

**UNTAR untuk INDONESIA**

# Enterprise Technologies and Big Data Business Intelligence

## BIG DATA – TK13025

# Online Transaction Processing (OLTP)

- OLTP adalah sistem perangkat lunak yang memproses data berorientasi transaksi.
- Istilah “transaksi online” mengacu pada penyelesaian suatu aktivitas secara real-time dan tidak diproses secara batch.
- Sistem OLTP menyimpan data operasional yang dinormalisasi.
- Data ini adalah sumber umum dari data terstruktur dan berfungsi sebagai masukan untuk banyak proses analitik.
- Hasil analisis Big Data dapat digunakan untuk menambah data OLTP yang disimpan dalam basis data relasional.



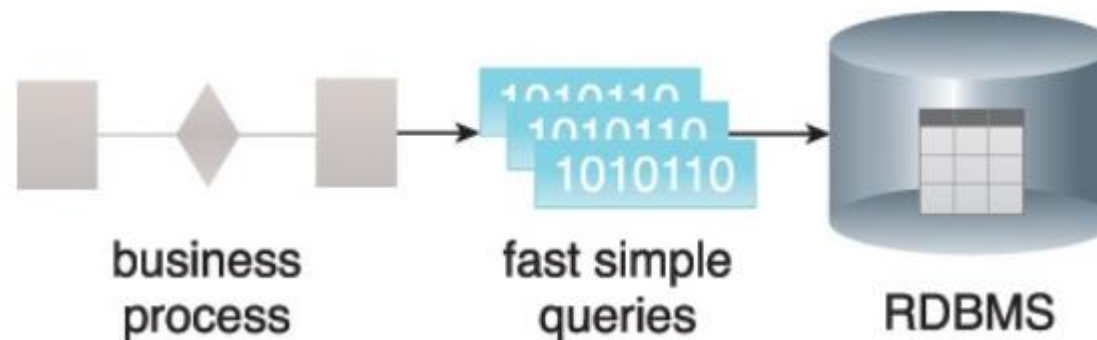
**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

# Online Transaction Processing (OLTP)

- Sistem OLTP, misalnya sistem point of sale, menjalankan proses bisnis untuk mendukung operasi perusahaan.
- Sistem OLTP melakukan operasi basis data sederhana untuk menyediakan waktu respons sub-detik.
- Kueri yang didukung oleh sistem OLTP terdiri dari operasi insert, delete dan update sederhana dengan waktu respons sub-detik.
- Contohnya termasuk sistem reservasi tiket, perbankan dan sistem point of sale.



**UN IAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

# Online Analytical Processing (OLAP)

- Sistem pemrosesan analitik online (OLAP) digunakan untuk memproses kueri analisis data.
- OLAP merupakan bagian integral dari business intelligence, penambangan data, dan proses pembelajaran mesin.
- Hal tersebut relevan dengan Big Data karena dapat berfungsi sebagai sumber data serta sink data yang mampu menerima data.
- Mereka digunakan dalam analitik diagnostik, prediktif dan preskriptif.
- Sistem OLAP melakukan kueri kompleks yang berjalan lama terhadap database multidimensi yang strukturnya dioptimalkan untuk melakukan analitik tingkat lanjut.



**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

# Online Analytical Processing (OLAP)

- Sistem OLAP menyimpan data historis yang dikumpulkan dan didenormalisasi untuk mendukung kemampuan pelaporan yang cepat.
- Sistem OLAP menggunakan database yang menyimpan data historis dalam struktur multidimensi dan dapat menjawab pertanyaan kompleks berdasarkan hubungan antara berbagai aspek data.
- Gambar di bawah ini menunjukkan sistem OLAP menggunakan database multidimensi.



# Extract Transform Load (ETL)

- Extract Transform Load (ETL) adalah proses memuat data dari sistem sumber ke sistem target.
- Sistem sumber dapat berupa database, flat file, atau aplikasi.
- Sistem target dapat berupa database atau sistem penyimpanan lainnya.
- ETL mewakili operasi utama di mana data warehouses diumpankan.
- Solusi Big Data mencakup set fitur ETL untuk mengonversi data dari berbagai jenis.



