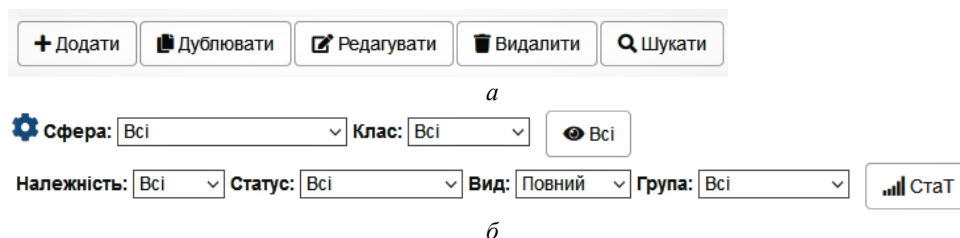


Рис. 5. Функціональний (а) та інструментальний (б) блоки меню інтерфейсу ЕГ

Назва терміну	Англ. термін	Визначення	Джерело
104 Придбання	Acquisition	Процес отримання системи, продукту або послуги [ДСТУ ISO/IEC/IEEE 15288:2016] Process of obtaining a system, product or service [ISO/IEC/IEEE 15288:2015, IDT]	ДСТУ ISO/IEC/IEEE 15288:2016
106 Угода	Agreement	Взаємне визнання положень й умов щодо робочих відносин Приклад. Контракт, меморандум про домовленість [ДСТУ ISO/IEC/IEEE 15288:2016]	ДСТУ ISO/IEC/IEEE 15288:2016

Рис. 6. Інформаційний блок ЕГ

використання терміну та рівня метамоделі термінів, перегляду термінологічних записів у різних розрізах та формах подання таблиці, перегляду статистики голосування експертів по конкретному терміну, тощо.

Інформаційний блок (рис. 6) являє собою таблицю термінологічних записів термінів з усіма категоріями. Кожен рядок таблиці є термінологічним записом і містить інформацію щодо одного терміну. За замовчуванням на екрані розгортається таблиця термінів з назвами термінів та їх категорій в повному вигляді (обов'язкові, додаткові та адміністративні категорії) і має 15 полів.

Для перегляду всіх термінологічних записів необхідно здійснювати перехід по сторінках таблиці. Кількість записів і сторінок представлення термінологічних записів залежить від налаштувань таблиці, встановлених в інструментальному блоці в меню «Шестерня». Всі назви термінів таблиці мають гіперпосилання, при натисканні на які експерту надається можливість працювати не з таблицею усіх термінологічних записів, а з таблицею, в якій надано термінологічні записи, що стосуються тільки обраного терміну (всі альтернативні пропозиції по даному терміну).

Експорт контенту ЕГ реалізований у форматах .doc(.docx), .pdf, .html.

Оцінка голосувань, які здійснювали експерти упродовж дня, виконується автоматично в нічний час один раз на добу. Така оцінка виконується послідовно для кожного пула (куща) альтернативних пропозицій термінологічних записів з однаковою назвою терміну.

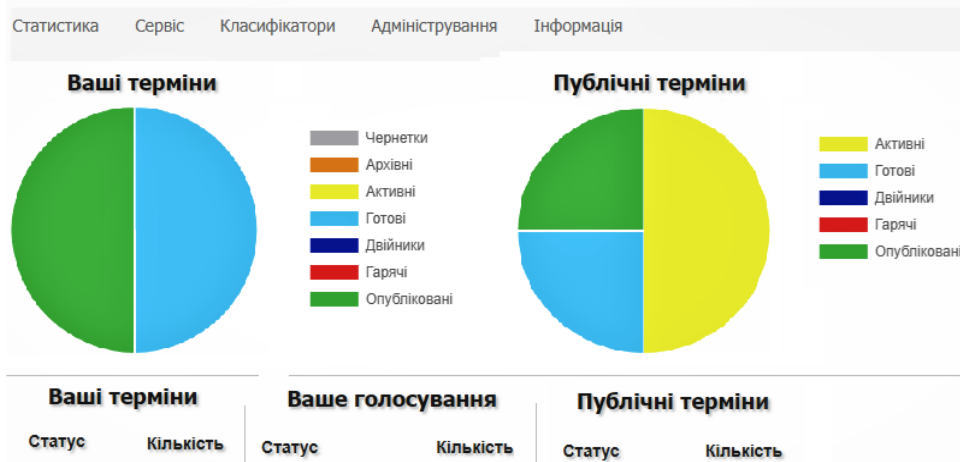


Рис. 7. Діаграмна форма «Статистика»

На сторінці «Статистика» можливо переглянути кругові діаграми зі статистикою персональних термінів і статистикою всіх термінів (рис. 7). На цій же сторінці присутня статистика й у табличній формі (на рис. 7 не показана). Натиснувши на будь-який статус у таблиці, експерт може перейти на сторінку термінів з відповідним статусом.

