**АНОТАЦІЯ**

Дипломний проект включає пояснювальну записку (50 с., 15 рис. 3 табл.).

Об’єктом розробки є створення програмних засобів автоматичного формування словника словосполучень української мови.

Метою дипломного проекту є вивчення теорії прикладної та комп’ютерної лінгвістики, створення програмних засобів автоматичного формування словника словосполучень української мови на базі великого електронного словника української мови (ВЕСУМ) та дослідження статистики використання слів. Базою даних слугує PostgreSQL.

В ході розробки дипломного проекту:

* було проведено аналіз вже існуючих словників та словників словосполучень різних мов;
* розглянуто методи аналізу текстів;
* розроблено програму для скачування текстів українською мовою з Wikipedia;
* проаналізовано статистику використання слів;
* розроблено зручний для користувача інтерфейс.

Програмне забезпечення розроблено задля допомоги при підборі влучного словосполучення та тлумаченні слова через контекст.

Ключові слова: комп’ютерна лінгвістика, словник словосполучень, українська мова, токенізація, JavaScript, Node.js, PostgreSQL.

**ABSTRACT**

This diploma project includes an explanatory note (50 pages, 15 pictures, 3 tables).

The object of development is the creation of software that automatically forms a Ukrainian collocation dictionary.

The purpose of the diploma project is to study the theory of applied and computational linguistics, to create software for automatic formation of the Ukrainian collocation dictionary on the basis of the Ukrainian dictionary and research on word usage statistics. PostgreSQL is used as a database.

During the development of the diploma project:

* already existing dictionaries and collocation dictionaries of different languages ​​ was analyzed;
* methods of text analysis was reviewed;
* a program for downloading texts in Ukrainian from Wikipedia was developed;
* word usage statistics was analyzed;
* user-friendly interface was developed.

The software was developed to help in selecting the right phrases and interpreting words through context.

Keywords: computational linguistics, collocation dictionary, Ukrainian language, tokenization, JavaScript, Node.js, PostgreSQL.