

АНОТАЦІЯ ВИБІРКОВОЇ ДИСЦИПЛІНИ

Назва дисципліни / освітнього компонента	«Сховища та простори даних»
Освітня програма	«Інженерія програмного забезпечення», «Комп'ютерні науки»
Компонент освітньої програми	Вибірковий
Загальна кількість кредитів та кількість годин для вивчення дисципліни	3,0 кредити / 90 годин
Вид підсумкового контролю	залік
Мова викладання	українська
Викладач	Турбал Юрій Васильович, д.т.н., професор, професор кафедри інформаційних технологій та моделювання
CV викладача на сайті кафедри	https://kitm.rshu.edu.ua
E-mail викладача	y.v.turbal@nuwm.edu.ua
Консультації	Згідно з графіком консультацій

МЕТА ТА ЗАВДАННЯ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Навчальна дисципліна «Сховища та простори даних» належить до вибірових компонентів циклу професійної підготовки для першого (бакалаврського) рівня вищої освіти за спеціальностями 121 Інженерія програмного забезпечення та 122 Комп'ютерні науки. Вона присвячена сучасним концепціям організації та архітектур сховищ даних, принципам їх моделювання та супроводження. Освітній компонент формує у здобувачів вищої освіти розуміння багатовимірних моделей даних, концепцій Business Intelligence та практичних інструментів для реалізації сховищ даних. **Актуальність** вивчення дисципліни визначається тим, що сховища даних є основою аналітичних систем і підтримки прийняття рішень у сучасних організаціях, а знання про їх побудову та супровід є важливою компетентністю майбутніх фахівців у галузі інформаційних технологій.

Вибірковий компонент вивчається після освоєння таких дисциплін: «Алгоритми і структури даних», «Бази даних та інформаційні системи».

Мета викладання дисципліни. Сформувати у здобувачів вищої освіти знання та практичні навички розроблення логічних і фізичних моделей сховищ даних, проектування, реалізація та супроводження сховищ даних, а також використання сучасних CASE-засобів і Microsoft SQL Management Studio для аналітики й управління життєвим циклом даних.

Цілі навчання:

- ознайомлення з архітектурами та концепціями сховищ даних;
- формування навичок розроблення логічних та фізичних моделей сховищ даних;
- розвиток умінь створювати багатовимірні моделі даних;
- усвідомлення ролі Business Intelligence систем у сучасному бізнесі;
- практичне освоєння CASE-засобів та SQL-середовищ для супроводження сховищ даних;

- навчання аналізу та вибору оптимальних технологій для конкретних бізнес-задач.

Завдання дисципліни:

- вивчення сучасних архітектур сховищ даних;
- ознайомлення з принципами моделювання та багатовимірного аналізу даних;
- практичне освоєння інструментів Microsoft SQL Management Studio та CASE-засобів;
- формування навичок супроводження та оптимізації сховищ даних;
- розвиток компетентності у виборі технологій для бізнес-аналітики.

Очікувані результати навчання. Сформовані у здобувачів вищої освіти знання про багатовимірні моделі даних та моделі сховищ даних; володіння концепціями, методами та технологіями опрацювання даних, зокрема Business Intelligence систем; здатність застосовувати CASE-засоби та SQL-середовища для реалізації сховищ даних.

Студенти мають знати:

- сучасні концепції та архітектури сховищ даних;
- принципи моделювання сховищ та багатовимірних моделей даних;
- концепції Business Intelligence систем;
- сучасні CASE-засоби та SQL-середовища для супроводження сховищ даних.

Студенти мають вміти:

- розробляти логічні та фізичні моделі сховищ даних;
- створювати багатовимірні моделі для аналітики;
- використовувати Microsoft SQL Management Studio для реалізації та супроводу сховищ;
- застосовувати CASE-засоби для побудови моделей;
- аналізувати та обирати оптимальні технології для бізнес-аналізу.

ЗМІСТ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Тема 1. Основні поняття про сховища даних. Концепція систем складування даних. OLTP vs OLAP. Концепція сховищ даних. Моделі даних сховищ даних. Сценарій функціонування сховищ даних. Налаштування середовища Microsoft SQL Server. Встановлення та налаштування середовища розроблення та підтримки моделі сховища даних.

