

**ДЕРЖАВНИЙ ТОРГОВЕЛЬНО-ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ**

**СИСТЕМА УПРАВЛІННЯ ЯКІСТЮ**

**Система забезпечення якості освітньої діяльності та якості вищої освіти**  
*сертифікована на відповідність ДСТУ ISO 9001:2015 / ISO 9001:2015*

**Кафедра комп'ютерних наук та інформаційних систем**

**ЗАТВЕРДЖЕНО**

вченою радою ДТЕУ

(пост. п. 6 від «28» 03 2024 р.)

Ректор



Анатолій МАЗАРАКІ

**ТЕХНОЛОГІЇ СТВОРЕННЯ РОЗПОДІЛЕНИХ  
БАЗ ДАНИХ ТА ЗНАНЬ/  
TECHNOLOGIES FOR CREATING DISTRIBUTED  
DATABASES AND KNOWLEDGE**

**ПРОГРАМА /  
COURSE SUMMARY**

**Київ 2024**

**Розповсюдження і тиражування без офіційного дозволу ДТЕУ заборонено**

Автори: Г. Т. САМОЙЛЕНКО, кандидат фізико-математичних наук, доцент,  
А.В. СЕЛІВАНОВА, старший викладач

Програму розглянуто і затверджено на засіданні кафедри комп'ютерних наук та інформаційних систем 05 березня 2024 р., протокол № 27

Рецензенти: Т.О. ФІЛІМОНОВА, кандидат фізико-математичних наук,  
доцент,  
С.В. ВОРОНОВ, Генеральний директор Українсько-швейцарського підприємства «Міжнародний комп'ютерний клуб України», кандидат фізико-математичних наук

**ТЕХНОЛОГІЇ СТВОРЕННЯ РОЗПОДІЛЕНИХ  
БАЗ ДАНИХ ТА ЗНАНЬ/  
TECHNOLOGIES FOR CREATING DISTRIBUTED  
DATABASES AND KNOWLEDGE**

**ПРОГРАМА /**

**COURSE SUMMARY**

## ВСТУП

Програма дисципліни «Технології створення розподілених баз даних та знань» призначена для здобувачів другого рівня вищої освіти галузі знань 12 «Інформаційні технології», спеціальності 122 «Комп'ютерні науки», освітньої програми «Комп'ютерні науки», та спеціальності 126 «Інформаційні системи та технології», освітньої програми «Інформаційні системи та технології».

Програму підготовлено з урахуванням вимог Стандартів вищої освіти України та відповідних освітньо-професійних програм підготовки магістрів.

Розроблена програма складається з таких розділів:

1. Мета, завдання та предмет дисципліни.
2. Передумови вивчення дисципліни як вибіркової компоненти освітньої програми.
3. Результати вивчення дисципліни.
4. Зміст дисципліни.
5. Список рекомендованих джерел.

### 1. МЕТА, ЗАВДАННЯ ТА ПРЕДМЕТ ДИСЦИПЛІНИ

*Метою* вивчення дисципліни «Технології створення розподілених баз даних та знань» є надання поглиблених знань та практичних навичок щодо теорії та практики побудови й використання баз даних.

*Завданням* вивчення дисципліни «Технології створення розподілених баз даних та знань» є засвоєння методів створення розподілених баз даних та знань, технологій їх проектування, наповнення та підтримання в робочому стані.

*Предметом* вивчення дисципліни є методи та технології проектування баз даних.

### 2. ПЕРЕДУМОВИ ВИВЧЕННЯ ДИСЦИПЛІНИ ЯК ВИБІРКОВОЇ КОМПОНЕНТИ ОСВІТНЬОЇ ПРОГРАМИ

*Знання:*

- принципів побудови розподілених систем обробки інформації;
- методів роботи в розподілених системах обробки інформації;
- способів введення та зберігання даних в інформаційних системах;
- особливостей використання клієнтських і серверних технологій при проектуванні розподілених баз даних;
- ролі розподілених систем обробки інформації в сфері професійної діяльності;
- переваг та недоліків розподілених СУБД;
- стратегічних цілей визначення і розподілу фрагментів розподілених баз даних;
- різних способів розміщення даних у системі;
- схем фрагментації глобальних відносин;

- переваг і недоліків синхронної й асинхронної реплікації.

