

Data Warehouse & Business Intelligence Glossary

...can analyze and ...
...the mix of activities
...performed in terms of the
...a **knowledge** base for
...proprietary knowledge ...
...peers as a means of
...communication

Notice

I have sourced most of these glossories from website and combined them into a super duper 23 pages of glossaries pertaining to Data Warehouse.

I have also amended and corrected some of the glossaries to make sure that are easier to read and where possible use simple business jargons.

Enjoy!!

Subra Suppiah,
Data Warehouse Practicioner

A

Activity A set of tasks related either by topic, dependencies, data, common skills, or deliverables. The next level of organization below a phase.

Application Gap A recorded variance between some aspect of a business requirement and the capability or features of the application system that are necessary to satisfy the requirement.

Application System An automated collection of business functions, entities, modules, technology platforms, and documentation that performs a specified set of business functions.

Approach A particular way of putting a method or task to use, including risk factors, tips from experience, recommendations, and additional advice.

Ad Hoc Query: A database search that is designed to extract specific information from a database. It is ad hoc if it is designed at the point of execution as opposed to being a “canned” report. Most ad hoc query software uses the structured query language (SQL).

Aggregation: The process of summarizing or combining data. It can be used to speed up query performance. Facts are summed up for selected dimensions from the original fact table. The resulting aggregate table will have fewer rows, thus making queries that can use them run faster.

Attribute: Attributes represent a single type of information in a dimension. For example, year is an attribute in the Time dimension.

B

Backup and Recovery Strategy A storage and recovery strategy that protects against business information loss resulting from hardware, software, or network faults.

Baseline 1. A starting point or condition against which future changes are measured. 2. A named set of object versions which fixes a configuration at a particular point in time. A baseline normally represents a milestone or key deliverable of a project.

Budget A plan for determining in advance the expenditure of time, money, etc.

Business An enterprise, commercial entity, or firm in either the private or public sector, concerned with providing products or services to satisfy customer requirements.

Business Aim A statement of business intent measured subjectively; for example, to move up market, or to develop a sustainable level of growth; usually strategic or tactical with a 3-5 year horizon.

Business Area The set of business processes within the scope of a project.

Business Constraint Any external, management, or other factor that restricts a business or system development in terms of resource availability, dependencies, timescales, or some other factor.

Business Function Something an enterprise does, or needs to do, in order to achieve its objectives.

Business Goal A statement of business intent.

Business Location A uniquely identifiable geographic location, site, or place from which one or more business units may wholly or partly operate.

Business Meta data The information whereby users can understand and access the data warehouse. It focuses on

what data is in the warehouse, how it was transformed, the source, and the timeliness of the data.

Business Object A physical or logical object of significance to a business; for example, a sales order, department,

assembly, item, balance, or invoice. A business object is analogous to a class in object-oriented terminology.

Business Organization Type A classification of a business organization into one of several functional categories. Each business organization type has a distinct set of business requirements. All the business organizations of a certain type will typically require similar applications and system capabilities. A given site may house one or more business organization types. Since business organizations may be related in a hierarchy, a high level business organization may be composed of several business organizations of different types. For the purposes of application architecture analysis and design, it is generally useful to decompose the hierarchy of business organizations until it is composed of atomic organization types.

Business Priority A statement of the level or urgency of important business needs.

Business Process The complete response that a business makes to an event. A business process entails the execution of a sequence of one or more process steps. It has a clearly defined deliverable or outcome. A Business Process is defined by the business event that triggers the process, the inputs and outputs, all the operational steps required to produce the output, the sequential relationship between the process steps, the business decisions that are part of the event response, and the flow of material and/or information between process steps.

Business Process Reengineering (BPR) The activity by which an enterprise reexamines its goals and how it achieves them, followed by a disciplined approach of business process redesign. A method that supports this activity.

Business Rule A rule under which an organization operates. A policy or decision that influences the process step.

Business System A combination of people and automated applications organized to meet a particular set of business objectives.

