

Анотація навчальної дисципліни вільного вибору здобувача вищої освіти

Дисципліна:	«OLAP технології»
Викладач:	Мороз І.П., к.ф.-м.н., доцент
E-mail:	igor_moroz@yahoo.com
Кількість кредитів:	3
Мова викладання:	українська
Вид контролю:	залік
Місце у структурно-логічній схемі:	вивчається в 2 семестрі другого (магістерського) рівня вищої освіти за спеціальністю 122 Комп'ютерні науки

Вступ

В останні роки у світі оформилася низка нових концепцій зберігання та аналізу корпоративних даних:

- сховища даних, або склади даних (Data Warehouse);
- оперативна аналітична обробка (On-Line Analytical Processing, **OLAP**);
- інтелектуальний аналіз даних – ІАД (Data Mining).

Онлайн(оперативна)-аналітична обробка (OLAP) - це програмна технологія, яку можна використовувати для аналізу бізнес-даних з різних точок зору. Організації збирають та зберігають дані з кількох джерел даних, таких як веб-сайти, додатки, інтелектуальні лічильники та внутрішні системи. OLAP об'єднує та групує ці дані за категоріями, щоб забезпечити практичну інформацію для стратегічного планування. Наприклад, роздрібний торговець зберігає дані про всі товари, такі як колір, розмір, вартість і місцезнаходження. Роздрібний торговець також збирає дані про покупки клієнтів, такі як назва замовлених товарів та загальна вартість продажу в іншій системі. OLAP об'єднує набори даних, щоб відповісти на такі питання, як кольорові продукти, більш популярні або як розміщення продукту впливає на продаж.

В OLAP дані подаються в системі координат з N осями (N вимірному просторі). Оскільки OLAP не накладає обмежень на кількість осей, застосовується термін «гіперкуб», що означає відображення даних в просторі необмеженої вимірності.

OLAP - є інструментом для аналізу великих обсягів даних. У процесі взаємодії з OLAP-системою користувач може здійснювати гнучкий перегляд інформації, отримувати довільні зрізи даних та виконувати аналітичні операції деталізації, згортки, наскрізного розподілу, порівняння у часі.

Метою викладання навчальної дисципліни «**OLAP технології**» є формування у здобувачів освіти необхідних компетенцій для проектування, розробки та впровадження сховищ даних і OLAP-систем.

Програма навчальної дисципліни

МОДУЛЬ 1. СХОВИЩА ДАНИХ (СД)

ТЕМА 1. ВСТУП У СД. Визначення основних понять. Підходи Біла Інмона та Ральфа Кімбала до створення СД. Етапи та учасники типового процесу створення сховища даних.

ТЕМА 2. МОДЕЛЮВАННЯ ДАНИХ ДЛЯ СД. Рівні моделювання даних у СД. Інструменти моделювання. Методи створення даних: ER-моделі, багатовимірне моделювання, DataVault, використання індустріальних моделей даних.

ТЕМА 3. ПРОЦЕС ІНТЕГРАЦІЇ ДАНИХ. ETL-процеси: методи проектування та інструменти реалізації. Інтеграція даних. Види інтеграції: консолідація, федералізація. Етапи життєвого циклу створення системи інтеграції даних. Способи взаємодії із джерелами даних. Інструменти для створення систем інтеграції даних. Типові рішення про створення ETL-процесів для СД, створених із застосуванням підходів DataVault, Багатовимірне моделювання. Архітектура ETL-процесів для індустріальних моделей IBM. Популярні формати даних (ODBC, JDBC, XML, JSON, CSV).

МОДУЛЬ 2. БАГАТОВИМІРНІ МОДЕЛІ ДАНИХ, OLAP-КУБИ

ТЕМА 1. ОСНОВИ OLAP. Напрямки розвитку сучасних систем та технологій обробки даних. Сучасні концепції, що лежать в основі методів бізнес-інтелекту (BI). Технології Microsoft підтримки BI-систем. OLAP та OLTP.

ТЕМА 2. OLAP-СИСТЕМИ ТА OLAP-ЗАСОБИ. Структура OLAP-кубу (гіперкубу). Виміри. Ієрархія вимірів OLAP-кубів. Операції над гіперкубом. Зріз, обертання, консолідація та деталізація. Схема даних сховища. Таблиця фактів. Основні типи фактів. Таблиці вимірів. Архітектура OLAP-систем. Клієнтські OLAP-засоби. Серверні засоби OLAP. Продукти Oracle для OLAP та бізнес-аналізу. Microsoft SQL Server Analysis Services. Технічні аспекти багатовимірного зберігання даних. Методи реалізації багатовимірної моделі – MOLAP, ROLAP, HOLAP.

ТЕМА 3. БАГАТОМІРНИЙ АНАЛІЗ ДАНИХ ЗА ДОПОМОГОЮ СЛУЖБИ SQL SERVER ANALYSIS SERVICES. Можливості служби SSAS. Компоненти BI-рішення Microsoft. Уніфікована багатовимірна модель (Unified Dimensional Model, UDM). Масштабованість та продуктивність. Подання джерела даних. Інтеграція із Microsoft Office. Інструменти керування службою SSAS. Типи проектів для Analysis Services. Планування та архітектура SSAS. Логічна та фізична архітектури. Архітектура компонентів служб SSAS. Архітектура програмування SSAS.

ТЕМА 4. ЗАПИТИ ДО БАГАТОВИМІРНИХ БАЗ ДАНИХ. Ключові поняття багатовимірних виразів. Вимірювання бази даних. Атрибут виміру. Елемент. Міра. Вимірювання мір. Група мір. Ключовий атрибут виміру бази даних. Атрибут гранулярності. Вимірювання кубу. Ієрархія атрибуту. Зв'язок атрибутів. Властивість елемента. Комірка кубу. Простір кубу. Вкладений куб. Кортєжі. Набори. Основні поняття про запити до багатовимірних баз даних. Синтаксис базової інструкції SELECT. Основні поняття про сценарії багатовимірних виразів.

ТЕМА 5. ВИКОРИСТАННЯ СЛУЖБ INTEGRATION SERVICES ЗІ СЛУЖБАМИ ANALYSIS SERVICES. Можливості Integration Services для роботи з OLAP. Злиття даних із різнорідних сховищ даних. Заповнення сховищ даних та вітрин даних. Очищення та стандартизація даних. Архітектура служб SSIS. Пакет SSIS. Елементи потоку керування. Елементи потоку даних.