



## Силабус освітнього компонента

Програма навчальної дисципліни

# Програмування, бази даних і знань

Шифр та назва спеціальності

122 – Комп'ютерні науки

Інститут

ННІ КНІТ

Навчально науковий інститут комп'ютерних наук та інформаційних технологій

Освітня програма

Комп'ютерні науки

Кафедра

Кафедра управління проєктами в інформаційних технологіях

Рівень освіти

Магістр

Тип дисципліни

Обов'язкова

Семестр

2

Мова викладання

Українська

## Викладачі, розробники



**Лисенко Антон Олександрович**

[anton.lysenko@khpі.edu.ua](mailto:anton.lysenko@khpі.edu.ua)

Кандидат технічних наук, асистент кафедри управління проєктами в інформаційних технологіях НТУ «ХПІ»

Досвід роботи – понад 11 років. Автор 26 наукових та навчально-методичних праць. Провідний лектор з дисциплін: «Основи web-технологій», «Основи Python для Data Science», «Стек технологій», «Програмування, бази даних і знань»

[Детальніше про викладача на сайті кафедри](#)

## Загальна інформація

### Анотація

Курс надає студентам фундаментальні знання та практичні навички в розробці, управлінні та оптимізації баз даних. Програма охоплює сучасні методи роботи з реляційними та нереляційними базами даних. Студенти дізнаються про основні принципи побудови баз даних, розробку SQL-запитів, нормалізацію даних, а також про сучасні технології, що використовуються для аналізу та зберігання великих обсягів даних.

### Мета та цілі дисципліни

Метою курсу є надання студентам знань та навичок, необхідних для створення та управління базами даних та системами управління знаннями. У результаті навчання студенти зможуть:

- розуміти принципи побудови баз даних та основи їх архітектури;
- створювати та підтримувати реляційні та нереляційні бази даних;
- оптимізувати бази даних для забезпечення високої продуктивності та надійності;
- розробляти та виконувати складні SQL-запити для витягування та маніпулювання даними;
- інтегрувати бази даних з додатками та сервісами для автоматизації бізнес-процесів;
- розуміти принципи забезпечення безпеки даних та застосовувати відповідні методи захисту.

Курс орієнтований на підготовку фахівців, здатних ефективно працювати з сучасними інформаційними системами та застосовувати отримані знання на практиці.

### **Формат занять**

Лекції, лабораторні роботи, самостійна робота, консультації. Підсумковий контроль – іспит.

### **Компетентності**

ЗК2 Здатність застосовувати знання у практичних ситуаціях.

СК1 Усвідомлення теоретичних засад комп'ютерних наук.

СК2 Здатність формалізувати предметну область певного проекту у вигляді відповідної інформаційної моделі.

СК4 Здатність збирати і аналізувати дані (включно з великими), для забезпечення якості прийняття проектних рішень.

СК5 Здатність розробляти, описувати, аналізувати та оптимізувати архітектурні рішення інформаційних та комп'ютерних систем різного призначення.

СК6 Здатність застосовувати існуючі і розробляти нові алгоритми розв'язування задач у галузі комп'ютерних наук.

СК7 Здатність розробляти програмне забезпечення відповідно до сформульованих вимог з урахуванням наявних ресурсів та обмежень

СК9 Здатність розробляти та адмініструвати бази даних та знань.

### **Результати навчання**

РН1 Мати спеціалізовані концептуальні знання, що включають сучасні наукові здобутки у сфері комп'ютерних наук і є основою для оригінального мислення та проведення досліджень, критичне осмислення проблем у сфері комп'ютерних наук та на межі галузей знань.

РН2 Мати спеціалізовані уміння/навички розв'язання проблем комп'ютерних наук, необхідні для проведення досліджень та/або провадження інноваційної діяльності з метою розвитку нових знань та процедур.

РН6 Розробляти концептуальну модель інформаційної або комп'ютерної системи.

РН9 Розробляти алгоритмічне та програмне забезпечення для аналізу даних (включно з великими).

РН10 Проектувати архітектурні рішення інформаційних та комп'ютерних систем різного призначення.