Business Unit Part of an organization treated for any purpose as a separate entity within the parent organization. Examples include a department or distribution center.

Buy-In The process whereby stakeholders, e.g. end-users as well as executives, come to see the goals of an organizational or change effort as their own; a key component in achieving change successfully.

C

CASE Tools A set of integrated Computer-Aided Systems Engineering (CASE) and application development tools that assist in software development; for example, analyzing business requirements, designing applications, generating application code, etc.

Category A structure for organizing, typing, and representing a dimension of a multidimensional cube.

Central Repository Location of a collection of documentation, customizations, modifications, or enhancements designed to alleviate the recreation of successfully completed work.

Change A deviation from a currently established baseline.

Change Management 1. The complete set of processes employed on a project to ensure that changes are implemented in a visible, controlled and orderly fashion. 2. The activity, or set of activities, undertaken to govern systematically the effects of organizational change.

Change Request 1. A request for a change to the required behavior of a system, usually from a user as a result of

reviewing current behavior. 2. The mechanism by which a change is requested, investigated, resolved and approved;

Client/Server A type of technical architecture that links many personal computers or workstations (clients) to

one or more large processors (servers). Clients generally manage the user interface, possibly with some local data.

Servers usually manage multiple-access databases, including ensuring data integrity and other invariants.

Cluster A means of storing data together from multiple tables, when the data in those tables contains common information accessed concurrently.

Column A means of implementing an item of data within a table. It can be in character, date, or number format, and be optional or mandatory.

Computer-Aided Systems Engineering (CASE) The combination of graphical, dictionary, generator, project management, and other software tools to assist computer development staff engineer and maintain high-quality systems, within the framework of a structured method.

Computer Network An interconnected group of computers.

Critical Success Factor (CSF) A business event, dependency, product, or other factor that, if not attained, would seriously impair the likelihood of achieving a business objective.

Custom Code Coding added to a packaged application or module generated by a CASE tool to implement functionality that the application or generator has not provided.

Catalog: A component of a data dictionary that describes and organizes the various aspects of a database such as its folders, dimensions, measures, prompts, functions, queries and other database objects. It is used to create queries, reports, analyses and cubes.

Cross Tab: A type of multi-dimensional report that displays values or measures in cells created by the intersection of two or more dimensions in a table format.

Conformed Dimension:A dimension that has exactly the same meaning and content when being referred from different fact tables

Cube: A multi-dimensional matrix of data that has multiple dimensions (independent variables) and measures (dependent variables) that are created by an Online Analytical Processing System (OLAP). Each dimension may be organized into a hierarchy with multiple levels. The intersection of two or more dimensional categories is referred to as a cell.

D

Dashboard: A data visualization method and workflow management tool that brings together useful information on a series of screens and/or web pages. Some of the information that may be contained on a dashboard includes reports, web links, calendar, news, tasks, e-mail, etc. When incorporated into a DSS or EIS key performance indicators may be represented as graphics that are linked to various hyperlinks, graphs, tables and other reports. The dashboard draws its information from multiple sources applications, office products, databases, Internet, etc.

Data Access The process of accessing the data warehouse database objects using various tools such as analysis, reporting, query, statistical, and data mining.

Data Acquisition The process of extracting, transforming, and transporting data from the source systems and external data sources to the data warehouse database objects.

Data Administration The process of managing the institutional data in order to provide reliable, accurate, secure and accessible data to meet strategic and management needs at all levels of the enterprise. It is the purpose of this process to improve the accuracy, reliability, and security of the institution's data; reduce data redundancy; provide ease of access, assuring that data are easily located, accessible once located, and clearly defined; and to provide data standards. It is also the purpose of the Data administration function to educate the user community on institutional data policies and to encourage the responsible use of data.

Data Aggregation The process of redefining data into a summarization based on some rules or criteria. Aggregation may also encompass de-normalization for data access and retrieval.

