



Гетерогенні розподілені бази даних

Силабус освітнього компонента

Реквізити навчальної дисципліни

Рівень вищої освіти	Третій (PhD)
Галузь знань	12 Інформаційні технології
Спеціальність	126 Інформаційні системи та технології
Освітня програма	Інформаційні системи та технології
Статус дисципліни	Вибіркова
Форма навчання	Очна (денна)
Рік підготовки, семестр	2 курс, весняний семестр
Обсяг дисципліни	5 кредита ECTS/150 годин (18 годин – Лекції, 18 годин – Лабораторні, самостійна робота: 114 год.)
Семестровий контроль/ контрольні заходи	Залік, бліц-опитування, календарний контроль.
Розклад занять	Згідно розкладу на весняний семестр поточного навчального року (rozklad.kpi.ua)
Мова викладання	Українська
Інформація про керівника курсу / викладачів	Лектор: к.т.н., Крилов Євген Володимирович, ekrylov1964@gmail.com , моб. +38(050)387-16-45 Лабораторні роботи: к.т.н., доцент, Крилов Євген Володимирович, ekrylov1964@gmail.com , моб. +38(050)387-16-45
Розміщення курсу	https://campus.kpi.ua

Силабус освітнього компонента «Гетерогенні розподілені бази даних» відповідає спеціальності 126 – Інформаційні системи та технології. При проходженні даної дисципліни, студенти познайомляться з поняттям «Гетерогенні розподілені бази даних», та детальніше вивчатимуть методи та засоби створення різноманітних структур гетерогенних баз даних. На лабораторних заняттях опанують етапи проектування гетерогенних розподілених баз даних. В курсі передбачений контроль якості отриманих знань у вигляді експрес-контрольних. На лекціях проводитимуться практичні групові заняття.

Метою навчальної дисципліни є формування та закріплення у студентів наступних компетентностей: ЗК 03: Здатність розробляти проекти та управляти ними; ФК 08: Здатність проводити інтелектуальний аналіз даних.

Предмет навчальної дисципліни: Для задоволення потреб обробки інформації, що зберігається в різноманітних базах даних, які об'єднуються в одну гетерогенну розподілену систему в науковій, банківській, фізико-технічній сферах і набуття досвіду та умінь використання та створення систем розподіленої обробки даних; для використання та розроблення програмного забезпечення з використанням сучасних засобів зберігання та обробки даних.

Програмні результати навчання, на формування та покращення яких спрямована дисципліна: *ПРН 07:* Проектувати та досліджувати цілісні системи Інтернету речей (в тому числі кінцеві пристрої, мережеві технології, хмарні платформи, реалізацію обміну та аналізу даних), проводити інтелектуальний аналіз цифрових масивів даних для вирішення конкретних практичних науково-прикладних задач; *ПРН 08:* Розробляти програмне забезпечення інформаційних систем у відповідності з принципами сервіс-орієнтованої архітектури розподілених програмних систем, проводити реінжиніринг прикладного інформаційного забезпечення; *ПРН 11:* Здатність використовувати сучасні методи та інструменти для проведення інтелектуального аналізу даних, включаючи здатність визначати й інтерпретувати ключові закономірності, виявляти складні зв'язки та здійснювати компетентну оцінку результатів аналізу для вирішення проблемних ситуацій в інформаційних системах.

Пререквізити та постреквізити дисципліни (місце в структурно-логічній схемі навчання за відповідною освітньою програмою)

Пререквізити: Володіння знаннями із дисциплін «Розподілені системи обробки даних», «Паралельні та розподілені обчислення», «Бази даних» та досвід роботи.

Постреквізити: в результаті навчання аспіранти зможуть: використовувати та організовувати розподілену обробку даних, аналізувати вимоги до сучасних гетерогенних баз даних, розв'язувати задачі проектування та експлуатації розподілених баз даних, розробляти нові архітектури систем з високою швидкістю обробки даних, адмініструвати розподілені бази даних, застосовувати теоретичні знання для розв'язання практичних задач обробки даних.

1. Зміст навчальної дисципліни

Лекційні заняття

Лекція 1. Концепції та основи побудови розподілених баз даних.

Лекція 2. Розробка архітектури розподілених баз даних.

Лекція 3. Побудова гетерогенних розподілених систем.

Лекція 4. Технології управління гетерогенними системами.

Лекція 5. Технології обробки великих масивів даних в розподілених базах даних.

Лекція 6. Управління гетерогенними розподіленими базами даних.

