

Яблочков Микита Мирчович

Український мовно-інформаційний фонд НАН України

ЕФЕКТИВНА ТЕХНОЛОГІЯ ДІДЖИТАЛІЗАЦІЇ БАГАТОМОВНИХ ФРАЗЕОЛОГІЧНИХ СЛОВНИКІВ

Невеликі за обсягом лексикографічні праці, що виходять на ринок тільки в форматі друкованої книги, в сучасних реаліях мають обмежене коло користувачів й не реалізують свій інформаційний потенціал. Описується ефективна та економічна технологія діджиталізації багатомовних фразеологічних словників.

Ключові слова: *фразеологія, словники, діджиталізація, ефективність*

Small-scale lexicographic works, which are marketed only in the format of a printed book, in modern realities have a limited number of users, and do not realize their informational potential. Efficient and economical technology for the digitization of multilingual phraseological dictionaries is described.

Keywords: *phraseology, dictionaries, digitalization, efficiency*

Словниковий ринок постійно поповнюється невеликими за обсягом друкованими лексикографічними працями. Такий формат звужує можливе коло користувачів і не дає можливостей оснастити словник сучасним інформаційно-дослідницьким інструментарієм.

Саме створення такого комп'ютерного інструментарію стало мотивацією для розроблення веб-версії [1] «Лексикону польської та української активної фразеології» [2]. В результаті, крім вебсайту була розроблена і реалізована технологія діджиталізації багатомовних фразеологічних словників з певною структурою їхніх лексикографічних систем [3]. Особливістю технології є повторне використання програмних доробків для інших словників, а також можливість інтегрування цифрових версій в єдину систему – ВЛЛ [4] (віртуальну лексикографічну лабораторію) перекладної фразеології.

Сам «Лексикон» представляє собою об'єднання в одну книгу польсько-українського та українсько-польського фразеологічних словників разом з загальним індексом фразеологізмів.

Створена електронна версія зберігає оригінальний дизайн візуалізації словникових статей та реалізує три різні способи доступу: 1) за алфавітом – імітація «гортання сторінок» словника; 2) за фразеологічним індексом з урахуванням мови та діапазону (укрїнсько-польський, або польсько-український, або обидва); 3) повнотекстовий пошук. Приклад входу за індексом продемонстровано на Рис. 1.

LEKSYKON
 aktywnej frazeologii polskiej i ukraińskiej

ЛЕКСИКОН
 польської та української активної фразеології

Вхід: Пошук по індексу

Пошуковий запит: [] [] []

Параметр пошуку: Містить

Мова індексу: Польська

Діапазон

Польсько-український словник

Українсько-польський словник

potrzebny

powidło

powiedzieć

powieka

powietrze

powlewać

powtarzać

pożądać

pozostawić

prac

pracować

pranie

prawda

prawo

próżny

prac brudne pieniądze 'wtedy, gdy ktoś legalizuje pieniądze uzyskane nielegalnie' Od dawna wydaje mi się, że jej chłopak nie prowadzi żadnej restauracji, tylko pierze brudne pieniądze.	відміти / відмивати гроші 'тоді, коли хтось легалізує фінансові засоби, які отримані незаконно' Виявилось, що на фірмі, де я вже рік працював, просто відмивали гроші.
prac brudy w (/ we własnym) domu 'wtedy, gdy ktoś nie rozpowiada o prywatnych, nieprzyjemnych sprawach' Dlaczego wszystkim opowiadasz o rozwodzie Iwony? Już tyle razy ci mówiłam, że brudy należy prac we własnym domu.	не виносити сміття з хати 'тоді, коли хтось не розповідає про неприємні особисті речі' В інтерв'ю актор розповів, що не виноситиме сміття з хати і не виставлятиме своє особисте життя напоказ.
відміти / відмивати гроші 'тоді, коли хтось легалізує фінансові засоби, які отримані незаконно' Чиновник відмиває гроші через фіктивну фірму.	prac brudne pieniądze 'wtedy, gdy ktoś legalizuje pieniądze uzyskane nielegalnie' W tej restauracji nigdy nikogo nie ma, a ona wciąż istnieje, pewnie ktoś tam pierze brudne pieniądze.
не виносити сміття з хати 'тоді, коли хтось не розповідає про неприємні особисті речі' Співачка намагалася не виносити сміття з хати і всіляко заперечувала проблеми у сім'ї.	prac brudy w (/ we własnym) domu 'wtedy, gdy ktoś nie rozpowiada o prywatnych, nieprzyjemnych sprawach' Brudy trzeba prac we własnym domu, a nie rozpowiadać wszystkim o naszych sprawach!

O książce
 Wstęp
 Kwalifikatory
 O autorach
 Pro книгу
 Вступ
 Ремарки
 O авторах

Рис. 1. Вебверсія «Лексикону». Вхід «за індексом»

На першому з трьох етапів діджиталізації текст словникових статей був зконвертований із RTF-формату в HTML-формат із збереженням форматування, що дозволяє виокремити структурні елементи статей.

На другому етапі була побудована база даних «Лексикону» (у вигляді класів словникових статей та індексу) на основі структур, виділених на попередньому етапі. Ми використали базу даних LiteDB (NoSQL БД документарного типу), що дозволяє структуровані тексти словникових статей зберігати в явному, цілісному виді. Побудована БД є самостійним цифровим об'єктом.

Створення веб-версії лексикону (третій етап) було розділено на дві відносно незалежні частини: 1) створення сервісної частини додатку (реалізація механізмів доступу до цифрового словника); 2) створення вебсторінки «Лексикону» (користувацький інтерфейс).

Завдяки використанню технології WEB-API і сильно структурованої моделі передачі даних, доступ до даних та функціоналу словника є незалежним від реалізації користувацької частини. Це дозволяє повторно використовувати сервісну частину додатку для інших словників, створюючи чи модифікуючи тільки сам користувацький інтерфейс, а також дає можливість для інтегрування баз даних різних словників.

Наступним кроком планується діджиталізація Українсько-Польського [5] та Польсько-Українського [6] словників еквівалентів слова.

Література

1. «Лексикон польської та української активної фразеології»: вебсайт. URL: <https://services.ulif.org.ua/lexicon/> (дата звернення 01.02.2020)
2. Тимошук Р., Сосновський В., Яскот М., Ганошенко Ю. «Лексикон польської та української активної фразеології»: словник. KJV Digital TOB, м. Варшава, 2018. 312 с.
3. Широков В. А. «Комп'ютерна лексикографія»: монографія. «Наукова думка», м. Київ, 2011. 351 с.
4. Надутенко М. В. «Віртуалізовані лексикографічні системи та їх застосування у прикладній лінгвістиці»: автореф. дис. канд. тех. наук, НАН України, Нац. б-ка України ім. В.І. Вернадського, м. Київ, 2016, 22 с.
5. Лучик А., Антонова О., Дубровська І. «Українсько-Польський словник еквівалентів слова»: словник. Укр. мов.-інформ. фонд НАН України, Нац. ун-т «Києво-Могилян. акад.», Ін-т славістики Пол. акад. наук. - К. : [б. в.], 2011. - 311 с.
6. Лучик А., Антонова О. «Польсько-Український словник еквівалентів слова»:словник. Укр. мовно-інформ. фонд НАН України, Нац. ун-т «Києво-Могилян. акад.», Ін-т славістики Пол. акад. наук. - К. : [б. в.], 2013. - 243 с.