Database A collection of data, usually in the form of tables or files, under the control of a database management

system.

Database Architecture The collective application and database instances that comprise the complete system.

Database Administrator A person (or group of people) responsible for the maintenance and performance of a database.

Database Index A mechanism to locate and access data within a database. An index may quote one or more columns and be a means of enforcing uniqueness on their values.

Database Instance One set of database management processes and an allocated area in memory for managing those processes. Typically, a database instance is associated with

one database. Note that a database instance may process data for one or more applications.

Database Management System (DBMS) A software environment that structures and manipulates data, and ensures data security, recovery, and integrity.

Data Cleansing The transformation of data in its current state to a pre-defined, standardized format using packaged software or program modules.

Data Definition The specification of a data element to be maintained. The specification includes datatype, size, and rules about processing: for example, derivation and validation

Data Dictionary A part of a database that holds definitions of data elements, such as tables, columns, and views.

Data Extraction The process of pulling data from operational and external data sources in order to prepare the source data for the data warehouse environment.

Dataflow Diagramming A technique for expressing the significant dataflows of a business system.

Data Integration The movement of data between two co-existing systems. The interfacing of this data may occur once every hour, once a day, etc.

Data Integrity The quality of the data residing in the database objects. The measurement which users consider when analyzing the value and reliability of the data.

Data Integrity Testing Verification that converted data is accurate and functions correctly within a single subsystem or application.

Data Map A technique for establishing a match, or balance, between the source data and the target data warehouse database object. This technique identifies data shortfalls and recognizes data issues.

Data Mart A data warehouse data class organized for a business functional area or department. The database contains data summarized at multiple levels of granularity and may be designed using relational or multidimensional database structures.

Data Mart Data Model The logical representation of the specific information requirements organized around a department of functional area.

Data Migration The movement of data from one database to another database -- but not necessarily to a working application or subsystem tables.

Data Model A representation of the specific information requirements of a business area; see also ENTITY

RELATIONSHIP MODEL.

Data Mining: The process of researching data marts and data warehouses to detect specific patterns in the data sets. Data mining may be performed on databases and multi-dimensional data cubes with ad hoc query tools and OLAP software. The queries and reports are typically designed to answer specific questions to uncover trends or hidden relationships in the data.

Data Partitioning A technique to improve application performance or security by splitting tables across multiple locations.

Data Registry The master copy of the data associated with a business object. Several databases may share access to a common data registry to ensure consistency and eliminate redundant entries across multiple applications and

databases. An example of a data registry would be a shared customer master. All updates and changes would be made to the customer master data registry and then propagated to subscribing sites. All systems requiring customer information would interface with the customer data registry.

Data Registry Interface An interface that transfers data registry data between similar or dissimilar applications.

Data Replication The copying of data to and from sites to improve local service response times and availability; frequently employed as part of a backup and recovery strategy.

Datastore A temporary or permanent storage concept for logical data items used by specified business functions and processes.

Data Scrubbing The process of manipulating or cleaning data into a standard format. This process may be done in conjunction with other data acquisition tasks.

Data Source An operational system, third-party organization, or external system that provides the data to support the information requirements of the client. The data source is accessed during the data acquisition process.

Data System of Record For an item that is populated across multiple systems (like social security number) name the source system.

Data Transfer The physical movement of data between applications, perhaps across sites.

Data Transformation The process of redefining data based on some predefined rules. The values are redefined based on a specific formula or technique.

Data Translation The process of redefining data in a manner differing between its original representation and its final representation.

Data Transportation The process of moving and loading the transformed data into the data warehouse database objects.

Data Warehouse An enterprise structured repository of subject-oriented, time-variant, historical data used for information retrieval and decision support. The data warehouse stores atomic and summary data. The data warehouse is the source data stored in the data marts.

Data Warehousing:The process of designing, building, and maintaining a data warehouse system.

Data warehouse administrator (DWA) A person or group of people that administer and manage a data warehouse

Data Warehouse Data Model The logical representation of the historical information requirements structured for the enterprise.