РН11 Створювати нові алгоритми розв'язування задач у сфері комп'ютерних наук, оцінювати їх ефективність та обмеження на їх застосування.

РН12 Проектувати та супроводжувати бази даних та знань.

РН13 Оцінювати та забезпечувати якість інформаційних та комп'ютерних систем різного призначення.

РН14 Тестувати програмне забезпечення.

### **Обсяг дисципліни**

Загальний обсяг дисципліни 150 год. (5 кредитів ECTS): лекції – 32 год., лабораторні роботи – 32 год., самостійна робота – 86 год.

### **Передумови вивчення дисципліни (пререквізити)**

Основи програмування, Алгоритми та структури даних, Об'єктно-орієнтоване програмування, Інтелектуальний аналіз даних, Сучасні методи математичного та комп'ютерного моделювання, Основи наукових досліджень,

### **Особливості дисципліни, методи та технології навчання**

Презентації лекцій з дисципліни, методичне забезпечення для виконання лабораторних робіт в електронному вигляді. Тести для поточного та підсумкового контролю знань і вмінь студентів.

# Програма навчальної дисципліни

## Теми лекційних занять

### Тема 1. Введення у дисципліну

Історія розвитку та передумови появи СУБД. Що таке дані? Що таке інформаційна система? Фізична та логічна незалежність даних. Поняття моделі даних та відмінності між ними. Реляційна модель даних та основні операції реляційної алгебри. Архітектура СУБД та її функції. Поняття ідентифікації, навігації, масової та штучної обробки даних. Мови баз даних, основні типи даних у популярних СУБД та їх особливості. Поняття схеми бази даних та версіонування схеми БД.

### Тема 2. Введення у проектування БД

Життєвий цикл бази даних. Інфологічне проектування, включаючи сутності, атрибути, зв'язки та агрегацію сутностей. Концептуальна ER-модель та предикатне подання моделі. Логічне проектування бази даних, процес створення схеми бази даних та визначення необхідних обмежень цілісності. Проблеми неякісного проектування баз даних, аномалії та їх різновиди. Нормалізація баз даних та розгляд нормальних форм. Поняття information\_schema та особливості фізичного проектування баз даних.

### Тема 3. Моделі даних

Які бувають моделі даних та чим вони відрізняються? Що включає в себе модель даних як інструмент? Недоліки реляційної моделі. Особливості моделі даних SQL: дублікати, невизначені значення, обмеження. Стандарти SQL. Постреляційна модель даних. Об'єктно-орієнтовані моделі даних та їх обмеження. Особливості структур даних в об'єктно-орієнтованих базах даних. Об'єктно-реляційні СУБД. Модель даних PostgreSQL: особливості та розширення. Слабко структуровані (semi-structured) дані. OLTP та OLAP системи та їх особливості. Багатовимірна модель даних.

### Тема 4. Огляд мови SQL

Історія розвитку мови SQL. Потужність мов СУБД. Поняття декларативності та імперативності. Декларативність мови SQL. Програмування запитів: Thinking Sets. Групи операторів мови SQL: DDL, DML, DCL, TCL. Огляд основних операторів та типів SQL. Порядок обробки запитів у СУБД. Вбудовані функції та оператори PostgreSQL. Робота з мовами та Collations. Використання Views. Корелюючі та некорелюючі підзапити. З'єднання та їх типи. Групування та віконні функції. Рекурсивні запити. Методи виконання перехресних запитів. Методи зберігання деревоподібних структур.

### Тема 5. Розробка додатків, взаємодія з СУБД у мовах програмування

Об'єктно-реляційна втрата відповідності. Міфи програмної інженерії. Інтерфейси для спілкування з базами даних у різних мовах програмування (DB API, JDBC, ODBC, ADO.NET тощо). Огляд популярних архітектурних патернів для взаємодії з СУБД (Data Mapper, Active Record). Виконання SQL-команд. Асинхронний доступ до СУБД. Об'єктно-реляційні відображення. Знайомство з ORM на прикладі використання SQLAlchemy. Огляд архітектурних патернів, які використовуються в ORM-рішеннях (IdentityMap, UnitOfWork). Стратегії впровадження ORM у проект. Використання розширень PostgreSQL.

