

# Metode Data Warehouse

Oleh  
S. Maryam | Pussainsa LAPAN

Pesatnya perkembangan teknologi informasi telah memberikan perubahan mendasar dalam struktur, operasi, dan manajemen pada sebuah organisasi. *Data warehouse* adalah suatu konsep dan kombinasi teknologi yang memfasilitasi organisasi untuk mengelola dan memelihara data historis yang diperoleh dari sistem atau aplikasi operasional. Pemakaian teknologi *data warehouse* hampir dibutuhkan oleh semua organisasi. *Data warehouse* memungkinkan integrasi berbagai macam jenis data dari berbagai aplikasi atau sistem. *Data warehouse* menjamin mekanisme akses satu pintu bagi pengelola untuk memperoleh informasi, dan menganalisisnya untuk pengambilan keputusan.

## Data, Informasi, dan Basis Data

Data adalah fakta, gambar atau suara yang mungkin atau tidak berhubungan atau berguna bagi tugas tertentu. Data terdiri dari fakta-fakta dan angka yang secara relatif tidak berarti bagi pemakai. Sedangkan informasi adalah data yang sudah diproses atau data yang memiliki arti. Disini kita dapat melihat bahwa data merupakan suatu bentuk keterangan-keterangan yang belum diolah atau dimanipulasi sehingga belum begitu berarti bagi sebagian pemakai. Sedangkan informasi merupakan data yang sudah diolah sehingga memiliki arti.

Sementara itu, basis data adalah

kumpulan data yang saling berhubungan yang disimpan secara bersama sedemikian rupa dan tanpa pengulangan (redundansi) yang tidak perlu, untuk memenuhi berbagai kebutuhan. Basis data merupakan tempat penyimpanan data yang saling berhubungan secara logika, sehingga bisa digunakan untuk mendapatkan suatu informasi yang diperlukan oleh suatu organisasi atau perusahaan. Sedangkan data yang diperoleh suatu organisasi atau perusahaan umumnya didapat dari kegiatan operasional sehari-hari atau hasil dari transaksi. Dari perkembangan model basis data, kemudian muncul yang disebut dengan *data warehouse*.

## Pengertian Data Warehouse

*Data warehouse* adalah koleksi data yang mempunyai sifat berorientasi pada subjek, terintegrasi, *time-variant*, dan bersifat tetap dari koleksi data dalam mendukung proses pengambilan keputusan manajemen. Menurut Poe<sup>[5]</sup>, *data warehouse* merupakan basis

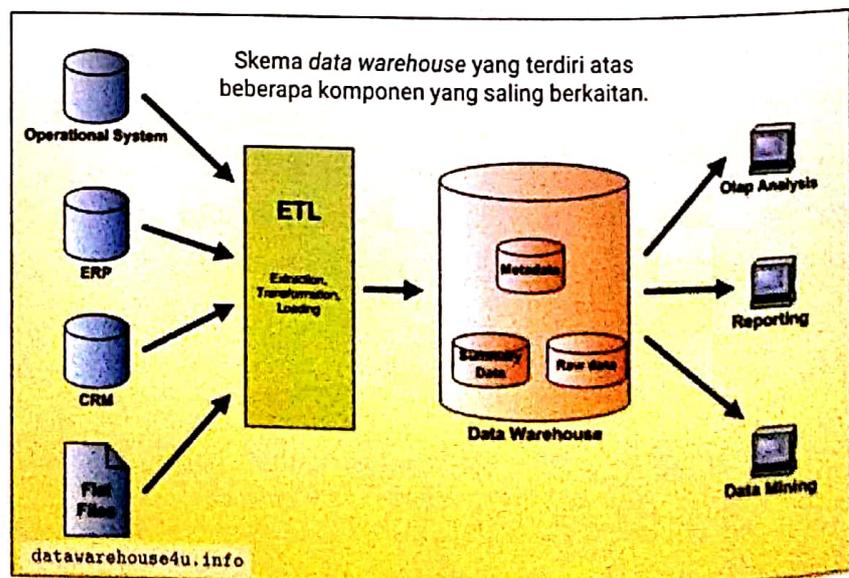
data yang bersifat analisis dan *read only* yang digunakan sebagai pondasi dari sistem penunjang keputusan. Jadi *data warehouse* merupakan metode dalam perancangan basis data, yang menunjang DSS (*Decision Support System*) dan EIS (*Executive Information System*).