Подальший розвиток методики експертного наповнення і супроводу ЕГ вбачається у побудові такого системного формалізованого опису термінологічного ресурсу глосарію, який би дозволяв аналізувати взаємозв'язки понять. Це має створити умови експертам продукувати несуперечливі визначення термінів для понять і відношень між ними, визначення інших термінів, які пов'язані з цим терміном. Таким формалізованим описом контенту глосарію може бути його онтологічна модель.

Порівнюючи отримані результати з іншими рішеннями, слід зазначити, що на відміну від порядку формування термінологічних ресурсів існуючих ЕГ у даному випадку запропоновані система критеріїв відбору термінів до глосарію, алгоритм схвального голосування експертів для оцінки та вибору «найкращих» варіантів визначення термінів ЕГ та розроблена відповідна технологія роботи експертів. Створений та апробований веб-інструментарій підтримки роботи експертної групи реалізує технологію експертного наповнення і супроводу ЕГ, дозволяє організувати колективну роботу експертів у комфортному віддаленому режимі роботи. Його функціональність на основі алгоритму схвального голосування дозволяє приводити різні визначення одного й того ж терміну до спільного, забезпечує постійне удосконалювання понятійного апарату та підтримку термінологічного ресурсу в актуальному стані. Для пересічних користувачів створюються умови зручності та оперативності пошуку інформації.

Висновки

Формування єдиного інформаційного простору, що є актуальним для різних сфер діяльності, передбачає забезпечення інтеперабельності різних складових — організаційних, юридичних, фінансових, а також інформації комп'ютерних систем. Для цього необхідна побудова ієрархії стандартів, регламентів та інших нормативно-технічних документів, а також формування та підтримка в актуальному стані термінологічного апарату, який використовується у практичних процесах. Одним із шляхів вирішення зазначеної проблеми інтеперабельності є створення ЕГ.

Теоретичне значення проведеного дослідження пов'язане з визначенням нових підходів до створення таких глосаріїв, які полягають у застосуванні експертної діяльності для оцінки та вибору «найкращих» варіантів визначення термінів на основі процедур схвального голосування та запропонованої системи критеріїв.

Практичне значення роботи визначається створенням на основі отриманих теоретичних результатів та апробованим веб-інструментарієм підтримки роботи експертної групи, який має універсальний характер та може застосовуватись в різних галузях.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. *Nesterenko O., Trofymchuk O.* Patterns in forming the ontology based environment of information-analytical activity in administrative management // *Eastern-European Journal of Enterprise Technologies*, 2019, № 5/2 (101), pp. 33—42. <http://doi.org/10.15587/1729-4061.2019.180107>
2. *Lupenko S.A., Pasichnyk V.V., Kusanets N.E.* Axiomatic-deductive strategy for IT discipline content formation // *Information Technologies and Learning Tools*, 2019, Vol. 73, № 5, pp. 149—160. <https://doi.org/10.33407/itlt.v73i5.2536>
3. *Navigli R., Velardi P.* GlossExtractor: A web application to automatically create a domain glossary // *10th Congress of the Italian-Association-for-Artificial-Intelligence. Artificial intelligence and human-oriented computing*, 2007, 4733, pp. 339—349. https://doi.org/10.1007/978-3-540-74782-6_30
4. *Кульчицький І.М.* Електронний глосарій з архівістики // *Зб. наук. праць «Інформаційні системи та мережі»*. Відп. редактор В.В. Пасічник. Львів: Вид-во Львівської політехніки, 2009, вип. 653, № 1, с. 148—156.
5. *Fabrychna Y.* Bilingual e-glossary of foreign languages and cultures teaching // *12th International Conference on Open Education as a Way to a Knowledge Society (DisCo)*. (June 26-27, 2017). Microsoft, Prague, Czech Republic, 2017, pp. 264—272.
6. *Plechko A., Chukhno T., Nikolaieva T., Apolonova L., Leleka T.* Anglicisms in the Field of Information Technology: Analysis of Linguistic Features. *International Journal of Computer Science and Network Security*, 2022, 22, no. 4, p. 183—92. <https://doi.org/10.22937/IJCSNS.2022.22.4.23>
7. *Elgendi E., Shaffer C.A.* Dynamic concept maps for etextbook glossaries: design and evaluation // *Frontiers in computer science*, 2020, Vol. 2, Article 7. <http://doi.org/10.3389/fcomp.2020.00007>

8. Jordan I., Heil E., Keding G. Uniform access to technical terms in nutritional and agricultural sciences. Research projects: an interactive glossary // *Ernahrungs umschau*, 2021, 68 (10), pp. M572-M577.
9. Nesterenko O., Netesin I., Polischuk V., Selin Y. Multifunctional Methodology of Expert Evaluation Alternatives in Tasks of Different Information Complexity // *Proceedings of 2021 IEEE 3rd International Conference on Advanced Trends in Information Theory (ATIT)*, 2021, pp. 226—231. <http://doi.org/10.1109/ATIT54053.2021.9678742>
10. Широков В.А. Елементи лексикографії. Київ: Довіра, 2005, 304 с.
11. Гнатієнко Г.М., Снитюк В.С. Експертні технології прийняття рішень. Київ: ТОВ «Маклаут», 2008, 444 с.

