

Методичні вказівки для самостійної роботи з дисципліни «Методи та засоби інтеграції даних» для студентів денної форми навчання, рівень підготовки – магістр. Напрямок підготовки – 122 «Комп'ютерні науки» / Козловська В.П.. – Одеса, ОДЕКУ, 2017. – 12 с.

## ЗМІСТ

ПЕРЕДМОВА .....	4
ПРОГРАМА КУРСУ «МЕТОДИ ТА ЗАСОБИ ІНТЕГРАЦІЇ ДАНИХ» .....	5
ОРГАНІЗАЦІЯ ПОТОЧНОГО ТА ПІДСУМКОВОГО КОНТРОЛЮ ЗНАНЬ .....	6
ЛІТЕРАТУРА.....	12

## ПЕРЕДМОВА

Метою дисципліни «Методи та засоби інтеграції даних» є формування у студентів теоретичних знань та практичних навичок, необхідних для роботи у будь-яких адміністративних, наукових та виробничих підрозділах, в яких використовують сховища даних, а також здійснюють обслуговування та розробку інформаційних систем та сховищ даних.

Згідно з поставленою метою до задач дисципліни входить:

- Ознайомлення студентів з основними поняттями теорії сховищ даних, їх призначення, загальні риси та відмінність від баз даних;
- Надання студентам знань про основні види сховищ даних;
- Надання студентам знань про задачі, які вирішуються за допомогою сховищ даних;
- Ознайомлення студентів з основними поняттями інтеграції даних у сховищах даних;
- Ознайомлення студентів з основними методами інтеграції даних;
- Ознайомлення студентів з основними технологіями та засобами інтеграції даних.

Після освоєння цієї дисципліни студент **повинен знати**: призначення сховищ даних; види сховищ даних, переваги та недоліки різних видів сховищ даних; основні групи користувачів сховищ даних; задачі, які вирішуються за допомогою сховищ даних; поняття інтеграції; види інтеграції даних; методи інтеграції даних; технології інтеграції даних; технології та засоби інтеграції даних; програмні продукти, що втілюють технології та засоби інтеграції даних.

За результатами навчання студент **повинен вміти**: обрати вид сховища даних, що найбільш відповідає задачам, що вирішуються в організації або установі; вибрати програмні продукти для інтеграції даних, які відповідають потребам організації та мають прийнятний рівень співвідношення ціна/якість.

Ця дисципліна базується на знаннях та навичках, отриманих під час вивчення дисципліни “Організація баз даних та знань”.

У процесі самостійного вивчення курсу студент повинен керуватися його програмою і вивчити за конспектом лекцій та літературою, що рекомендована викладачем, відповідний теоретичний матеріал.

## **ПРОГРАМА КУРСУ «МЕТОДИ ТА ЗАСОБИ ІНТЕГРАЦІЇ ДАНИХ»**

Курс містить 2 лекційних та два практичних змістовні модулі.

Лекційні модулі – ЗЛ1, ЗЛ2.

### **ЗЛ1. Сховища даних**

**Тема 1.** Сховище даних як об'єкт для інтеграції даних. Поняття сховища даних.

**Тема 2.** Складові сховища даних.

**Тема 3.** Підвиди сховищ даних.

### **ЗЛ2. Інтеграція даних**

**Тема 4.** Характеристики інтеграції даних

**Тема 5.** Методи інтеграції даних

**Тема 6.** Технології інтеграції даних

**Тема 4.** Засоби інтеграції даних

Практичні модулі – ЗП1, ЗП2.

### **ЗП1. Робота зі сховищами даних**

**Тема 1.** Організація сховища даних.

**Тема 2.** Запити на вибірку даних зі сховища даних.

### **ЗП2. Засоби інтеграції даних в СУБД MS SQL Server та Oracle**

**Тема 3.** Засоби інтеграції даних в MS SQL Server

**Тема 4.** Засоби інтеграції даних в Oracle

Для самостійної теоретичної підготовки з тем модулю ЗЛ1 студенти використовують літературу: С.1 – 61 [1], С. 191 – 226 [2], [3], С. 1227 – 1238 [4].