**UNTAR**  
Universitas Tarumanagara

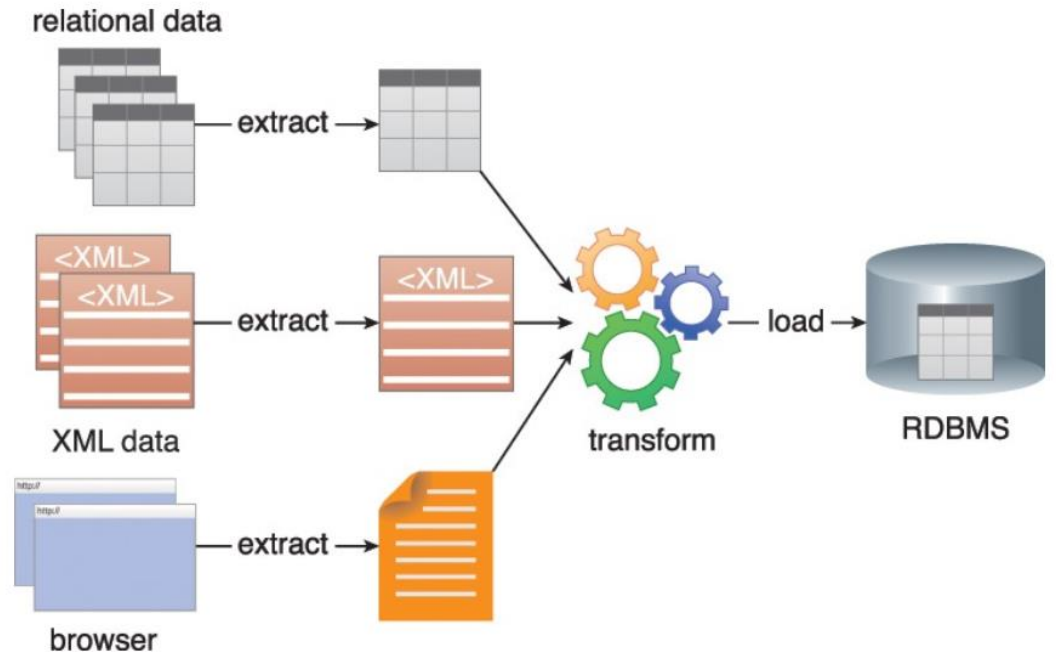


**UNTAR untuk INDONESIA**



# Extract Transform Load (ETL)

- Data yang diperlukan diperoleh atau diekstraksi dari sumbernya, setelah itu ekstrak dimodifikasi atau ditransformasikan dengan penerapan aturan.
- Data dimasukkan atau dimuat ke dalam sistem target.
- Proses ETL dapat mengekstrak data dari berbagai sumber dan mengubahnya untuk dimuat ke dalam sistem target tunggal.



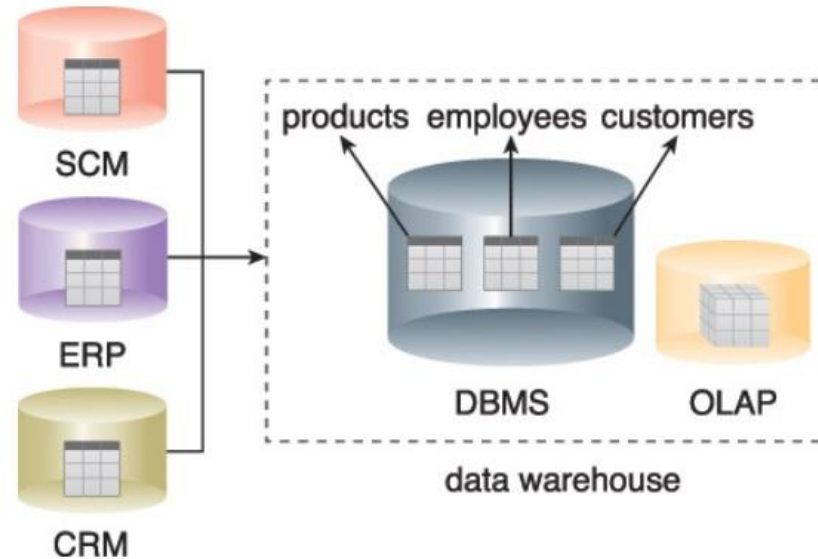
**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

# Gudang Data (Data Warehouses)

- Sebuah gudang data (*data warehouse*) adalah pusat, repositori perusahaan
- Gudang data banyak digunakan oleh BI untuk menjalankan berbagai kueri analitik, dan mereka biasanya berinteraksi dengan sistem OLAP untuk mendukung kueri analitik multidimensi-lebar yang terdiri dari data historis dan saat ini.
- *Batch* memuat data secara berkala ke dalam gudang data dari sistem operasional seperti ERP, CRM, dan SCM.





