

УДК 004.89:004.65

ВІЗУАЛЬНИЙ ГЕНЕРАТОР СХЕМ РБД ЗАСОБАМИ ШІ

Кашпур І.В., Гриньова О.Є.

e-mail: iryna.kashpur@nure.ua, olena.hrynova@nure.ua

Харківський національний університет радіоелектроніки, каф. ШІ

м. Харків, Україна

This paper presents a Visual Generator for RDB Schemes (VGRDS) using AI. The system employs two AI models: M_1 for generating new schemas based on natural language descriptions, and M_2 for searching library schemas. VGRDS optimizes database design by automating SQL code generation, schema optimization, and providing a visual interface. The tool reduces development time, improves schema quality, and supports multiple DBMS. Future developments include a classifier for the subschema library, enabling users to store, categorize, and reuse database schema fragments efficiently.

Проектування баз даних (БД) є важливим етапом у розробці інформаційних систем. Для проектування БД “з нуля” необхідно, щоб розробник створив структуру (схему) та реалізував її в кодї SQL. За допомогою класичних методів аналізу та документування ER-моделі можуть бути створені повністю вручну, що вимагає значних зусиль та часу, а також самотійного написання SQL-коду, що підвищує ризик помилок та створення неефективних запитів. Розробники постійно стикаються з необхідністю створення схожих структур, таких як схеми аутентифікації чи каталоги продуктів, замість повторного використання готових рішень.

Для оптимізації процесу проектування БД доцільно використовувати інструменти автоматизованого проектування, або використовувати бібліотеки готових SQL-скриптів, або застосовувати шаблони проектування для типових структур даних. Можливість зберігати та ділитися підсхемами, публічними для всіх або приватними (для особистого використання), підвищить ефективність роботи розробників БД, забезпечуючи швидкий доступ до перевірених та оптимізованих структур даних. Тож розробка інструменту, який дозволить генерувати та класифікувати підсхеми за предметними галузями, а також повторно використовувати готові бібліотеки підсхем у своїх проєктах, є актуальною задачею.

Візуалізація підсхем (таблиць та зв'язків між ними) допомагає розробникам краще розуміти структуру бази даних та ефективніше працювати з нею. Візуальне середовище розробки дає можливість легко генерувати та зберігати нові підсхеми, підключати існуючу бібліотечну підсхему для створення комплексних рішень. Підтримка генерації коду на основі запитів природною мовою є значним кроком вперед у зручності та доступності проектування БД. Використання моделей машинного навчання

для перетворення запитів природною мовою в SQL-код спрощує процес для розробників з різним рівнем підготовки [1]. Саме тому розробка інструменту, що поєднує візуальне проектування, бібліотеку підсхем та генерацію коду з природної мови, є надзвичайно актуальним завданням.

Автоматизація проектування БД дозволяє підвищити продуктивність та якість моделювання [2]. Більшість існуючих інструментів не підтримують автоматичну генерацію SQL-коду або вимагають складних налаштувань, або небезкоштовні. Запропонований інструмент дозволяє користувачам створювати ER-моделі, відповідно до визначених правил нормалізації, в інтерактивному інтерфейсі та автоматично генерувати SQL-код. Важливим аспектом є написання запитів (prompts) природною мовою для генерації структур БД, що допомагає уникати поширених помилок при проектуванні баз даних. Це включає автоматичне визначення первинних та зовнішніх ключів, аналіз зв'язків між сутностями та генерацію підказок щодо нормалізації схеми.

Запропонована архітектура візуального генератора схем для реляційних баз даних (ВГБД). Система працює з двома моделями штучного інтелекту: M_1 – генератор нової схеми за описом на природній мові, та M_2 – пошук бібліотечної схеми за описом на природній мові (рисунок 1).

