

АВТОМАТИЗАЦІЯ ЛІНГВІСТИЧНИХ ДОСЛІДЖЕНЬ НА БАЗІ ЦИФРОВОГО ТЕКСТУ ТЛУМАЧНОГО СЛОВНИКА ІСПАНСЬКОЇ МОВИ

Купріянов Є. В.

*Національний технічний університет
«Харківський політехнічний інститут», м. Харків*

На сьогодні цифрове словникарство перетворилося у багатоаспектну галузь, що являє собою: а) прикладну наукову дисципліну на стику мовознавства та інформатики, що вивчає проблеми застосування методів та технологій інформаційної науки у теорії та практиці створення широкого спектру лексикографічних систем; б) комп'ютерну індустрію, яка швидко розвивається внаслідок того, що лексикографування знань відрізняється від ефективних способів їх виявлення та розподілу.

У цьому контексті ми розглядаємо текст слоника, головним чином, не як довідкову систему, а як засіб передання лінгвістичних знань. Передусім, це стосується фундаментальних національних тлумачних словників. З огляду на це, постає проблема забезпечення цих словників відповідним інструментарієм. Реалізація інструментарію на належному рівні можлива лише в цифровому середовищі.

Одним з найбільш ефективних інструментів для роботи зі словниками в цифровому середовищі є віртуальні лексикографічні лабораторії (ВЛЛ). На відміну від електронних словників, вони призначені переважно для професійних лінгвістів. У роботі висвітлено питання автоматизації лінгвістичних досліджень засобами віртуальної лексикографічної лабораторії для тлумачного словника іспанської мови «Diccionario de la lengua española, 23^a ed.» (DLE 23). За допомогою теорії лексикографічних систем була розроблена формальна модель DLE 23, використовуючи теорію лексикографічних систем проф. В. А. Широкова. На основі моделі визначено структуру бази даних та елементи інтерфейсу ВЛЛ. Інтерфейс ВЛЛ DLE 23 має такі переваги: 1) складання інвентаризації мовних одиниць у словнику або у зразку; 2) проведення лінгвістичних досліджень на основі тексту DLE 23 з метою виявлення лексико-семантичних, етимологічних, граматичних і лінгвопрагматичних властивостей іспанської мови; і 3) побудова вторинних лексикографічних об'єктів або під словників на основі DLE 23, наприклад: словник морфем, омонімів, розмов тощо.

Розроблена ВЛЛ DLE 23 доступна на сайті Українського мовно-інформаційного фонду НАН України (<https://services.ulif.org.ua:44359/>) і дає змогу:

- статистичні дослідження як по всьому словнику, так і по окремій вибірці словникових статей;
- аналізувати іспанську лексику та тексти словникових статей для вилучення різних лінгвістичних фактів, зафіксованих у DLE 23 (семантика, етимологія, граматика, використання слів);
- створювати вторинні лексикографічні об'єкти (або похідні словники), що описують окремий лексичний склад іспанської мови.