

KnowledgeBaze Data Warehouse Glossary

Access Path:

The path chosen by a database management system to retrieve the requested data.

Ad-Hoc Query:

Any query that cannot be determined prior to the moment the query is issued. A query that consists of dynamically constructed SQL, which is usually constructed by desktop-resident query tools.

Ad-Hoc Query Tool:

An end-user tool that accepts an English-like or point-and-click request for data and constructs an ad-hoc query to retrieve the desired result.

Administrative Data:

In a data warehouse, the data that helps a warehouse administrator manage the warehouse. Examples of administrative data are user profiles and order history data.

Aggregate Data:

Data that is the result of applying a process to combine data elements. Data that is taken collectively or in summary form.

Alerts:

A notification from an event that has exceeded a pre-defined threshold.

Atomic Data:

Data elements that represent the lowest level of detail. For example, in a daily sales report, the individual items sold would be atomic data, while rollups such as invoice and summary totals from invoices are aggregate data.

Authorization Request:

A request initiated by a consumer to access data for which the consumer does not presently have access privileges.

Authorization Rules:

Criteria used to determine whether or not an individual, group, or application may access reference data or a process.

Base Tables:

The normalized data structures maintained in the target warehousing database. Also known as the detail data.

Bi-directional Extracts:

The ability to extract, cleanse, and transfer data in two directions among different types of databases, including hierarchical, networked, and relational databases.

Braking Mechanism:

A software mechanism that prevents users from querying the operational database once transaction loads reach a certain level.

Bulk Data Transfer:

A software-based mechanism designed to move large data files. It supports compression, blocking and buffering to optimize transfer times.

Business Architecture:

One of the four layers of an information systems architecture. A business architecture describes the functions a business performs and the information it uses.

Business Data:

Information about people, places, things, business rules, and events, which is used to operate the business. It is not metadata. (Metadata defines and describes business data.)

Business Drivers:

The people, information, and tasks that support the fulfillment of a business objective.

Business Model:

A view of the business at any given point in time. The view can be from a process, data, event or resource perspective, and can be the past, present or future state of the business.

Business Transaction:

A unit of work acted upon by a data capture system to create, modify, or delete business data. Each transaction represents a single valued fact describing a single business event.

CASE:

Computer Aided Software Engineering.

CASE Management:

The management of information between multiple CASE "encyclopedias," whether the same or different CASE tools.

Catalog:

A component of a data dictionary that contains a directory of its DBMS objects as well as attributes of each object.

Central Warehouse:

A database created from operational extracts that adheres to a single, consistent, enterprise data model to ensure consistency of decision-support data across the corporation. A style of computing where all the information systems are located and managed from a single physical location.

Change Data Capture:

The process of capturing changes made to a production data source. Change data capture is typically performed by reading the source DBMS log. It consolidates units of work, ensures data is synchronized with the original source, and reduces data volume in a data warehousing environment.

Classic Data Warehouse Development:

The process of building an enterprise business model, creating a system data model, defining and designing a data warehouse architecture, constructing the physical database, and lastly populating the warehouses database.

Client/Server:

A distributed technology approach where the processing is divided by function. The server performs shared functions -- managing communications, providing database services, etc. The client performs individual user functions -- providing customized interfaces, performing screen to screen navigation, offering help functions, etc.

Client/Server Processing:

A form of cooperative processing in which the end-user interaction is through a programmable workstation (desktop) that must execute some part of the application logic over and above display formatting and terminal emulation.

Collection:

A set of data that resulted from a DBMS query.

Communications Integrity:

An operational quality that ensures transmitted data has been accurately received at its destination.

Consumer:

An individual, group or application that accesses data/information in a data warehouse.

Consumer Profile:

Identification of an individual, group or application and a profile of the data they request and use: the kinds of warehouse data, physical relational tables needed, and the required location and frequency of the data (when, where, and in what form it is to be delivered).

