



Інтелектуальні бази та сховища даних

Робоча програма навчальної дисципліни (Силабус)

1. Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Першій (бакалаврській)</i>
Галузь знань	<i>12 Інформаційні технології</i>
Спеціальність	<i>121 Інженерія програмного забезпечення</i>
Освітня програма	<i>Інженерія програмного забезпечення інтелектуальних кібер-фізичних систем в енергетиці</i>
Статус дисципліни	<i>Вибіркові освітні компоненти</i>
Форма навчання	<i>очна(денна) /змішана</i>
Рік підготовки, семестр	<i>3 курс, осінній семестр</i>
Обсяг дисципліни	<i>4 кредитів, 120 годин, з яких 54 години аудиторних (36 год лекції, 18 год практичні), 66 годин становить самостійна робота</i>
Семестровий контроль/ контрольні заходи	<i>залік / МКР</i>
Розклад занять	<i>1 лекція (2 години) 1 раз на тиждень; 1 практичне заняття (2 години) 1 раз на 2 тижні</i>
Мова викладання	<i>Українська</i>
Інформація про керівника курсу / викладачів	<i>Лектор: старший викладач, Дацюк Оксана Антонівна, doka70@ukr.net Практика: : старший викладач, Дацюк Оксана Антонівна, doka70@ukr.net, тел. 097-524-90-54</i>
Розміщення курсу	<i>Електронний кампус</i>

2. Програма навчальної дисципліни

1. Опис навчальної дисципліни, її мета, предмет вивчення та результати навчання

У курсі розглядаються принципи побудови систем, орієнтованих на аналіз даних, різні моделі даних, які використовуються для побудови сховищ даних. Також розглянуті питання побудови систем на основі сховищ даних, організація збору та зберігання даних у сховищі, технологія інтелектуального аналізу даних та інші питання.

Метою навчальної дисципліни є формування та закріплення у студентів наступних компетентностей:

(ЗК 1) Здатність до абстрактного мислення, аналізу та синтезу. (ЗК 2) Здатність застосовувати знання у практичних ситуаціях (ЗК 3) Здатність спілкуватися державною мовою як усно, так і письмово (ЗК 4) Здатність вчитися і оволодівати сучасними знаннями (ЗК 6) Здатність до пошуку, оброблення та аналізу інформації з різних джерел (ЗК 7) Здатність працювати в команді

(ФК 1) Здатність ідентифікувати, класифікувати та формулювати вимоги до програмного забезпечення (ФК 2) Здатність брати участь у проектуванні програмного забезпечення, включаючи проведення моделювання (формальний опис) його структури, поведінки та процесів функціонування (ФК 3) Здатність розробляти архітектури, модулі та компоненти програмних систем (ФК 7) Володіння знаннями про інформаційні моделі даних, здатність створювати програмне забезпечення для зберігання, видобування та опрацювання даних. (ФК 11) Здатність реалізувати фази та ітерації життєвого циклу програмних систем інформаційних технологій на основі відповідних моделей і підходів розробки програмного забезпечення (ФК 14) Здатність до алгоритмічного та логічного мислення (ФК 16) Володіти скриптовими та декларативними мовами програмування.

Предмет навчальної дисципліни – теоретичні підходи та практичні навички з організації сховищ даних, пов'язаних із аналітичною обробкою даних в інформаційних системах

Програмні результати навчання, на формування та покращення яких спрямована дисципліна.

Згідно з вимогами програми навчальної дисципліни студенти після засвоєння кредитного модуля мають продемонструвати такі результати навчання:

(РН 1) Аналізувати, цілеспрямовано шукати і вибирати необхідні для вирішення професійних завдань інформаційно-довідникові ресурси і знання з урахуванням сучасних досягнень науки і техніки. (ПРН 3) Знати основні процеси, фази та ітерації життєвого циклу програмного забезпечення. (ПРН 4) Знати і застосовувати професійні стандарти і інші нормативно-правові документи в галузі інженерії програмного забезпечення.

