

МЕТОДОЛОГІЯ РОЗРОБКИ ТА ПРИНЦИПИ ПОБУДОВИ КОМАНДНОГО ТЕЗАУРУСУ

TEAM THESAURUS DEVELOPMENT METHODOLOGY AND ITS BUILDING PRINCIPLES

Кармазіна Л.Л.,

orcid.org/0000-0002-8639-9678

кандидат технічних наук, доцент,

доцент кафедри філології та перекладу

Українського державного хіміко-технічного університету

В статті розглядається систематичний підхід до побудови командного тезаурусу з метою ефективного здійснення процесів збору, обробки та використання знань у колективі. В даному дослідженні командний тезаурус є зібраним та організованим словником термінів, який використовується в спілкуванні та роботі команди або групи людей над спільним проектом. Основна мета його створення полягає в уніфікації термінології та понять в команді, що дозволяє ефективно спілкуватися та передавати знання між її членами. Командний тезаурус являє собою результат поєднання тезаурусів різних видів, таких як тезаурус предметної галузі, тезаурус інформаційного ресурсу, тезаурус експерта та тезаурус користувача. Спершу команді необхідно визначити, що може бути об'єктами тезаурусу. Це можуть бути терміни, поняття, скорочення, показники, фрази, методики, технології тощо. Після визначення об'єктів тезаурусу, команда повинна уважно вивчити їх і визначити, які з них є ключовими та потребують бути включеними до тезаурусу. Після того, як команда визначить терміни, які можуть бути включені до командного тезаурусу, необхідно їх зібрати та визначити взаємозв'язки між ними. Налагодження зв'язків між термінами дозволяє швидко знаходити пов'язану інформацію та переходити між поняттями, покращити організацію знань та взаємозв'язків між ними, допомагає визначити і зафіксувати структуру знань, що є важливим для їх подальшої організації та збереження. В статті описані декілька формальних вимог до визначень термінів у тезаурусі, а також наведені правила користування тезаурусом, які можуть бути встановлені в команді. Розробка командного тезаурусу є необхідним етапом у колективній роботі, оскільки це дозволяє забезпечити єдину термінологію та покращити спілкування між членами команди.

Ключові слова: тезаурус, командний тезаурус, термін, зв'язки між термінами, правила використання термінів в командному тезаурусі.

The article considers a systematic approach to building a team thesaurus aimed at effective collection, processing, and use of knowledge in a team. In this study, a team thesaurus is a collected and organized dictionary of terms used in communication and work of a team or group of people on a joint project. Its main purpose is to standardize the terminology and concepts within the team, which allows effective communication and knowledge sharing among its members. The team thesaurus is the result of combining various types of thesaurus, such as a subject field thesaurus, an information resource thesaurus, an expert thesaurus, and a user thesaurus. First, the team needs to determine what the thesaurus objects might be. These can be terms, concepts, abbreviations, indicators, phrases, techniques, technologies, etc. After identifying the thesaurus objects, the team should carefully examine them and determine which ones are key and need to be included in the thesaurus. After the team has identified the terms that can be included in the team thesaurus, it is necessary to collect them and determine the relationships between them. Establishing connections between terms allows for quick access to related information and enables easy navigation between concepts, improving the organization of knowledge and their relationships, and helps to determine and document the structure of knowledge, which is important for its further organization and preservation. The article describes several formal requirements for defining terms in a thesaurus, as well as rules for using the thesaurus that can be established within a team. Developing a team thesaurus is a necessary step in collaborative work, as it allows to ensure uniform terminology and improve communication between team members.

Key words: thesaurus, team thesaurus, term, relationships between terms, rules for using terms in a team thesaurus.

Постановка проблеми. Зараз створення інформаційних систем, що базуються на знаннях в конкретній галузі, є важливим і необхідним для організацій. Це пов'язано з тим, що діяльність організації залежить від наявності знань, які є найціннішим нематеріальним ресурсом, що належить організації [1]. Однак, сучасні інформаційні системи не є досить досконалими для вирішення питання з управління знаннями в організаціях, які б дозволяли ефективно збирати, обробляти та використовувати знання, що є володінням органі-

зації. Більшість існуючих інформаційних систем здатні збирати та зберігати дані, але недостатньо уваги приділяється їх обробці та використанню з метою створення нових знань. Для ефективного управління знаннями необхідно розробити методологію створення командного тезаурусу, який би дозволяв стандартизувати терміни та поняття в галузі, сприяв би збільшенню обсягу знань та підвищенню їх якості.