Cooperative Processing:

A style of computer application processing in which the presentation, business logic, and data management are split among two or more software services that operate on one or more computers. In cooperative processing, individual software programs (services) perform specific functions that are invoked by means of parameterized messages exchanged between them.

Copy Management:

The analysis of the business benefit realized by the cost of expenditure on some resource, tool, or application development.

Critical Success Factors:

Key areas of activity in which favorable results are necessary for a company to reach its goal.

Crosstab:

A process or function that combines and/or summarizes data from one or more sources into a concise format for analysis or reporting.

Currency Date:

The date the data is considered effective. It is also known as the "as of" date or temporal currency.

Data:

Items representing facts, text, graphics, bit-mapped images, sound, analog or digital live-video segments. Data is the raw material of a system supplied by data producers and is used by information consumers to create information.

Data Access Tools:

An end-user oriented tool that allows users to build SQL queries by pointing and clicking on a list of tables and fields in the data warehouse.

Data Analysis and Presentation Tools:

Software that provides a logical view of data in a warehouse. Some create simple aliases for table and column names; others create data that identify the contents and location of data in the warehouse.

Data Consumer:

An individual, group, or application that receives data in the form of a collection. The data is used for query, analysis, and reporting.

Data Custodian:

The individual assigned the responsibility of operating systems, data centers, data warehouses, operational databases, and business operations in conformance with the policies and practices prescribed by the data owner.

Data Dictionary:

A database about data and database structures. A catalog of all data elements, containing their names, structures, and information about their usage. A central location for metadata. Normally, data dictionaries are designed to store a limited set of available metadata, concentrating on the information relating to the data elements, databases, files and programs of implemented systems.

Data Element:

The most elementary unit of data that can be identified and described in a dictionary or repository which cannot be subdivided.

Data Extraction Software:

Software that reads one or more sources of data and creates a new image of the data.

Data Flow Diagram:

A diagram that shows the normal flow of data between services as well as the flow of data between data stores and services.

Data Loading:

The process of populating the data warehouse. Data loading is provided by DBMS-specific load processes, DBMS insert processes, and independent fastload processes.

Data Management:

Controlling, protecting, and facilitating access to data in order to provide information consumers with timely access to the data they need. The functions provided by a database management system.

Data Management Software:

Software that converts data into a unified format by taking derived data to create new fields, merging files, summarizing and filtering data; the process of reading data from operational systems. Data Management Software is also known as data extraction software.

Data Mapping:

The process of assigning a source data element to a target data element.

Data Mining:

A technique using software tools geared for the user who typically does not know exactly what he's searching for, but is looking for particular patterns or trends. Data mining is the process of sifting through large amounts of data to produce data content relationships. This is also known as data surfing.

Data Model:

A logical map that represents the inherent properties of the data independent of software, hardware or machine performance considerations. The model shows data elements grouped into records, as well as the association around those records.

Data Modeling:

A method used to define and analyze data requirements needed to support the business functions of an enterprise. These data requirements are recorded as a conceptual data model with associated data definitions. Data modeling defines the relationships between data elements and structures.

Data Owner:

The individual responsible for the policy and practice decisions of data. For business data, the individual may be called a business owner of the data.

Data Partitioning:

The process of logically and/or physically partitioning data into segments that are more easily maintained or accessed. Current RDBMS systems provide this kind of distribution functionality. Partitioning of data aids in performance and utility processing.

Data Pivot:

A process of rotating the view of data.

Data Producer:

A software service, organization, or person that provides data for update to a system-of-record.

Data Propagation:

The distribution of data from one or more source data warehouses to one or more local access databases, according to propagation rules.

Data Replication:

The process of copying a portion of a database from one environment to another and keeping the subsequent copies of the data in sync with the original source. Changes made to the original source are propagated to the copies of the data in other environments.

Data Scrubbing:

The process of filtering, merging, decoding, and translating source data to create validated data for the data warehouse.

Data Store:

A place where data is stored; data at rest. A generic term that includes databases and flat files.

Data Surfing

: See Data Mining.

Data Transfer:

The process of moving data from one environment to another environment. An environment may be an application system or operating environment. See Data Transport.