# Gudang Data (Data Warehouses)

- Data yang berkaitan dengan beberapa entitas bisnis dari sistem operasional yang berbeda secara berkala diekstraksi, divalidasi, diubah dan dikonsolidasikan ke dalam satu database denormalisasi.
- Dengan impor data berkala dari seluruh perusahaan, jumlah data yang terkandung dalam gudang data tertentu akan terus meningkat.
- Hal ini menyebabkan waktu respons kueri yang lebih lambat untuk tugas analisis data.
- Untuk mengatasi kekurangan ini, gudang data biasanya berisi database yang dioptimalkan, yang disebut database analitis, untuk menangani tugas pelaporan dan analisis data.
- Database analitik dapat eksis sebagai DBMS terpisah, seperti dalam kasus database OLAP.



**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

# Data Marts

- Data mart adalah bagian dari data yang disimpan di gudang data yang biasanya dimiliki oleh departemen, divisi, atau lini bisnis tertentu.
- Data warehouses dapat memiliki beberapa data mart.
- Data seluruh perusahaan dikumpulkan dan entitas bisnis kemudian diekstraksi.
- Entitas khusus domain disimpan ke dalam gudang data melalui proses ETL.



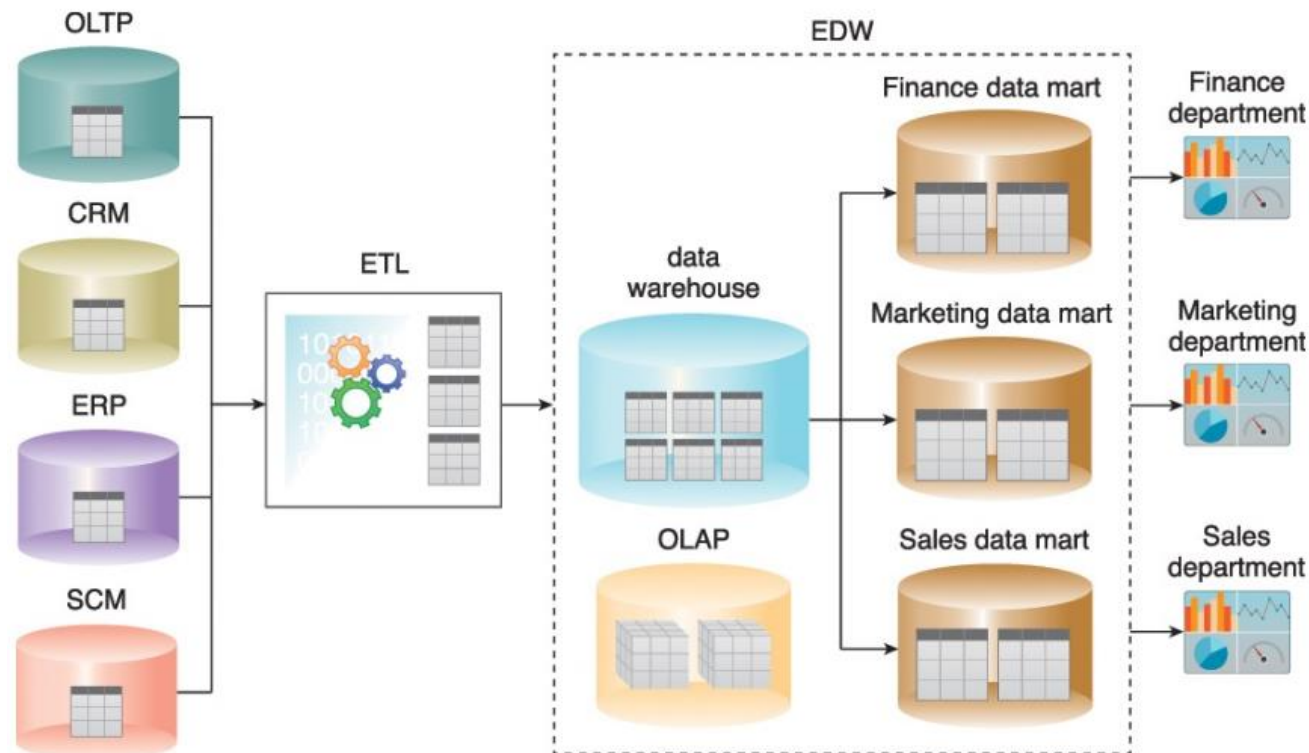
**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

# Data Marts

- Versi tunggal "kebenaran" data warehouse didasarkan pada data yang dibersihkan, yang merupakan prasyarat untuk laporan yang akurat dan bebas kesalahan, sesuai dengan output yang ditampilkan di sebelah kanan.