(ПРН 8) Вміти розробляти людино-машинний інтерфейс (ПРН 9) Знати та вміти використовувати методи та засоби збору, формулювання та аналізу вимог до програмного забезпечення. (ПРН 10) Проводити передпроектне обстеження предметної області, системний аналіз об'єкта проектування. (ПРН 11) Вибирати вихідні дані для проектування, керуючись формальними методами опису вимог та моделювання. (ПРН 12) Застосовувати на практиці ефективні підходи щодо проектування програмного забезпечення. (ПРН 13) Знати і застосовувати методи розробки алгоритмів, конструювання програмного забезпечення та структур даних і знань. (ПРН 15) Мотивовано обирати мови програмування та технології розробки для розв'язання завдань створення і супроводження програмного забезпечення (ПРН 18) Знати та вміти застосовувати інформаційні технології обробки, зберігання та передачі даних.

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

Отримані знання дозволять студентам використовувати методи інформаційного моделювання при вивченні інших інженерних дисциплін, виконанні курсових і дипломних робіт.

Пререквізити дисципліни.

Матеріал курсу базується на знаннях, які студенти отримали при вивченні основ традиційного процедурного програмування з курсів «Алгоритми та структури даних» «Об'єктно-орієнтоване програмування», «Бази даних», «Компоненти програмної інженерії». В курсі використовуються знання окремих розділів дискретної математики (теорія графів, теорія множин, та ін.), які даються в курсі «Комп'ютерна дискретна математика».

Постреквізити дисципліни.

У структурно-логічній схемі навчання кредитний модуль розміщений у 6 семестрі. Набуті студентами знання можуть бути використані для вивчення дисциплін, пов'язаних з моделюванням та розробкою програмного забезпечення автоматизованих систем, що використовують нереляційні бази даних для збереження інформації. Це такі дисципліни, як «Побудова

масштабованих систем обробки даних у реальному часі», «Проектування експертних систем», «Хмарні технології» «Інтелектуальний аналіз даних», «Дипломне проектування».

3. Зміст навчальної дисципліни

Розділ 1. Інтелектуальні інформаційні системи

Тема 1.1 Business Intelligence

Тема 1.2 Темпоральні бази даних

Тема 1.3 Системи класу OLTP

Розділ 2. Сховища даних

Тема 2.1 Проектування структури сховища даних

Тема 2.2 Інтеграція даних

Розділ 3. Аналіз даних

Тема 3.1 on-line analytical processing

Тема 3.2 Візуалізація даних

4. Навчальні матеріали та ресурси

Основна література

1. J. Kunigk, I. Buss, P. Wilkinson, L. George, *Architecting Modern Data Platforms*, O'Reilly Media, Inc., 2019, 633 p.
2. E. Sciore, *Database Design and Implementation: Second Edition*, Springer Nature, 2020, 468 p.
3. *Overview of Online Analytical Processing (OLAP)*. Microsoft [Електронний ресурс] – Режим доступу до ресурсу: <https://support.microsoft.com>
4. P. Baltzan, *Business Driven Information Systems*, McGraw-Hill Education, 2021, 809 p.
5. A. Simon, *Data Lakes For Dummies*, John Wiley & Sons, Inc., 2021, 387 p. 5. *Data Management Body of Knowledge. Second Edition*, Technics Publications, 2017, 778 p.
6. *Сховища даних. Підручник* · Автор: Пасічник В. В., Шаховська Н. Б. · Видавництво: Магнолія, · 2021, 496 с. ·

