



УДК [004.7-047.72]:656.2

**DESIGN, CREATION AND OPTIMIZATION OF THE SCIENTIFIC  
DEVELOPMENT DATABASE OF THE DEPARTMENT  
ПРОЕКТУВАННЯ, СТВОРЕННЯ ТА ОПТИМІЗАЦІЯ БАЗИ ДАНИХ  
НАУКОВИХ РОЗРОБОК КАФЕДРИ**

**Пахомова В. М. / Pakhomova V. M.***к.т.н., доц. / c.t.s., as. prof.*

ORCID: 0000-0002-0022-099X

**Хрестян А. В. / Hrestyan A. V.***здобувач 2-го ступеня / applicant of the 2<sup>st</sup> degree*

ORCID: 0000-0002-1774-5282

*Український державний університет науки і технологій,**Україна, Дніпро, вул. Лазаряна, 2, 49010**Ukrainian State University of Science and Technology,**Ukraine, Dnipro, Lazaryan St., 2, 49010*

**Анотація.** Спроектована реляційна база даних наукових розробок кафедри ЕОМ з використанням даних репозитарія ДІТ з науковими розробками та даних про науково-педагогічний склад кафедри. Підвищення продуктивності роботи бази даних пов'язано з досягненням оптимальної швидкості роботи з базою даних, що потребує як оптимізації самої бази даних, так і оптимізації програмних конструкцій SQL. При проектуванні бази даних наукових розробок кафедри використаний метод нормальних форм, результати якого співпали з результатами проектування за графічним методом «Сутність-Зв'язок», що заснований на використанні діаграм ER-типу. Проведені дослідження призначення простих та композитних індексів, а також оптимізація операторів SQL. Визначено, що час виконання зменшується приблизно у два рази при оптимізації програмних конструкцій SQL.

**Ключові слова:** збиткове дублювання, проектування, оптимізація бази, нормальні форми, діаграма ER-типу, SQL, оптимізація конструкцій, продуктивність.

**Вступ**

**Постановка проблеми.** У відкритому доступі можна знайти дані про наукові розробки науковців і викладачів деяких кафедр, що відносяться до різних вищих навчальних закладів, хоча це займе відповідний час на пошук. Але використовуючи систему управління базами даних і маючи навички створення конструкцій SQL (Structured Query Language), можна спроектувати власну базу даних, яка скорочує пошук, що підтверджує актуальність теми.

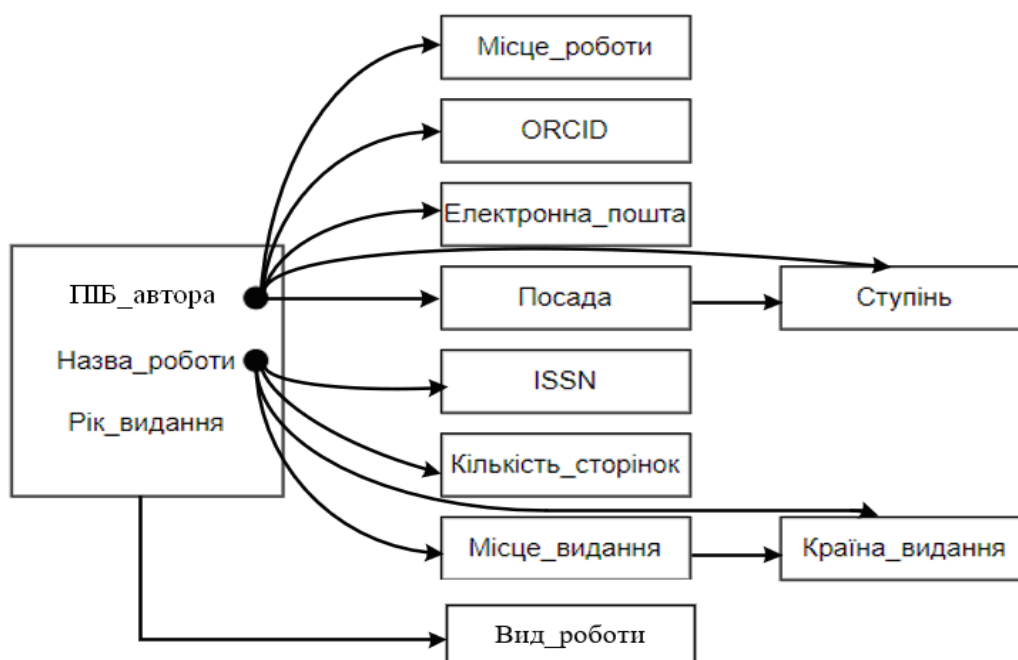
**Аналіз останніх досліджень.** Відомо, що тип бази даних визначається на основі різновиду моделі подання даних (ієрархічна, мережева, реляційна, пост реляційна, багатомірна, об'єктно-орієнтована та інші), яка лежить в її основі. На сучасному етапі існують різні додатки для створення бази даних, а саме: Altibase; Microsoft Access; MySQL Workbench; Oracle; PostgreSQL; SQL Azure та інші. Проведений аналіз останніх досліджень і публікацій [1-3] виявив наступне: 1) розповсюджене використання на сучасному етапі реляційних баз даних; 2) існування широкого спектру програмних додатків по створенню та обробці реляційних баз даних; 3) відсутність єдиного підходу щодо підвищення ефективності створеної бази даних, і став підставою для розробки власної методики по проектуванню та оптимізації створеної бази даних.