Data Transformation:

Creating "information" from data. This includes decoding production data and merging of records from multiple DBMS formats. It is also known as data scrubbing or data cleansing.

Data Transport:

The mechanism that moves data from a source to target environment. See Data Transfer.

Data Warehouse:

An implementation of an informational database used to store sharable data sourced from an operational database-of-record. It is typically a subject database that allows users to tap into a company's vast store of operational data to track and respond to business trends and facilitate forecasting and planning efforts.

Data Warehouse Architecture:

An integrated set of products that enable the extraction and transformation of operational data to be loaded into a database for end-user analysis and reporting.

Data Warehouse Architecture Development:

A KnowledgeBaze DW service program that provides an architecture for a data warehouse that is aligned with the needs of the business. This program identifies and designs a warehouse implementation increment and ensures the required infrastructure, skill sets, and other data warehouse foundational aspects are in place for a Data Warehouse Incremental Delivery.

Data Warehouse Engines:

Relational databases (RDBMS) and Multi-dimensional databases (MDBMS). Data warehouse engines require strong query capabilities, fast load mechanisms, and large storage requirements.

Data Warehouse Incremental Delivery:

A KnowledgeBaze DW program that delivers one data warehouse increment from design review through implementation.

Data Warehouse Infrastructure:

A combination of technologies and the interaction of technologies that support a data warehousing environment.

Data Warehouse Management Tools:

Software that extracts and transforms data from operational systems and loads it into the data warehouse.

Data Warehouse Network:

An industry organization for know-how exchange. KnowledgeBaze DW was the first vendor member of the Data Warehouse Network.

Data Warehouse Network:

An integrated network of data warehouses that contain sharable data propagated from a source data warehouse on the basis of information consumer demand. The warehouses are managed to control data redundancy and to promote effective use of the sharable data.

Data Warehouse Orientation:

A KnowledgeBaze DW program that provides an orientation to business and technical management of opportunities and approaches to data warehousing. The Orientation program encompasses a high level examination of solutions to business problems, return on investment, tools and techniques as they relate to data warehouse implementation. In addition, the program's objective is to assist customers in determining their readiness to proceed with data warehousing and to determine the appropriate data warehouse for their environment.

Database Schema:

The logical and physical definition of a database structure.

DBA:

Database Administrator.

Decentralized Database:

A centralized database that has been partitioned according to a business or end-user defined subject area. Typically ownership is also moved to the owners of the subject area.

Decentralized Warehouse:

A remote data source that users can query/access via a central gateway that provides a logical view of corporate data in terms that users can understand. The gateway parses and distributes queries in real time to remote data sources and returns result sets back to users.

Decision Support Systems (DSS):

Software that supports exception reporting, stop light reporting, standard repository, data analysis and rule-based analysis. A database created for end-user ad-hoc query processing.

Delta Update:

Only the data that was updated between the last extraction or snapshot process and the current execution of the extraction or snapshot.

Denormalized Data Store:

A data store that does not comply to one or more of several normal forms. See Normalization.

Derived Data:

Data that is the result of a computational step applied to reference or event data. Derived data is the result either of relating two or more elements of a single transaction (such as an aggregation), or of relating one or more elements of a transaction to an external algorithm or rule.

Desktop Applications:

Query and analysis tools that access the source database or data warehouse across a network using an appropriate database interface. An application that manages the human interface for data producers and information consumers.

DRDA:

Distributed Relational Database Architecture. A database access standard defined by IBM.

Drill Down:

A method of exploring detailed data that was used in creating a summary level of data. Drill down levels depend on the granularity of the data in the data warehouse.

DSS:

See Decision Support System.

DWA:

Data Warehouse Administrator.

Dynamic Dictionary:

A data dictionary that an application program accesses at run time.

Dynamic Queries:

Dynamically constructed SQL that is usually constructed by desktop-resident query tools. Queries that are not pre-processed and are prepared and executed at run time.

EIS:

Executive Information System.