### Тема 6. Транзакції

Поняття транзакції. Вимоги до транзакційних систем, принципи ACID (Atomicity, Consistency, Isolation, Durability). Фіксація та обрив транзакцій. Журнал транзакцій. Контрольні точки (checkpoints). SAVEPOINTS. Стратегії ізоляції у транзакційних системах (песимістична та оптимістична). Реалізація принципу ізолюваності методом MVCC (Multi-Version Concurrency Control). Блокування (locks). Проблеми паралелізму, виникнення взаємоблокувань (deadlocks). Особливості реалізації ACID у рішеннях NoSQL. Рівні ізолюваності транзакцій (read uncommitted, read committed, repeatable read, serializable). Огляд аномалій, які можуть виникати при використанні різних рівнів ізолюваності транзакцій (втрачене оновлення, "брудне" читання, неповторюване читання, "фантомне" читання, неузгоджений запис, аномалія тільки читаючої

транзакції, аномалії серіалізації). Методи впровадження розподілених транзакцій. ХА протоколи та персистентні черги.

### Тема 7. Тригери та процедури

Мови програмування, які підтримуються найпопулярнішими СУБД. Особливості функцій та процедур у різних СУБД. Тригери: що це таке, для чого використовуються, їх переваги та недоліки. Створення та налаштування тригерів у PostgreSQL та інших популярних СУБД. Використання тригерів для забезпечення цілісності даних та автоматизації завдань. Повідомлення (notifications) у СУБД: механізми взаємодії та їх використання. Приклади застосування тригерів та процедур у реальних проектах. Рекомендації щодо використання тригерів для забезпечення продуктивності та цілісності даних.

### Тема 8. Індекси. Профілювання та оптимізація

Поняття індексу та його призначення у базах даних. Типи індексів у PostgreSQL та їх характеристики. Огляд найпопулярніших типів індексів (B-Tree, Hash, GIN, GiST, бітова карта) та їх особливості: коли та як їх варто використовувати. Поняття часткового індексу, функціонального індексу, кластерного індексу та покриваючого індексу. Протоколювання запитів: навіщо це потрібно та як це робити. Методи профілювання та оптимізації баз даних. Використання індексів для оптимізації продуктивності: приклади, поради та найкращі практики. Аналіз планів виконання запитів та оптимізація запитів за допомогою індексів. Типові помилки при індексуванні та способи їх уникнення.

### Тема 9. Оптимізація запитів та структури даних

Що таке план виконання запиту і як його читати. Розгляд основних етапів виконання запиту: розбір, оптимізація, виконання. Аналіз та інтерпретація планів виконання запитів для виявлення вузьких місць та потенційних проблем у продуктивності. Поняття нормалізації та денормалізації даних: переваги та недоліки кожного підходу. Методи оптимізації структури даних для підвищення продуктивності запитів. Оптимізація доступу до даних, зокрема зменшення часу блокувань та зменшення конфліктів між транзакціями. Методи оптимізації для різних типів запитів: селективні запити, агрегації, об'єднання та підзапити. Використання індексів для прискорення виконання запитів та зменшення витрат на введення/виведення. Профілювання запитів та виявлення "важких" запитів. Оптимізація на рівні додатку (клієнта СУБД): використання кешування, попереднє підготовлення запитів, мінімізація кількості запитів до бази даних. Використання з'єднань та підзапитів для зменшення кількості звернень до бази даних. Балансування навантаження та використання розподілених баз даних для підвищення масштабованості. Стратегії зменшення часу очікування та покращення взаємодії між додатком та базою даних. Рекомендації щодо моніторингу та постійного покращення продуктивності баз даних та запитів.