Лекція 7. Моделі даних NoSQL. Пост реляційна та графова модель.

Лекція 8. Багатомірна та KEY Value моделі.

Лекція 9. Колонко-орієнтована та документо - орієнтована моделі.

Лабораторні заняття

1. Створення технічного завдання на розробку гетерогенної бази даних (вибір моделі даних, проектування структури БД, формування запитів).
2. Дослідження реляційної та нереляційної структури гетерогенних розподілених БД
3. Робота з графовими розподіленими базами даних.
4. Робота з документо-орієнтованими розподіленими базами даних
5. Робота з колонковими розподіленими базами даних.
6. Робота з ключ/значення розподіленими базами даних
7. Реалізація проекту з обраною гетерогенною базою даних

2. Навчальні матеріали та ресурси

Базова література:

1. Технології проектування No Sql баз даних. Конспект лекцій./ Крилов Євген Володимирович Анікін Володимир Костянтинівич/ Навчальний посібник Електронне мережеве видання Гриф надано Методичною радою КПП ім. Ігоря Сікорського (протокол №3 від 07.12.2023 р.) за поданням Вченої Ради ФІОТ (протокол № 2 від 19.09.2023 р.), як навчальний посібник для здобувачів ступеня магістра першого року навчання.
<https://ela.kpi.ua/handle/123456789/63578>

Допоміжна література:

1. V. Nikitin, E. Krylov, Y. Kornaga, V. Anikin COMBINED INDEXING METHOD IN NOSQL DATABASES Міжвідомчий науково-технічний збірник «Адаптивні системи автоматичного управління».— Київ: Національний технічний університет України “Київський політехнічний інститут”. 2021 Том 1(38) стор. 3-9. DOI: <https://doi.org/10.20535/1560-8956.38.2021.232948> (Фахове видання)
2. Використання транзакційного годинника для пришвидшення процесу узгодження даних в розподілених системах / E. Krylov . V. Nikitin/ Вісник Хмельницького національного університету.
3. MongoDB от теории к практике. Руководство по установке кластера mongoDB/ <https://habr.com/ru/hub/nosql/>
4. Варіанти забезпечення суворої узгодженості в NoSql / Крилов Є.В., Белоус Р.В., Нікітін В.А. / V International Scientific and Practical Conference «SCIENTIFIC RESEARCHES AND METHODS OF THEIR CARRYING OUT: WORLD EXPERIENCE AND DOMESTIC REALITIES» 17.02.2023 | Vinnytsia, UKR - Vienna, AUT (Online) Google Scholar

Навчальний контент

3. Методика опанування навчальної дисципліни (освітнього компонента)

Лекційні заняття

№ з/п	Назва теми лекції та перелік основних питань (перелік дидактичних засобів, посилання на літературу та завдання на СРС)
1	Лекція 1. Концепції та основи побудови розподілених баз даних <i>Визначення розподіленої бази даних , історія назви , розвиток ідеї , Коротка історія баз даних NoSQL .</i> Дидактичні матеріали: Презентація PowerPoint, комп'ютер, проектор.
2	Лекція 2. Розробка архітектури розподілених баз даних <i>Визначення базової архітектури розподіленої бази даних.</i> Дидактичні матеріали: Презентація PowerPoint, комп'ютер, проектор.
3	Лекція 3. Побудова гетерогенних розподілених систем <i>Основні особливості гетерогенних систем.</i> Дидактичні матеріали: Презентація PowerPoint, комп'ютер, проектор.
4	Лекція 4. Технології управління гетерогенними системами <i>Базові Технології управління гетерогенними системами.</i> Дидактичні матеріали: Презентація PowerPoint, комп'ютер, проектор.
5	Лекція 5. Технології обробки великих масивів даних в розподілених базах даних <i>Поняття великого об'єму даних, використання в розподілених системах .</i> Дидактичні матеріали: Презентація PowerPoint, комп'ютер, проектор.
6	Лекція 6. Управління гетерогенними розподіленими базами даних <i>Основні поняття управління гетерогенними розподіленими базами даних .</i> Дидактичні матеріали: Презентація PowerPoint, комп'ютер, проектор.
7	Лекція 7. Моделі даних No SQL. Пост реляційна та графова моделі. <i>Пост реляційна модель. Таблиця накладних. Таблиця товарів.</i> <i>Об'єктноорієнтована модель. Приклад об'єктноорієнтованої бази даних. Клас об'єкта. Інкапсуляція. Наслідування. Поліморфізм. Недоліки об'єктно-орієнтованих моделей.</i> Дидактичні матеріали: Презентація PowerPoint, комп'ютер, проектор.
8	Лекція 8. Багатомірні та KEY Value моделі. <i>Приклад тримірної моделі. Зріз. Обертання. KEY Value. Redis, Dynamo, Riak - деякі приклади баз даних сховищ ключів і значень.</i> Дидактичні матеріали: Презентація PowerPoint, комп'ютер, проектор.