Додаткова література

1. D. Petkovic, *Microsoft SQL Server 2019: A Beginner's Guide. Seventh Edition*, McGraw Hill Professional, 2020, 896 p.
2. *Handbook of Big Data Analytics. Volume 1: Methodologies*, The Institution of Engineering and Technology, 2021, 390 p.
3. *Handbook of Big Data Analytics. Volume 2: Applications in ICT, security and business analytics*. – The Institution of Engineering and Technology, 2021. – 419 p.
4. A. Meier, M. Kaufmann, *SQL & NoSQL Databases: Models, Languages, Consistency Options and Architectures for Big Data Management*, Springer, 2019, 229 p.
5. J. O. Padallan, *Distributed Database Architecture*, Arcler Press, 2021, 266 p.
6. *PostgreSQL 16.1 Documentation*. The PostgreSQL Global Development Group [Електронний ресурс] – Режим доступу до ресурсу: <https://www.postgresql.org/docs/current/index.html>
7. *Застосування технології OLAP на прикладі даних підсистеми Відносини з клієнтами* [Електронний ресурс] – Режим доступу до ресурсу: <https://ispro.ua/uploads/pages/April2021/oKQwYuMchQxoflXoAP4.pdf>.
8. Черняк О. І., Захарченко П. В. *Інтелектуальний аналіз даних: Підручник*. Київ, 2014.

Навчальний контент

Лекційні заняття

№ з/п	Назва теми лекції та перелік основних питань (перелік дидактичних засобів, посилання на інформаційні джерела)
Розділ 1. Інтелектуальні інформаційні системи	
Тема 1.1 Business Intelligence	
Лек1	Поняття інтелектуальних систем і технологій. Мета і задачі інтелектуальних систем управління. Основні властивості. Архітектура інтелектуальної інформаційної системи. Відмінності інтелектуальних систем управління від традиційних систем. Технології обробки великих даних.
Лек2	BI системи та Інтелектуальний аналіз даних Business Intelligence – інтелектуальний аналіз даних та бізнес-аналітика Збереження даних. Інтеграція даних. Аналіз даних. Представлення даних в сховищі BI-системи (data warehouse, DW) Data Warehouse (DWH), Data Lake (DL) та Data Swamp (DS). Аналітичний шар DWH.
Тема 1.2 Темпоральні бази даних	
Лек3	Темпоральні бази даних Характерні особливості темпоральних баз даних. Основні принципи та особливості побудови темпоральних баз даних. Темпоральні розширення в стандарті SQL
Лек4	Методи представлення даних у темпоральних баз даних. Ненормалізовані бази даних Багатовимірні бази даних Розглянете підходи до проектування Кімбала та Інмана.
Тема 1.3 Системи класу OLTP	
Лек5	Оперативна і аналітична обробка даних Системи класу OLTP, їх призначення та особливості. Проблеми, що виникають при застосуванні систем класу OLTP та традиційних баз даних при аналітичній обробці даних та підтримці прийняття рішень.
Розділ 2. Сховища даних	
Тема 2.1 Проектування структури сховища даних	
Лек6	Основні поняття сховищ даних Визначення сховища даних. Властивості сховищ даних. Засоби роботи зі сховищами даних.
Лек7	Проектування структури сховища даних Категорії даних у сховищах даних, їх класифікація. Схеми зв'язування даних у сховищах даних, їх особливості.
Лек8	Озера даних. Вітрини даних. Кіоски даних. Призначення та основні особливості озер даних.