End User Data:

Data formatted for end-user query processing; data created by end users; data provided by a data warehouse.

End User Mindset:

"Give me what I say I want, then I can tell you what I really want." To build a successful data warehouse, end users must be able to explore the possibilities.

Enterprise:

A complete business consisting of functions, divisions, or other components used to accomplish specific objectives and defined goals.

Enterprise Data:

Data that is defined for use across a corporate environment.

Enterprise Modeling:

The development of a common consistent view and understanding of data elements and their relationships across the enterprise.

Entity Identification:

The identification of the entities involved in the subject area. Entity identification is the process of giving data entities unique data elements by which they can be identified.

Entity Relationship Diagramming:

A process that visually identifies the relationships between data elements.

Event Analysis:

A process of analyzing notifications and taking action based on the notification content.

Event-Based Execution Rules:

The process of identifying those tasks that must be successfully executed to completion, or the system events that must occur, before a given task is to be triggered for processing.

Event Data:

Data about business events (usually business transactions) that have historic significance or are needed for analysis by other systems. Event data may exist as atomic event data and aggregate data.

Executive Information Systems (EIS):

Tools programmed to provide canned reports or briefing books to top-level executives. They offer strong reporting and drill-down capabilities. Today these tools allow ad-hoc querying against a multi-dimensional database, and

most offer analytical applications along functional lines such as sales or financial analysis.

Extendibility:

The ability to easily add new functionality to existing services without major software rewrites or without redefining the basic architecture.

Extract Date:

The date data was extracted.

Extract Frequency:

The latency of data extracts, such as daily versus weekly, monthly, quarterly, etc. The frequency that data extracts are needed in the data warehouse is determined by the shortest frequency requested through an order, or by the frequency required to maintain consistency of the other associated data types in the source data warehouse.

Extract Specification:

The standard expectations of a particular source data warehouse for data extracts from the operational database system-of-record. A system-of-record uses an extract specification to retrieve a snapshot of shared data, and formats the data in the way specified for updating the data in the source data warehouse. An extract specification also contains extract frequency rules for use by the Data Access environment.

Fastload:

A technology that typically replaces a specific DBMS load function. A fastload technology obtains significantly faster load times by preprocessing data and bypassing data integrity checks and logging.

FIFO:

A method of posting a transaction in first-in-first-out order. In other words, transactions are posted in the same order that the data producer entered them.

Filters:

Saved sets of chosen criteria that specify a subset of information in a data warehouse.

Frequency:

The timing characteristics of the data.

Functional Data Warehouse:

A warehouse that draws data from nearby operational systems. Each functional warehouse serves a distinct and separate group (such as a division), functional area (such as manufacturing), geographic unit, or product marketing group.

Gateway:

A software product that allows SQL-based applications to access relational and non-relational data sources.

Global Business Models:

Provides access to information scattered throughout an enterprise under the control of different divisions or departments with different databases and data models. This type of data warehouse is difficult to build because it requires users from different divisions to come together to define a common data model for the warehouse.

Hash:

Data allocated in an algorithmically randomized fashion in an attempt to evenly distribute data and smooth access patterns.

Historical Database:

A database that provides an historical perspective on the data.

Host-Driven:

A processing method in which the host computer controls the session. A host-driven session typically includes terminal emulation, front ending or client/server types of connections. The host determines what is displayed on the desktop, receives user input from the desktop, and determines how the application responds to the input.

Householding:

A methodology of consolidating names and addresses.

Immediate Processing:

Processing that occurs at the time the request for processing is made. Data may be requested and updated in an immediate mode.

Impact Analysis:

Identifying the impact of change on an object to its related objects.

Increment:

Data warehouse implementation can be broken down into segments or increments. An increment is a defined data warehouse implementation project that has a specified beginning and end. An increment may also be referred to as a departmental data warehouse within the context of an enterprise.

Info-Glut in Cyberspace:

Too much data! (30+ million electronic mailboxes, 7000 CD-ROMs with 650 Megs, 5000+ on-line databases, 500 cable channels, etc.)