Аналіз останніх досліджень і публікацій. На сьогодні існує велика необхідність в розробці

таких тезаурусів, які б на рівні світових стандартів задовільнили потреби інформатизованих користувачів в систематизації термінів в межах певної предметної галузі. Сучасний рівень розвитку інформаційних технологій та досвід складання тезаурусів (словників) дозволяють створювати не тільки загальномовні, а й вузькопрофільні тезауруси. Вагомий внесок в теорію побудови тезаурусу зробили Ш. Баллі, Шведова, Л. В. Щерба, Ю. М. Караулов, Л. Г. Бабенко, Ю. Д. Апресян, О. С. Баранов, та ін. До плеяди українських науковців, які займалися дослідженням проблематики тезаурусів належать В. А. Широков, Н. П. Дарчук, В. В. Дубічинський, І. М. Гетьман, Н. В. Сніжко, М. Д. Сніжко, А. Гладун та ін. Щодо популярних тезаурусів, які функціонують та розвиваються в інформаційній мережі, це Принстонський тезаурус WordNet, MindNet, FrameNet, VerbNet, HowNet, ConceptNet, EuroWordNet, та багато інших [2] Кожен з цих тезаурусів має свої переваги та недоліки, а також використовується у певних галузях знань. Наприклад, WordNet є одним з найбільш популярних тезаурусів, який зосереджений на лексичному аналізі мови та містить значення слів та їх синоніми, MindNet досліджує ментальні концепти та їх зв'язки, FrameNet орієнтується на семантику рамок (frames), а VerbNet фокусується на граматичних та семантичних відношеннях між дієсловами. Ці тезауруси стали популярними завдяки своїй високій якості та точності, що дозволяє використовувати їх у різних галузях, наприклад, в природних мовах обробки, інформаційному пошуку, машинному навчанні та багатьох інших. Однак, варто зазначити, що ці тезауруси не є універсальними і не підходять для всіх мов та галузей знань, тому дослідження та розробка нових тезаурусів є важливою задачею в області управління знаннями.

Постановка завдання. Командні тезауруси є важливим інструментом для управління знаннями в організації, але на жаль, системне дослідження та розробка командних тезаурусів певної предметної галузі поки що не отримують достатньої уваги. Системне дослідження командних тезаурусів передбачає не лише вивчення їх структури та змісту, а й визначення принципів їх побудови, підходів до відбору термінів та понять, а також методів контролю якості. На сьогоднішній день відсутні єдині стандарти створення командних тезаурусів, тому вироблення методології є кроком до розв'язання проблеми управління знаннями в організації. Завданням даного дослідження є розробка систематичного підходу до побудови командного тезаурусу з метою ефек-

тивного здійснення процесів збору, обробки та використання знань у колективі, що дозволить позитивно вплинути на діяльність організації в цілому.

Виклад основного матеріалу. Однією з основних причин поганого комунікаційного процесу в команді є семантичний бар'єр. Це пов'язано з тим, що не всі учасники команди мають однаковий досвід та розуміння предмета, що призводить до використання різного понятійного апарату. Для досягнення взаєморозуміння в команді необхідно мати спільний тезаурус, який визначатиме терміни та їх відношення в межах даної області знань.

Створення командного тезаурусу є важливим етапом в управлінні знаннями та підвищенні ефективності комунікації між членами команди. Це є важливою задачею для будь-якої організації, що має багато співробітників та проєктів. Але перед тим, як перейдемо до командного тезаурусу, дамо загальне визначення поняттю «тезаурус».

Флорентійський учений Брунето Латіні був першим, хто використав термін «тезаурус» в назві своєї систематизованої енциклопедії «Книга про скарб». Цей термін відповідав семантиці слова «скарбниця» або «багатство» [3].

Термін «тезаурус» походить від грецького слова «θησαυρός» (скарб, запас) і використовується в лінгвістиці для опису словника, що містить лексичні одиниці конкретної мови разом з їхніми семантичними значеннями та зв'язками між ними. У відміню від термінологічних словників, енциклопедій і глосаріїв, які зазвичай розташовують слова за алфавітом, тезаурус має тематичну структуру. Розміщення слова в тезаурусі визначається його значенням, семантикою та ступенем близькості до інших слів, з якими воно утворює семантичні групи [4].