	<p>Відмінності між озерами даних та сховищами даних. Технології сховища даних та кіоски даних Вітрини даних. Представлення даних вітрин.</p>
Лек9	<p>Простори даних Ефективне зберігання і обробка великих обсягів даних. Організація багаторівневих довідників метаданих. Забезпечення інформаційної безпеки СД.</p>
<p>Тема 2.2 Інтеграція даних</p>	
Лек10	<p>Засоби інтеграції та інтероперабельності Сховища даних Методи інтеграції корпоративних даних. Завантаження даних до озер даних та сховищ даних. ETLпроцедури та ELT-процедури. Призначення та основні особливості. Відмінності між ETL та ELT.</p>
Лек11	<p>Моніторинг ETL-процесів. Агрегація даних Інструменти, що забезпечують отримання даних із зовнішніх джерел (extract), їх перетворення (transform) та очищення, завантаження (load) даних в сховище або в іншу базу.</p>
<p>Розділ 3. Аналіз даних</p>	
<p>Тема 3.1 on-line analytical processing</p>	
Лек12	<p>OLAP-системи Формати зберігання даних в OLAP-кубах (ROLAP, MOLAP, HOLAP). Загальна характеристика, переваги та недоліки.</p>
Лек13	<p>Технології OLAP Визначення OLAP. Тест FASMI. Правила Кодда для OLAP-систем. Концепція багатомірної моделі даних. Поняття багатомірного простору (гіперкуба). Основні характеристики гіперкубу. Основні операції над гіперкубом.</p>
Лек14	<p>Особливості реалізації сховищ даних Варіанти реалізації сховищ даних. Концепція Corporate Information Factory. Концепція Data Warehouse Bus.</p>
Лек15	<p>Ключові індикатори ефективності в аналітичних системах Призначення, основні особливості ключових індикаторів ефективності. Поняття про керування корпоративною ефективністю.</p>
Лек16	<p>Аналіз даних на основі OLAP-кубів Засоби створення сховищ даних</p>
<p>Тема 3.2 Візуалізація даних</p>	
Лек17	<p>Графічні засоби представлення даних. Основні види візуалізації даних. Звіти, графіки, діаграми. Візуалізація даних через інформаційні панелі (dashboards)</p>

	<p>Популярні інструменти бізнес-інтелекту для сприйняття даних.</p> <p>Робота з Looker Studio</p> <p>Основні види графіків, які можна побудувати в Looker Studio</p> <p>Найкращі практики візуалізації даних</p> <p>Робота з Tableau. Візуалізація, фільтрація результатів</p> <p>Зберігання результатів аналізу до дашбордів та побудова історії на основі даних</p>
--	---

Виконання студентами практичних робіт з комп'ютерного практикуму можливе у середовищі будь-якої СКБД.

Практичні заняття

№ з/п	Перелік практичних робіт
1	Створення оперативної бази даних
2	Створення сховища даних
3	Створення ETL-процедур для завантаження даних до сховища даних
4	Створення багатомірної бази даних та засобів роботи із нею
5	Розробка та застосування ключових індикаторів ефективності

Самостійна робота студента

Самостійна робота здобувача є основним засобом засвоєння навчального матеріалу у вільний від навчальних занять час і включає: опрацювання навчального матеріалу, підготовку до лекцій та інших видів навчальних занять, виконання індивідуальних завдань, підготовку кваліфікаційної роботи, науково-дослідну роботу тощо. [Положення про організацію освітнього процесу в КПІ ім. Ігоря Сікорського <https://osvita.kpi.ua/sites/default/files/files/Положення%20про%20освітній%20процес.pdf>]

Метою виконання індивідуального завдання студентами є отримання практичних навичок у роботі зі сховищами даних.

У якості індивідуального завдання кожному студенту пропонується розробити концептуальну модель БД та виконання завдань з комп'ютерного практикуму створеній БД.

№ з/п	Вид самостійної роботи	Кількість годин СРС
1	Підготовка до практичних занять	50
2	Підготовка до МКР	10
3	Підготовка до заліку	6

Для закріплення матеріалу проводиться модульна контрольна робота у вигляді тестових завдань або самостійних завдань. Дані завдання носять практичний характер і можуть виконуватися як у письмовому вигляді, так і на комп'ютерах.

5. Політика навчальної дисципліни (освітнього компонента)

Відвідування лекцій, а також відсутність на них, не оцінюється.

Викладання матеріалу на лекційних заняттях матеріал по проектуванню сховищ даних чергується з матеріалом з практичною роботою з даними.