Для самостійної теоретичної підготовки з тем модулю ЗЛ2 студенти використовують літературу: С.62 – 121 [1], С. 253 – 285 [2], С. 1173 – 1194 [4].

Для самостійної теоретичної підготовки з тем модулю ЗП1 студенти використовують літературу: С. 194 – 197 [2], С. 1235 – 1240, 1256 – 1262 [4].

Для самостійної теоретичної підготовки з тем модулю ЗП2 студенти використовують літературу: С. 323 – 343, 572 – 624 [5], С. 1256 – 1262, 1279 – 1288 [4].

## ОРГАНІЗАЦІЯ ПОТОЧНОГО ТА ПІДСУМКОВОГО КОНТРОЛЮ ЗНАНЬ

Поточна оцінка рівня знань студентів з дисципліни “Методи та засоби інтеграції даних” здійснюється за модульною системою. Семестровий контроль проводиться у вигляді заліку.

Для організації поточного контролю рівня набутих студентами знань, вмінь та навичок весь теоретичний курс поділено на 2 лекційних змістовних модуля та 2 практичних змістовних модуля, які відповідають розділам робочої програми дисципліни.

Засобами контролю знань студентів по темам лекційних модулів є дві контрольних роботи у формі комп’ютерних тестів та індивідуальне завдання у формі реферату.

Розподіл балів за виконання завдань на СРС по лекційним модулям наводиться у табл. 1.

Таблиця 1 – Розподіл балів за виконання заходів СРС по модулям ЗЛ1, ЗЛ2

Змістовний модуль	Форма контролю	Максимальна сума балів
ЗЛ1	Комп’ютерний тест	22
ЗЛ2	Комп’ютерний тест	18
Реферат	Захист реферату	40
<b>Разом</b>		<b>80</b>

Засобами контролю знань та вмінь студентів по темам практичних є усне опитування.

Розподіл балів за виконання завдань на СРС по практичним модулям наводиться у табл. 2.

Таблиця 1 – Розподіл балів за виконання заходів СРС по модулям ЗП1, ЗП2

Змістовний модуль	Форма контролю	Максимальна сума балів
ЗП1, Тема 1	Усне опитування	5
ЗП1, Тема 2	Усне опитування	5
ЗП2, Тема 3	Усне опитування	5
ЗП2, Тема 4	Усне опитування	5
<b>Разом</b>		<b>20</b>

# ІНДИВІДУАЛЬНЕ ЗАВДАННЯ

## ТЕМИ РЕФЕРАТІВ

1. Сховища даних: призначення, структура, порівняння з базами даних – загальні риси та відмінності.
2. Реляційні та багатовимірні сховища даних: призначення, переваги та недоліки.
3. Види сховищ та вітрин даних.
4. Системи бізнес-аналітики та BI (Business Intelligence systems).
5. Технології інтеграції даних, що базуються на методі «Консолідація даних».
6. Технології інтеграції даних, що базуються на методі «Федералізація даних».
7. Технології інтеграції даних, що базуються на методі «Розповсюдження даних».
8. Мова XML як засіб обміну слабо структурованими даними.
9. Методи інтеграції СУБД у середовище Web.
10. OLAP системи: призначення, порівняння зі сховищами даних, інструменти OLAP.

# ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ СТУДЕНТІВ ТА ПІДГОТОВКИ ДО КОМП'ЮТЕРНОГО ТЕСТУВАННЯ

## Питання до тем модулю ЗЛ1 «Сховища даних»

1. Сховища даних за Інмоном – це :
2. Сховища даних за Кімболом – це:
3. Узагальнене визначення сховища даних – це
4. Вимоги до сховища даних - це:
5. Питання, на які можна відповідати за допомогою сховищ даних:
  - Наскільки точно різні підрозділи компанії виконують встановлений бюджет?
  - Які тенденції витрат за різними підрозділами, статтями бюджету?
  - Наскільки вчасно надходять платежі?- відносяться до:
6. Питання, на які можна відповідати за допомогою сховищ даних:
  - Наскільки виконаний план з продажу?
  - Які ключові показники продуктивності компанії в поточному періоді?
  - Які тенденції зміни ключових показників продуктивності компанії з часом?- відносяться до:
7. Питання, на які можна відповідати за допомогою сховищ даних:
  - Які продукти (послуги) дають найбільший прибуток?
  - Яка комбінація підрозділів і товарів (послуг) просуває бізнес?
  - Які сегменти ринку дають найбільший прибуток?
  - Який набір продуктів купують найприбутковіші клієнти?- відносяться до:
8. Питання, на які можна відповідати за допомогою сховищ даних:
  - Хто з торгових представників краще за всіх продає товари (послуги)?
  - Які з партнерів забезпечують найбільший прибуток?
  - Які продукти, групи продуктів найкраще продає певний партнер?
  - Які тенденції зміни обсягів продажу через партнерів?