Dari definisi yang dijelaskan, dapat disimpulkan *data warehouse* adalah basis data yang saling berinteraksi yang dapat digunakan untuk *query* dan analisis, bersifat orientasi subjek, terintegrasi, *time-variant*, tidak berubah yang digunakan untuk membantu para pengambil keputusan.

## Peristilahan Pada Data Warehouse

**Data Mart** adalah suatu bagian pada data warehouse yang mendukung pembuatan laporan dan analisis data pada suatu unit, bagian atau operasi pada suatu perusahaan.

**On-Line Analytical Processing (OLAP)** merupakan suatu proses basis data yang menggunakan tabel fakta dan dimensi untuk



dapat menampilkan berbagai macam bentuk laporan, analisis, dan query dari data yang berukuran besar.

**On-Line Transaction Processing**

(OLTP) adalah suatu proses penyimpanan data mengenai kegiatan operasional transaksi sehari-hari.

**Dimension Table**

adalah tabel yang berisikan kategori dengan ringkasan data mendetail yang dapat dilaporkan. Seperti laporan laba pada tabel fakta dapat dilaporkan sebagai dimensi waktu (yang berupa perbulan, perkuartal dan pertahun).

**Fact Table**

merupakan tabel yang umumnya mengandung angka dan data history dimana key (kunci) yang dihasilkan sangat unik, karena key tersebut terdiri dari foreign key (kunci asing) yang merupakan primary key (kunci utama) dari beberapa dimensi tabel yang berhubungan.

**DSS**

merupakan sistem yang menyediakan informasi kepada pengguna yang menjelaskan bagaimana sistem ini dapat menganalisis situasi dan mendukung suatu keputusan yang baik.

**Karakteristik Data Warehouse**

Data warehouse merupakan sekumpulan data yang memiliki karakter berorientasi-subjek,

terintegrasi, rentang waktu, non-volatile (bersifat tetap selamanya), ringkas, tidak ternormalisasi, serta data berasal dari berbagai sumber.

**Berorientasi subjek**

, artinya data warehouse dirancang untuk menganalisis data berdasarkan subyek-subyek tertentu dalam organisasi, bukan pada proses atau fungsi aplikasi tertentu. Data warehouse diorganisasikan disekitar subjek utama dari suatu organisasi/ lembaga dan tidak diorganisasikan pada area aplikasi utama (pelayanan pengguna). Hal ini disebabkan oleh kebutuhan data warehouse untuk menyimpan data yang bersifat sebagai penunjang suatu keputusan, daripada berupa aplikasi yang berorientasi data. Jadi dengan kata lain, data yang disimpan adalah berorientasi kepada subjek bukan terhadap proses.

**Terintegrasi**

, artinya data warehouse dapat menyimpan data yang berasal dari sumber terpisah kedalam suatu format yang konsisten dan saling terintegrasi satu dengan lainnya. Data tidak bisa dipecah karena merupakan suatu kesatuan yang menunjang keseluruhan konsep dari data warehouse. Syarat integrasi sumber data harus konsisten dalam penamaan dan ukuran variabel, konsisten dalam