# BI tradisional

- BI tradisional terutama menggunakan analitik deskriptif dan diagnostik untuk memberikan informasi tentang peristiwa historis dan terkini.
- Ini bukan "cerdas" karena hanya memberikan jawaban atas pertanyaan yang dirumuskan dengan benar.
- Merumuskan pertanyaan dengan benar membutuhkan pemahaman tentang masalah dan isu bisnis dan data itu sendiri.
- BI melaporkan KPI yang berbeda melalui:
  - ✓ laporan ad-hoc
  - ✓ dashboards



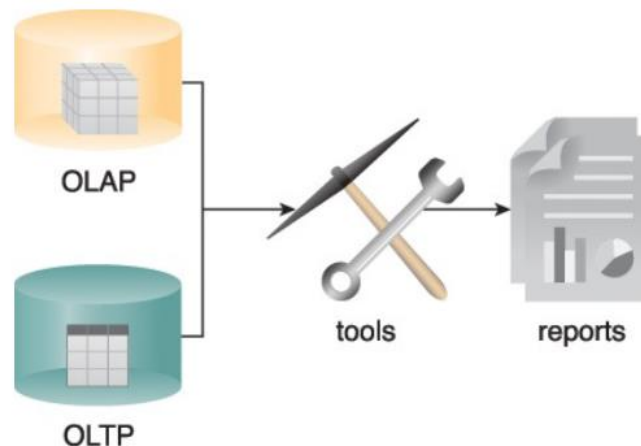
**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

# Ad-hoc Reports

- Pelaporan ad-hoc adalah proses yang melibatkan pemrosesan data secara manual untuk menghasilkan laporan yang dibuat khusus.
- Fokus laporan ad-hoc biasanya pada area bisnis tertentu, seperti pemasaran atau manajemen rantai pasokannya.
- Laporan ubahsuaian yang dihasilkan terperinci dan sering bersifat tabular.
- Sumber data OLAP dan OLTP dapat digunakan oleh alat BI untuk pelaporan dan dasbor ad-hoc.



# Dashboards

- Dasbor memberikan pandangan menyeluruh tentang area bisnis utama.
- Informasi yang ditampilkan di dasbor dihasilkan secara berkala dalam *real-time* atau *near real-time*.
- Penyajian data di dasbor bersifat grafis, menggunakan *bar charts*, *pie charts*, dan *gauges*.



**UNTAR**  
Universitas Tarumanagara

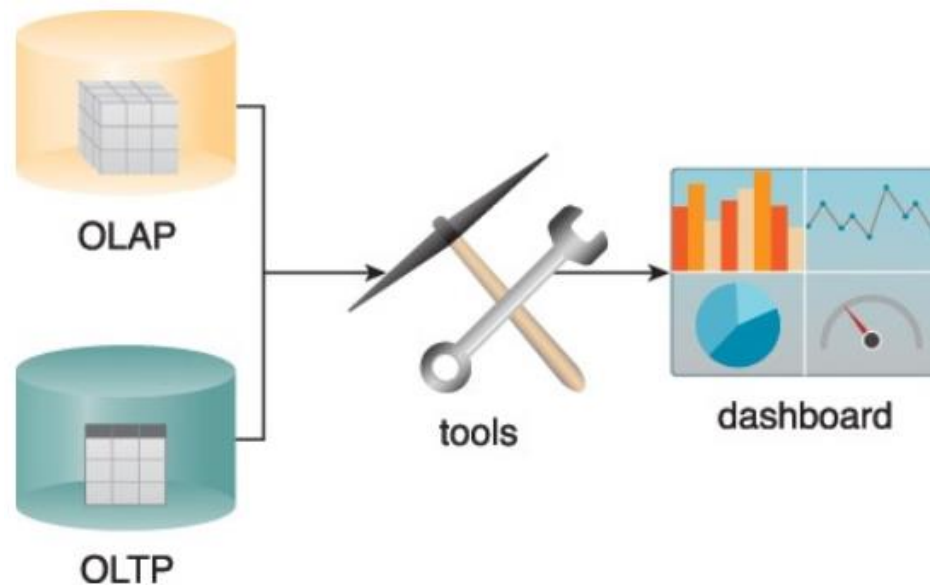


**UNTAR untuk INDONESIA**



# Dashboards

- BI menggunakan OLAP dan OLTP untuk menampilkan informasi di dasbor.