В даному дослідженні під «командним тезаурусом» будемо розуміти зібраний та організований словник термінів, що використовуються в певній команді або групі людей для спілкування та роботи над спільним проєктом. Основна мета командного тезаурусу полягає в тому, щоб уніфікувати термінологію та поняття в команді, що дозволить ефективно спілкуватися та передавати знання між членами команди. При цьому, тезаурус повинен бути створений у зручній формі для використання, з врахуванням специфіки робочого процесу та потреб користувачів.

Окрім того, командний тезаурус може бути корисним інструментом для навчання нових членів команди, допомагаючи їм швидко ознайомитися з термінологією та поняттями відповідної предметної галузі. Також, з його допомогою

можна підвищити якість документації та забезпечити її стандартизацію, що сприятиме збереженню знань та передачі їх між співробітниками при зміні складу команди.

Командний тезаурус є результатом поєднання тезаурусів різних видів, які використовуються в організації: тезаурус предметної галузі, тезаурус інформаційного ресурсу, тезаурус експерта, тезаурус користувача. Зокрема, він базується на тезаурусі предметної галузі, який містить терміни та поняття, що відносяться до конкретної професійної галузі знань. Також використовується тезаурус інформаційного ресурсу, що містить терміни та поняття, пов'язані з конкретними джерелами інформації. Тезаурус експерта включає терміни та поняття, які використовуються фахівцем в даній галузі. Тезаурус користувача містить терміни та поняття, які належать тезаурусам учасників команди проєкту. Поєднання цих тезаурусів дозволяє створити комплексний тезаурус для забезпечення спільної мови між учасниками команди, що сприяє підвищенню ефективності їх роботи.

Далі визначимо основні принципи та правила створення командного тезаурусу. Спершу команді необхідно визначити, що може бути об'єктами тезаурусу. Це можуть бути терміни, поняття, скорочення, показники, фрази, методики, технології тощо (далі «терміни»).

Після визначення об'єктів тезаурусу, команда повинна уважно вивчити їх і визначити, які з них є ключовими та потребують бути включеними до тезаурусу. Важливо також визначити критерії відбору термінів і провести контроль їх якості. Необхідно забезпечити стандартизацію термінології та дотримання міжнародних стандартів, щоб забезпечити зрозумілість термінів та їхню узгодженість.

Після того, як команда визначить терміни, які можуть бути включені до командного тезаурусу, необхідно їх зібрати та визначити взаємозв'язки між ними. Для цього можуть використовуватись різні методи, наприклад, інтерв'ювання експертів в галузі, аналіз наукових досліджень, перегляд літератури та інших джерел.

Наступним етапом є встановлення зв'язків між термінами. Для прикладів візьмемо галузь інформаційних технологій (ІТ). Основні типи зв'язків, які можна встановлювати між термінами тезаурусу, включають:

1) родо-видові відношення: це гіперон-гіпонімічна група, до якої входять терміни, об'єднані відношеннями ієрархічної супряженості (наприклад, програмування: Java, Python, C++, JavaScript, і т.д.);

2) синонімічні відношення: група термінів, що мають близьке або тотожне значення, наприклад, «кібербезпека» та «інформаційна безпека»;

3) антонімічні відношення: наприклад, терміни, що мають протилежні значення в ІТ сфері: front-end та back-end, open-source та proprietary, encryption та decryption;

4) асоціативні відношення: зв'язки, що виникають між термінами через спільні асоціації або схожість. Наприклад, терміни «браузер» та «Інтернет» можуть мати асоціативний зв'язок, оскільки браузери використовуються для перегляду веб-сторінок, які доступні через Інтернет;

5) відношення «частина та ціле»: зв'язки між термінами, які вказують на те, що один елемент є частиною іншого. Наприклад, «клавіатура» є частиною «комп'ютера», «модуль пам'яті» є частиною «сервера», «графічний чіп» є частиною «відеокарти»;

6) відношення «використання та застосування»: зв'язки між термінами, які вказують на те, що один елемент використовується для застосування іншого. Наприклад, «програмне забезпечення» використовується для застосування на комп'ютері. Або, «база даних» використовується для зберігання та організації інформації в ІТсистемах;

7) відношення «операція – процедура»: такі відношення відображають залежність між двома елементами, де один елемент є необхідним для виконання дії з іншим елементом. Так, в ІТ вони використовуються для опису залежності між елементами в процесі роботи з програмним забезпеченням. Наприклад, можна описати зв'язок між програмою для збереження даних та операцією збереження даних, що виконується цією програмою. Для додавання нового користувача до системи необхідно виконати операцію «створення облікового запису», яка створює новий запис в базі даних користувачів;

