



Virtual Internship Experience

Data Modeling Best Practices

Data Engineer VIX

A. Data Modeling Technique

Pemodelan data muncul pada 1960-an seiring dengan meningkatnya penggunaan database. Ini memungkinkan organisasi untuk membawa konsistensi, pengulangan dan pengembangan disiplin untuk pemrosesan dan manajemen data. Berikut ini adalah beberapa teknik data modeling yang paling banyak digunakan.

1. Hierarchical Data Modeling

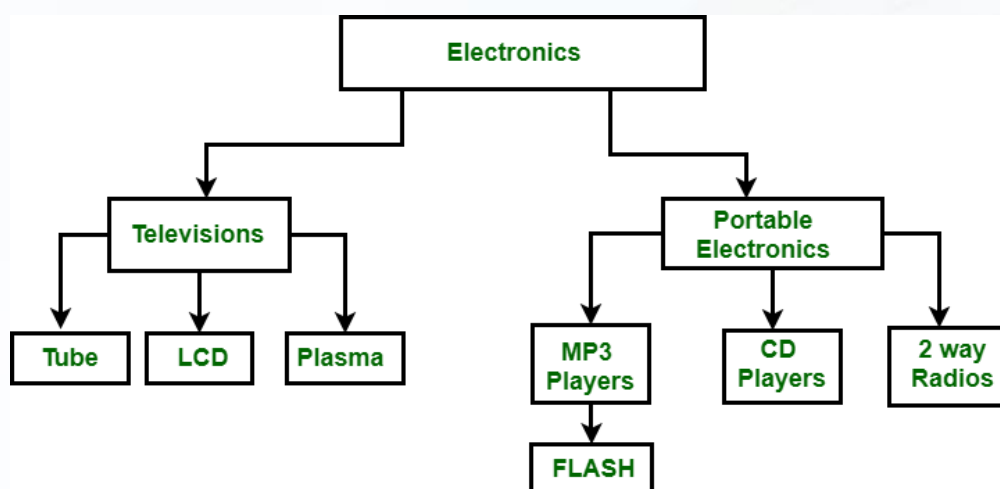


Fig. 3. Contoh Hierarchical Data Model

<https://www.geeksforgeeks.org/difference-between-hierarchical-and-network-data-model/>

Hierarchical Data Model mengatur data dalam susunan seperti pohon dari catatan induk dan anak. Catatan anak hanya dapat memiliki satu orang tua, menjadikannya metode pemodelan *one-to-many*. Pendekatan Hierarchical yang paling terkenal adalah Information Management System (IMS) IBM. IMS masih tersedia dan digunakan oleh banyak organisasi walaupun telah muncul metode serupa yang lebih baru. Metode hierarchical juga digunakan saat ini dalam XML, yang secara resmi dikenal sebagai Extensible Markup Language.

2. Entity Relationship Data Modeling

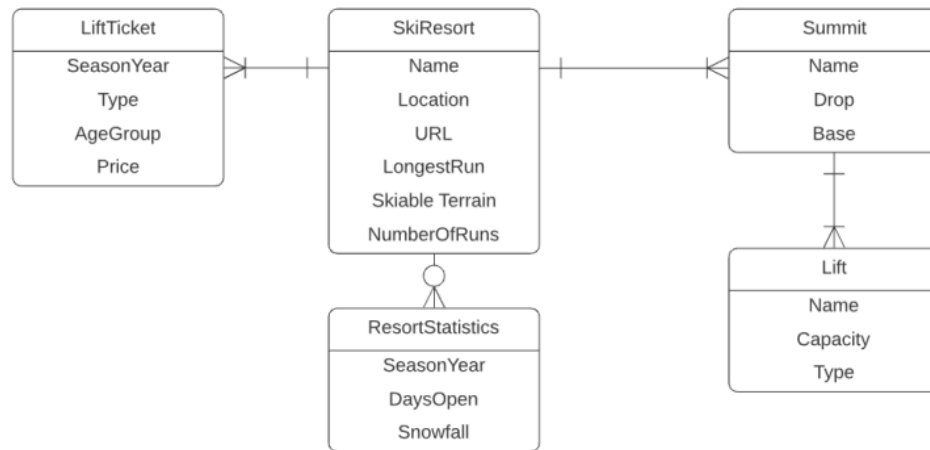


Fig. 4. Contoh Entity Relationship Data Model

<https://why-change.com/2020/11/15/how-to-create-a-conceptual-data-model/>

Variasi dari model relasional yang juga dapat digunakan dengan tipe database lain, model entity-relationship diagram (ERD) yang memetakan entitas secara visual, atributnya, dan hubungan antara entitas yang berbeda. Misalnya, atribut entitas data karyawan dapat mencakup nama belakang, nama depan, tahun bekerja, dan data relevan lainnya. Model ER memberikan pendekatan yang efisien untuk pengambilan data dan proses pembaruan, membuatnya sangat cocok untuk aplikasi pemrosesan transaksi.

3. Relational Data Modeling

Relational Data Modeling dibuat sebagai alternatif yang lebih fleksibel daripada model hierarkis. Pertama kali dijelaskan dalam makalah teknis tahun 1970 oleh peneliti IBM Edgar F. Codd, Relational Data Modeling memetakan hubungan antara elemen data yang disimpan dalam tabel berbeda. Relational Data Modeling mengatur panggung untuk pengembangan basis data relasional, dan penggunaannya secara luas menjadikannya teknik pemodelan data yang dominan pada pertengahan 1990-an.