### Тема 11. Масштабування СУБД.

Що таке реплікація? Принципи поділу підходів до реплікації: відмінності між фізичною та логічною реплікацією. Поняття узгодженості реплік: які є підходи до забезпечення узгодженості даних у реплікованих системах. Варіанти реалізації реплікації, їх переваги та недоліки. Синхронна та асинхронна реплікація: порівняння, випадки використання, вплив на продуктивність та узгодженість даних. Як особливості архітектури СУБД впливають на реалізацію реплікації. Види масштабування систем: вертикальне та горизонтальне масштабування, їх переваги та обмеження. Масштабування СУБД на кластер, балансування навантаження між вузлами. Секціонування (partitioning): основні принципи, види, та застосування для підвищення продуктивності та керованості даних. Шардінг: що це таке, види шардінгу, переваги та проблематика, приклади реалізації. Розгляд історичних та актуальних даних, передумови появи OLAP систем: основні концепції, переваги OLAP для аналітики та бізнесу.

### Тема 12. Повнотекстовий пошук.

Повнотекстовий пошук: що це таке і де застосовується. Основні проблеми повнотекстового пошуку: оцінка релевантності результатів, врахування словоформ, багатомовність, швидкість обробки запитів. Інструменти для побудови повнотекстового пошуку, зокрема засоби

повнотекстового пошуку в PostgreSQL, використання `ts_vector` та `ts_query` для побудови внутрішнього подання документів та пошукових запитів. Побудова повнотекстового індексу, обробка документів, зберігання та пошук. Застосування конфігурацій для різних мов, налаштування та використання словників у PostgreSQL.

### Тема 13. Підтримка слабо-структурованих (semi-structured) даних у базах даних.

Зберігання слабо-структурованих даних у форматах JSON та JSONB. Використання JSON у PostgreSQL: функції для перетворення JSON у таблиці та навпаки, підтримка XML. Створення індексів для полів JSON та XML, використання індексів для прискорення запитів. Операції з JSON: перевірка наявності підмножин, отримання елементів за ключами, використання GIN індексів для JSONB. Приклади запитів з використанням JSON: пошук за ключами, фільтрація за значеннями, оптимізація запитів за допомогою індексів.

### Тема 14. Адміністрування СУБД. Налаштування та моніторинг.

Огляд важливих параметрів роботи СУБД: налаштування конфігураційних файлів, оптимізація параметрів для підвищення продуктивності та стабільності роботи. Резервні копії: методи створення, інструменти для резервування, стратегії резервування даних. Відновлення після відмов: плани відновлення, тестування відновлювальних процесів. Табличні простори: що це таке, як їх використовувати для організації та оптимізації зберігання даних. Огляд та використання LVM (Logical Volume Manager): принципи роботи, переваги використання для адміністрування великих об'ємів даних. Моніторинг СУБД: важливість моніторингу, інструменти для моніторингу роботи СУБД, основні показники для відстеження. Системний каталог PostgreSQL: що це таке, як використовувати для адміністрування та моніторингу. Очищення застарілих даних у базі даних: методи, інструменти, рекомендації щодо підтримки чистоти даних.

### Тема 15. Безпека даних

Принципи найменших привілеїв: основи забезпечення мінімально необхідного доступу для користувачів та процесів. Безпека на рівні бази даних: підходи та методи забезпечення безпеки даних в СУБД. Авторизація у СУБД, користувачі та ролі: створення, управління та адміністрування користувачів і ролей, призначення привілеїв. Привілеї та робота з ними: різновиди привілеїв, їх призначення та обмеження доступу. Проксі-авторизація: механізми та приклади використання для безпечного доступу до даних. Row Level Security (RLS): політики безпеки на рівні рядків, налаштування політик для контролю доступу до даних на рівні рядків у таблицях. Реєстрація подій та змін: методи і інструменти для аудиту дій користувачів та змін у базі даних. Шифрування паролів та інших даних: принципи, методи шифрування, забезпечення безпечного зберігання та передачі конфіденційної інформації. Поняття SQL-Injection: що це таке, як виявляти вразливості, методи захисту від SQL-ін'єкцій. Інструменти автоматизованого пошуку вразливостей: огляд інструментів для аналізу безпеки баз даних та виявлення потенційних загроз.