**Метою статті** є розробка методики по проектуванню та оптимізації бази даних наукових розробок кафедри електронних обчислювальних машин (ЕОМ).

**Постановка задачі.** Початковими даними для проектування бази даних є дані репозитарія Дніпровського інституту інфраструктури і транспорту (ДІТ) з науковими розробками (Назва\_роботи, Вид\_роботи, Кількість\_сторінок, Рік\_видання, Місце\_видання) та даними про науково-педагогічний склад кафедри ЕОМ (ПІБ\_автора, Посада\_автора, Ступінь, Місце\_роботи, ORCID).

**1. Проектування реляційної бази даних за методом нормальних форм.** У складеному початковому відношенні (ступінь – 13, потужність – 160), що знаходиться в 1НФ, проглядалося збиткове дублювання даних, від якого необхідно було позбутися. Графічна інтерпретація залежностей між атрибутами початкового відношення представлена на рисунку 1.

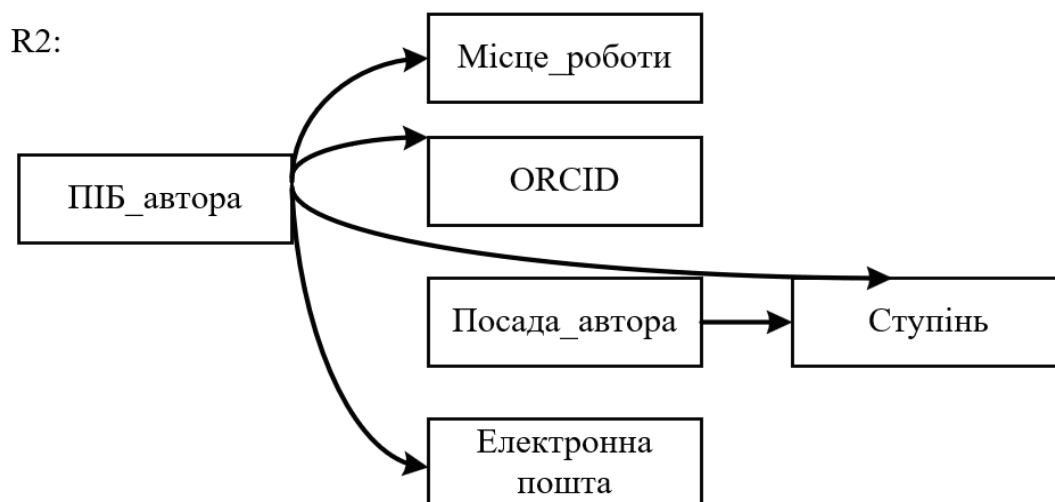


**Рисунок 1 – Залежності між атрибутами початкового відношення**  
Авторська розробка

Мета першого етапу нормалізації – позбавлення від часткових залежностей. Для переведення початкового відношення із 1НФ до 2НФ необхідно скласти наступні проєкції: R1; R2 та R3 (рисунок 2, рисунок 3, рисунок 4 відповідно).