*Вміння:*

- використовувати клієнтські і серверні технології побудови та експлуатації розподілених баз даних;
- розділяти відношення на фрагменти та розподіляти фрагменти по вузлах;
- забезпечувати прозорість у розподілених СУБД;
- виявляти і розв'язувати конфлікти в транзакціях;
- організовувати резервне копіювання;
- організовувати способи відновлення розподілених даних.

### 3. РЕЗУЛЬТАТИ ВИВЧЕННЯ ДИСЦИПЛІНИ

Дисципліна «Технології створення розподілених баз даних та знань», як вибіркова компонента освітньо-професійної програми, забезпечує оволодіння студентами загальними та фаховими компетентностями і досягнення ними програмних результатів навчання за відповідними освітніми програмами:

*Комп'ютерні науки (ОС «Магістр»)*

<b>Номер в освітній програмі</b>	<b>Зміст компетентності</b>	<b>Номер теми, що розкриває зміст компетентності</b>
<i>Загальні компетентності за освітньо-професійною програмою</i>		
ЗК 2	Здатність до абстрактного мислення, аналізу та синтезу.	1,2,3,4,5,10,11,12,13,14
ЗК 5	Здатність вчитися й оволодівати сучасними знаннями.	1,2,3,4,8,9,10,11,12,13,14
<i>Фахові компетентності за освітньо-професійною програмою</i>		
СК 07	Здатність розробляти програмне забезпечення відповідно до сформульованих вимог з урахуванням наявних ресурсів та обмежень.	3,8,10
СК 09	Здатність розробляти та адмініструвати бази даних та знань.	7,8,9,10,11,12
СК 11	Здатність ініціювати, планувати та реалізовувати процеси розробки інформаційних та комп'ютерних систем та програмного забезпечення, включно з його розробкою, аналізом, тестуванням, системною інтеграцією, впровадженням і супроводом.	5,6,7,8,9,10,11,12
<i>Програмні результати навчання за освітньо-професійною програмою</i>		
РН 2	Мати спеціалізовані уміння/навички розв'язання проблем комп'ютерних наук, необхідні для проведення досліджень та/або провадження інноваційної діяльності з метою розвитку нових знань	7,8,9,10,11,12

	та процедур.	
PH 12	Проектувати та супроводжувати бази даних та знань.	5,6,7,8,9,10,11,12
<i>Інформаційні системи та технології (ОС «Магістр»)</i>		
<i>Загальні компетентності за освітньо-професійною програмою</i>		
ЗК 01	Здатність до абстрактного мислення, аналізу та синтезу.	1,2,3,4,10,11,12,13,14
ЗК 03	Здатність спілкуватися з представниками інших професійних груп різного рівня (з експертами з інших галузей знань/видів економічної діяльності).	1,2,3,4,5,6,10,11,12,13,14
<i>Фахові компетентності за освітньо-професійною програмою</i>		
СК 01	Здатність розробляти та застосувати ІСТ, необхідні для розв'язання стратегічних і поточних задач.	2,4,5,6,7,8,9
СК 03	Здатність проектувати інформаційні системи з урахуванням особливостей їх призначення, неповної/недостатньої інформації та суперечливих вимог.	5,6,7,8,9,10,11,12
<i>Програмні результати навчання за освітньо-професійною програмою</i>		
PH 03	Приймати ефективні рішення з проблем розвитку інформаційної інфраструктури, створення і застосування ІСТ.	9,10,11,12,13,14
PH 09	Розробляти і використовувати сховища даних, здійснювати аналіз даних для підтримки прийняття рішень.	8,9,10,11,12
PH 10	Забезпечувати якісний кіберзахист ІСТ, планувати, організовувати, впроваджувати та контролювати функціонування систем захисту інформації.	8,9,10,11,12,13,14

