

ДОКУМЕНТАЛЬНІ КОМУНІКАЦІЇ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

Богдан Кияк, Наталія Соколовська, Тетяна Соколовська

Створення електронного англо-українського словника уніфікованої термінології фундаментальної науки

У статті розглянуто проблеми створення термінологічних словників фундаментальної науки; опрацьований великий масив ключових слів та виразів, що були вживані у проектах Державного фонду фундаментальних досліджень. Створені електронні українсько-англійський і англо-український словники активної наукової лексики, відкриті для поповнення. Словники будуть розміщені на ВЕБ-сторінці ДФФД і можуть бути використані авторами наукових проектів при підготовці конкурсних запитів та для аналізу пріоритетних напрямів фундаментальних досліджень.

Рівень зрілості суспільства визначається станом розвитку його інформаційної структури. Тому пріоритетним напрямом є створення високоорганізованої інфраструктури. Це забезпечує якісно нові технології і високий ступінь інформаційного обслуговування. Щоб створити таку ефективну інфраструктуру, необхідно здійснити інформаційне наповнення мереж, сформувати національні банки знань, електронно-довідкові й навчальні системи [1]. Програма широкого доступу користувачів до інформаційних ресурсів пов'язана не тільки з одностороннім виходом у світову мережу, а й з інтеграцією в неї власних ресурсів. Вищевказане неможливе без створення потужних інформаційно-довідкових систем, інформаційних серверів, баз даних і знань, насамперед у галузі освіти й науки, де не лише споживатимуться нові технології, а й продукуватимуться нові знання. У науково-освітній інформаційно-аналі-

тичній системі, яка може бути використана для безперервного спостереження за об'єктом будь-якої природи і ступеня поширеності інформації про цей об'єкт, моніторинг може здійснюватися і через автоматизоване створення частотних словників ключових слів, формування і обробку числових БД, виведення результатів статистичного аналізу в графічному вигляді, придатному для інтелектуального аналізу та інтерпретації. Доцільно також реалізувати ймовірносний статистичний (факторний) аналіз частот спільного вживання ключових слів з різних форм згортання інформації про дослідження (тема, анотація, реферат, повний текст), що стане придатним для первинної диференціації наукових досліджень, які закінчуються, розробляються, та оригінальних, котрі зароджуються.

Сьогодні є актуальною проблема включення вітчизняного інтелектуального потенціалу в систему наукових комунікацій. Щоб її розв'язати, слід поряд з традиційними джерелами все ширше використовувати електронні інформаційні ресурси (ІР). Один з відомих проектів електронної бібліотеки DLI [2] попри створення засобів для зберігання та організації даних в електронній формі із забезпеченням доступу до них через комп'ютерні мережі покликаний удосконалити зв'язки між фундаментальними дослідженнями й інноваційними технологіями. Пошук інформації в повнотекстових документах може реалізуватися через використання ключових слів, словників, що розкривають семантичні та синтаксичні зв'язки при пошуку термінів. Розглядаючи лінгвістичне забезпечення електронних бібліотек, автори [2–5] констатують його недостатню розробленість і необхідність ширшого залучення до вирішення цих питань мовознавців, бібліографознавців, спеціалістів у галузі інформатики.

Кияк Богдан Романович, к. ф.-м. н., начальник відділу фундаментальних досліджень і гуманітарної політики Міністерства освіти і науки України, виконавчий директор ДФФД,

Соколовська Наталія Валентинівна, студентка Національного технічного університету «Київський політехнічний інститут»,

Соколовська Тетяна Єремівна, головний спеціаліст відділу фундаментальних досліджень і гуманітарної політики Міністерства освіти і науки України, координатор бібліотеки ДФФД.

© Кияк Б. Р., Київ, 2001
© Соколовська Н. В., Київ, 2001
© Соколовська Т. Є., Київ, 2001

Мовознавці-термінологи стверджують, що існує мова науки, в межах якої міститься мова техніки [6]. Існує й поділ лексики на загальноживану, загальнонаукову та вузькофахову. Головною перешкодою при перекладі наукової літератури може бути відсутність у мові перекладу відповідних лексичних еквівалентів. Тому виникає необхідність у пристосуванні мови до конкретної галузі та стандартизації відповідної термінології.

Словотворча архітектура термінологічних одиниць йде шляхом лексичної модернізації та стандартизації структури мови. «Кількість термінів майже в кожній науці зростає ..., проте до цього часу майже нічого не робиться для того, щоб нові терміни були суворо вмотивовані та зрозумілі всім представникам даної науки» [7].