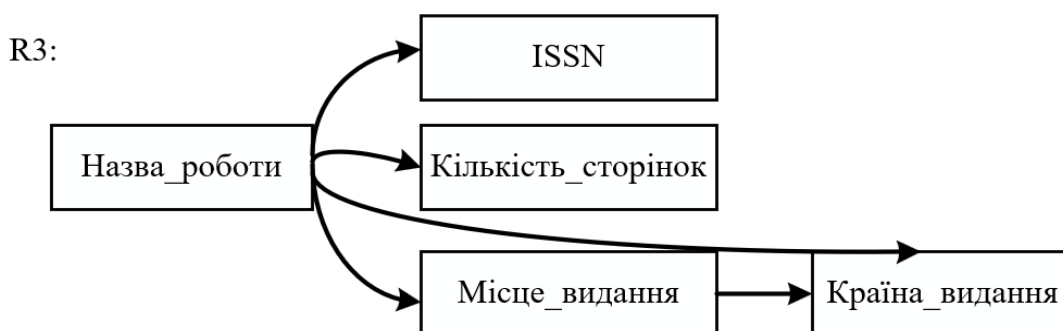


**Рисунок 2 – Залежності між атрибутами відношення R1**  
Авторська розробка



**Рисунок 3 – Залежності між атрибутами відношення R2**

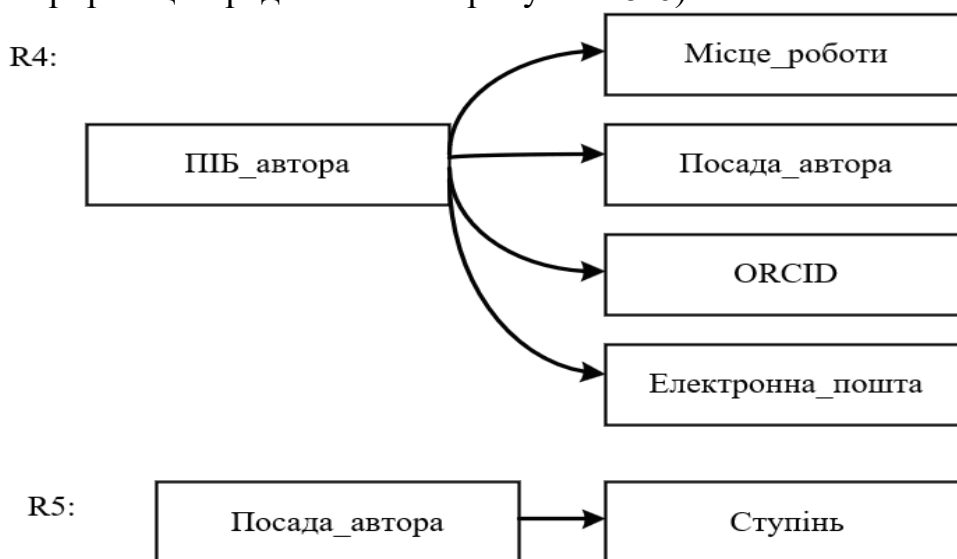
*Авторська розробка*



**Рисунок 4 – Залежності між атрибутами відношення R3**

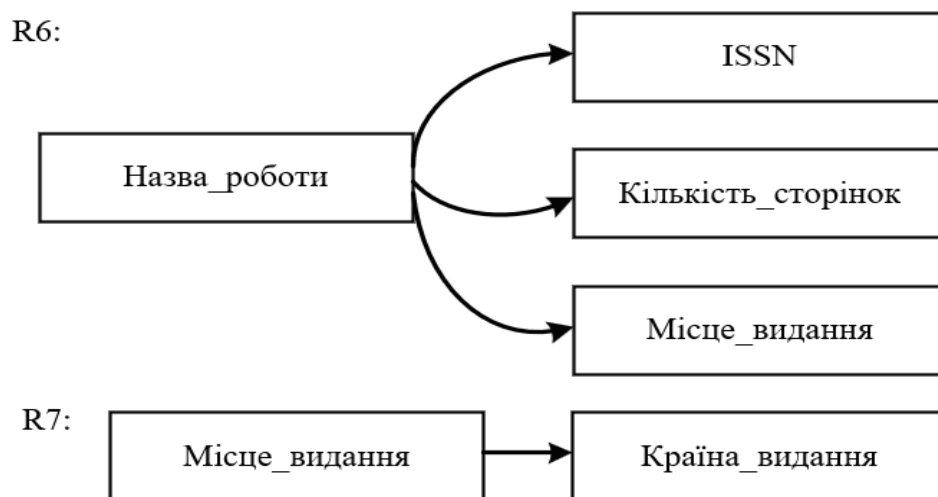
*Авторська розробка*

Але у відношеннях R2 і R3 залишилось збиткове дублювання даних. Мета другого етапу нормалізації – позбавлення від транзитивних залежностей. Для відношень R2 і R3 складемо наступні проекції: R4, R5, R6, R7 (відповідні графічні інтерпретації представлені на рисунках 5-6).