Data Warehouse Integration The process on reconciling each data warehouse increment with the strategic data warehouse architecture.

Data Warehouse Method (DWM) A structured method for full life-cycle custom development data warehouse projects.

Data Validation The process of ensuring correct data based on error and exception handling rules. This process directly impacts data integrity.

Decision Support System (DSS) An application primarily used to consolidate, summarize, or transform transaction data to support analytical reporting and trend analysis.

Deliverable A tangible, measurable output of a task.

De-normalization A database design activity that restructures a database by introducing derived data, replicated data, and/or repeating data to tune an application system and increase performance.

Dependency The relationship of one task to another where the start or end date of the second task (successor) is constrained by the start or end date of the first (predecessor).

Derived Attribute A value that is derived by some algorithm from the values of other attributes; for example, profit, which is the difference between revenue and expense.

Derived Column A value derived by some algorithm from the values of other columns; see also DERIVED ATTRIBUTE, DERIVED DATA ITEM , and DERIVED FIELD.

Derived Data Item A value derived by some algorithm from the values of other data items; for example, profit, which is the difference between revenue and expense.

Dimension A multidimensional structure which represents a side of a multidimensional cube. Each dimension represents a different category, such as region, time, product type. **Discovery** The evaluation and validation of the implemented data warehouse increment, experiences and lessons learned, and scope for next increment to be developed.

Dimensional Model: A type of data modeling suited for data warehousing. In a dimensional model, there are two types of tables: dimensional tables and fact tables. Dimensional table records information on each dimension, and fact table records all the "fact", or measures.

Dimension Table A table that contains discrete values (usually a countable text field like school or degree). Also see fact table. Imagine viewing a spreadsheet. The row and column names would be the dimensions and the numeric data within would be the facts.

Distributed Database A database that is physically located on more than one computer processor. It is connected via some form of communications network. An essential feature of a true distributed database is that users or programs work as if they had access to the whole database locally.

Distributed Processing The ability to have several computers working together in a network, where each processor runs different activities for a user, as required.

Domain A set of business validation rules, format constraints, and other properties that apply to a group of attributes or database columns; for example: a list of values, a range, a qualified list or range, or any combination of these.

Drill-Down :The ability of a data-mining tool to move down into increasing levels of detail in a data mart, data warehouse or multi-dimensional data cube. Data analysis to a child attribute

Drill Up:The ability of a data-mining tool to move back up into higher levels of data in a data mart, data warehouse or multi-dimensional data cube. Data analysis to a parent attribute.

Drill Across: Data analysis across dimensions.

Drill Through: Data analysis that goes from an OLAP cube into the relational database.

E

Element A thing of significance about which information is recorded; a component at the most useful, basic level.

Element Type Any element held in the repository is classified as a particular type. Examples of element type are entity, attribute, program module, process, table, diagram, text, softbox. Occurrences or instances of these are called elements.

End User see **USER** .

End-User Layer The user interface and layout of the multidimensional structures designed for the data access tools. This includes the customization of the tools for the users.

Enterprise A group of departments, divisions, or companies which make up an entire business.

Enterprise Support Systems The set of all computer-based systems, documents, and procedures used in support of business enterprise operations.

Enterprise Technical Architecture (ETA) A series of rules, guidelines, and principles used by an organization to direct the process of acquiring, building, modifying, delivering, and integrating Information Technology resources throughout the enterprise. These resources can include equipment, software, business processors, protocols, standards, methodologies, IT organizational structures and more.

Entity A thing of significance, whether real or imagined, about which information needs to be known or held. It is implemented in a database as one or more tables.

Entity Relationship Diagram (ERD) A diagram that pictorially represents entities, the relationships between them and the attributes used to describe them.

Entity Relationship Model A type of data model. Part of the business model that consists of many Entity Relationship Diagrams.

