Data Modeling Introduction

SkyEss Techno Solutions Pvt. Ltd.

What is The Need For Data Modeling?

SkyEss Techno Solutions Pvt. Ltd.

- Data Modeling is A Process of Analysis And Design, But Not A Formula To Follow With Straight Forward Steps.
- Data Modeling Refers To The Process of Following A Series of Steps To Fix The Business Data in A Structured And Organized Manner.
- Data Modeling is The Key Component For Successful Implementation of Data Storage OR Data Management Environments With Quality, Consistency, Accuracy And Integrity.
- Data Modeling Provides An Advanced Information Regarding The Anomalies That Can Arise in The System During The Business Process Transactions Prior To The Failures.
- Data Modeling Acts As A Warning Signal To The System That is Being Developed For The Areas of Failure That May Be Encountered in The Process of System Development.
- Data Modeling Makes The Over All System Consistent And Gives A Synchronized Look
  - From Client To The Analyst
  - From Analyst To The Architect
  - From Architect To The Administrator
  - From Administrator To The Developer
  - From The Developer To The Integration Engineer
  - From Integration Engineer To The Testing Engineer
  - From Testing Engineer To The Deployment Engineer
  - From Deployment Engineer To Maintenance Engineer

What Finally A Data Model Serves?

SkyEss Techno Solutions Pvt. Ltd.

- Serves As A True Representation of Some Aspects OR All of The Aspects That Exist in The Real World.
- A Model Enables Crystal Clear Communication About The Real World Aspects Narrated By The Analyst OR Architect.
- Serves As A Blueprint To Shape And Construct The Proposed Structures in The Real World With Balanced Thoughts Among The Team Members.

Some Questions We Should Answer Definitely

SkyEss Techno Solutions Pvt. Ltd.

Is Data Modeling A Technique OR Science?

- Data Modeling is A Technique That is Followed By Experience With Participation in The System OR Domain.

Is Data Model A Tool OR Software?

- A Data Model is A Critical Tool For Establishing Communication Among The Users Participating in The Project Development.

Data Model Actually Serves Who?

- A Data Model Serves The Users Who Are Part of The System of Development And Usage Starting From The Business User To The Core Developer.

How We Should Treat A Data Model, A Blue Print OR Frame Work?

- A Data Model Serves Both As A Blueprint And Frame Work For Design And Developing The Database System For The
  - Component Developers
  - Module Integrators
  - Deployment Engineers
  - Testing And Quality Analysts
  - Installation And Configuration Engineers
  - System And Database Administrators
  - Debugging Specialists
  - Project Managers

Where We Should Place A Data Model in Application Development?

- The Data Model is Placed Between Requirements Definition Phase And Database System Design And Implementation.

What Does Data Modeling Techniques Explore?

- Data Modeling is A Technique That Should Be Used For Exploring The Data Structures Needed To Support Business Operations Existing in A Business Organization.

What is Expected To Be Recorded By A Data Model?

- A Data Model Must Record And Indicate The
  - Content of The Business Data
  - Shape And Structure of The Business Process
  - Size of The Business Data
  - Rules To Be Applied Upon The Data Elements
  - Scope of The Usage of The Data Element
  - Cardinality And State of The Data Element
- All The Above Components Should Be Used Throughout The Scope of The Various Business Processes of The Organization.

Data Model is A Conceptual OR Physical Representation?

- Data Model is A Conceptual Representation OR Replica of The Data Structures Required in The Actual Database System.

What Are The Main Focussed Components of The Data Model?

- The Main Focused Components of The Data Model Are
  - What Data is Required?
  - How The Data Should Be Organized?

Does The Data Model Reflects The Operations of The Business?

- Data Model in Reality Does Not Necessarily Reflect The Operations Expected To Be Performed on The Data.

Can We Apply The Data Modeling Techniques At Various Levels?

- Yes, Data Modeling Techniques Can Be Applied To Represent The Information Requirements At Various Levels of The Business Environment.

Is Data Model is Independent of Hardware And Software?

- Yes, Data Model is Independent of Hardware And Software Techniques That Are Applied, At The Highest Level of Conceptual Model, Hence it is Considered Generic in Nature.

Does The Data Model Represents A Database Approach Type?

- No, A Data Model Does Not Represents The Type of Database Approach.
- The Data Model is Designed Independent of
  - An Object-relational Database
  - A Relational Database
  - A Hierarchical Database
  - A Network Database

Are There Any Levels That Follow Down The Data Model?