**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

# Dashboards

- Data warehouses dan data mart berisi informasi yang terkonsolidasi dan tervalidasi tentang entitas bisnis di seluruh perusahaan.
- BI tradisional tidak dapat berfungsi secara efektif tanpa data mart karena berisi data yang dioptimalkan dan dipisahkan yang diperlukan BI untuk tujuan pelaporan.
- Tanpa data mart, data perlu diekstraksi dari gudang data melalui proses ETL secara ad-hoc setiap kali kueri perlu dijalankan.
- Hal ini meningkatkan waktu dan upaya untuk mengeksekusi kueri dan menghasilkan laporan.



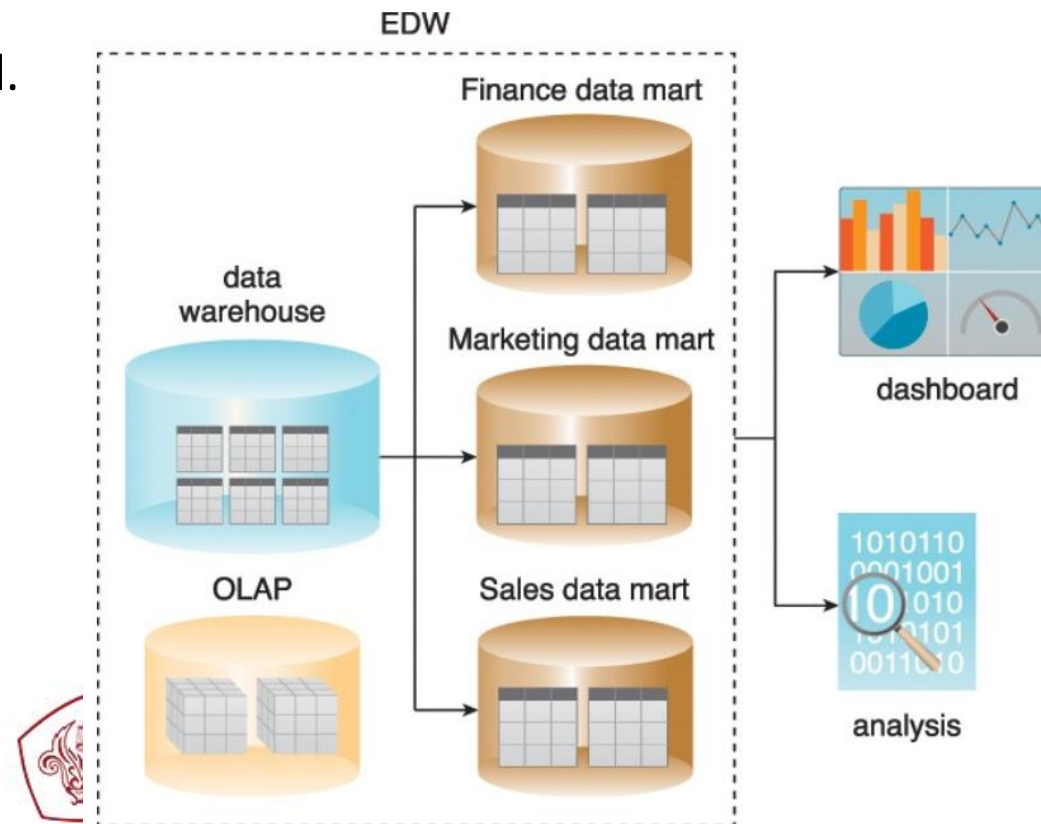
**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

# Dashboards

- BI tradisional menggunakan gudang data dan data mart untuk pelaporan dan analisis data karena mereka memungkinkan kueri analisis data yang kompleks dengan beberapa gabungan dan agregasi untuk diterbitkan.
- Gambar di bawah ini contoh BI tradisional.