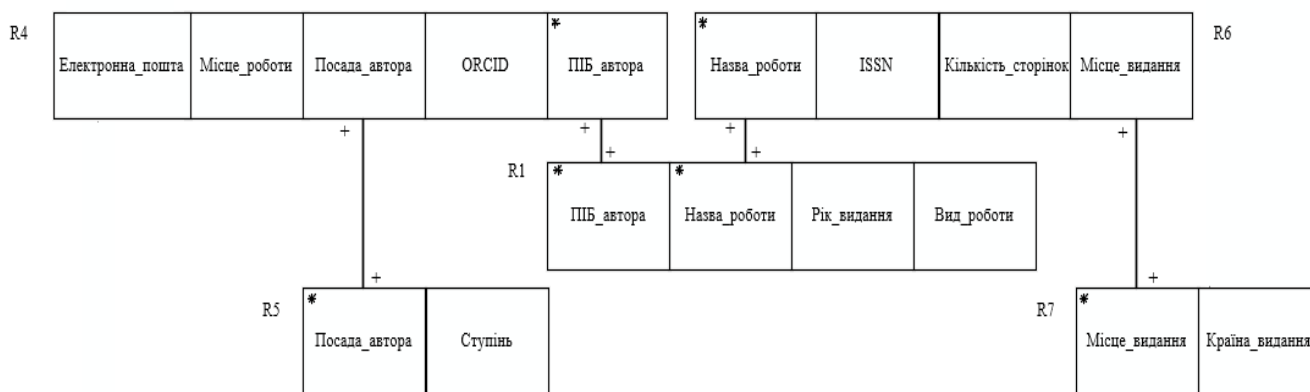
**Рисунок 5 – Залежності між атрибутами відношення R4 та R5**

*Авторська розробка*



**Рисунок 6 – Залежності між атрибутами відношення R6 та R7**  
 Авторська розробка

Загальна структура спроектованої бази даних представлена на рисунку 7.



**Рисунок 7 – Структура бази даних за методом нормальних форм**  
 Авторська розробка

## 2. Проектування реляційної бази даних за методом «Сутність-Зв’язок».

На першому етапі проектування визначені сутності та зв’язки між ними. Виділено наступні сутності: АВТОР (ПІБ\_автора); РОЗРОБКА (Назва\_роботи, Рік\_видання); ВИДАННЯ (Місце\_видання); ПОСАДА (Посада\_автора). Виділено між сутностями наступні зв’язки: АВТОР ПИШЕ РОЗРОБКУ; АВТОР ЗАЙМАЄ ПОСАДУ; РОЗРОБКА ДРУКУЄТЬСЯ ВИДАННЯМ.

На другому етапі проектування побудовано діаграми ER-типу з урахуванням всіх сутностей та зв’язків між ними. Загальна діаграма ER-типу початкового відношення представлена на рисунку 8.

На третьому етапі сформований набір відношень з використанням діаграми ER-типу. Зв’язок АВТОР ПИШЕ РОЗРОБКУ задовольняє умовам правила № 6, відповідно до якого отримано три відношення: АВТОР (ПІБ\_автора, Посада\_автора); РОЗРОБКА (Назва\_роботи, Місце\_видання); ПИШЕ (ПІБ\_автора, Назва\_роботи, Рік\_видання). Зв’язок АВТОР ЗАЙМАЄ ПОСАДУ задовольняє умовам правила № 4, відповідно до якого отримано два відношення: АВТОР (ПІБ\_автора, Посада\_автора); ПОСАДА (Посада\_автора).



Зв'язок РОЗРОБКА ДРУКУЄТЬСЯ ВИДАННЯМ задовольняє умовам правила № 4, відповідно до якого також отримано два відношення: РОЗРОБКА (Назва\_роботи, Місце\_видання); ВИДАННЯ (Місце\_видання).