#### **4. ЗМІСТ ДИСЦИПЛІНИ**

##### **Тема 1. Розподілені бази даних.**

Основні принципи, покладені в основу теорії баз даних. Поняття розподіленої бази даних (Distributed DataBase - DDB). Основна задача систем управління розподіленими базами даних. Система управління розподіленими базами даних. Однорідні та неоднорідні розподілені бази даних. Принцип К.Дж. Дейта. Властивості розподілених баз даних згідно К.Дж. Дейту: Локальна автономія. Незалежність від центрального вузла. Безперервні операції. Прозорість розміщення. Прозора фрагментація Прозорість тиражування. Обробка розподілених запитів. Незалежність від обладнання. Незалежність від операційних систем. Прозорість мережі. Незалежність від баз даних. Переваги і недоліки розподілених СУБД (РСУБД).

##### **Список рекомендованих джерел**

*Основний: 1 [с.6-28]*

*Додатковий: 2 [с.1-26]*

*Інтернет-ресурси: 1,2*

## **Тема 2. Архітектура програмно-технічних засобів розподілених СУБД.**

Властивості архітектури. Розподіленість. Неоднорідність. Автономність. Два типи РСУБД залежно від типу програмного забезпечення: однорідні та неоднорідні. Різновиди архітектури. Клієнт-серверна архітектура. Архітектура з багатьма незалежними серверами. Архітектура із взаємодіючими серверами. Архітектура однорангової мережі. Розподілене зберігання даних. Два основні механізми розподіленого зберігання даних - фрагментація; реплікація. Поділ розподілених архітектур баз даних по типам: системи недублюючого розбиття (при великому об'ємі даних, що часто змінюються), системи частикового дублювання (при невеликому об'ємі даних, що часто змінюються), системи повного дублювання при невеликому об'ємі даних, що рідко змінюються).

### **Список рекомендованих джерел**

*Основний: 1 [с.6-28], 2 [с.42-54]*

*Додатковий: 2 [с.1-26]*

*Інтернет-ресурси: 1*

## **Тема 3. Основні проблеми створення розподілених баз даних**

Основні проблеми створення розподілених баз даних: Фрагментація даних і розподілення по комп'ютерам. Створення глобального каталогу з інформацією про кожний фрагмент БД та його місце в мережі. (Каталог може бути на одному вузлі або бути розподіленим). Організація обробки запитів (синхронізація кількох запитів до одних і тих же даних, виключення аномалій видалення і оновлення одних і тих же даних, що знаходяться на різних вузлах, оптимізація послідовності кроків по обробці запитів і т.д.).

### **Список рекомендованих джерел**

*Основний: 1 [с.29-68]*

*Додатковий: 2 [с.1-26]*

*Інтернет-ресурси: 1*

## **Тема 4. Процеси побудови розподілених баз даних.**

Принципи побудови розподілених баз даних. Мінімізація інтенсивності обміну даними. Оптимальне розміщення серверних та клієнтських додатків у мережі. Декомпозиція даних на сегменти, що використовуються часто і рідко (для правильного настроювання реплікації – розміщення даних, що часто використовуються на АРМ кінцевих користувачів). Збереження копій даних і виконання дій по підтримці цілісності розподіленої ІС. Паралельні процеси. Методи побудови розподілених баз даних «зверху вниз» і «знизу нагору». Розподілені обчислення. Двомірне подання даних кінцевому користувачеві. Багатомірне подання

при описі структур даних. Розклад транзакцій Транзакції, властивості транзакцій.