- Yes, The Next Level Followed Down The Line in The Conceptual Data Model is Logical Model, Which Relates To The Particular Type of Database To Be Followed, Whether it is Should Be Relational, Hierarchical, Network OR Object Oriented.
- The Next Level of The Logical Model is A Physical Model Relating To The Particular Database Management System Software That Can Be Considered.

The Final Big Picture of Data Model Existence

SkyEss Techno Solutions Pvt. Ltd.

Real World
Information System

Requirements
Collection And
Definition

Data Model

Database System

Communication
Tool

Database Blue
Print

End User Cluster

Software Developers

What Happens Once The Data is Structured OR Modeled?

SkyEss Techno Solutions Pvt. Ltd.

- The Business Data Once Structured is Generally Transformed And Implemented into A Database Management System.
- The Main Motto Behind Developing Database Management Systems is to Manage Vast Amounts of Both Structured And Unstructured Data.

Structured Data

- It is Any Data That is Referred in The Business Environment of The Client in Standard Format Collected in The Form of
  - Alphabets
  - Numbers
  - Special Characters

Unstructured Data

- It is Any Data That is Referred in The Business Environment of The Client in Standard Format Collected in The Form of
  - Word Processing Documents
  - e-mail Messages
  - Pictures OR Images
  - Digital Video
  - Audio Files
- Data Modeling Theory Provides Standards And Principles For Organizing And Managing The Structured Data, But Not Unstructured Data.
- The Data Model That is Developed on The Bias of Structured Data is Developed And Implemented Using Relational Database Management System Software.

What is The Focus Area of The Data Modeling?

SkyEss Techno Solutions Pvt. Ltd.

- When We Are Developing A Software, The Process of The Software Project Will Concentrates on The Design of Data Models Called As "Conceptual Data Model" OR "Logical Data Model".
- Once The Project is On The Process of Development, The Data Model is Usually Referenced As "Physical Data Model".
- "Logical Data Model" And "Physical Data Model" Are Two Different Instances in The Process of Software Development, And Represent Two Unique Ways of Describing The Data Model.

Focus of The Logical Model

- Logical Model Focuses on The Basic Features of The Business Process And Process Integrations Among The Different Divisions of The Business OR The Enterprise.
- Logical Model is Always Designed Outside The Boundaries of Any Specific Implementation of The Business OR Enterprise.

Focus of The Physical Model

- Physical Model Focuses on The Physical Features of The Business Process And The Enterprise Implementation in Real Time.
- Physical Model Concentrates on The Specification of The Database Management System Software That Has The Ability of Hosting The Data Model's Features.
- Physical Model is More Specific on The Technology And The Architecture That Should Be Applied To Reach Execution.

What is Actually Managed By Data Model?

SkyEss Techno Solutions Pvt. Ltd.

- Data Model Actually Describes The Structure of The Data Within The Confined Boundaries of A Particular Domain.
- Data Model Actually Specifies A Special "Grammatical Rules" To Be Followed For The Business Domains And its Own Private Artificial Language.

Representations of The Data Models

- Data Model Represents Different Entity Classes.
- Data Model Represents The Specifics Behind The Information
- The Relationship Among The Differing Entities And Attributes.

Note

- The Data That is Represented By The Data Model May Be Represented in A Different Fashion on The Actual Computer System.

What A Conceptual Data Model Represents?

- A Conceptual Data Model Describes The Semantics of A Particular Subject Area in The Business Process.
- The Conceptual Data Model in Reality is Basically A Collection of Assertions About The Type of Information That is Being Used By A Company.
- In Conceptual Model The Entity Classes Are Named Using Natural Language Rather Than Using Technical Jargon.
- In Conceptual Model We Do Not Follow The Concrete Assertions About The Subject Area in The Business.
- The Conceptual Model Speaks Only About The Natural Relations That Exist in The System For The Basic Business Process.

Types of Data Modeling Techniques

- IDEF1X(Integration Definition For Information Modeling)
- IE(Information Engineering Modeling)
- Entity-Relationship Modeling
- Bachman Diagrams
- Richard Barker's Notation(Used By Oracle Corporation)
- Object Role Modeling (Nijssen's Information Analysis Method)
- Business Rules Approach
- Object-Relationship Modeling
- RM/T(Relational Model/Tasmania)
- Chen's Notation
- Data Vault Modeling
- Extended Backus–Naur form
- Relational Model/Tasmania

Data Modeling Tools

- Oracle SQL Developer
- Datanamic DeZign
- Toad Data Modeler
- CA ERwin
- ARIS
- Oracle Designer
- Visio Microsoft
- SILVERRUN
- GNU Ferret
- Mogwai ER-Designer
- MySQL Workbench
- PowerDesigner
- ER/Studio
- DbSchema
- Database Visual Architect
- Smart Draw

Let Us Catch The Concept of Data Life Cycle

SkyEss Techno Solutions Pvt. Ltd.