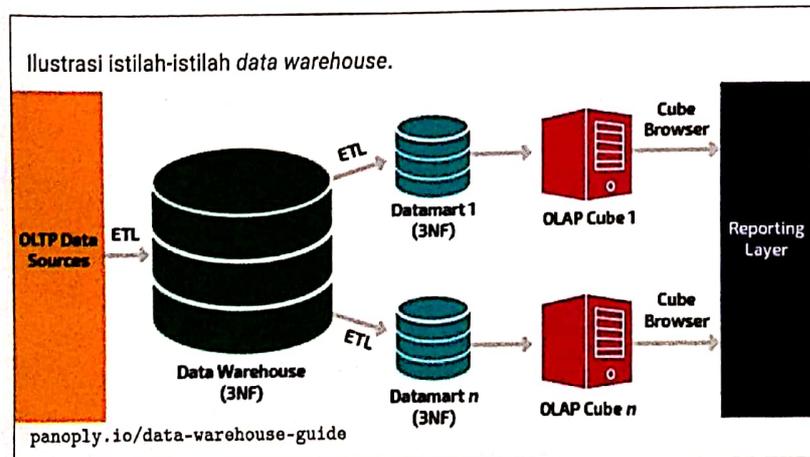
struktur pengkodean dan dalam atribut fisik. Contoh pada lingkungan operasional terdapat berbagai macam aplikasi yang mungkin dibuat oleh developer yang berbeda. Oleh karena itu, mungkin dalam aplikasi tersebut ada variabel yang memiliki maksud yang sama tetapi nama dan formatnya berbeda. Variabel tersebut harus dikonversi menjadi nama yang sama dan format yang disepakati bersama. Dengan demikian tidak ada lagi kerancuan karena perbedaan nama dan format. Barulah data tersebut bisa dikategorikan sebagai data yang terintegrasi karena kekonsistennannya.