4. Network Data Modeling

Network Data Modeling merupakan salah satu pilihan pemodelan data yang populer di database mainframe yang tidak banyak digunakan sekarang.

Network Data Modeling memperluas Hierarchical Data Modeling dengan memungkinkan anak dihubungkan ke satu atau lebih induk.

5. Object Oriented Data Modeling

Pendekatan Object Oriented Data Modeling mirip dengan metode ER dalam cara mewakili data, atribut dan hubungan, tetapi abstrak entitas menjadi objek. Objek berbeda yang memiliki atribut dan perilaku yang sama dapat dikelompokkan ke dalam class, dan class baru dapat mewarisi atribut dan perilaku yang sudah ada. Tetapi Object Oriented Data Modeling tetap menjadi teknologi khusus untuk aplikasi tertentu, yang membatasi penggunaan pemodelan berorientasi objek.

B. Tahap Data Modeling

Setelah kalian memilih diagram berdasarkan kategori Data Model yang tersedia sebelumnya, untuk melakukan sebuah data model ada tiga tahap bentuk data model yang dirancang, mulai dari bentuk *Conceptual Model*, *Logical Model*, hingga *Physical Model*.

1. Conceptual Model

Pada tahap ini, data diidentifikasi sesuai dengan data yang dibutuhkan atau transaksi bisnis yang terjadi pada perusahaan. Data akan diidentifikasi dalam bentuk model yang belum teridentifikasi hubungannya antara satu data dengan data lain.

Tujuan dari dibentuknya model ini adalah memberikan gambaran secara umum mengenai data yang dibutuhkan dan dihasilkan dari proses bisnis yang terjadi pada perusahaan tanpa menjelaskan secara detail karakteristik flow data tersebut. Selain mendefinisikan gambaran umum, tujuan Conceptual Model antara lain :

- Mendefinisikan entitas bisnis utama
- Mendefinisikan hubungan antara entitas
- Menerapkan terminologi dalam data yang konsisten
- Mendukung pengetahuan bisnis bagi perusahaan

Untuk mencapai tujuan tersebut dan mendapatkan data yang diinginkan, campur tangan dari stakeholder terkait adalah salah satu faktor penting. Stakeholder yang biasanya terlibat antara lain adalah : Business Executive, CEO, Business Analyst, atau pihak yang memiliki andil serupa.

2. Logical Model

Setelah mendefinisikan data yang digunakan secara umum pada tahap Conceptual Model, tahap selanjutnya adalah mendefinisikan data lebih spesifik, khususnya sisi teknisnya, tahap inilah yang disebut sebagai tahap Logical Model. Logical Model, sekilas tampak seperti tahap Conceptual Model, hanya saja pada tahap ini atribut data akan lebih di spesifikkan walaupun belum menentukan secara eksplisit. Selain itu, hubungan antar entitas yang telah didefinisikan pada tahap Conceptual Model akan dijelaskan secara lebih spesifik pada tahap ini,

Untuk mempermudah, berikut adalah karakteristik dari Logical Model :

- Logical Model dapat menggambarkan kebutuhan data untuk setiap entitas. Namun, dirancang untuk berintegrasi mulus dengan Logical Model lainnya jika diperlukan.
- Logical Model dapat dikembangkan dan dirancang secara independen dari sistem manajemen basis data. Jenis sistem manajemen basis data tidak terlalu mempengaruhinya.
- Atribut data berisi tipe data yang didefinisikan secara tepat.
- Dalam Logical Model, tidak ada *primary key* atau *secondary key* yang ditentukan.
- Logical Model dapat berbentuk seperti representasi grafis dari kebutuhan informasi dari area bisnis. Namun bukan database atau sistem manajemen database itu sendiri.
- Logical Model tidak tergantung pada perangkat penyimpanan data apa pun.
- Logical Model harus dirancang untuk tidak bergantung pada teknologi, agar tidak terpengaruh oleh perubahan teknologi yang cepat.

3. Physical Model

Tahap akhir adalah mengimplementasikan Logical Model pada Database Management System (DBMS) yang digunakan. Tentunya tiap DBMS memiliki kapasitas, ketentuan, dan arsitektur yang berbeda beda, disinilah tahap Physical Model diterapkan. Pada tahap ini struktur database atau sistem file yang akan digunakan untuk menyimpan dan mengelola data didefinisikan. Termasuk tabel, kolom, *fields*, *indexes*, *constraints*, *triggers*, hardware, data access, dan elemen DBMS lainnya. Physical Model digunakan untuk membuat desain dan menghasilkan skema basis data.

Referensi

<https://www.tibco.com/reference-center/what-is-a-logical-data-model>
<https://why-change.com/2020/11/15/how-to-create-a-conceptual-data-model/>
<https://www.techtarget.com/searchdatamanagement/definition/data-modeling>
<https://whatagraph.com/blog/articles/data-modeling-techniques>
<https://www.tutorialcup.com/dbms/physical-data-models.htm>
<https://analystanswers.com/what-is-data-modeling-a-visual-introduction-with-examples/>