- відносяться до:

9. Питання, на які можна відповідати за допомогою сховищ даних:

- Які сегменти ринку забезпечують найбільший прибуток?
- Які клієнти дають найбільший прибуток?
- Які властивості характерні клієнтам, що забезпечують найбільший прибуток?

- відносяться до:

10. Питання, на які можна відповідати за допомогою сховищ даних:

- Яка вірогідність відгуку певного сегменту ринку на нову пропозицію?
- Яка вартість проведення маркетингової кампанії через заданий канал?
- Який канал проведення маркетингових кампаній є найефективнішим?

- відносяться до:

11. Питання, на які можна відповідати за допомогою сховищ даних:

- Наскільки швидшим або повільнішим стало постачання продуктів (послуг) в певний сегмент ринку?
- Які основні причини відмови від продукту (послуги)?
- Які клієнти найближчим часом, можливо, перестануть користуватися послугами компанії?

- відносяться до:

12. Питання, на які можна відповідати за допомогою сховищ даних:

- Наскільки клієнти задоволені якістю товарів (послуг)?
- Як задоволеність клієнтів змінюється з часом?
- Які продукти постачають вчасно, які - із запізненням?
- Чи мають певні клієнти або продукти неприпустимо довгий термін постачання?

- відносяться до:

13. Перерахуйте відмінності сховищ даних від реляційних баз даних

14. Дайте визначення типів метаданих у сховищі даних за Кімболом

15. Перерахуйте учасників сховища даних

16. Перерахуйте засоби завантаження сховища даних

17. Перерахуйте засоби моніторингу сховища даних



18. Перерахуйте засоби створення і розвитку сховища даних
19. Дайте визначення способу опису пасивного використання метаданих
20. Дайте визначення способу опису активного використання метаданих
21. Оберіть тип запиту до сховища даних за визначенням: комбінує кубби, які мають одне або декілька загальних вимірів
22. Оберіть тип запиту до сховища даних за визначенням: узагальнює значення вимірів:
23. Оберіть тип запиту до сховища даних за визначенням: повертає тільки ті результати, які з'являються у верхній або нижній частині впорядкованого певним чином списку
24. Оберіть тип запиту до сховища даних за визначенням: дає користувачам можливість побачити дані, згруповані за іншим виміром
25. Оберіть тип ієрархії за визначенням: ієрархія, в якій кількість рівнів визначена її структурою і незмінна, і кожна гілка ієрархічного дерева містить об'єкти кожного з рівнів
26. Оберіть тип ієрархії за визначенням: ієрархія, в якій кількість рівнів визначена її структурою і постійна, проте деякі гілки ієрархічного дерева можуть не містити об'єкти якого-небудь рівня

### **Питання до тем модулю ЗЛ2 «Інтеграція даних»**

1. Назвіть найбільш поширений спосіб інтеграції даних. З точки зору реалізації це найпростіший спосіб, але має найбільше недоліків.
2. Процес приведення даних електронних таблиць двох БД в ідентичний стан. Цей процес заснований на поняттях "видавець" і "передплатник". Видавцем є сервер публікації, тобто сервер, що відправляє інформацію. Передплатником є відповідно сервер, що приймає дані – сервер підписки. Як називається технологія інтеграції даних, який відповідає даний опис?
3. Технологія інтеграції даних, яка представляється як зручний засіб інтеграції додатків, що дозволяє легко реалізувати міжплатформене взаємодію. Ця технологія не може розглядатися як загальний підхід до інтеграції додатків з кількох причин, зокрема, вона непридатна для обробки великих обсягів інформації, оскільки всі дані перетворюються в формат XML, що веде до збільшення обсягу даних і навантаження на систему в цілому. Як називається технологія інтеграції даних, який відповідає даний опис?