Information:

Data that has been processed in such a way that it can increase the knowledge of the person who receives it. Information is the output, or finished goods," of information systems. Information is also what individuals start with before it is fed into a Data Capture transaction processing system.

Information Consumer:

A person or software service that uses data to create information.

Information Needs Analysis:

The identification and analysis of the needs for information required to satisfy a particular business driver.

Information Systems Architecture:

The authoritative definition of the business rules, systems structure, technical framework, and product backbone for business information systems. An information systems architecture consists of four layers: business architecture, systems architecture, technical architecture, and product architecture.

Information Warehouse:

IBM's approach to data warehousing that supports the implementation of either functional, central or decentralized warehouses.

Intelligent Agent:

A software routine that waits in the background and performs an action when a specified event occurs. For example, agents could transmit a summary file on the first day of the month or monitor incoming data and alert the user when certain transactions have arrived.

Interviews:

A procedure to obtain prioritized information needed to generate warehouse increments.

Inverted File Indexes:

A more efficient method to access data in an ad-hoc or analysis environment. It maintains indexes to all values contained in an indexed field. Those values, in turn, can be used in any combination to identify records that contain them, without actually scanning them from disk.

Journal File:

A file that contains update activity for rollback and data recovery purposes. Examples of update activity are commit checkpoints, as well as "before" and "after" operational database images. A journal file may be used to construct snapshot information for the data warehouse.

Local Access Database (LAD):

A database that serves individual systems and workgroups as the end point for shared data distribution. LADs are the "retail outlets" of the data warehouse network. They provide direct access to the data requested by specific systems or desktop query services. Data is propagated to LADs from data warehouses according to orders for subsets of certain shared data tables and particular attributes therein, or subsets of standard collections. This data is usually located on a LAN server. If servers are not available and the data is static, it may be located on the user's desktop. See Data Warehouse Network.

Local Directory:

A data dictionary propagated from the repository to the desktop containing metadata used for developing desktop applications and for generating transactions. A local directory is also used to bind definitions of local data structures used by desktop applications to the data requested from servers.

Location Transparency:

A mechanism that keeps the specific physical address of an object from a user. The physical location is resolved within the system so that operations can be performed without knowledge of the actual physical location.

Logical Data Model:

Actual implementation of a conceptual module in a database. It may take multiple logical data models to implement one conceptual data model.

Magic Arrow:

An arrow used in marketing materials that gives the illusion of an integrated and automated process.

Meta Muck:

An environment created when metadata exists in multiple products and repositories (DBMS catalogs; DBMS dictionaries; CASE tools warehouse databases; end user tools; and repositories).

Metadata

or **Meta Data:**

Metadata is data about data. Examples of metadata include data element descriptions, data type descriptions, attribute/property descriptions, range/domain descriptions, and process/method descriptions. The repository environment encompasses all corporate metadata resources: database catalogs, data dictionaries, and navigation services. Metadata includes things like the name, length, valid values, and description of a data element. Metadata is stored in a data dictionary and repository. It insulates the data warehouse from changes in the schema of operational systems.

Metadata Synchronization:

The process of consolidating, relating and synchronizing data elements with the same or similar meaning from different systems. Metadata synchronization joins these differing elements together in the data warehouse to allow for easier access.

Methodology:

A system of principles, practices, and procedures applied to a specific branch of knowledge.

Mid-Tier Data Warehouses:

To be scalable, any particular implementation of the data access environment may incorporate several intermediate distribution tiers in the data warehouse network. These intermediate tiers act as source data warehouses for geographically isolated sharable data that is needed across several business functions.

Middleware:

A communications layer that allows applications to interact across hardware and network environments.

Mini Marts:

A small subset of a data warehouse used by a small number of users. A mini mart is a very focused slice of a larger data warehouse.

MIP-O-Suction:

A query that consumes a high percentage of CPU cycles.

MIPS:

An acronym for millions of instructions per second. MIPS is mistakenly considered a relative measure of computing capability among models and vendors. It is a meaningful measure only among versions of the same processors configured with identical peripherals and software.

