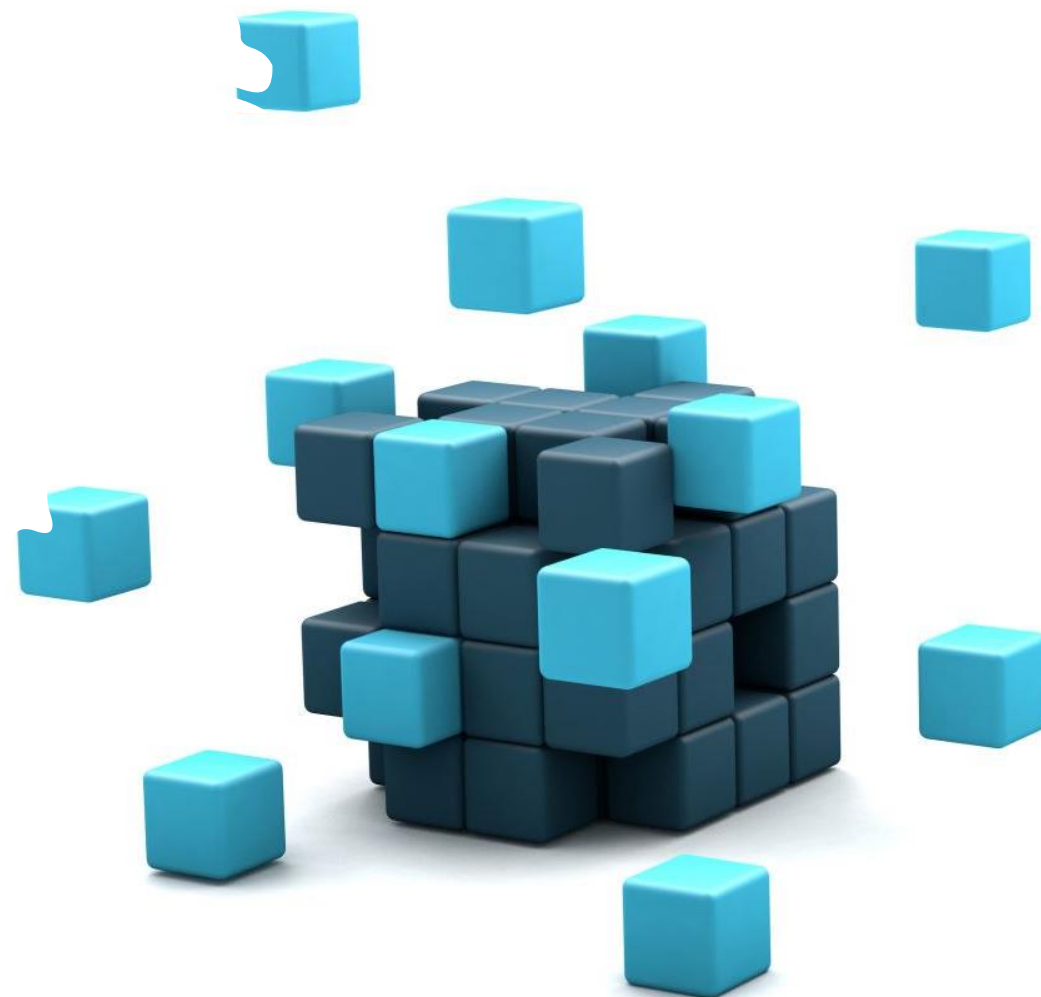


ETL процес

«ОРГАНІЗАЦІЯ СХОВИЩА
ДАНИХ»



Основні цілі і задачі процесу *ETL*

- ETL, (extraction, transformation, loading) – комплекс методів, що реалізують процес перенесення вихідних даних з різних джерел в аналітичний додаток, підтримує його сховище даних.
- Для перенесення вихідних даних з різних джерел в СД слід використовувати спеціальний інструментарій, який повинен отримувати дані з джерел різного формату, перетворювати їх в єдиний формат, підтримуваний СД, а при необхідності – проводити очищення даних від факторів, що заважають коректно виконувати їх аналітичну обробку.

Отримання оперативних даних в *ETL*

На цьому етапі дані витягуються з одного або декількох джерел і готуються до перетворення відповідно визначеному регламенту.

Процедуру вилучення можна реалізувати двома основними способами:

1. Витяг даних за допомогою спеціалізованих програмних засобів
2. Витяг даних засобами тієї системи, в якій вони зберігаються

Очищення даних в *ETL*

Розглядають 2 етапи

1. Первинне очищення - очищення даних перед їх завантаженням в сховище. Носить переважно технічний характер, основне завдання - підготувати дані до завантаження в сховище.
2. Вторинне очищення (або додаткове) - є користувальницької, воно спрямовано на підготовку даних до вирішення конкретної аналітичної задачі.

Перетворення даних в *ETL*

Мета – підготовка даних до розміщення в СД і приведення їх до виду, найбільш зручному для подальшого аналізу

В процесі перетворення даних в рамках ETL найчастіше виконуються наступні операції:

- ✓ перетворення структури даних;
 - ✓ агрегування даних;
 - ✓ переведення значень;
 - ✓ створення нових даних.

Завантаження даних в сховище

- ✓ Першими в процесі завантаження даних в СД зазвичай завантажуються таблиці вимірів, які містять сурогатні ключі і іншу описову інформацію, необхідну для таблиць фактів. При завантаженні таблиць вимірів потрібно і додавати нові записи, і змінювати існуючі.
- ✓ При завантаженні таблиці фактів нова інформація зазвичай додається в кінець таблиці, щоб не змінювати існуючі дані.

A large, irregular green brushstroke shape serves as a background for the text. It has a textured, hand-painted appearance with varying shades of green and some white highlights.

Кінець презентації

ДЯКУЮ ЗА УВАГУ!