Серед завдань, які реалізує в своїй діяльності Державний фонд фундаментальних досліджень (ДФФД), поряд з основною метою — грантовою підтримкою конкурсних наукових проєктів, — не менш важливою функцією є інформаційне забезпечення в сфері науки та освіти. Створення електронного словника уніфікованої термінології фундаментальної науки дасть можливість використовувати англо-українську базу ключових слів усіх наукових напрямів при формуванні запитів для наступних конкурсів Фонду, швидко знаходити точний переклад термінів, а в подальшому на підставі аналізу поповнення словника робити деякі висновки щодо стану і шляхів розвитку фундаментальної науки в Україні порівняно з такими самими напрямками фундаментальної науки в інших державах.

Виходячи з довгострокових пріоритетів соціально-економічного, науково-технічного і національно-культурного розвитку, в нашій країні виконується державна програма автоматизації бібліотек.

Актуальним залишається створення лінгвістичної бази автоматизованих бібліотечних систем і насамперед — словника уніфікованої лексики для складання предметних рубрик [4], що має виконувати комунікативну, верифікаційну, енциклопедичну і прототипну функції.

Точність і повнота пошуку необхідної інформації в існуючому розмаїтті є одними з найактуальніших проблем сьогодення, а результативність пошуку безпосередньо залежить від якості пошукових систем. У процесі проведення конкурсів ДФФД та створення пошукової БД за проєктами-грантами й підсумковими звітами виникла необхідність автоматичної обробки авторських текстів з метою реалізації автоматизованого пошуку проєктів за ключовими словами (для призначення відповідних за

спеціалізацією експертів) та для інших завдань, що вирішуються в роботі з великими масивами інформації. Всі ці питання пов'язані зі створенням інформаційно-пошукової мови (ІПМ), яка повинна мати достатню семантичну потужність і бути спроможною описувати поняття з необхідною повнотою і точністю та забезпечувати однозначність кожного запису. Важливим аспектом її є відкритість для поповнення та коригування лексичних засобів.

Сукупність лексичних одиниць ІПМ становить її словник [11]. У даному разі лексичні одиниці представлені дескрипторами (ключовими словами) і словосполученнями, які визначають найважливіші топіки, відбиті в текстових документах. Створюваний нами словник інформаційно-пошукової мови є словником дескрипторів (ключових слів) усіх напрямів фундаментальної науки.

Безпосередньо основою для підготовки й роботи над словником стала вибірка наукових термінів українською та їх відповідників англійською мовою з конкретного фактичного матеріалу БД за 780 звітами наукових і науково-технічних проєктів 4-го та 5-го конкурсів Фонду за такими науковими напрямками: математика, механіка, інформатика; фізика і астрономія; хімія; наукові основи нових перспективних технологій; біологія; наука про Землю та Всесвіт і проблеми навколишнього середовища; фундаментальні аспекти гуманітарних наук; фундаментальні соціально-економічні проблеми; проблеми прогнозування розвитку науки та освіти. За розробленими електронними формами звітів кожен автор проєкту подавав до 10 ключових слів українською та їх переклад англійською мовою. Певна частина словника відповідає окремому конкретному (з перерахованих) науковому напрямку.

Словник віддзеркалює сучасний стан фундаментальних знань щодо широкого спектра наук. У ньому найширше представлено наукову лексику, використовувану українськими науковцями в останні роки (1997—2000). Вона пов'язана з фундаментальною наукою, а отже, ілюструє діапазон даних від основних понять до розвитку нових орієнтирів — пріоритетів фундаментальної науки. Робоча картотека налічує понад 5,0 тис. термінів українською мовою та стільки ж їхніх відповідників англійською.

Таблиця містить також підсумковий кількісний показник загальнонаукової і термінологічно-вживаної лексики проєктів запитів нового (сьомого) конкурсу ДФФД.

