

Міністерство освіти і науки України  
Чернівецький національний університет  
імені Юрія Федьковича

## **Технології розробки розподілених баз даних**

**Методичні рекомендації та  
завдання для лабораторних робіт**

Чернівці  
Чернівецький національний університет  
2024

**УДК 004.65 (076.5)**  
**T384**

Рекомендовано до друку Вченою радою факультету математики та інформатики  
Чернівецького національного університету імені Юрія Федьковича  
( протокол № 4 від «23» жовтня 2024 року )

Укладачі:

Олександр Матвій, кандидат фізико-математичних наук, доцент  
кафедри математичного моделювання;

Василь Мельник, кандидат фізико-математичних наук, асистент  
кафедри математичного моделювання;

Ігор Черевко, доктор фізико-математичних наук, професор кафедри  
математичного моделювання.

**T384 Технології розробки розподілених баз даних:** методичні  
рекомендації та завдання для лабораторних робіт. Укл.: Олександр  
Матвій, Василь Мельник, Ігор Черевко – Чернівці: Чернівецький  
національний університет, 2024. – 27 с.

Методичні рекомендації містять завдання для лабораторних робіт та  
теоретичні відомості, необхідні для їх виконання, з дисципліни «Технології  
розробки розподілених баз даних».

Для студентів, що здобувають освіту в галузі знань „Інформаційні  
технології” для спеціальностей 122 - “Комп’ютерні науки”, 124 -  
“Системний аналіз”, 113 - ”Прикладна математика” та інших  
спеціальностей, що пов’язані із вивченням комп’ютерних інформаційних  
технологій.

**УДК 004.65 (076.5)**

## Вступ

Сучасний світ інформаційних технологій швидко розвивається, і розробка програмного забезпечення та управління даними стають одними з найбільш затребуваних і динамічних галузей. Останні десятиліття характеризуються стрімким зростанням обсягу даних, необхідністю їх обробки в реальному часі та зростаючими вимогами до стійкості й масштабованості систем. Ці виклики та потреби стимулювали розвиток нових підходів у побудові розподілених баз даних, архітектурі систем і взаємодії між компонентами великих розподілених середовищ.

Дисципліна “Технології розробки розподілених баз даних” призначена для вивчення основ проєктування, організації та управління реляційними і нереляційними базами даних у розподілених середовищах. Основними завданнями дисципліни є ознайомлення студентів з принципами побудови масштабованих систем, а також з основними аспектами синхронізації та узгодження даних.

Методичні вказівки містять завдання до лабораторних робіт, теоретичні відомості, які необхідні для їх виконання, а також практичні приклади, які допоможуть студентам самостійно засвоїти здобуті знання та застосувати їх для побудови стійких, ефективних і масштабованих систем. У даних методичних вказівках розглядаються як класичні підходи, такі як нормалізація даних і транзакційні системи, так і новітні тенденції, зокрема використання NoSQL, мікросервісної архітектури та системи обміну повідомленнями, такі як Apache Kafka. Особлива увага приділена питанням забезпечення стійкості та продуктивності, а також принципам транзакційності (ACID) і алгоритмам координації, таким як 2PC і 3PC.

Матеріали методичних вказівок можуть бути корисними студентам різних спеціальностей всіх форм навчання, як для самостійного вивчення теоретичного матеріалу, так і в якості допоміжного засобу при організації дистанційного навчання.

## **Мета та завдання навчальної дисципліни** “Технології розробки розподілених баз даних”

Метою вивчення дисципліни «Технології розробки розподілених баз даних» є отримання знань, навичок і вмінь у проєктуванні, організації та управлінні розподіленими базами даних, а також програмуванні прикладних систем обробки даних.

Завданнями дисципліни є:

- ознайомлення студентів з теоретичними основами класичної та сучасної теорії розподілених баз даних;
- вивчення важливих практичних аспектів створення та управління розподіленими базами даних.

В результаті вивчення дисципліни студенти повиненні

### **Знати:**

- типи моделей розподілених даних, особливості їх організації, переваги і недоліки;
- етапи проєктування розподілених баз даних;
- основні компоненти графових баз даних та їх призначення;
- спеціалізовані мови запитів для управління даними в графових базах даних, зокрема SPARQL;
- принципи роботи систем клієнт-сервер та проміжного програмного забезпечення для графових баз даних;
- технології ODBC та ADO.NET для забезпечення доступу до зовнішніх даних;
- сучасні перспективні напрями досліджень у сфері технологій розподілених баз даних.

### **Уміти:**

- формулювати задачу для конкретної предметної області з використанням сучасного проміжного програмного забезпечення;
- створювати розподілену базу даних, організувати пошук та маніпулювання даними за допомогою мови запитів SPARQL;
- організувати забезпечення надійності систем зберігання та обробки розподілених даних.

# Основні поняття і визначення теорії розподілених баз даних

## 1. Реляційні бази даних. Нормалізація

### Визначення реляційної бази даних

Реляційна база даних — це тип бази даних, яка зберігає та організовує дані у вигляді таблиць. Таблиці (або реляції) складаються з рядків (записів) і стовпців (полів). Кожна таблиця описує певну сутність (наприклад, клієнти, продукти тощо), а кожен рядок представляє окремий екземпляр цієї сутності [1-2].

Основна ідея реляційних баз даних полягає у використанні зв'язків між таблицями для організації даних без надмірного дублювання, що забезпечує ефективне зберігання та запити до даних.

### Нормалізація

Нормалізація — це процес організації даних у базі даних з метою мінімізації надмірності та запобігання аномаліям під час операцій вставки, оновлення та видалення. Основна мета нормалізації полягає в тому, щоб зберігати дані у вигляді окремих логічних частин, забезпечуючи, щоб кожен факт зберігався тільки один раз [1].

Нормалізація включає кілька послідовних етапів, які називаються нормальними формами (NF), кожен з яких додає додаткові вимоги до структури даних.

Перша нормальна форма (1NF) є основою нормалізації бази даних і має на меті забезпечити структурованість даних у таблицях. Основні вимоги 1NF:

- Атомарність значень.  
Кожне поле (стовпець) таблиці повинно містити лише одне значення, тобто значення не повинні бути розділені на кілька частин. Наприклад, у полі "Телефон" не повинно бути кількох номерів, а лише один.
- Унікальність записів.  
Кожен запис (рядок) у таблиці повинен бути унікальним, що зазвичай досягається за допомогою первинного ключа. Це означає, що не може бути двох однакових рядків у таблиці.