Data Life Cycle

- The Different Stages The Data Goes Through in An Organization in Association To The Different Business Processes is Called As Data Life Cycle.

Data Life Cycle Stages

Stage 1

- A Need For Data Arises To Perform The Various Business Processes of An Organization.

Stage 2

- A Determination is Made About Exactly What Data is Needed.

Stage 3

- Gathering of The Identified Data Takes Place From The Business Process of The Organization.

Stage 4

- The Gathered Data From The Business Environment Gets Stored into The Database System.

Stage 5

- Data is Manipulated By Reading The Data From Storage Devices, Combining The Data in Various Desired Ways, And Changing The Data As Per The Business Needs.

Stage 6

- After A Some Time Part of The Data Gets Archived And Stored Elsewhere As Backup.

Stage 7

- After Some Time, Part of The Data Completes its Usefulness, Such Data Gets Deleted From The Database System.

Who Should Perform Data Modeling?

SkyEss Techno Solutions Pvt. Ltd.

- Data Modeling is A Culture of Various Skill Sets Integrated into One Place, Hence Data Modeling Can Be Done By Personalities Who Are Experts in Various Skills.
- Depending on The Size And Complexity of The Database System, One OR More Persons Are Entrusted With The Responsibility of Creating The Data Models.
- Data Modeling Process Demands Different Skills And Training At Various Levels in The Process of The System Design.

Primary Responsibility of Data Modeler

- Data Modelers Should Model And Describe The Real World Process of The Business Domain That is of Interest To The Organization To Achieve The Goals And Purposes of The Business System.

Functions of The Data Modeler

Function 1 → Scanning Current Details

- The Data Modeler Scans And Captures The Details of Current State of The Data System At The Enterprise Level OR Subject Level.
- New Models Are Built By Looking At The Current Data Structures That Have Been Scanned And Recorded With Details.

Function 2 → Designing The Architecture

- The Data Modeler Can Be Considered As An Architect Who is Responsible For Designing The New Data Model.
- Data Modeler Puts Together All The Pieces of The Architecture in An Integrated Frame Following The Data Flow And Rules As Existing in The Real System.

Function 3 → Documenting And Maintaining Meta-Data

- The Data Modeler Can Be Considered Like A Librarian And He The Final Authority For The Custody of The "Data About The Data" of The Organization.
- The Data Modeler is A Tremendous Source of Information About The Data Structures And Elements of The Business System, Current And Proposed in Future.

Function 4 → Providing Advice And Consultation

- Data Modeler is A Personality With In-Depth Knowledge About The Composition of The Data System of An Organization Specific OR All of The Business Processes.
- The Data Modeler is An Expert For Final Consultation And Conclusions Upon The Ambiguities That Are Arising in The Business System in The Process of Development.

The Basic Purpose of A Data Model

- Serve As A Communication Channel With The Users OR Domain Experts.
- Act As A Blueprint And Frame Work For The Proposed Database System For The Organization.

Purpose 1 → Communication Tool

What it Should Contain?

- Sufficient Information To Review its Components
- Should Describe The Various Parts of The Business System With Clarity
- Explain The Different Connections of The Business Data
- Make Users Understand The Ultimate Data Required in The System
- Make End User Clear With System That is Being Built For The Business

What it Should Not Contain?

- The Data Model Should Not Contain Any Complexities in Terms of The Data Structures When Presented To The End Users.
- Data Model Should Not Indicate Any Physical Storage Considerations.
- Should Not Reveal How Data Structures Are Laid Out OR Perceived By Analysts And Programmers.

Final Goal

- The Data Model Must Just Represent A Conceptual Portrayal of The Information Requirements in Human Terms.
- The Data Model Must Be A Business Representation Using A High Level of Ideas on System State And Process Flow.
- Should Have Sufficient Information To Establish Clear Communication With The Domain Experts.

Purpose 2 → Blue Print And Framework

What it Should Contain?

- Should Step Through The Business Components, One By One, And Should Be Useful To Design And Create The Database System.
- Should Include Details of All The Data Structures.
- Should Indicate And Include All of The Relationships.
- Should Represent How Data is Viewed By Analysts And Programmers.
- Should Bear Connections To How Database Vendors View Business Data And Design Their Database Products.
- Should Have Information For Building The Database System And Determine How Data Will Be Stored on Physical Storage Device.
- Should Provide Details For, How Data Will Be Accessed And Used
- Should Provide More Clear Details For How Data is Viewed By Programmers And Analysts.