4. Технологія інтеграції даних, яка використовує проміжні сервери в інтеграційному середовищі, шлюзи, які здійснюють обробку потоків даних і повідомлень, а також розподіл даних між додатками, що мають різні інтерфейси. Як називається технологія інтеграції даних, який відповідає даний опис?
5. Якого виду конфлікти можуть виникати при вирішенні задачі інтеграції метаданих в системах інтеграції структурованих даних при використанні різних моделей даних для різних джерел?
6. Якого виду конфлікти можуть виникати при вирішенні задачі інтеграції метаданих в системах інтеграції структурованих даних при виборі різних рівнів абстракції для моделювання подібних сутностей реального світу?
7. Якого виду конфлікти можуть виникати при вирішенні завдання інтеграції метаданих в системах інтеграції структурованих даних при поданні одних і тих же сутностей в різних джерелах різними структурами даних?
8. Поява якого виду конфліктів можлива при вирішенні завдання інтеграції метаданих в системах інтеграції структурованих даних з використанням в різних схемах різної термінології, яка призводить до омонімів і синонімів в іменуванні різних моделей даних для різних джерел?
9. Системи інтеграції даних можуть забезпечувати інтеграцію даних на фізичному, логічному і семантичному рівні. Як називається найбільш проста інтеграція даних, яка зводиться до конверсії даних з різних джерел в необхідний єдиний формат їх фізичного представлення?
10. Системи інтеграції даних можуть забезпечувати інтеграцію даних на фізичному, логічному і семантичному рівні. Як називається інтеграція даних, яка передбачає можливість доступу до даних, що містяться в різних джерелах, в термінах єдиної глобальної схеми, яка описує їх спільне подання з урахуванням структурних і, можливо, поведінкових властивостей даних?
11. Системи інтеграції даних можуть забезпечувати інтеграцію даних на фізичному, логічному і семантичному рівні. Як називається інтеграція даних, яка забезпечує підтримку єдиного уявлення даних з урахуванням їх семантичних властивостей в контексті єдиної онтології предметної області?
12. Визначте технологію інтеграції за означенням: це технологія, за допомогою якої організація добивається централізації та оптимізації інтеграції корпоративних застосувань, зазвичай використовуючи ті або інші форми технології оперативної доставки інформації, яка керується зовнішніми подіями.

13. Визначте технологію інтеграції за означенням: це технологія, яка перетворить дані (зазвичай за допомогою їх пакетного опрацювання) з операційного середовища, що включає гетерогенні технології, в інтегровані, дані, що узгоджуються між собою, придатні для використання в процесі підтримки прийняття рішень.
14. Визначте технологію інтеграції за означенням: це технологія для інтеграції в режимі реального часу незіставних типів даних з численних джерел як всередині, так і за межами корпорації.
15. Дайте визначення поняття "Інтеграція даних".
16. Дайте визначення поняття "Консолідація даних".
17. Дайте визначення поняття "Федералізація даних".
18. Дайте визначення поняття "Розповсюдження даних".

## **ЛІТЕРАТУРА**

1. Козловська В.П. Методи та засоби інтеграції даних. Конспект лекцій. Електронна версія. – Одеса, ОДЕКУ, 2016 – 121 с.
2. Пасічник В.В., Шаховська Н.Б.. Сховища даних: Навчальний посібник. – Львів: «Магнолія 2006», 2008 – 496 с.
3. Дейт К. Дж. Основы будущих систем баз данных: третий манифест / Дейт К. Дж., Дарвен Х. Перевод под ред. С.Д. Кузнецова. – Л.; Янус, 12004 – 26 с.
4. Конноли, Томас, Бегг, Каролин. Базы данных. Проектирование, реализация и сопровождение. Теория и практика. 3-е издание.: Пер. с англ. – М.: Издательский дом «Вильямс», 2003. – 1440 с.
5. Мамаев Е.В. Microsoft SQL Server 2000. – СПб.: БХВ-Петербург, 2004. – 1260 с.: мул.