9	Лекція 9. Колонко - орієнтована та документно - орієнтована моделі. Column-based. Document_based. Amazon SimpleDB, CouchDB, MongoDB, Riak, Lotus Notes, MongoDB - популярні системи СУБД, створені на основі документів. Дидактичні матеріали: Презентація PowerPoint, комп'ютер, проектор.
---	---

Лабораторні заняття

№	Назва лабораторної роботи	Кількість ауд. годин
1	Лабораторна робота 1. Створення технічного завдання на розробку програмного забезпечення з використанням гетерогенних розподілених не реляційних баз даних (вибір моделі даних, проектування структури БД, формування запитів).	2
2	Лабораторна робота 2. Дослідження реляційної та не реляційної структур гетерогенних розподілених БД (аспірант повинен спочатку розробити реляційну структуру БД, побачити недоліки, вибрати модель нереляційної БД, спроектувати структуру нереляційної БД)	2
3	Лабораторна робота 3. Робота з графовими розподіленими базами даних (Amazon Neptune, Neo4J, Infinite Graph, OrientDB, FlockDB)	4
4	Лабораторна робота 4. Робота з документо-орієнтованими базами даних (MongoDB)	2
5	Лабораторна робота 5 Робота з ключ/значення розподіленими базами даних (Redis, Dynamo, Riak)	4
6	Лабораторна робота 6. Робота з колонковими розподіленими базами даних. (HBase, Cassandra, Hypertable.)	2
7	Лабораторна робота 7. Розробка та презентація проєкту	2

4. Самостійна робота студента/аспіранта

№ з/п	Назва теми, що виноситься на самостійне опрацювання	Кількість годин СРС
1	Адміністрування розподілених баз даних	20
2	Розв'язання задач проектування та експлуатації розподілених баз даних	20
3	Методи оцінки продуктивності розподілених систем	20
4	Аналіз вимог до сучасних гетерогенних баз даних	20
5	Використання та організація розподіленої обробки даних	34
	Загалом	114

Політика та контроль

5. Політика навчальної дисципліни (освітнього компонента)

Система вимог, які ставляться перед аспірантом:

- відвідування лекційних та лабораторних занять є обов'язковою складовою вивчення матеріалу;

- на лекції викладач користується власним презентаційним матеріалом; відпрацьовує практичну частину на віртуальній машині з встановленою відповідною РСУБД; використовує гугл-диск для викладання матеріалу поточної лекції, додаткових ресурсів, лабораторних робіт та інше; викладач відкриває доступ до певної директорії гугл-диска для скидання електронних лабораторних звітів та відповідей на МКР;
- на лекції заборонено відволікати викладача від викладання матеріалу, усі питання, уточнення та ін. студенти задають в кінці лекції у відведений для цього час;
- лабораторні роботи захищаються у два етапи – перший етап: студенти виконують завдання на допуск до захисту лабораторної роботи; другий етап – захист лабораторної роботи. Бали за лабораторну роботу враховуються лише за наявності електронного звіту;
- заохочувальні бали виставляються за: активну участь на лекціях; участь у факультетських та інститутських олімпіадах з навчальних дисциплін, участь у конкурсах робіт, підготовка оглядів наукових праць; презентацій по одній із тем СРС дисципліни тощо. Кількість заохочуваних балів не більше 10;
- штрафні бали виставляються за: невчасну здачу лабораторної роботи. Кількість штрафних балів не більше 10.

6. Види контролю та рейтингова система оцінювання результатів навчання (РСО) (очна форма)

Протягом семестру аспіранти виконують 7 лабораторних робіт. Максимальна кількість балів за кожну лабораторну роботу: 10 балів.

Бали нараховуються за: - якість виконання лабораторної роботи: 0-5 балів;

- відповідь під час захисту лабораторної роботи: 0-3 бали;

- своєчасне представлення роботи до захисту: 0-2 бали.

Критерії оцінювання якості виконання:

4-5 балів – робота виконана якісно, в повному обсязі;

1-3 бали – робота виконана в повному обсязі, але містить незначні помилки;

0 балів – робота виконана не в повному обсязі, або містить суттєві помилки.