Мета роботи — отримати і вивчити поданий інформаційний матеріал, що зумовлює розв'язання таких окремих завдань:

База наукових термінів електронного словника

№	Назва напрямку	Кількість проектів	Кількість поданих авторами слів (укр. мова)
1	математика, механіка, інформатика	85	664
2	фізика та астрономія	233	1615
3	хімія	91	598
4	наукові основи нових перспективних технологій	127	827
5	біологія	90	513
6	наука про Землю та Всесвіт і проблеми навколишнього середовища	60	357
7	фундаментальні аспекти гуманітарних наук	22	179
8	фундаментальні соціально-економічні проблеми	56	484
9	проблеми прогнозування розвитку науки та освіти	16	111
	Всього за проектами 4-5-го конкурсів Фонду	780	5348
	Отримано за проектами 7-го конкурсу	2671	близько 25 тис. ключових слів

- ✓ визначення понять термінів фундаментальної науки;
- ✓ з'ясування шляхів розвитку і становлення термінів фундаментальної науки;
- ✓ виявлення конкретних словотвірних моделей українською та англійською мовами, щоб дати можливість авторам наукових публікацій легше розуміти основи фундаментальної науки, встановити зв'язки між напрямками, а також застерегти науковців від помилок у використанні наукових українських та англомовних термінів;
- ✓ вивчення закономірностей сполучуваності словотвірних форматів.

У процесі роботи над словником усі ключові слова, використані в звітах, перевірялися й уточнювалися відповідно до стандартних словників. Необхідною виявилась також перевірка дескрипторів, поданих авторами-науковцями. Доводилося шукати відповіді на питання, чи несуть вказані ключові слова в тексті необхідне змістове навантаження в інформаційному пошуку, чи подані слова є термінами, тобто словосполученнями з єдністю звукового знака та співвіднесеного з ним відповідного поняття в системі даної галузі науки та техніки [6]. У результаті аналізу кількість ключових слів, узятих

з проектів, суттєво зменшилася. Після перевірки потребували граматичного коригування й англійські варіанти ключових слів. До деяких українських термінів були подані точніші англійські еквіваленти. Перевірка перекладу термінології виявилася нелегким завданням. Адекватний переклад термінів ускладнювався розбіжностями в історії їх виникнення, у вживанні, наявності синонімів та полісемії.

Отже, змістом запропонованого словника наукової термінології є особлива група слів, кількість яких у розвинених мовах дуже велика. Якщо ж вважати термінами й групу загально-вживаних лексичних одиниць, які є термінами лише контекстуально в наукових текстах (*вода* — в хімії, *дірка* — у фізиці та ін.), тоді можна припустити, що кількість термінів у мові в багато разів перевищує чисельність загально-вживаних слів.

У більшості наукових праць вивчення словотвірних процесів здійснюється на базі термінологічної лексики, оскільки остання є найбільш збагачуваною частиною словникового складу мови [6, 11].

Науковці використовують у своїх підсумкових публікаціях термінологічні одиниці, які надходять до мови різними шляхами:

- ❖ через запозичення з рідної мови, в тому числі з діалектизмів, регіоналізмів, професіоналізмів; часто мають місце такі стилістичні засоби, як переосмислення, метонімічне перенесення;
- ❖ через запозичення з інших мов, як правило, з мови країни, розвиток котрої досяг певного прогресу в тій чи іншій галузі.

У мовах європейсько-американського ареалу дуже поширеним джерелом творення нових лексичних одиниць є насамперед класичні мови — давньогрецька та латинська. Використання греко-латинських компонентів є особливим типом виникнення термінів, а їх поєднання в мові дуже продуктивне (minicomputer, alphanumeric, mini-based, retroactively).

Зустрічаються терміни, з кількома греко-латинськими елементами (multi-minicomputer, megamini). На підставі поданих авторами проєктів ключових слів можна зробити висновок, що терміносистеми української мови активно послуговуються цими компонентами.

У поданих (у конкурсних проєктах і звітах) наукових термінах широко використовується процес словотворення, виникнення одних одиниць мови на основі інших. Такий процес має два потенційні шляхи розвитку — розгортання (створення багатокомпонентних термінів) та згортання (створення компактнішої одиниці). Дери́ваційні процеси поширилися в англійській мові; розвиваються вони і в українській (data — database — database system, аналогічно в українській: дані — база даних — система бази даних).

У кожній мові діє закон мовної економії. Що стосується терміносистем, то їм особливо притаманне стискання форми вираження (але ні в якому разі не змісту) лексичної одиниці (chain of blocks = block-chaining, cubic number = cube, setpoint adjustments = settings). Не всі науковці для варіанта перекладу використовували одиниці, які виникли нещодавно. Попередня форма вираження може іноді вважатися архаїчною чи навіть змінити своє смислове значення. У такому разі його вживання призводить до помилкового розуміння.