Let Us Classify The Information Levels

SkyEss Techno Solutions Pvt. Ltd.

Level 1 → Conceptual Level

- Conceptual Level OR Stable Information Level is The Highest Level Which Consists of General Ideas About The Information Content.
- Conceptual Level Contains The Description of Application of Business Domain in Terms of Human Concepts.
- The End Users Can Exactly Understand What is The Business OR Data System That is About To Be Developed.
- The Conceptual Data Model Portrays
  - Base Type Business Objects
  - Constraints on The Business Objects, Their Characteristics
  - Any Derivation Rules
- The Conceptual Data Model is Independent of All Physical Considerations.
- The Conceptual Data Model Hides All Complexities About The Data Structures From The Users Through Levels of Abstraction.
- The Conceptual Data Model Serves As An Excellent Tool For Communication With The Domain Experts OR End Users.

Level 2 → External Level

- External Level Consists of Fragments of The Entire Conceptual Model, Each Fragment Can Be Considered As A Mini-Conceptual Model.
- External Data Model Contains Representation of A Particular Segment of Information Requirements As Applicable To Only One User Group.
- All Individual External Data Models When Aggregated Will Arrive At The Comprehensive Conceptual Data Model For The Entire Organization.
- External Data Model Enables The Database Practitioners To Separate Out The Conceptual Data Model By Individual User Groups And Allocates Data Access Authorizations Appropriately.

Level 3 → Logical Level

- In Logical Level, The Domain Concepts And Their Relationships Are Explored Further.
- Logical Level Accommodates More Details About The Information Content As Identified in The Business System.
- Logical Level Never Contains, Storage And Physical Considerations of The Business Data, And Also Never Contains Any Details of A Specific DBMS Software.
- Logical Model Provides The Representation on The Type of Database Implementation Whether it is
  - Relational Database
  - Hierarchical Database
  - Network Database
  - Object Relational Database
  - Object Oriented Database
  - Web-Enabled Database
- The Type of Database Model That is Perceived Controls The Change of The Content of Objects And The Structural Components in The Model.
- The Data Model Confirmed At The Logical Level Will Be Finally Used in The Ultimate Construction of The Database System.

Level 4 → Internal OR Physical Level

- Internal OR Physical Level Deals With The Implementation of The Database on Secondary Storage Devices.
- This Level Applies The Concept of Considerations For
  - Storage Management of The Data
  - Access Management of The Data
  - Database Performance Management For Operations

SkyEss Techno Solutions Pvt. Ltd.

- In Internal OR Physical Model The Intricate And Complex Details of The Particular Database Are Defined And Discussed.
- The Intricacies And The Limitations of The Particular DBMS Software Are Taken into Account At The Physical Level.
- The Physical Data Model Should Represent The Details of Implementation At The Hardware And Software Levels.
- The Data Model At This Level is Primarily Intended To Be Like A Blueprint OR Frame Work For Implementation, Hence Cannot Be Used As A Means For Communication With The Users.
- The Physical Data Model Represents The Information Requirements in Terms of Different Requirements That Arise Upon
  - Files And Their Architecture
  - The Management Standards of Data Blocks
  - The Technical Specifications of Storage For Data Records
  - The Index Management Standards of The Data Records For Performance And Optimization Management.
  - The Technical Standards To Be Implemented For File Organizations
  - The Composition of Clustering Techniques To Be Followed Upon The Common Data
  - The Security And Sharable Aspects Upon The Database Objects
  - The Concepts of Auto Generation Aspects To Be Followed For Key Values
  - The Type of Constraints And Integrity Management To Be Followed
  - The User And Privilege Management Aspects
  - The Profile Management Issues Essential At The Front End Level
  - The Component Segregation in The Concept of Client Server Architecture

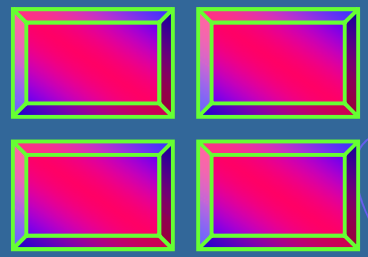
The Final Picture of Data Model Levels

SkyEss Techno Solutions Pvt. Ltd.

Information Levels

Database Models

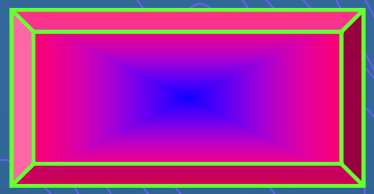
External Level



View of Data System By Individual User Groups From Outside. Used As Communication Tool With Each User Group.