# Big Data BI

- Big Data BI dibangun di atas BI tradisional dengan bertindak berdasarkan data seluruh perusahaan yang telah dibersihkan dan dikonsolidasikan di data warehouse dan menggabungkannya dengan sumber data semi-terstruktur dan tidak terstruktur.
- Ini terdiri dari analitik prediktif dan preskriptif untuk memfasilitasi pengembangan pemahaman kinerja bisnis di seluruh perusahaan.
- Sementara analisis BI tradisional umumnya berfokus pada proses bisnis individu, analisis BI Big Data berfokus pada beberapa proses bisnis secara bersamaan.
- Ini membantu mengungkapkan pola dan anomali di seluruh lingkup yang lebih luas di dalam perusahaan.
- Ini juga mengarah pada penemuan data dengan mengidentifikasi wawasan dan informasi yang mungkin sebelumnya tidak ada atau tidak diketahui.



**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

# Big Data BI

- Big Data BI memerlukan analisis data tidak terstruktur, semi terstruktur, dan terstruktur yang berada di data warehouse (gudang data) perusahaan.
- Ini memerlukan “next-generation” data warehouse yang menggunakan fitur dan teknologi baru untuk menyimpan data bersih yang berasal dari berbagai sumber dalam satu format data seragam.
- Penggabungan gudang data tradisional dengan teknologi baru ini menghasilkan gudang data hibrid.
- Gudang ini bertindak sebagai gudang seragam dan pusat dari data terstruktur, semi-terstruktur, dan tidak terstruktur yang dapat menyediakan alat Big Data BI dengan semua data yang diperlukan.
- Ini menghilangkan kebutuhan alat Big Data BI untuk terhubung ke beberapa sumber data untuk mengambil atau mengakses data.