Event An occurrence in a business's environment to which that business must respond; see also **BUSINESS SYSTEM** and **EVENT RESPONSE** .

Executive Information System (EIS) A reporting application targeted for use by executives. Usually such applications have extremely user-friendly, graphical interfaces with a small local data store derived from connection to a data warehouse. It is often used synonymously with decision support system.

Extensibility The ability to add new components, technology, and increments to the data warehouse solution. This is a critical capability for the data warehouse architecture and technical architecture.

External Data Source A system or data file provided by an organization external to the client. This includes parent companies or subsidiaries, alliances, partners, and data brokers.

Extraction, Transformation and Loading (ETL) Tool: Software that is used to extract data from a data source like a operational system or data warehouse, modify the data and then load it into a data mart, data warehouse or multi-dimensional data cube.

F

Fact Table Usually the primary table that contains measurements (mostly numeric data like grade, wages, etc.). Also see dimension table. Imagine looking at a spreadsheet. The row and column names would be the dimensions and the numeric data within would be the facts.

Feasibility The mechanism for balancing business constraints with technology constraints to produce a cost-effective solution.

Feedback Response, including corrections, additions, and approval, elicited from users, stakeholders, sponsors, and others, to any deliverable or deliverable component.

Feedback Session A meeting organized to present work in progress in order to gain feedback.

Field A means of implementing an item of data within a file. It can be in character, date, number, or other format and be optional or mandatory.

File Transfer Protocol (FTP) The physical movement of data files between applications, often across sites.

Focus Area 1. A group of associated activities that define and establish program level deliverables, i.e. standards, configuration, and processes. These deliverables are used by multiple projects creating commonality and reusability across the program. 2. A team of individuals working within the Program Office framework for a common family of processes. These focus areas could also be referred to as Program Office Projects. 3. A scoped area of the client organization with common responsibilities and information requirements. The focus area provides the scope for incremental development efforts and may be referred to as the INCREMENT . The focus area can overlap business functions or may reside in a specific business function.

Focus Group A small group selected to provide opinions and responses to topics or issues presented in a group setting; an assessment technique.

Foreign Key One or more columns in a relational database table that implement a many-to-one relationship that the table in question has with another table or with itself.

Format The type of data that an attribute or column may represent; for example, character, date, number, sound, or image.

G

Gantt Chart A scheduling tool used to display the status of a project's tasks. A Gantt chart shows each task's duration as a horizontal line. The ends of the lines correspond to the task's start and end dates.

Gap Analysis 1. The process of determining, documenting, and approving the variance between business requirements and system capabilities in terms of packaged application features and technical architecture. 2. The process of determining and evaluating the variance or distance between two items' properties being compared.

Generator A mechanism for transforming the specification of a module into executable program code, also known as a code generator.

Granularity A term used to describe how finally broken down a fact is in a table. For example, we might have wages individually recorded per employee in one table but we might have another table with wages aggregated by department.

Group Interview Any session where users, stakeholders, or sponsors collectively discuss the requirements, priorities, design, or implementation of a business solution system.

Guideline Text that provides instructions and advice for performing a task and suggests possible approaches.

Granularity: The level of detail in a data store or report

H

Hardware Node A computer on a network; for example, clients and servers.

Help Desk A support system designed to assist end users with technical and functional questions and problems.

Hierarchy: The organization of data, e.g. a dimension, into a outline or logical tree structure. The strata of a hierarchy are referred to as levels. The individual elements within a level are referred to as categories. The next lower level in a hierarchy is the child; the next higher level containing the children is their parent.

I

Impact Analysis The process of understanding the complete effect of a particular change.

Implementation The installation of an increment of the data warehouse solution that is complete, tested, operational, and ready. An implementation includes all necessary software, hardware, documentation, and all required data.

Implementation Questionnaire A tool you use to collect business and system information during a business baseline interview. It consists of a pre-built set of questions organized by business function that are to be supplemented by the analyst with relevant company terms and other characteristics before use in driving the interview.