**Рисунок 8 – Діаграма ER-типу початкового відношення**

*Авторська розробка*

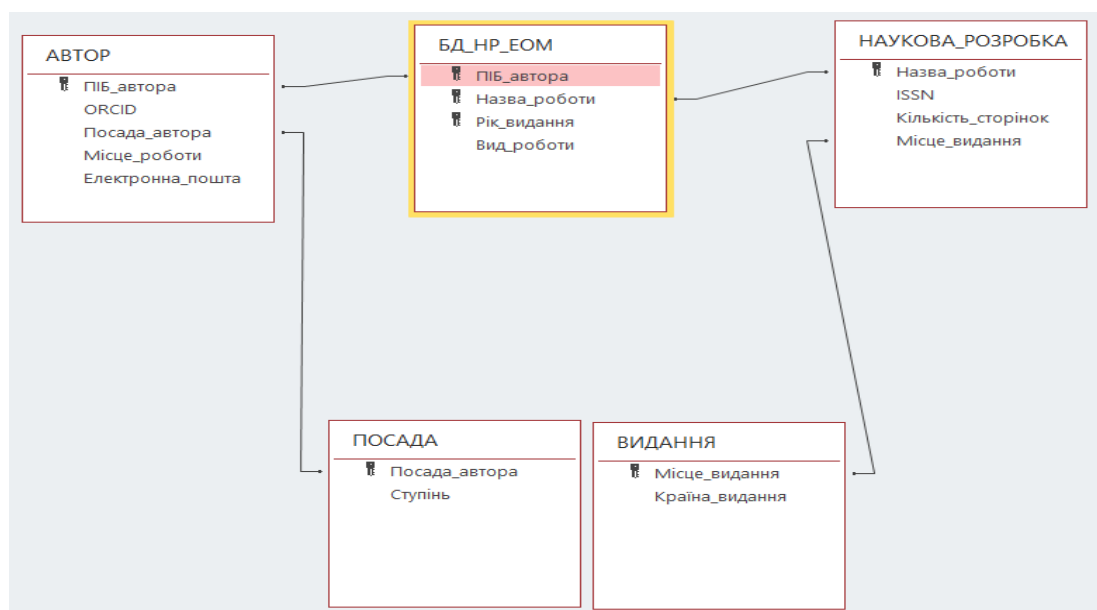
На четвертому етапі додані неключові атрибути, які не були обрані в якості ключових раніше, та призначення їх одному з відношень з тією умовою, щоб відношення відповідали вимогам нормальної форми Бойса-Кодда (БКНФ). Після додавання неключових атрибутів відношення мають наступний вигляд: АВТОР (ПІБ\_автора, Посада\_автора, ORCID, Місце\_роботи, Електронна\_пошта); РОЗРОБКА (Назва\_роботи, Місце\_видання, ISSN, Кількість\_сторінок); ПИШЕ (ПІБ\_автора, Назва\_роботи, Рік\_видання, Вид\_роботи); ПОСАДА (Посада\_автора, Ступінь); ВИДАННЯ (Місце\_видання, Країна\_видання). Усі результуючі відношення відповідають вимогам БКНФ. Таким чином, отримано ті ж відношення, що і при проектуванні бази даних методом нормальних форм: відношення АВТОР співпадає з відношенням R4; відношення РОЗРОБКА співпадає з відношенням R6; відношення ПИШЕ співпадає з відношенням R1; відношення ПОСАДА співпадає з відношенням R5; відношення ВИДАННЯ співпадає з відношенням R7.

**3. Створення в Microsoft Access бази даних з використанням SQL.** За допомогою конструкцій CREATE TABLE та INSERT INTO створено п'ять таблиць, а саме: БД\_НР\_ЕОМ (замість ПИШЕ; 4 атрибути та 161 кортеж); АВТОР (5 атрибутів та 51 кортеж); НАУКОВА\_РОЗРОБКА (4 атрибути та 94 кортежів); ПОСАДА (2 атрибути та 5 кортежів); ВИДАННЯ (2 атрибути та 20 кортежів). Схема даних створена в Microsoft Access (рисунок 9).