#### **Список рекомендованих джерел**

*Основний: 1 [с.29-68], 2 [с.49-58]*

*Додатковий: 2 [с.1-26]*

*Інтернет-ресурси: 1*

#### **Тема 5. Обробка розподілених запитів. Засоби SQL для розподілених систем.**

Загальна характеристика мови. Реляційна алгебра як основа SQL. Поділення мови SQL на чотири частини: оператори визначення даних (англ. Data Definition Language, DDL) оператори маніпуляції даними (англ. Data Manipulation Language, DML) оператори визначення доступу до даних (англ. Data Control Language, DCL) оператори управління транзакціями (англ. Transaction Control Language, TCL). Засоби пошуку даних. Засоби маніпулювання даними. Операції над схемою бази даних. Віртуальні таблиці та індекси. Транзакції. Обробка розподілених запитів. Переваги та недоліки SQL.

#### **Список рекомендованих джерел**

*Основний: 1 [с.29-68]*

*Додатковий: 2 [с.1-26]*

*Інтернет-ресурси: 1*

#### **Тема 6. Розподіл даних.**

Централізоване розміщення даних. Роздільне (фрагментоване) розміщення даних. Розміщення з повною реплікацією. Розміщення з вибірковою реплікацією. Горизонтальна фрагментація. Вертикальна фрагментація. Змішана фрагментація. Методи поділу відносин на фрагменти й розподіл фрагментів по вузлах. Оптимальний розподіл фрагментів за вузлами мережі. Додаткова інформація, яка стосується розподілу фрагментів за вузлами мережі.

#### **Список рекомендованих джерел**

*Основний: 1 [с.71-80], 2 [с.381-400]*

*Додатковий: 2 [с.1-26]*

*Інтернет-ресурси: 1*

#### **Тема 7. Реплікація.**

Визначення. Переваги. Механізми реплікації. Сервери: видавець, дистриб'ютор і передплатник. Два методи відновлення даних передплатників: реплікація за запитом та примусова реплікація. Моделі реплікації. Реплікація моментальних знімків. Реплікація транзакцій. Топологія реплікацій. Функції служби реплікації. Схеми володіння даними. Збереження цілісності транзакцій. Моментальні знімки таблиць. Тригери баз даних. Виявлення та розв'язання конфліктів.

#### **Список рекомендованих джерел**

*Основний: 1 [с.71-80]*

*Додатковий: 2 [с.1-26]*

*Інтернет-ресурси: 1,2*

### **Тема 8. Проектування розподілених реляційних баз даних.**

Критерії побудови розподілених баз даних: Всебічний аналіз інформаційних потреб предметної області з виявленням об'ємів х даних, що зберігаються, їх складності, достовірності, взаємозв'язку. Моделювання мережевого трафіку при роботі розподілених баз даних з різними моделями реплікації даних. Кластеризація елементів даних і програм їх обробки з метою - добитися максимальної автономності та слабкої пов'язаності кластерів. Прив'язка кластерів даних до ймовірносних користувачів чи АРМ. Підтримка еталонної копії даних і обмеження реплікаційного механізму. Розробка і реалізація правил приведення локальних і центральної бази даних в несуперечливий стан. Спеціалізовані програмні модулі для створення розподілених баз даних від фірм-розробників програмного забезпечення ГІС. Модуль ArcSDE (Spatial Database Engine) фірми ESRI, який дозволяє зовнішнім користувачам працювати з просторовими базами даних. Модуль ArcGIS, що дозволяє працювати із зовнішніми базами даних. Модуль SDE, який забезпечує сучасні функції створення розподілених баз даних, підтримки різних моделей просторових даних).

#### **Список рекомендованих джерел**

*Основний: 1 [с.81-96]*

*Додатковий: 2 [с.1-26]*

*Інтернет-ресурси: 1*

### **Тема 9. Організація баз знань**

Загальна характеристика баз знань. Базові поняття. Продукційна модель знань. Семантична модель знань. Моделі подання знань. Інтелектуальний аналіз даних. Склад і структура баз знань. Засоби побудови та опрацювання баз знань. Поняття сховищ даних та їх місце в системах оброблення даних. Поняття віртуального, корпоративного, глобального сховища даних та кіоска (вітрини) даних. Поняття залежного та незалежного кіоска (вітрини) даних. Архітектура сховищ даних. Технологія ETL (extract, transform and load) та характеристика її основних фаз. Віртуалізація сховищ. Характеристика основних ознак сховищ даних: предметна орієнтація, інтегрованість, підтримка хронології, незмінність і мінімальна надлишковість. Вимоги до вибору моделей побудови сховищ даних. Характеристика просторової моделі побудови сховища даних і її різновиди моделі „зірка” та „сніжинка”. Характеристики основних елементів багатовимірного представлення даних