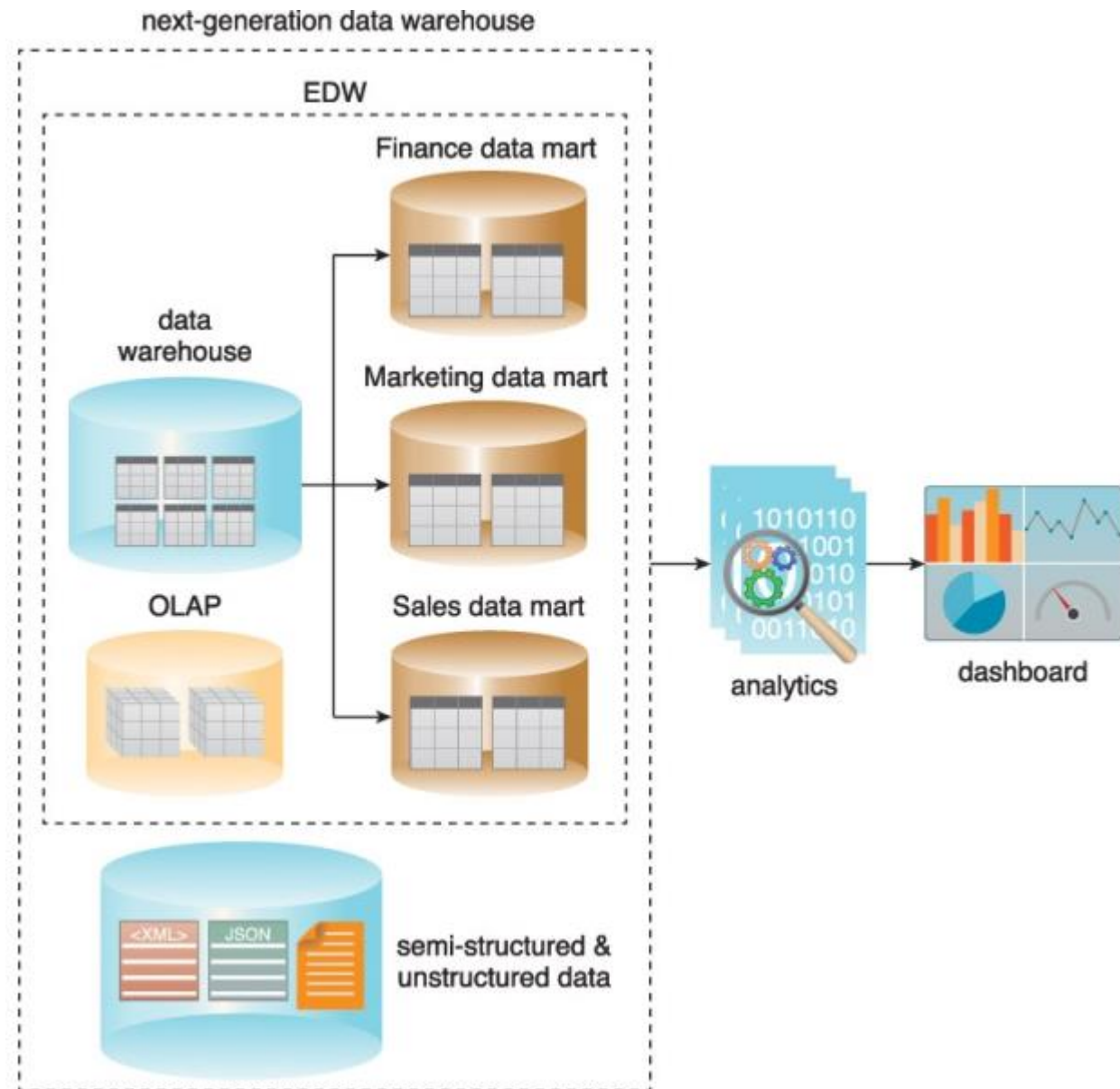


**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

Next-generation data warehouse menetapkan lapisan akses data standar di berbagai sumber data.





# Traditional Data Visualization

- Visualisasi data adalah teknik di mana hasil analisis dikomunikasikan secara grafis menggunakan elemen seperti bagan (charts), peta (map), kisi data (data grids), infografis (infographics), dan peringatan (alerts).
- Merepresentasikan data secara grafis dapat mempermudah untuk memahami laporan, melihat tren, dan mengidentifikasi pola.
- Visualisasi data tradisional sebagian besar menyediakan bagan dan grafik statis dalam laporan dan dasbor, sedangkan alat visualisasi data kontemporer bersifat interaktif dan dapat memberikan tampilan data yang diringkas dan terperinci.
- Hal ini dirancang untuk membantu orang-orang yang tidak memiliki keterampilan statistik dan/atau matematika untuk lebih memahami hasil analisis tanpa harus menggunakan spreadsheet.
- Alat visualisasi data tradisional meminta data dari database relasional, sistem OLAP, gudang data, dan spreadsheet untuk menyajikan hasil analisis deskriptif dan diagnostik.



**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

# Visualisasi Data untuk Big Data

- Solusi Big Data memerlukan alat visualisasi data yang dapat terhubung dengan mulus ke sumber data terstruktur, semi-terstruktur, dan tidak terstruktur dan selanjutnya mampu menangani jutaan catatan data.
- Alat visualisasi data untuk solusi Big Data umumnya menggunakan teknologi analitik dalam memori yang mengurangi latensi yang biasanya dikaitkan dengan alat visualisasi data tradisional berbasis disk.
- Alat visualisasi data canggih untuk solusi Big Data menggabungkan analisis data prediktif dan preskriptif serta fitur transformasi data.
- Alat-alat ini menghilangkan kebutuhan akan metode pra-pemrosesan data, seperti ETL.



**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

# Visualisasi Data untuk Big Data

- Alat ini juga menyediakan kemampuan untuk terhubung langsung ke sumber data terstruktur, semi-terstruktur, dan tidak terstruktur.
- Sebagai bagian dari solusi Big Data, alat visualisasi data tingkat lanjut dapat menggabungkan data terstruktur dan tidak terstruktur yang disimpan dalam memori untuk akses data yang cepat.
- Kueri dan rumus statistik kemudian dapat diterapkan sebagai bagian dari berbagai tugas analisis data untuk menampilkan data dalam format yang mudah digunakan, seperti di dasbor.



**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**

# Visualisasi Data untuk Big Data

- Fitur umum alat visualisasi yang digunakan dalam Big Data:
  - **Aggregation** – memberikan pandangan holistik dan ringkasan data di berbagai konteks.
  - **Drill-down** – memungkinkan tampilan mendetail dari data yang diinginkan dengan berfokus pada subset data dari tampilan ringkasan.
  - **Filtering** – membantu fokus pada kumpulan data tertentu dengan menyaring data yang tidak menarik secara langsung.
  - **Roll-up** – mengelompokkan data di beberapa kategori untuk menampilkan subtotal dan total.
  - **What-if analysis** – memungkinkan beberapa hasil divisualisasikan dengan memungkinkan faktor-faktor terkait diubah secara dinamis.



**UNTAR**  
Universitas Tarumanagara



**UNTAR untuk INDONESIA**