Increment The defined scope of the portion of the data warehouse selected for implementation. Each increment satisfies elements of the total data warehouse solution.

Increment Integration The process of reconciling the designs, modules, database objects, and data access components within a given increment. This effort is addressed during the data warehouse architecture process.

Incremental Development A technique for producing all or part of a production system based on an outline definition. The technique involves iterations of a cycle of build, refine, and review so that the correct solution gradually emerges. This technique can be difficult to control, but nonetheless is very useful when properly used; also called quick build and interactive development.

Information Access Model A model that depicts access to key process and organization information for reporting and/or security purposes. **Information Flow Model** A model that visually depicts information flows in the business between business functions, business organizations and applications.

Information Systems (IS) A system for managing and processing information, usually computer-based. Also, a functional group within a business that manages the development and operations of the business's information

systems.

Information Systems Strategies (ISS) A method that aligns information technology priorities with business strategies and defines the approach to take to achieve those goals.

Initial Deliverable A deliverable is initial if it is intended to be updated later. An initial deliverable is usually preliminary and its content changeable by a later task when more information is known.

Initial Load The first population of the production database installations using the data acquisition modules for extraction, transformation, and transportation.

Installation The loading of an instance of an application system that is complete, tested, operational, and ready. An installation includes all necessary software, hardware (including terminals, networks, etc.) and documentation, and includes all required data.

Integration Test A sequence of steps or set of procedures to verify the inter-operability of various system components.

Interface A linkage between systems which can be either automated (via software programs) or procedural (manual).

Issue A situation or concern which requires a resolution. Some issues, if not addressed, could adversely impact the success of a project.

Iterative Development The application of a cyclic, evolutionary approach to the development of requirements definition, design, or construction using prototyping and iterative build techniques.

K

Key A way of accessing something. Any set of columns used for retrieval of rows from a table.

Key Performance Indicator (KPI) A significant measure used on its own, or in combination with other key performance indicators, to monitor how well a business is achieving its quantifiable objectives.

Key Resource A person with a wide range of skills or experiences who can be effective in many types of tasks, or is critical to the completion of a specific task.

L

Legacy System An existing system repository of information and processes. **Link Test** A test to identify errors in linked modules of an application system. Link testing is an extension of module testing carried out on a number of levels of detail. Examples include, linked modules of a program, linked programs of a functional area or subsystem, and linked subsystems of the complete application system. The link test is usually a white box test.

Level: A tier or strata in a dimensional hierarchy. Each lower level represents an increasing degree of detail. Levels in a location dimension might include country, region, state, county, city, zip code, etc.

Live Implementation process has ended and the solution is put into production.

Logical Application Architecture A complete map of the application instances required to support the applications architecture.

Logical Data Warehouse Architecture The framework which outlines all of the data warehouse functions and components for the strategic data warehouse. This includes the data classes, relational and multidimensional databases, meta data repository, ETT components, warehouse management, and client and server processes.

Logical Naming Standards What the business community calls a field. Also see physical naming standards.

Logical System Design The task of designing a system to support business needs without making final decisions regarding the physical implementation. The same logical design should be appropriate for many physical implementations using, for instance, different versions of a database management system.

Look and Feel The appearance and behavior of a system facility as perceived by the end user. This includes the data, the layout, and the user interaction through menus, buttons, text editing, and other devices.

M

Mechanism 1. A particular technique or technology for delivering a function. Examples might be a telephone, a computer, or an electronic mail service. 2. Resources that enable or facilitate the step/sequence in a test scenario.

Merge-Purge The process of compiling multiple data records, retaining the desired data, and removing unwanted data. This process may be invoked during the data acquisition process.

Measure: A quantifiable variable or value stored in a multi-dimensional OLAP cube. It is a value in the cell at the intersection of two or more dimensions.

Member: One of the data points for a level of a dimension.