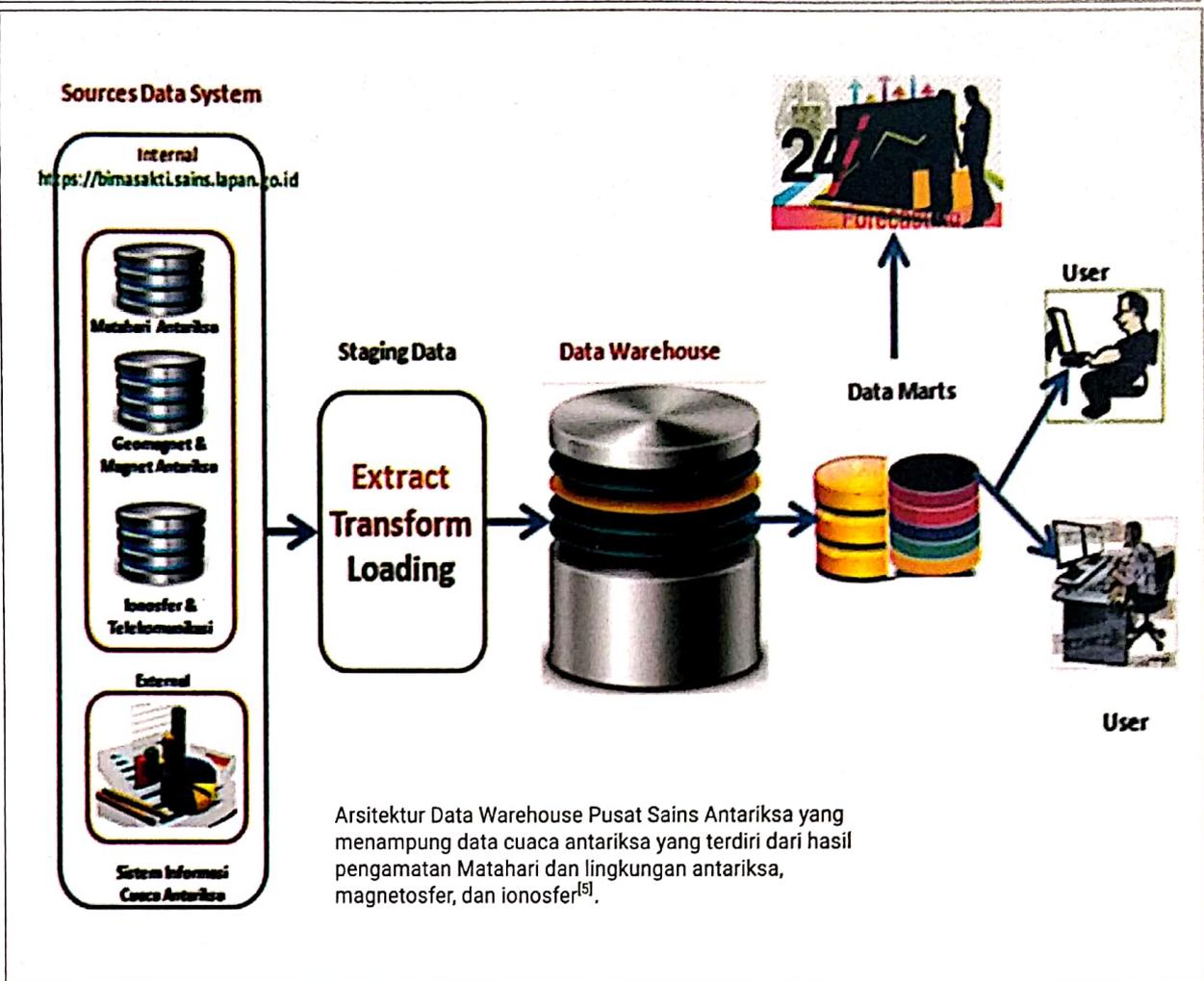
**Rentang Waktu**

seluruh data pada data warehouse dapat dikatakan akurat atau valid pada rentang waktu tertentu. Untuk melihat interval waktu yang digunakan dalam mengukur keakuratan suatu data warehouse dapat digunakan cara yang paling sederhana, yaitu dengan menyajikan data warehouse pada rentang waktu tertentu, misalnya antara 5 sampai 10 tahun ke depan. Cara yang kedua, dengan menggunakan variasi perbedaan waktu yang disajikan dalam data warehouse secara implisit, eksplisit dan snapshot.

**Non-Volatile**

maksudnya data pada data warehouse tidak diperbaharui secara real time tetapi diperbaharui dari sistem operasional secara reguler. Data yang baru selalu ditambahkan sebagai suplemen bagi basis data itu sendiri sebagai sebuah perubahan. Basis data tersebut secara kontinu menyerap data baru kemudian secara bertahap disatukan dengan data sebelumnya. Berbeda dengan basis data operasional yang dapat melakukan pembaharuan, sisipan dan penghapusan terhadap data. Pada data





warehouse hanya ada dua kegiatan memanipulasi data yaitu *loading* data (mengambil data) dan akses data. Mengakses *data warehouse* seperti melakukan *query* atau menampilkan laporan yang dibutuhkan.

**Ringkas**, *data warehouse* menyediakan ringkasan data operasional yang sederhana dan mudah dipahami oleh pihak manajemen.

**Tidak ternormalisasi**, data dalam sebuah *data warehouse* biasanya tidak dinormalisasi sehingga basis datanya sangat memungkinkan terjadinya redundansi.

**Data dari berbagai sumber**, data yang diperoleh berasal dari berbagai sumber internal dan eksternal.

### Kegunaan Data Warehouse

Kegunaan *data warehouse* mempermudah pembuatan aplikasi DSS (*Decision Support System*) dan EIS (*Executive Information System*) untuk keperluan:

**Pembuatan laporan**. Dengan menggunakan *query* sederhana data untuk laporan harian, bulanan, tahunan atau jangka waktu lain yang diinginkan dapat diperoleh dengan mudah.

**On-Line Analytical Processing (OLAP)**, dengan adanya *data warehouse*, pengguna akan mudah mendapatkan semua informasi detail dan hasil ringkasan yang dibutuhkan untuk proses analisis. OLAP memanfaatkan konsep data multidimensi, memungkinkan para pengguna menganalisis data

sampai mendetail tanpa mengetikkan perintah SQL. Pada konsep multi dimensi, data dengan fakta yang sama dapat dilihat dengan menggunakan fungsi yang berbeda. Fasilitas lain yang terdapat pada OLAP adalah *roll-up* dan *drill-down*. *Drill-down* sendiri adalah kemampuan untuk melihat detail dari suatu informasi, sedangkan *roll-up* adalah kebalikannya.

**Proses informasi eksekutif**, *data warehouse* dapat membuat ringkasan informasi penting untuk keputusan organisasi, tanpa harus menjelajahi keseluruhan data. Dengan menggunakan *data warehouse* laporan dapat diringkas atau dirinci secara lengkap, sehingga mempermudah proses pengambilan keputusan.