8) відношення «Дивись... (про А дивись В)»: вказує на те, що елемент А має певну залежність або взаємозв'язок з елементом В, і для отримання більш повної картини про А, необхідно розглянути В. Наприклад, якщо веб-додаток має проблему з відображенням на певних пристроях, то розглядання його коду може не дати повної картини про причину проблеми. В цьому випадку можна використати відношення «Дивись... (про додаток дивись на пристрій)», щоб розглянути конкретний пристрій та його параметри, які можуть впливати на відображення додатка;

9) відношення «аспект (А розглянуто в аспекті В)»: використовується для опису

зв'язку між двома термінами, де один елемент (А) розглядається або оцінюється з точки зору іншого елемента (В). Це може бути корисним для розуміння певних аспектів інформації та допомагає зорієнтуватись в певній області знань. Наприклад, можна сказати, що «проблеми кібербезпеки (А) розглядаються в аспекті захисту даних (В)», щоб показати, як захист даних впливає на безпеку програмного забезпечення та інтернету загалом;

10) відношення «параметр (А характеризується В)»: для позначення залежності певної характеристики (параметру) від іншої характеристики (характеристики-характеризатора). Наприклад, у програмуванні параметр «швидкість обробки даних» може бути характеризуваний параметром «розмір буфера», що відповідає обсягу даних, який може бути оброблений програмою за один раз;

11) відношення «імплікація» відображає взаємозв'язок, який описується за допомогою вислову «якщо..., то...». У цьому відношенні один термін (назвемо його А) є умовою, яка веде до іншого терміну (назвемо його В), що виступає наслідком умови. Наприклад, в контексті програмування можна навести таке відношення: якщо змінна «х» менше за 10, то виконати певну дію. У цьому випадку змінна «х» виступає як умова, а дія – як наслідок виконання цієї умови.

Цей перелік відносин між термінами не охоплює всі можливі варіанти. Також, відношення між термінами можуть бути контекстуальними і залежати від контексту, в якому вони використовуються. Наприклад, термін «безпека» може мати різні відношення в залежності від контексту, такі як «захист від хакерських атак», «захист від вірусів» тощо. Отже, список відношень між термінами не є вичерпним і може бути доповнений новими відношеннями в залежності від галузі знань та контексту використання.

В тезаурусі, зв'язки між термінами, як інформаційними елементами бази даних, можуть бути прямими або зворотніми за своєю природою. Пряма залежність вказує на те, що в описовій статті терміну є згадка інших понять, а зворотна залежність означає наявність посилань на термін в інших статтях (ці зв'язки вказують на ступінь важливості термінів та їх взаємозалежність).

Налагодження зв'язків в тезаурусі дозволяє досягти щонайменше наступних важливих результатів:

– забезпечити однозначність термінів та понять, що використовуються в базі даних. Це допомагає уникнути непорозумінь між користу-

вачами бази даних, які можуть мати різні тлумачення певних термінів;

– вдосконалити пошук та зручність використання інформації. Налагодження зв'язків між термінами дозволяє швидко знаходити пов'язану інформацію та переходити між поняттями;

– покращити організацію знань та взаємозв'язків між ними. Налагодження зв'язків допомагає визначити і зафіксувати структуру знань, що є важливим для їх подальшої організації та збереження.

Окремим важливим аспектом в формуванні командного тезаурусу є складання описових статей до термінів. Так як добре написані описові статті допомагають уникнути непорозумінь між членами команди щодо значення та використання термінів.

Пропонується дотримуватись декількох формальних вимог до визначень термінів у тезаурусі:

– визначення повинно бути в однині, за винятком тих понять, які є множинними;

– визначення має пояснювати, що означає конкретне поняття, а не тільки вказувати на те, чим воно не є;

– визначення повинно мати форму описової фрази або речення та містити лише широко відомі скорочення;

– визначення повинно бути чітким і зрозумілим без використання визначень інших термінів або базових понять;

– визначення має бути включене до відповідного глосарію, а якщо потрібне додаткове визначення, то його можна додати як примітку після первинного визначення або як окремий запис у словнику;

– визначення має точно відображати сутність поняття та бути достатньо зрозумілим, щоб забезпечити його однозначну інтерпретацію;

– визначення повинно бути коротким і мати можливість використовуватися окремо;

– визначення термінів має бути без використання додаткової інформації, такої як пояснення функціонального використання або процедурних деталей.

– визначення має бути чітким і запобігати циклічним посиланням [2].