Meta data Also known as data about data is the information about the contents and uses of the data warehouse. Meta data is created by several components of the data warehouse and provides a business and technical view of the data warehouse solution.

Metric: A measured value. For example, total sales is a metric.

Method A structured organization of tasks, estimates, and guidelines that provide a systematic approach or discipline.

MIPS Millions of Instructions Per Second – a measure of computer processing capacity.

Module A logical program unit. Examples include: forms, reports, user exits, C programs, PL/SQL procedures, and database triggers.

Multidimensional Database A database management system in which data can be viewed and manipulated in multiple dimensions. Data is stored using multidimensional structures and is organized to support analytical operations such as drill-down, consolidation, slicing, and dicing.

Multi-Dimensional Online Processing (MOLAP): Software that creates and analyzes multi-dimensional cubes to store its information.

N

Node A single computer, group of computers, or mechanism for handling some communication traffic through a

particular point on a computer network.

Normalization A technique to eliminate data redundancy.

Non-Volatile Data: Data that is static or that does not change. In transaction processing systems the data is updated on a continual regular basis. In a data warehouse the database is added to or appended, but the existing data seldom changes.

Normalization: The process of eliminating duplicate information in a database by creating a separate table that stores the redundant information. For example, it would be highly inefficient to re-enter the address of an insurance company with every claim. Instead, the database uses a key field to link the claims table to the address table. Operational or transaction processing systems are typically “normalized”. On the other hand, some data warehouses find it advantageous to de-normalize the data allowing for some degree of redundancy.

O

Object Orientation (OO) The perspective that systems should be constructed from objects, which themselves may be aggregations of smaller objects.

On-Line Analytical Processing (OLAP) On-line retrieval and analysis of data to reveal business trends and statistics not directly visible in the data directly retrieved from a data warehouse. Also known as multidimensional analysis.

Operational Datastore (ODS) A database structure that is a repository for near real-time operational data rather than long term trend data.

Operational Data Source The current operational system which contains the source data to be extracted, transformed, and transported to the data warehouse database objects.

Operational System A system which supports the operations of the client's business. These systems may be mission critical systems or support systems.

P

Phase A grouping of activities that lead to a major project deliverable or milestone.
Phase Completion The project management tasks which conclude and secure client sign-off of a phase.

Physical Naming Standards Standard portion of a name assigned to a type of field (ex. Amt, flag, derived, etc.) Also see logical naming standards.

Pilot An initial project which will serve as a model or template for future projects.

Plan A scheme, method or design for the attainment of some objective or to achieve something.

Prerequisite Something needed by a task produced by a previous task or an external source.

Primary Index An index used to improve performance on the combination of columns most frequently used to access rows in a table.

Primary Key A set of one or more columns in a database table whose values, in combination, are required to be unique within the table.

Problem Report The mechanism by which a problem is recorded, investigated, resolved and verified.

Procedure A written set of steps that specifies how to carry out a business function. If the business function is system-assisted, its corresponding procedure will indicate how the application system carries out that business function.

Process 1. The sequential execution of functions triggered by one or more events. 2. A grouping of tasks within a method based on common functions or disciplines which lead to one or more key deliverables.

Process Flow The passing of execution of a process from one process step to the next. It may include the passing of information or materials from the first step to the second.

Process Flow Diagram A diagram which shows the triggering event(s), sequential flow of process steps, decision points, and deliverable or outcome of a single process.

Production Environment The database, equipment, documentation, and procedures used in support of live business operations.

Program 1. A set of coded instructions that a computer executes or interprets to perform an automated task. 2. An interrelated group of projects that are either being run concurrently or sequentially and that share a system goal. Individual projects may have different goals, however the combined set of projects will have a program goal.

Project Library 1. A system for storing, organizing and controlling all documentation produced or used by the project. 2. The physical location of all deliverables for a single project, plus administrative and support materials. An administrative office that all members of a team should have access to.

Project Life-cycle The organization of a project according to its three major parts: planning, execution and completion.