Отримано 05.06.22;
після доопрацювання 03.07.22

REFERENCES

1. Nesterenko, O. and Trofymchuk, O. (2019), “Patterns in forming the ontology-based environment of information-analytical activity in administrative management”, *Eastern-European Journal of Enterprise Technologies*, Vol. 5/2, no. 101, pp. 33-42, available at: <http://doi.org/10.15587/1729-4061.2019.180107>.
2. Lupenko, S.A., Pasiichnyk, V.V. and Kunanets, N.E. (2019), “Axiomatic-deductive strategy for IT discipline content formation”, *Information Technologies and Learning Tools*, Vol. 73, no. 5, pp. 149-160. <https://doi.org/10.33407/itlt.v73i5.2536>.
3. Navigli, R. and Velardi, P. (2007), “GlossExtractor: A web application to automatically create a domain glossary”, *Proceedings of the 10th Congress of the Italian-Association-for-Artificial-Intelligence. Artificial intelligence and human-oriented computing*, Vol. 4733, pp. 339-349, available at: https://doi.org/10.1007/978-3-540-74782-6_30.
4. Kulchytskyi, I.M. (2009), “Electronic glossary of archival science”, *Informatsiyeni systemy ta merezhi*, Vol. 653, no. 1, pp. 148-156.
5. Fabrychna, Y. (2017), “Bilingual e-glossary of foreign languages and cultures teaching”, *Proceedings of the 12th International Conference on Open Education as a Way to a Knowledge Society (DisCo)*. Microsoft, Prague, Czech Republic, June 26-27, 2017, pp. 264-272.
6. Plechko, A., Chukhno, T., Nikolaieva, T., Apolonova, L. and Leleka, T. (2022), “Anglicisms in the Field of Information Technology: Analysis of Linguistic Features”, *International Journal of Computer Science and Network Security*, Vol. 22, no. 4, pp. 183–92, available at: <https://doi.org/10.22937/IJCSNS.2022.22.4.23>.
7. Elgendi, E. and Shaffer, C.A. (2020), “Dynamic concept maps for etextbook glossaries: design and evaluation”, *Frontiers in computer science*, Vol. 2, available at: <http://doi.org/10.3389/fcomp.2020.00007>.
8. Jordan, I., Heil, E. and Keding, G. (2021), “Uniform access to technical terms in nutritional and agricultural sciences. Research projects: an interactive glossary”, *Ernahrungs umschau*, Vol. 68, no. 10, pp. M572-M577.
9. Nesterenko, O., Netesin, I., Polischuk, V. and Selin, Y. (2021), “Multifunctional Methodology of Expert Evaluation Alternatives in Tasks of Different Information Complexity”, *Proceedings of 2021 IEEE 3rd International Conference on Advanced Trends in Information Theory (ATIT)*, pp. 226-231, available at: <http://doi.org/10.1109/ATIT54053.2021.9678742>
10. Shyrokov, V.A. (2005), *Elementy leksykohrafiyi* [Elements of lexicography], Dovira, Kyiv, Ukraine.
11. Hnatiienko, H.M. and Snytiuk, V.Ie. (2008), *Ekspertni tekhnolohiyi pryynyattya rishen* [Expert decision making technologies], Kyiv, Ukraine.

Received 05.06.22;
after revision 03.07.22

О.В. Nesterenko, I.E. Netesin

EXPERT FORMATION OF E-GLOSSARY

A new methodological approach to the terminological systematics and support of the conceptual apparatus based on the electronic glossary is proposed. A metamodel of data and a glossary structure using the conceptual apparatus of set theory have been developed. A formalized description of procedures, as well as technology (methodology) of expert content and support of e-glossary content is given. Decision-making criteria have been developed for use by experts in including terms in the glossary and searching for their definitions. In contrast to the procedure for forming the terminological resources of existing e-glossaries, the approval voting algorithm of experts for evaluation and selection of the "best" options for determining the terms of the e-glossary has been implemented. Possibilities of web tools of support of technology of expert filling and support of e-glossary are shown. The proposed tools are universal and can be used in various fields. Further development of the methodological apparatus of the e-glossary is associated with the construction of an ontological model of the subject area, which will help experts to determine consistent terms for each concept and relationship, as well as for interrelated terms.

Key words: terminological resource, glossary, expert voting, web tools.

НЕСТЕРЕНКО Олександр Васильович, д-р техн. наук, доцент, ст. дослідник, зав. кафедри інформаційних технологій Міжнародного європейського університету. В 1975 р. закінчив Київський політехнічний інститут. Область наукових досліджень — інформаційні технології підтримки прийняття рішень, програмна інженерія, кібербезпека.

НЕТЕСІН Ігор Євгенович, канд. фіз.-мат. наук провідний науковий співробітник Українського наукового центру розвитку інформаційних технологій. У 1985 р. закінчив Московський державний університет. Область наукових досліджень — інформаційні технології підтримки прийняття рішень, технології програмування.