Критерії оцінювання відповіді:

3 бали – відповідь повна, добре аргументована;

1-2 бали – в цілому відповідь правильна, але має недоліки або незначні помилки;

0 балів – немає відповіді або відповідь неправильна.

Критерії оцінювання своєчасності представлення роботи до захисту:

2 бали – робота представлена до захисту не пізніше вказаного терміну;

0 балів – робота представлена до захисту пізніше вказаного терміну.

Максимальна кількість балів за виконання та захист лабораторних робіт: $10 \text{ балів} \times 7 \text{ лабораторних робіт} = 70 \text{ балів}$.

Протягом семестру на лекціях відбувається опитування за темою поточного заняття.

Максимальна кількість балів за опитування, яку можна отримати протягом семестру: 30 балів. Рейтингова шкала з дисципліни дорівнює: $R = R_c = 70 \text{ балів} + 30 \text{ балів} = 100 \text{ балів}$.

Календарний контроль:

проводиться двічі на семестр як моніторинг поточного стану виконання вимог силабусу.

На першій атестації (8-й тиждень) студент отримує «зараховано», якщо його поточний рейтинг не менше 12 балів (50 % від максимальної кількості балів, яку може отримати студент до першої атестації).

На другій атестації (14-й тиждень) студент отримує «зараховано», якщо його поточний рейтинг не менше 20 балів (50 % від максимальної кількості балів, яку може отримати студент до другої атестації).

Семестровий контроль: залік:

Умови допуску до семестрового контролю:

За семестрового рейтингу (RC) не менше 60 балів та зарахуванні усіх лабораторних робіт аспірант отримує залік «автоматом» відповідно до таблиці (Таблиця відповідності рейтингових балів оцінкам за університетською шкалою).

В іншому разі він має виконати залікову контрольну роботу. Необхідною умовою допуску до виконання залікової контрольної роботи є виконання і захист лабораторних робіт. Аспірант може спробувати підвищити свою оцінку шляхом написання залікової контрольної роботи, при цьому його бали, отримані за семестр, анулюються («жорстка» система оцінювання).

Склад та критерії оцінювання залікової контрольної роботи:

Завдання на залікову контрольну роботу складається з 7 запитань – 5 теоретичних та 2 практичних. Відповідь на кожне теоретичне запитання оцінюється 10 балами, а відповідь на практичне запитання оцінюється 25 балами.

Критерії оцінювання кожного теоретичного запитання контрольної роботи:

9-10 балів – відповідь правильна, повна, добре аргументована;

7-8 балів – відповідь правильна, розгорнута, але не дуже добре аргументована;

5-6 балів – в цілому відповідь правильна, але має недоліки;

3-4 балів – у відповіді є незначні помилки;

1-2 бали – у відповіді є суттєві помилки;

0 балів – немає відповіді або відповідь неправильна.

Критерії оцінювання практичного запитання контрольної роботи:

24-25 балів – відповідь правильна, розрахунки виконані у повному обсязі;

21-23 бали – відповідь правильна, але не дуже добре підкріплена розрахунками;

17-20 балів – в цілому відповідь правильна, але має недоліки;

12-16 балів – у відповіді є незначні помилки;

1-11 бали – у відповіді є суттєві помилки;

0 балів – немає відповіді або відповідь неправильна.

Максимальна кількість балів за залікову контрольну роботу: $10 \text{ балів} \times 5 \text{ теоретичних запитань} + 25 \text{ балів} \times 2 \text{ практичних запитань} = 100 \text{ балів}$.

Таблиця 1. Переведення рейтингових балів до оцінок за університетською шкалою

Кількість балів	Оцінка
100-95	Відмінно
94-85	Дуже добре
84-75	Добре
74-65	Задовільно
64-60	Достатньо
Менше 60	Незадовільно
Не виконані умови допуску	Не допущено

7. Додаткова інформація з дисципліни (освітнього компонента)

Передбачена можливість закривати частину лабораторного та лекційного матеріалу шляхом здобування сертифікатом по online курсам відповідних розділів та тем дисципліни;

Робочу програму навчальної дисципліни (силабус):

Складено доцент кафедри ICT, к. т. н., Крилов Євген Володимирович.

Ухвалено кафедрою ICT (протокол № 16 від 12.06.2024 р.).

Погоджено Методичною комісією факультету¹ (протокол № 10 від 21.06.2024 р.).