### Тема 16. Введення у NoSQL СУБД

Що таке NoSQL СУБД? Передумови розвитку NoSQL систем: історія та причини появи NoSQL баз даних, особливості та основні відмінності від традиційних реляційних СУБД. Моделі цілісності (consistency models): огляд різних моделей забезпечення цілісності даних у NoSQL системах. Принципи NoSQL СУБД - BASE (Basic Availability, Soft State, Eventual Consistency): основні принципи, які відрізняють NoSQL від ACID реляційних баз даних. Види NoSQL СУБД: огляд основних категорій NoSQL баз даних та їх застосування. Сховища типу ключ-значення: принципи роботи, переваги, приклади використання. Документо-орієнтовані БД: особливості, приклади, знайомство з MongoDB. BigTable бази даних: структура, застосування, переваги та недоліки. Графові БД: що це таке, приклади використання, переваги для роботи з графовими даними. Підхід Стоунбрейкера та NewSQL бази даних: огляд концепції та сучасних рішень, що поєднують переваги реляційних і NoSQL систем.

## Теми практичних занять

Не передбачено програмою навчальної дисципліни

## Теми лабораторних робіт

### Тема 1. Інфологічне проектування бази даних

**Мета роботи:** Вивчення основ інфологічного проектування БД шляхом створення ER-моделі.

**Зміст:** Виділення сутностей, атрибутів та зв'язків, класифікація атрибутів, визначення підтипів та супертипів, формалізація зв'язків, класифікація сутностей та побудова ER-діаграми для заданої предметної області.

### Тема 2. Логічне проектування бази даних

**Мета роботи:** Виконати етап логічного проектування реляційної бази даних та створити спроектовану базу даних засобами СУБД.

**Зміст:** Організація даних, виділених на етапі інфологічного проектування, у форму, прийняту в обраній СУБД. Відображення об'єктів предметної області у реляційні таблиці, визначення атрибутів, первинних та зовнішніх ключів, нормалізація відношень до третьої нормальної форми, завдання декларативних обмежень цілісності. Створення та наповнення таблиць, написання SQL-запитів для вибірки даних, перевірка працездатності запитів.

### Тема 3. Реалізація взаємодії з базою даних із використанням ORM

**Мета роботи:** Реалізувати взаємодію з базою даних із використанням ORM SQLAlchemy на мові програмування Python.

**Зміст:** Створення docker-контейнера з налаштуванням SQLAlchemy. Опис таблиць у вигляді моделей на Python відповідно до першого завдання. Написання скриптів для заповнення тестовими даними, забезпечення зв'язків між даними. Написання запитів до бази даних з використанням ORM, виведення результатів у консоль. Створення zip-архіву з вихідним кодом проекту. Оформлення звіту з описом мети, тексту завдання, обраної ORM, інструкцією із запуску у docker-контейнері та скріншотами результатів виконання запитів.

### Тема 4. Кешування даних

**Мета роботи:** Навчитися використовувати кешування даних, які часто використовуються в оперативній пам'яті, і підтримувати актуальний стан даних в кеші за допомогою тригерів та збережених процедур..

**Зміст:** Реалізація кешування даних з використанням Memcached. Налаштування PostgreSQL для взаємодії з Memcached і передача даних до кешу. Створення REST-сумісного API для запиту даних через Nginx. Написання тригерів для автоматичної синхронізації даних між базою даних та кешем. Розробка та тестування збережених процедур для підтримки цілісності та узгодженості даних.

### Тема 5. Розробка програмного додатку над базою даних

**Мета роботи:** Реалізувати програмний додаток, що працює зі створеною базою даних, використовуючи обрані засоби розробки додатків.

**Зміст:** Розробка програмного додатку для додавання, модифікації та видалення інформації з бази даних, забезпечуючи при цьому відповідність даних специфікаціям предметної області. Підтримка цілісності бази даних за допомогою валідації даних та виконання операцій у межах транзакцій. Заповнення бази даних достатньою кількістю інформації для демонстрації запитів, включаючи специфічні запити для обраної предметної галузі. Розробка додатку із використанням технологічного стеку: Python, SQLAlchemy, Bottle, nginx, Docker. Опис процесу налаштування системи для розробки додатку. Створення звіту про виконання лабораторних робіт 1-5.