MPP:

Massive Parallel Processing. The "shared nothing" approach of parallel computing.

Multi-dimensional Database (MDBS and MDBMS):

A powerful database that lets users analyze large amounts of data. An MDBS captures and presents data as arrays that can be arranged in multiple dimensions.

Normalization:

The process of reducing a complex data structure into its simplest, most stable structure. In general, the process entails the removal of redundant attributes, keys, and relationships from a conceptual data model.

Object:

A person, place, thing, or concept that has characteristics of interest to an environment. In terms of an object-oriented system, an object is an entity that combines descriptions of data and behavior.

Object Description:

All the properties and associations that describe a particular object.

ODBC:

Open Database Connectivity. A standard for database access co-opted by Microsoft from the SQL Access Group consortium.

OLAP:

On-Line Analytical Processing.

OLTP:

On-Line Transaction Processing. OLTP describes the requirements for a system that is used in an operational environment.

Operational Database:

The database-of-record, consisting of system-specific reference data and event data belonging to a transaction-update system. It may also contain system control data such as indicators, flags, and counters. The operational database is the source of data for the data warehouse. It contains detailed data used to run the day-to-day operations of the business. The data continually changes as updates are made, and reflect the current value of the last transaction.

Operational Data Store (ODS):

An ODS is an integrated database of operational data. Its sources include legacy systems and it contains current or near term data. An ODS may contain 30 to 60 days of information, while a data warehouse typically contains years of data.

Order:

A message sent to data access services which triggers the delivery of required data. There are three types of orders: select order, transform order, and propagate order.

Parallelism:

The ability to perform functions in parallel.

Population:

See Data Loading and Data Replication.

Product Architecture:

One of the four layers of an information systems architecture. It describes standards to be followed in each portion of the technical architecture and vendor-specific tools and services to apply in developing and running applications.

Production Data:

Source data which is subject to change. It is a data capture system, often on a corporation's mainframe.

Propagated Data:

Data that is transferred from a data source to one or more target environments according to propagation rules. Data propagation is normally based on transaction logic.

Protocol:

A set of conventions that govern the communications between processes. Protocol specifies the format and content of messages to be exchanged.

QFH:

Query From Hell.

Quality Assurance:

The process of ensuring a correct result.

Query:

A (usually) complex SELECT statement for decision support. See Ad-Hoc Query or Ad-Hoc Query Software.

Query Governor:

A facility that terminates a database query when it has exceeded a predefined threshold.

Query Response Times:

The time it takes for the warehouse engine to process a complex query across a large volume of data and return the results to the requester.

Query Tools:

Software that allows a user to create and direct specific questions to a data base. These tools provide the means for pulling the desired information from a database. They are typically SQL-based tools and allow a user to define data in end-user language.

RDBMS:

Relational Database Management System.

RDBMS Concurrency:

Overlapping, concurrent execution of code segments.

Redundancy:

The storage of multiple copies of identical data.

Redundancy Control:

Management of a distributed data environment to limit excessive copying, update, and transmission costs associated with multiple copies of the same data. Data replication is a strategy for redundancy control with the intention to improve performance.

Reference Data:

Business data that has a consistent meaning and definition and is used for reference and validation (Process, Person, Vendor, and Customer, for example). Reference data is fundamental to the operation of the business. The data is used for transaction validation by the data capture environment, decision support systems, and for representation of business rules. Its source for distribution and use is a data warehouse.

Refresh Technology:

A process of taking a snapshot from one environment and moving it to another environment overlaying old data with the new data each time.

Replicated Data:

Data that is copied from a data source to one or more target environments based on replication rules. Replicated data can consist of full tables or rectangular extracts.

Repository Environment:

The Repository environment contains the complete set of a business's metadata. It is globally accessible. As compared to a data dictionary, the repository environment not only contains an expanded set of metadata, but can be implemented across multiple hardware platforms and database management systems (DBMS).