Розвиток терміносистем відбувається дуже швидко. Певна кількість лексичних одиниць відмирає, бо поняття, які вони відображали, є неактуальними, а деякі з таких термінів набувають нового значення. Тому при аналізі словникової бази деякі лексичні одиниці були вилучені, а деякі — додані. Лексичні одиниці даної ІПМ створили два великі пласти:

- ✓ терміни, що мають відповідні еквіваленти (слова, які повністю збігаються зі словом оригіналу) чи аналоги (слова, які знахо-

- дять шляхом аналізу синонімічного ряду);
- ✓ терміни, що не мають безпосередніх відповідників: відносимо в цю категорію і неологізми авторів-науковців; для таких слів були знайдені адекватні заміни за допомогою лексичних перекладацьких трансформацій (транскрипції, транслітерації, калькування) чи описового перекладу, використовуючи розроблені рекомендації [12].

Планується, що словник існуватиме в електронній формі, тобто у вигляді БД, для забезпечення постійного користування та ефективного пошуку термінів. Можливим стане швидке внесення змін, поповнення і його модернізація. Словник в електронній формі наочно представлятиме інформацію, матиме дружній інтерфейс і буде доступним користувачам різної кваліфікації; буде виставлений на Web-сторінці ДФФД (www.mon.gov.ua) для загального користування та пошуку необхідної інформації українськими, а також зарубіжними науковцями. На базі даного словника можлива розробка і створення дескрипторного словника із зафіксованими в ньому парадигматичними відношеннями — інформаційно-пошукового тезауруса.

Стрімкий розвиток інформаційних і комунікаційних технологій значною мірою вплинув на більшість сфер людського життя. Електронна пошта, лазерні диски, бази даних на відстані, мережа для пошуку джерел інформації, перевірка правопису і граматики автоматизованими програмами-коректорами — такі інформаційні способи спілкування і використання електронних ресурсів можуть бути поєднані з традиційними способами [13—15]. Необхідно продовжити аналіз і систематизацію наукових та технологічних ідей, розробок, що були виконані та виконуються вітчизняними науковцями, забезпечити вільний доступ до інформаційного банку широкого кола зацікавлених користувачів, постійно поповнювати БД новою відкритою інформацією, щоб створити в Україні ринок науково-технічної продукції, інтегрувати наявний потенціал у світовий науковий, інноваційний та інформаційний простір, сприяти розвитку освітніх програм і міжнародної співпраці. Завдання полягає не тільки у створенні банку наукових ідей вітчизняних учених, публікацій за результатами виконання розробок, оцінці їхнього індексу цитувань, а й в активізації термінотворчого процесу в середовищі фундаментальної науки.

1. Згуровський М., Сергієнко І. Інформаційні технології у сучасному суспільстві // Вісн. НАН України. — 2000. — № 12. — С. 9—16.

2. Павлуша І. Електронні бібліотеки: зарубіжний досвід, питання розробки української концепції // Бібл. вісн. — 1999. — № 4. — С. 13—24.
3. Стеллецкий В. И., Аветисов М. А. Представление словарей и справочников в электронной библиотеке знаний // 36. наук. пр. восьмої міжнар. конф. «Крим 2001». — «Бібліотеки та асоціації в світі, що змінюються: нові технології та нові форми співробітництва». — Судак. — 2001. — Т. 1. — С. 327—329.
4. Збанацька О. М. Створення Словника предметних рубрик НПБ України для пошуку в електронному каталозі // Наук.-техн. б-ки в єдиному інформаційному просторі України: Матер. наук. міжнар. конф. — К: ДНТБ України. — 2000. — С. 119—122.
5. Кияк Б. Р., Соколовська Т. Є., Соколовська Н. В. Створення електронних українсько-англійського та англо-українського словників ключових слів та виразів фундаментальної науки // 36. наук. пр. восьмої міжнар. конф. «Крим 2001» ... — С. 525—528.
6. Д'яков А. С., Кияк Т. Р., Куделько З. Б. Основи терміноутворення. — К.: Вид. дім «КМ Academia». — 2000. — 216 с.
7. Будагов Р. А. Человек и его язык. — М: Изд-во Моск. ун-та, 1974. — 262 с.
8. Російсько-український словник наукової термінології: Математика. Фізика. Техніка. Науки про Землю та Космос / За ред. В. В. Гейченко, В. М. Завірюхіна, О. О. Зеленюк та ін. — К.: Наук. думка, 1998.
9. Російсько-український словник наукової термінології: Біологія. Хімія. Медицина / За ред. С. П. Васер, І. О. Дудка, В. І. Єрмоленко, М. Д. Зерова, А. Я. Ільченко, О. К. Усатенко. — К.: Наук. думка, 1996.
10. Російсько-український словник наукової термінології: Суспільні науки / За ред. Й. Ф. Андерис, С. А. Воробйова, М. В. Кравченко та ін. — К., 1994.
11. Штерн І. Б. Вибрані топіки та лексикон сучасної лінгвістики. — К.: «АртЕк». — 1998. — 335 с.
12. Рекомендації щодо транслітерування літерами української абетки власних назв, поданих англійською, французькою, німецькою та італійською мовами // Інтелектуальна власність, УІПВ. — 2001. — № 5/6. — С. 63—92.
13. Сюнтюрєнко О. В. Программы РФФИ в области развития информационной инфраструктуры науки и образования // 36. наук. пр. шостої міжнар. конф. «Крим 99». — «Бібліотеки та асоціації в світі, що змінюються: нові технології та нові форми співробітництва». — Судак. — 1999. — С. 29—33.
14. Кияк Б. Р., Соколовська Т. Є. Щодо специфіки створення бібліотеки на базі фундаментальних наукових публікацій під егідою Державного Фонду Фундаментальних досліджень // Там само. — С. 74—76.
15. Новиков В. Д. Опыт РФФИ создания научной электронной библиотеки // Вестн. ФФИ. — Минск. — 2000. — № 4. — С. 63—84.