Combine

Conceptual Level



View of Data System By Entire User Community. Used As Communication Tool With Whole Set of All User Group.

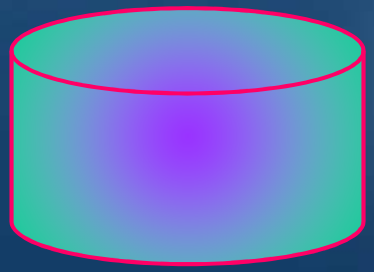
Map Components

Logical Level

View of Data System By According To Specific Database Methodology. Used Moe As A Database Blueprint.

Transform

Physical Level



View of Data System With Details of Hardware/Storage And Specific Database Methodology. Used Exclusively As A Database Blueprint.

Let Us Summarize The Content

SkyEss Techno Solutions Pvt. Ltd.

- The Goal of The Data Model is To Make Sure That The All Data Objects Required By The Database Are Completely And Accurately Represented.
- The Data Model Uses Easily Understood Notations And Natural Language, Hence it Can Be Reviewed And Verified For Correctness By The End-Users.
- The Data Model is Also Detailed Enough To Be Used By The Database Developers To Use As A "Blueprint OR Frame Work" For Building The Physical Database.
- The Information Contained in The Data Model Will Be Used To Define The Relational Tables, Primary And Foreign Keys, Stored Procedures, And Triggers Required in Implementing The System in Digital Form.
- Data Modeling is The Act of Exploring Data-Oriented Structures Exactly As Demanded by The Corresponding Database Management Software.
- Data Modeling Can Be Used For Deriving High-Level Conceptual Data Models As Well As Physical Data Models.
- Data Modeling Helps Us To Identify Entity Types And Data Attributes Assigned To Entity Types Along With The Corresponding Datatypes And Constraints.
- Associations Between Entities Are Called As Relationships And With in The Data Modeling Inheritance, Composition, And Aggregation Are All Applicable Concepts.

- Traditional Data Modeling Focuses Solely on Data Alone But Do Not Explore The Behavior And Data Aspects of The Domain.
- The Focus on The Data And its Existential State Makes The Data Modelers Develop A Tendency To Be Much Better At Getting The Data "Right" From The Business System Than Object Modelers.

Why To Build OR Create A Data Model?

- To Avoid Redundancy of Data in An OLTP Database Management System There By Promising Consistency And Integrity.
- In Data Warehousing, Data From Source Systems Can Be Transformed As Per The Rules And Loaded into Target Tables.
- In Data Warehousing, You Can Do Data Profiling By Cleaning The Data From Source Systems And Load That into Data Warehouse Columns.
- In Data Warehousing Environment Same Column From Different Source System May Have Different Data Structure And Column Name Data Model Helps in Making Clear Distinction Among Such Columns.
- In Data Warehouse Environments, The Developers Can Create A Column As Per Standards of The Data Warehouse And Load The Data As Per The Necessity.
- In Data Warehousing Environment We Can Identify The Several Columns of Data That Helps in Predicting The Future Trend of Data Analysis, Used For Data Mining.
- In Data Warehousing OR In Data Mart, You Can Understand The Standards of Drill Down And Drill Up of The Data To A Certain Level of Aggregation That is Demanded.
- For Data Ware Housing Environment it Consolidates The Idea For Rollup And Cubing Combinations.

- Rewriting Data Models From Existing Systems That May Need To Change As Per The Standards of Reports.
- Helps in Correcting The Incorrect Data Modeling Techniques Applied in The Existing Systems.
- Helps in Generating The Data Model For A Data Base That Has No Data Model Specified Previously.

Advantages and Importance of Data Model

- The Data Model Makes Sure That All Data Objects Provided By The Functional Team Are Completely And Accurately Represented.
- Data Model is Detailed Enough To Be Used By The Technical Team For Building The Physical Database.
- The Information Contained in The Data Model Will Be Used To Define The Significance of Business, Relational Tables, Primary And Foreign Keys, Stored Procedures, And Triggers.
- Data Model Can Be Used To Communicate About The Business Within And Across Businesses.
- Data Model is A Single Component of Communication For Every One Who is Part of The System Development.
- Data Model Acts Like A Reference Tool Through Out The Project And Application Development.
- Data Model is The Source of Analyzing Any Bugs Specific To Operations in The Business System Implementation.
- Data Model itself Acts As A Document For The Development Team.