#### **Список рекомендованих джерел**

*Основний: 1 [с.81-96], 3 [с.120-168]*

*Додатковий: 2 [с.1-26]*

*Інтернет-ресурси: 1,2*

### **Тема 10. Технологія проектування сховищ даних**



Оперативна OLTP та аналітична OLAP обробка даних. OLTP(OnLine Transaction Processing): використання, вимоги, переваги і недоліки. OLAP(OnLine Analytical Processing): дія OLAP, правила для систем OLAP. Концепція сховища даних. Основні характеристики сховищ даних. Модель та 12 правил Кодда, яким повинен задовольняти програмний продукт класу OLAP. Тест FASMI. Властивості інформаційних сховищ. Основні елементи і операції OLAP. Типи OLAP. Багатомірний OLAP (MOLAP). Реляційний OLAP (ROLAP) Гібридні OLAP (HOLAP) Основні компоненти інформаційного сховища. Робота з гіперкубом. Реалізація сховищ і вітрин даних. Проблеми інтеграції даних. Неоднорідність програмного середовища. Розподілений характер організації. Підвищення вимог до безпеки даних. Необхідність наявності багаторівневих довідників метаданих. Наявні програмні реалізації. Архітектура OLAP-додатків. Реалізація серверів. Основи побудови банків даних БНД. Підходи до проектування сховищ даних. Автоматизація проектування сховищ даних.

#### **Список рекомендованих джерел**

*Основний: 1 [с.81-96], 3 [с.198-236]*

*Додатковий: 2 [с.1-26]*

*Інтернет-ресурси: 1, 2*

#### **Тема 11. Забезпечення безпеки РСУБД.**

Поняття адміністрування баз та сховищ даних. Основні функції адміністратора БД та сховищ даних. Поняття інформаційної безпеки баз і сховищ даних. Типові загрози безпеки та їх запобігання. Захист даних – функція адміністратора. Рівні захисту. Привілеї та їх визначення. Прозорість розподілу. Прозорість фрагментації. Прозорість розташування. Прозорість реплікації. Прозорість локального відображення. Прозорість іменування. Прозорість транзакцій. Прозорість паралельності. Прозорість відмов. Прозорість виконання. Прозорість використання СУБД. Засоби забезпечення конфіденційності баз даних. Ідентифікація та перевірка справжності користувачів. Засоби ідентифікації і аутентифікації об'єктів баз даних. Обліковий запис. Режими аутентифікації. Організація взаємодії СУБД і базової ОС. Управління ключами безпеки. Введення в мережевий обмін даних. Дискреційна модель управління доступом. Основні категорії користувачів. Використання схем для забезпечення безпеки. Поділ користувачів і схем. Види привілеїв. Рольова модель розмежування доступу. Концепція і реалізація механізму ролей. Мандатне управління доступом. Мітки конфіденційності. Рівні конфіденційності об'єктів і рівні довіри суб'єктів доступу. Примусовий контроль доступу.