## Самостійна робота

Опрацювання лекційного матеріалу, що включає перегляд та конспектування лекційних записів, вивчення додаткової літератури та ресурсів, рекомендованих викладачем. Підготовка до модульного тесту, що передбачає структурований повтор основних тем, що будуть охоплені тестом, виконання практичних завдань для закріплення матеріалу. Підготовка до іспиту, що включає глибоке вивчення всіх тем курсу, систематизацію знань, проходження додаткових тестів та роботу над помилками, виявленими під час модульного тестування.

## Література та навчальні матеріали

1. Методичні вказівки до виконання лабораторних робіт з дисципліни "Програмування, бази даних і знань" [електронний ресурс] : для студентів 122 спец. "Комп'ютерні науки" / уклад.: Лисенко А.О.; нац. техн. ун-т "ХПІ". – Харків, 2024. – 132 с. – URI: <https://repository.kpi.kharkov.ua/handle/KhPI-Press/72798>.
2. C.J. Date. An Introduction to Database Systems, 8th edition – Pearson Education, Inc., 2004. - 1024 p. - ISBN 0-321-19784-4.
3. Salahaldin Juba, Andrey Volkov. Learning PostgreSQL 10, Second Edition. A beginner's guide to building high-performance PostgreSQL database solutions – Packt Publishing, 2017. - 552 p. - ISBN 978-1-78839-201-3.
4. Anthony Molinaro. SQL Cookbook – O'Reilly Media, 2009. - 636 p. - ISBN 0-596-00976-3.
5. Jason Myers, Rick Copeland. Essential SQLAlchemy: Mapping Python to Databases, Second Edition – O'Reilly Media, 2015. - 258 p.
6. Henrietta Dombrovskaya, Boris Novikov, Anna Baillieikova. PostgreSQL Query Optimization – Apress, 2021. - 295 p. - ISBN 978-1-4842-6884-1.
7. Adrian Mouat. Using Docker: Developing and Deploying Software with Containers – O'Reilly Media, 2016. - 348 p. - ISBN 978-1-491-91576-9.

## Система оцінювання

### Критерії оцінювання успішності студента та розподіл балів

100% підсумкове оцінювання у вигляді іспиту та поточного оцінювання.  
40% іспит, 60% поточне оцінювання:  
Модульний контроль (35%)  
Лабораторні роботи (25%)  
Лабораторна робота №1 (5%)  
Лабораторна робота №2 (5%)  
Лабораторна робота №3 (5%)  
Лабораторна робота №4 (5%)  
Лабораторна робота №5 (5%)

### Шкала оцінювання

Сума балів	Національна оцінка	ECTS
90–100	Відмінно	A
82–89	Добре	B
75–81	Добре	C
64–74	Задовільно	D
60–63	Задовільно	E
35–59	Незадовільно (потрібне додаткове вивчення)	FX
1–34	Незадовільно (потрібне повторне вивчення)	F

## Норми академічної етики і політика курсу

Студент повинен дотримуватися «Кодексу етики академічних взаємовідносин та доброчесності НТУ «ХПІ»: виявляти дисциплінованість, вихованість, доброзичливість, чесність, відповідальність. Конфліктні ситуації повинні відкрито обговорюватися в навчальних групах з викладачем, а при неможливості вирішення конфлікту – доводитися до відома співробітників дирекції інституту. Нормативно-правове забезпечення впровадження принципів академічної доброчесності НТУ «ХПІ» розміщено на сайті: <http://blogs.kpi.kharkov.ua/v2/nv/akademichna-dobrochesnist/>

## Погодження

Силабус погоджено

Дата погодження, підпис  
27.08.2024р.

Завідувач кафедри  
Марина ГРИНЧЕНКО

27.08.2024р.

Гарант ОП  
Ігор КОНОНЕНКО