ФАКТИ НАУКОВОГО ЖИТТЯ. ХРОНІКА ПОДІЙ

Презентація книжкової продукції Видавничого дому «КМ Академія» (жовтень 2001 р.)

Сім років функціонує в структурі Національного університету «Кієво-Могилянська академія» Видавничий дім «КМ Академія» (директор В. Й. Соловйова), чия продукція користується серед наукової громадськості та широкого загалу зростаючим попитом і стає дедалі відомішою за кордоном.

У рамках тижня НаУКМА в жовтні поточного року було організовано презентацію видань дочірнього підприємства (ДП) одного з найстаріших вузів нашої країни. В розкішному Конгрегаційному залі університету в урочистій обстановці відомі вчені, викладачі столичних вищих навчальних закладів, колеги, представники бібліотек Києва, студенти вітали редакторів та поліграфістів, тобто тих, хто дає дорогу в світ науковим монографіям, підручникам, посібникам, словникам, «Науковим запискам» (виходять тематичними випусками за 12-ма галузями знань). За роки наполегливої праці, творчих пошуків «КМ Академією» підготовлено понад сто книжок. Започатковано серії: «Джерела мудрості», «Були, є і будемо вівіки», «Християнські філософи», «Книга НаУКМА».

Як підкреслила В. Й. Соловйова, група ентузіастів перетворилася нині в солідний високоінтелектуальний колектив, який чітко й усвідомлено веде свою творчу політику, спрямовану на постійне поліпшення якості роботи, розширення та урізноманітнення асортименту продукції, прагне йти шляхом, наслідуючи традиції потужної «Наукової думки», що залишається взірцем для всіх видавництв України та взагалі українського книгодрукування, закладеного на наших землях Іваном Федоровичем понад 400 років тому.

Оригінальним пунктом програми презентації універ-

ситетського видавництва стало й обговорення нової праці д. філос. н. С. Рябова «Політологія: словник понять і термінів», випущеного ДП НаУКМА, в котрому взяли участь д. і. н. О. Гарань, доц. О. Деменчук, д. ю. н. Т. Смирнова, головний редактор журналу «Політична думка» О. Дергачов, викладач Чернівецького університету, к. ю. н. В. Ферсанов та ін.

Про роботу над бібліографічними покажчиками розповіли директор Наукової бібліотеки університету Т. Ярошенко та її заступник Н. Козакова.

У планах видавництва НаУКМА — випуск енциклопедичної праці «Кієво-Могилянська академія в іменах XVII—XVIII ст.», книга Ю. Хорунжего «Українські меценати», заявлено три масштабні теми в серію «Християнські філософи», готується каталог «Іван Миколайчук: життя і творчість» та ін.

— Наука, — наголосив перший віце-президент, віце-президент з наукової роботи НаУКМА д. х. н. М. Брик, — як відомо, потребує фінансової підтримки. Навчальний заклад знайшов її та духовне взаєморозуміння з боку фундації родини Фещенко-Чопівських, завдячуючи яким вдалося видати кілька підручників.

До презентації було підготовлено виставку книжкової продукції Видавничого дому «КМ Академія», чимало примірників котрої вже надійшло в бібліотеки України.

*Ярослава Ясенко,
наш кор.*

© Ясенко Я. А., Київ, 2001