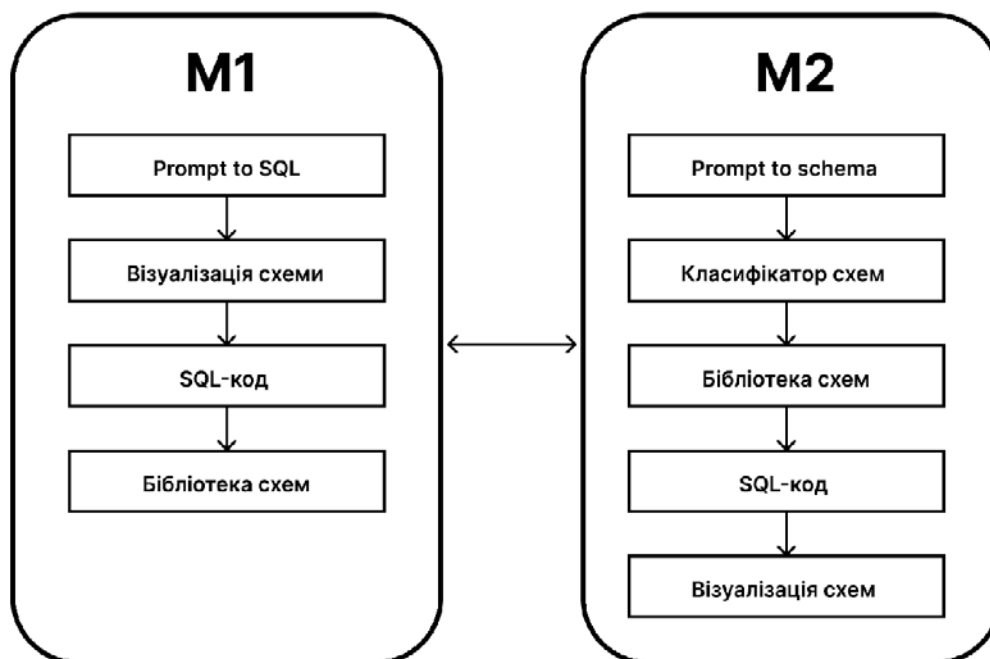


Рисунок 1 – Архітектура візуального генератора схем РБД

Модель M_1 може генерувати не лише SQL-код для трьох відомих СКБД (MSSQL, Postgre, MySQL), але й проводить його оптимізацію, що включає автоматичне додавання індексів для підвищення продуктивності запитів, використання оптимальних типів даних та усунення надлишкових зв'язків, що зменшує складність схеми (рисунок 2, 3).

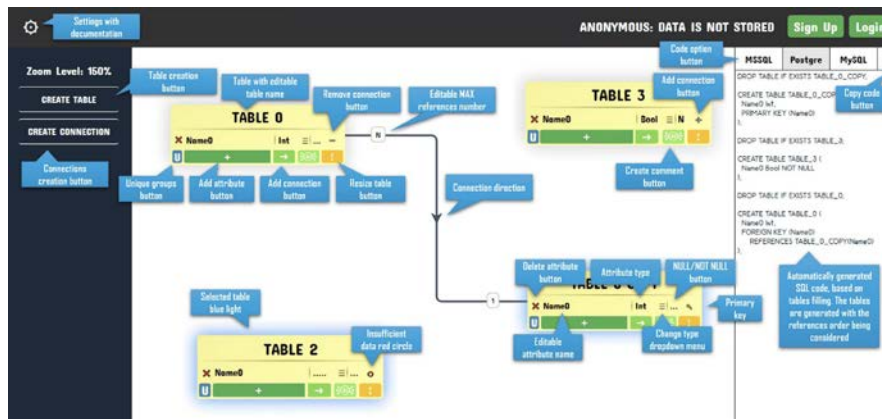


Рисунок 2 – Інтерфейс генератора бібліотечних схем, модель M₁

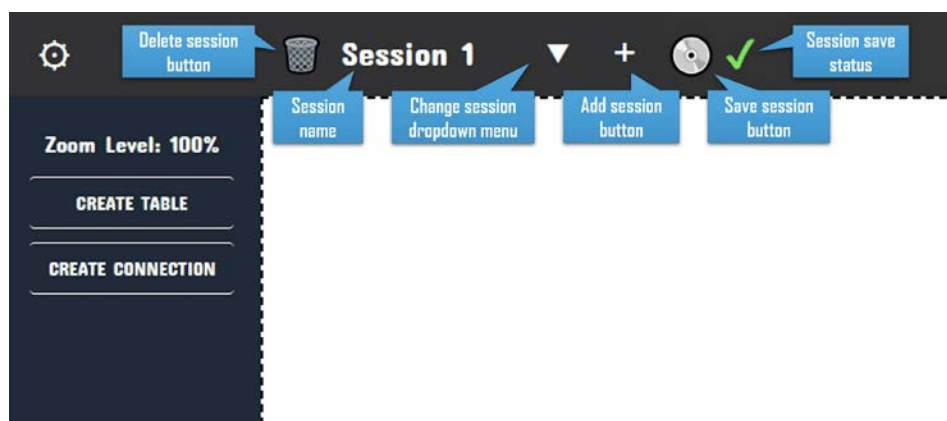


Рисунок 3 – Функціонал генератора бібліотечних схем, модель M₁

Тестування ВГБД показало, що час на створення ER-моделі та SQL-коду скоротився у 3 рази, оскільки усуває проблеми з генерацією таблиць та коду. Оцінка безпеки та валідності схем показала, що DUI (Dedicated User Interface) схеми вдвічі кращі за схеми, розроблені звичайними підходами.

Майбутній розвиток ВГБД передбачає розробку класифікатору бібліотеки підсхем, яка дозволить користувачам зберігати, класифікувати та повторно використовувати готові фрагменти схем БД. В перспективі планується що бібліотека буде забезпечена потужними інструментами пошуку та фільтрації (модель M₂) на природній мові.

Список використаних джерел:

1. Filatov, V., Semenets, V., & Zolotukhin, O. (2020). Data mining in relational systems. *Innovative Technologies and Scientific Solutions for Industries*, (3 (13)), 65–76. <https://doi.org/10.30837/itssi.2020.13.065> (дата звернення 25.02.2025).
2. Li, Y., et al. "Seq2SQL: Generating Structured Queries from Natural Language using Reinforcement Learning." *arXiv preprint arXiv:1709.00103* (2017).