Для підвищення ефективності роботи бази даних доречно використання в Microsoft Access наступних механізмів: відключення автосорест імен; налаштування автоматичного стиску та відновлення бази даних; відключення параметрів автосорест. Функція автосорест імен забезпечує працездатність об'єктів бази даних при перейменуванні інших об'єктів від яких вона залежить. З часом продуктивність бази даних може сповільнюватись через те, що залишається виділене місце для видалених або тимчасових об'єктів. Команда «Стискання та відновлення» видаляє це місце і допомагає швидше та ефективніше працювати з базою даних. Параметр автосорест за замовчуванням виправляє орфографію по мірі її написання, щоб підвищити ефективність



роботи бази даних потрібно вимкнути її. Чим більше параметрів функції autocorrect відключено, тим вища продуктивність роботи бази даних.



**Рисунок 9 – Схема даних, що створена в Microsoft Access**

*Авторська розробка*

**4. Оптимізація операторів SQL.** Для досягнення оптимальної швидкості роботи з базою даних необхідно, не тільки оптимізувати спроектовану базу даних (що зроблено при нормалізації відношень), так і оптимізувати оператори доступу до бази даних. З одного боку, погано оптимізована база даних потребує багато зайвих зусиль з оптимізації операторів SQL, з іншого боку, навпаки, гарна оптимізація бази даних не допоможе, якщо використані оператори без відповідної оптимізації. Проведено дослідження часу виконання конструкцій SELECT у наступних випадках: 1) зміна порядку використання таблиць у виразі FROM; 2) зміна розташування умови обмеження у виразі WHERE; 3) зміна послідовності розташування умов обмежування та зв'язування у виразі WHERE. За результатами дослідження час виконання конструкцій SQL зменшується приблизно в два рази за рахуванням наступних умов: розташування невеликих таблиць спочатку переліку у виразі FROM; використання найбільш обмеженої умови у виразі WHERE; розташування найбільш обмеженої умови перед умовою зв'язування у виразі WHERE. Крім того, для підвищення продуктивності роботи бази даних проведено додаткове дослідження призначення простих та композитних індексів.

### **Висновки**

1. На основі даних репозитарія ДІТ та даних про науково-педагогічний склад кафедри ЕОМ створено початкове відношення, ступінь якого 13, а потужність – 160. Проектування бази даних наукових розробок кафедри ЕОМ виконано за наступними методами: нормальних форм (класичний метод) та «Сутність-Зв'язок» (графічний метод). Отримані результати проектування за різними методами співпали, що свідчить про правильність проектування.





Створено в Microsoft Access з використанням SQL базу даних наукових розробок кафедри ЕОМ, яка містить зв'язані між собою п'ять відношень. Для підвищення ефективності роботи бази даних доречно використання наступних механізмів, що передбачені в Microsoft Access: відключення autocorrect імен; налаштування автоматичного стиску та відновлення бази даних; відключення параметрів autocorrect. Визначено, що час виконання зменшується приблизно у два рази при відповідному складанні оптимальних конструкцій SQL.

### Література:

1. Вдовичин Т. Я., Лазурчак Л. В. Проектування інформаційно-пошукових систем як засіб використання сучасних технологій // Вчені записки ТНУ імені В. І. Вернадського. Серія: Технічні науки. Том 33 (72). № 4. 2022. 66-71. URL: <https://doi.org/10.32838/2663-5941/2022.4/11>.
2. Гарбера І. В. Лінгвістична база даних «Концепт людина у фразеології східностепових українських говірок»: структура та функції // Лінгвістичні студії. Вип. 37. 2019. 123-130. DOI: 10.31558/1815-3070.2019.37.21.
3. Лакида П. І., Бідолах Д. І., Кузьович В. С. Просторова база даних урболандшафтів на прикладі зелених насаджень міста Бережани // Науковий вісник НЛТУ України. Том 30. № 4. 2020. 51-56. URL: <https://doi.org/10.1036930/40300409>.

**Abstract.** *A relational database of scientific developments of the Department of Computer Science was designed using the data of the DIIT repository with scientific developments and data on the scientific and pedagogical staff of the department. Improving the performance of the database is associated with achieving the optimal speed of working with the database, which requires both the optimization of the database itself and the optimization of SQL software constructs. When designing the database of scientific developments of the department, the method of normal forms was used, the results of which coincided with the results of designing according to the graphic method «Essence-Relation», which is based on the use of ER-type diagrams. Conducted studies on the assignment of simple and composite indexes, as well as optimization of SQL statements. It is determined that the execution time is reduced by approximately two times when optimizing the SQL program constructs.*

**Key words:** *excessive duplication, design, database optimization, normal forms, ER-type diagram, SQL, optimization of structures, performance.*