Як самостійну роботу студенти отримують завдання для комп'ютерного практикуму. Індивідуальні завдання здаються на практичних роботах. Студент особисто показує виконане завдання та захищає роботу. У якості захисту роботи студент показує знання та навички з практичної роботи. При захисті практичних робіт студент має продемонструвати розроблений програмний код та результати його виконання на тестах, які він підготував заздалегідь вдома, так і запропонованих викладачем. У випадку дистанційної форми навчання захист відбувається на відповідній конференції шляхом демонстрації екрана.

Політика та принципи академічної доброчесності визначені у розділі 3 Кодексу честі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського». Норми етичної поведінки Норми етичної поведінки студентів і працівників визначені у розділі 2 Кодексу честі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського». Детальніше: <https://kpi.ua/code>.

6. Види контролю та рейтингова система оцінювання результатів навчання (PCO)

***Поточний контроль:** вправи на лекційних заняттях, тестування, МКР, виконання та захист практичних робіт.*

***Календарний контроль:** провадиться двічі на семестр як моніторинг поточного стану виконання вимог силабусу.*

***Семестровий контроль:** залік.*

Рейтингова оцінка з навчальної дисципліни формується з балів, що отримуються за виконання практичних робіт та за модульну контрольну роботу. Семестровим контролем є залік.

***Лабораторна робота.** Робота складається з практичної та теоретичної частини. Практична частина складається з закінченого фрагменту завдання, виконаного середовищі СКБД (MS SQL / PostgreSQL). Теоретична частина – відповіді на додаткові питання за темою роботи.*

- правильна та раціонально виконана робота та повна відповідь на теоретичні питання – 10 балів;*
- не раціональна виконана робота, програмна частина завдання виконана неточно або не в повному обсязі та неповна відповідь на теоретичні питання – 8 балів;*
- не раціональна структура даних, не коректна робота з даними БД, не коректна використані команди роботи з БД, та нерозуміння теоретичного матеріалу – 5 балів;*
- не раціональна структура даних, не коректно застосовані команди обробки даних та незадовільна відповідь на теоретичні питання – 2 бали;*
- не виконана робота – 0 балів.*

Модульна контрольна робота.

Містить завдання з написання запитів для роботи зі сховищем даних.

- правильна та раціонально виконана робота та повна відповідь на теоретичні питання – 8-10 балів;*
- не раціональна виконана робота, програмна частина завдання виконана неточно або не в повному обсязі та неповна відповідь на теоретичні питання – 5-7 балів;*
- не раціональна структура даних, не коректно застосовані команди обробки даних та незадовільна відповідь на теоретичні питання – 1-4 балів;*
- не виконана робота – 0 балів.*

Заохочувальні бали:

- виконання додаткових завдань з підвищеною складністю – 2-5 балів;
- використання в лабораторних роботах самостійно вивчених тем – 2-5 балів.

Сума вагових балів контрольних заходів протягом семестру складає:

$$R_c = (50_{(\text{лаб.роб})} + 10_{(\text{к.р})}) * 1,67 = 100 \text{ балів}$$

Умови допуску до семестрового контролю: виконані та захищені лабораторні роботи написана контрольна робота.

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

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

Форма семестрового контролю – залік

Рейтинговий бал складається з балів, отриманих протягом семестра. Умовою допуску до заліку є стартовий рейтинг (r_c) не менше 40% від R , тобто 40 балів. Студенти, які набрали протягом семестра необхідну кількість балів (60 балів і більше) отримують рейтингову оцінку з дисципліни.

Робочу програму навчальної дисципліни (силабус): Інтелектуальні бази та сховища даних

Складено старшим викладач кафедри ІПЗЕ Дацюк О.А. .

Ухвалено кафедрою ІПЗЕ (протокол № 28 від 15.05.2023р.)

Погоджено Методичною комісією ННІАТЕ КПІ ім. Ігоря Сікорського (протокол № 9 від 26.05.2023р.)