#### **Список рекомендованих джерел**

*Основний: 1 [с.81-96]*

*Додатковий: 1 [с.327-355], 3[с.51-57]*

*Інтернет-ресурси: 1*

#### **Тема 12. Засоби контролю цілісності інформації**

Загрози безпеці розподілених систем. Розподілені транзакції. Методи розподілу даних. Технологія тиражування даних. Тиражування даних (Data Replication - DR). Реплікатор - спеціальний модуль СУБД - сервер тиражування даних. Переваги та недоліки DR-технології. Цілісність даних. Необхідність відновлення. Транзакції й відновлення. Функції відновлення: механізм резервного копіювання, файл журналу. Створення контрольних вузлів. Методи відновлення. Відновлення з використанням відкладеного відновлення. Відновлення з використанням негайного відновлення. Метод тінювого сторінкового обміну. Засоби забезпечення цілісності даних. Засоби контролю цілісності інформації. Засоби забезпечення доступності баз даних. Аудит системи безпеки бази даних.

#### **Список рекомендованих джерел**

*Основний: 1 [с.81-96]*

*Додатковий: 1 [с. 356-371], 3[с.51-57]*

*Інтернет-ресурси: 1*

### **Тема 13. Безпека розподілених інформаційних систем**

Основні поняття безпеки розподілених інформаційних систем. Криптографічний захист інформації. Захист інформації в каналах зв'язку. Інформаційна безпека в мережах. Захист від несанкціонованого доступу. Принципи організації безпеки комп'ютерних мереж. Методи та засоби забезпечення вимог політики безпеки комп'ютерної мережі. Захист приватної мережі. Етапи проектування комплексної системи захисту розподіленої інформаційної системи.

#### **Список рекомендованих джерел**

*Основний: 1 [с.81-96]*

*Додатковий: 1 [с. 407-451], 3[с.51-57], 4 [с.75-81]*

*Інтернет-ресурси: 1*

### **Тема 14. Інформаційна безпека підприємств і організацій**

Модель багаторівневого захисту розподілених інформаційних систем підприємств і організацій. Засоби забезпечення безпеки інформаційних систем. Безпека інформації на об'єктах. Канали витоку інформації. Системи інформаційної безпеки.

#### **Список рекомендованих джерел**

*Основний: 1 [с.81-96]*

*Додатковий: 1[с. 500-562], 3[с.51-57], 4 [с.75-81]*

*Інтернет-ресурси: 1*

## **5. СПИСОК РЕКОМЕНДОВАНИХ ДЖЕРЕЛ**

### **Основний:**

1. Ярцев В.П. Розподілені бази даних: навчальний посібник. К. ДУТ. 2018. 97с.

2. Берко А.Ю., Верес О.М., Пасічник В.В. Системи баз даних та знань. Книга 1. Організація баз даних та знань: навч. посібник. Львів : «Магнолія-2006». 2021. 440 с.
3. Берко А.Ю., Верес О.М., Пасічник В.В. Системи баз даних та знань. Книга 2. Системи управління базами даних та знань: навч. посібник. Львів : «Магнолія-2006». 2021. 584 с.

#### **Додатковий:**

1. Бобало Ю.Я., Горбатий І.В., Кіселичник М.Д. та ін.; за заг. ред. Ю.Я. Бобала, І.В. Горбатого. Інформаційна безпека: навч. посіб. / Ю.Я. Бобало, І.В. Горбатий, М.Д. Кіселичник та ін.; за заг. ред. Ю.Я. Бобала, І.В. Горбатого. – Львів : Львівська політехніка, 2019. – 580 с
2. *Самойленко Г.Т., Селіванова А.В. Технології створення розподілених баз даних та знань: методичні рекомендації до лабораторних занять. Київ: ДТЕУ. 2024. 26 с.*
3. Самойленко Г.Т., Юрченко Ю.Ю. Особливості математичного обґрунтування комплексної системи захисту даних медичного закладу. Математичні машини і системи. 2023. № 4. С. 51–57.
4. Юрченко Ю.Ю. Впровадження елементів системи захисту підприємства. Математичні машини і системи. 2023. № 2. С. 75–81

#### **Інтернет-ресурси:**

1. Official MySQL Database Site. URL: <https://www.mysql.com/>
2. Проектування розподілених баз даних. URL: <https://helpiks.org/7-49596.html>.

\*- Курсивом виділені джерела, що є в бібліотеці ДТЕУ