Roll Up Queries:

Queries that summarize data at a level higher than the previous level of detail.

RPC:

Remote Procedure Call.

Scalability:

The ability to scale to support larger or smaller volumes of data and more or less users. The ability to increase or decrease size or capability in cost-effective increments with minimal impact on the unit cost of business and the procurement of additional services.

Schema:

The logical and physical definition of data elements, physical characteristics and inter-relationships.

Securability:

The ability to provide differing access to individuals according to the classification of data and the user's business function, regardless of the variations.

SELECT:

A SQL statement (command) that specifies data retrieval operations for rows of data in a relational database.

Semantic Mapping:

The mapping of the meaning of a piece of data.

Server:

A service that provides standard functions for clients in response to standard messages from clients. Note: A commonly used definition of server also refers to the physical computer from which services are provided.

Slice and Dice:

A term used to describe a complex data analysis function provided by MDBMS tools.

SMP:

Symmetrical Multi-Processing. The "shared everything" approach of parallel computing.

Source Database:

An operational, production database or a centralized warehouse that feeds into a target database.

SQL:

Structured Query Language. A structured query language for accessing relational, ODBC, DRDA, or non-relational compliant database systems.

SQL-Compliant:

Conformity to ANSI standards for Structured Query Language specifications.

SQL Query Tool:

An end-user tool that accepts SQL to be processed against one or more relational databases.

Standard Query:

A stored procedure of a recently executed query. Technically, a standard query may be stored on the desktop as "canned " SQL and passed as dynamic SQL to the server database to execute. This is undesirable unless the stored query is seldom executed.

Static Query:

A stored, parameterized procedure, optimized for access to a particular data warehouse.

Stoplighting:

A technique using colored circles to identify the content of a data element. The colors are defined by a set of predefined thresholds.

Subject Oriented Databases:

Rather than build one massive, centralized data warehouse, most companies are building numerous subject-oriented warehouses to serve the needs of different divisions.

Summarization Tables:

These tables are created along commonly used access dimensions to speed query performance, although the redundancies increase the amount of data in the warehouse. See Aggregated Data.

Syntactic Mapping:

The mapping required to unravel the syntax of information.

Systems Architecture:

One of the four layers of the information systems architecture. The systems architecture represents the definitions and inter-relationships between applications and the product architecture.

Tactical Data Warehouse Development:

The process of selecting a portion of an enterprise and implementing a data warehouse. The process includes constructing a data model for the area, determining the data warehouse architecture, constructing the physical model, and populating the warehouse database. It also includes creating or buying the applications to access the data warehouse, prototyping the tactical warehouses (access definitions, views, etc.) and incorporating end-user feedback.

TANSTAAFL:

There Ain't No Such Thing As A Free Lunch! In developing a data warehouse, there is work involved, and there is no "free lunch."

Target Database:

The database in which data will be loaded or inserted.

Technical Architecture:

One of the four layers of the information systems architecture. The technical architecture defines and describes the interfaces, parameters and protocols used by the product and systems architecture.

Tool Encyclopedias:

Encyclopedias, repositories or dictionaries used by application development tools. The non-definable "repository" used by a tool.

Transformers:

Rules applied to change data.

Triggering Data:

Data that selects and loads data on a scheduled basis.

Unit of Work Consolidation:

The process of consolidating multiple updates to a single row image into a single update.

Update:

Not allowed in a data warehouse.

Versioning:

The ability for a single definition to maintain information about multiple physical instantiations.

VITAL:

An acronym for Virtually Integrated Technical Architecture Lifecycle.

VITAL Compliance:

Conformance to the design objectives and principles, distributed computing styles, development approaches, standards, and data-distribution and access techniques; functional compatibility with the VITAL technical architecture.

VLDB:

Very Large Databases.

Warehouse Business Directory:

Provides business professionals access to the data warehouse by browsing a catalog of contents.

Warehouse Technical Directory:

Defines and manages an information life cycle, a definition of warehouse construction, change management, impact analysis, distribution and operation of a warehouse.

-----#-----