Важливо, щоб описові статті були короткими, зрозумілими та чіткими, тому їх розробку можуть виконувати фахівці з галузі, які мають глибокі знання про термінологію та поняття, що використовуються в команді. Після створення тезаурусу та описових статей до термінів важливо регулярно оновлювати його, додавати нові терміни та поняття, оновлювати вже існуючі статті,

щоб вони відображали найновіші тенденції та зміни в галузі.

Крім того, необхідно визначити правила використання термінів в командному тезауусі, уникнути непорозумінь між членами команди, що можуть виникнути через різне тлумачення термінів або використання їх в різних значеннях. Наведемо деякі правила, які можуть бути встановлені в команді:

1. Використовуйте терміни, які вже зазначені в тезауусі, замість створення нових аналогічних термінів.

2. Визначте терміни, які використовуються в команді і додайте їх до тезауусу, якщо вони ще не включені.

3. Дотримуйтеся єдності термінології, використовуючи один термін для позначення одного поняття, якщо це можливо.

4. Уникайте дублювання термінів, які мають одне й те ж значення, та стандартизуйте терміни, які мають більше одного значення.

5. Дайте короткий опис кожному терміну, щоб забезпечити його точне тлумачення та запобігти неправильному використанню.

6. Перевіряйте і поповнюйте тезауус періодично, щоб забезпечити його актуальність та відповідність потребам команди.

9. Використовуйте тезауус як ресурс для уточнення термінології та вирішення спорів щодо тлумачення термінів.

10. Надавайте доступ до тезауусу всім членам команди та забезпечуйте, щоб він був завжди легко доступним та зрозумілим.

Необхідно також встановити норми використання скорочень та аббревіатур. Якщо в команді використовуються велика кількість аббревіатур та скорочень, необхідно включити в тезауус список аббревіатур та скорочень, що використовуються в команді, і визначити стандартні терміни,

що відповідають кожній з них. Наприклад, якщо в команді використовується аббревіатура "CRM" для позначення системи управління відносинами з клієнтами, то в тезауусі можна визначити, що стандартним терміном, що відповідає цій аббревіатурі, буде "система управління відносинами з клієнтами". Крім того, можна визначити правила використання аббревіатур та скорочень, наприклад, коли можна використовувати аббревіатуру, а коли краще використовувати повний термін, або, чи потрібно пояснювати значення аббревіатури при першому її використанні в тексті.

Для того, щоб командний тезауус був ефективним і корисним, необхідно забезпечити доступність до нього для всіх його учасників. Кожен член команди має мати можливість ознайомитися з тезауусом та користуватися ним під час роботи. Також важливо забезпечити процес оновлення та доповнення тезауусу відповідно до змін в галузі знань та потреб команди проекту. Це допоможе зберегти актуальність тезауусу та зробити його більш ефективним і корисним для команди.

Висновки. Створення командного тезауусу є важливим етапом у роботі команди над проектом, оскільки він дозволяє унормувати термінологію та покращити комунікацію між учасниками команди. В статті для побудови командного тезауусу пропонується дотримуватись наступних етапів: визначення термінів, які потрібно включити до тезауусу; налагодження відносин між термінами, що дозволяє досягти кількох важливих результатів, таких як вдосконалення пошуку та зручності використання інформації, покращення організації знань та взаємозв'язків між ними. Крім того, були визначені правила використання термінів в командному тезауусі, щоб забезпечити уникнення непорозумінь між членами команди.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Гладун А.Я., Рогушина Ю.В. Основи методології формування тезауусів з використанням онтологічного та мереологічного аналізу. *Штучний інтелект*. 2008. № 4. URL: <http://dspace.nbuv.gov.ua/bitstream/handle/123456789/7356/007-Gladun.pdf?sequence=3> (дата звернення: 20.02.2023).
2. Бісікало О.В., Яхимович О.В. Автоматизоване визначення лексичних онтологій з тезауусу технічного спрямування. *Оптико-електронні інформаційно-енергетичні технології*. 2016. № 1. URL: <https://oeipt.vntu.edu.ua/index.php/oeipt/article/view/437/434> (дата звернення: 21.02.2023).
3. Тур О. Теоретичні передумови тезауусного опису термінології. *Вісник ХДАК*. 2014. Вип. 45. С. 14–20.
4. Величко. Л. Тезауус як засіб семантизації наукових понять. *Біологія і хімія в рідній школі*. 2014. № 2. URL: https://lib.iitta.gov.ua/7065/1/Velychko_Bio_him_2_2014.pdf (дата звернення: 21.02.2023).