Тема 2. Архітектура сховищ даних. Підвиди сховищ даних. Архітектури сховищ даних. Підходи до організації робіт з проектування сховищ даних. Аналіз розроблених схем сховищ даних на прикладі аналізу навчальної бази даних (AdventureWorks).

Тема 3. Метод багатовимірного моделювання. Основні поняття методу багатовимірного моделювання. Багатовимірна модель. Факти. Ключі. Виміри. Основні схеми логічних моделей сховищ даних. Моделювання таблиць фактів та вимірів. Ієрархії вимірів. Створення таблиць вимірів та таблиці фактів, налаштування взаємозв'язків між ними засобами програмного середовища Microsoft SQL Server. Налаштування завантаження і аудиту завантажень даних до сховища даних засобами SQL Server Data Tools, керування шкідкодiєю сховищ даних; завантаження і аудиту завантажень.

Тема 4. Підтримка життєвого циклу сховищ даних. Огляд рішень відомих виробників для підтримки життєвого циклу сховища даних. Замкнуті та відкриті рішення. Об'єкти фізичної моделі сховища даних. Моделювання об'єктів фізичної моделі даних сховища даних. Розробка скриптів для створення об'єктів фізичної моделі даних сховища даних. Реалізація процесів витягання, перетворення та завантаження даних засобами програмного середовища Microsoft SQL Server. Розробка пакетів SSIS у SSDT; знайомство з потоком керування, потоком даних і диспетчерами з'єднань.

Тема 5. Інтеграція та агрегація даних. Системи інтеграції. Задачі процесу інтеграції даних. Методи інтеграції. Схема процесу консолідації даних. Поняття агрегату. Часткова и повна агрегація. Оператор агрегації. Процедури формування агрегатів. Агрегація розрідженого гіперкуба. Налаштування диспетчерів з'єднань засобами програмного середовища Microsoft SQL Server. Створення диспетчера з'єднань неструктурованого

файлу, OLE DB. Налаштування потоків управління та налаштування черговості з'єднань засобами програмного середовища Microsoft SQL Server. Визначення потоку управління; застосування пакета SSIS для обробки файлів. Налаштування потоків даних, елементів перетворень об'єктів та налаштування черговості з'єднань. Створення простих потоків даних; робота із перетвореннями потоку даних.

Тема 6. Метадані в сховищах даних. Простори даних. Поняття метаданих. Функції метаданих в сховищах даних. Логічна структура метаданих в сховищі даних. Стандарти метаданих. Вибір моделі метаданих для сховища даних. Причини появи просторів даних. Модель простору даних. Учасники простору даних. Задачі учасників простору даних. Особливості просторів даних. Дослідницькі проблеми. Перспективи розвитку.

Тема 7. Огляд аналітичних баз даних. Озера даних. Огляд найпоширеніших аналітичних баз даних. Призначення та особливості використання. Поняття озера даних. Архітектура озера даних. Сфери використання озер даних. Моделі озер даних. Технології озер даних. Порівняння технологій аналітичних баз даних, озер даних та сховищ даних. Ознайомлення із рішеннями поширених проблем сховищ даних в управлінні вимірами, що повільно змінюються SDC – Slowly Changing Dimensions. Реалізація механізму відслідковування змін даних (CDC) за допомогою засобів SSIS; використання потоку помилок.

Тема 8. Business Intelligence System та сховища даних. Поняття Business Intelligence System. Технології обробки даних. Архітектура систем бізнес-рівня. Засоби візуалізації даних. Створення більш надійних в експлуатації пакетів за допомогою налаштування транзакцій та контрольних точок; використання механізму обробки подій та застосування додаткових завдань у пакеті засобами програмного середовища Microsoft SQL Server. Увімкнення підтримки транзакцій і спостереження за відкатом транзакцій; встановлення та перегляд контрольних точок; застосування обробників подій. Створення динамічних пакетів з використанням параметрів фільтрації вихідного запиту та налаштувань конфігурації засобами програмного середовища Microsoft SQL Server.