Proof-of-Concept An approach used for demonstrating immediate business benefit, proving feasibility and critical

aspects of a solution, and to demonstrate value to the business. The proof-of-concept has a well-defined scope and set of objectives. The proof-of-concept may be the first incremental development effort for a data warehouse.

Prototype A facsimile of an end product used to demonstrate a concept rapidly, check feasibility, and/or gain acceptance.

Prototyping The construction of a partial system to demonstrate some aspect or aspects of the intended system behavior in order to gain user acceptance or to establish technical feasibility.

Q

Quality Review A review used to assess the quality of a deliverable in terms of fitness for purpose and adherence to defined standards and conventions.

Questionnaire A written or electronic survey instrument comprised of a series of questions, designed to measure a specific item or set of items.

R

Record In a non-relational database system, a record is an entry in a file, consisting of individual elements of information, which together provide full details about an aspect of the information needed by the system. Individual elements are held in fields and all records are held in files. An example of a record might be an employee. Every detail of the employee, for example, date of birth, department code, or full names will be found in a number of fields. In a relational system, record is an alternate word for row.

Referential Integrity Constraint Rules that specify the correspondence of a foreign key to the primary key of its

related table. For example, what should happen to the foreign key when the referenced primary key row is deleted.

Refresh The process for updating the data warehouse database objects with new data. The refresh process occurs after initial load on a scheduled basis and is monitored via warehouse management procedures.

Relational Database Management System (RDBMS) A database management system in which data can be viewed and manipulated in tabular form. Data can be sorted in any order and tables of information are easily related or joined to each other.

Relational Online Analytical Processing (ROLAP): OLAP software that employs a relational strategy to organize and store the data in its database

Relationship 1. What one entity has to do with another. 2. Any significant way in which two things of the same or different type may be associated.

Reporting Database A database used by reporting applications. Reporting databases are often duplicates of transaction databases used to off-load report processing from transaction databases.

Repository A mechanism for storing any information about the definition of a system at any point in its life-cycle. Repository services would typically be provided for extensibility, recovery, integrity, naming standards, and a wide variety of other management functions.

Replication: The process of copying data from one database table to another.

Request for Proposal The formal mechanism by which a company conveys its business requirements during the search for a new application system. Known as the RFP, this document drives the Pre-sales cycle and provides valuable information into the business requirements definition process of the implementation.

Resource Any persons, equipment, or material needed to perform a task(s).

Reverse Engineering The automatic creation of system specifications from existing code and data definitions.

Risk The potential of an adverse condition occurring on a project which will cause the project to not meet expectations. A risk requires management assessment and a strategy for its mitigation.

Role A skill set for resources assigned to a project.

Row An entry in a table that typically corresponds to an instance of some real thing, consisting of a set of values for all mandatory columns and relevant optional columns. A row is often an implementation of an instance of an entity.

S

Sample A statistically-significant subset selected and analyzed to estimate the characteristics of a larger group or population; a set of individuals within an organization assessed to provide information on the preferences, opinions, attitudes, and practices of the group they represent.

Scaleability The ability to increase volumes of data and numbers of users to the data warehouse solution. This is a critical capability for the data warehouse architecture and technical architecture.

Schema An information model implemented in a database. A schema may be a logical schema, which will define, for example, tables, columns, and constraints, but which may not include any optimization. It may be a physical schema that includes optimization, for example, table clustering.

Scope The boundaries of a project expressed in some combination of geography, organization, applications and/or business functions.

Scope Change A change to project scope. A scope change requires an adjustment to the project work plan, and nearly always impacts project cost, schedule or quality.

Scope Creep The common phenomenon where additional requirements are added after a project has started without reconsidering the resourcing or timescale of the project. Scope creep arises from the misapprehension that such small additions will not affect the project schedule.

Scoping Workshop A workshop, usually attended by the project sponsor and developers, with the objective of defining the boundaries of the scope for an intended project and prioritizing requirements within the scope.

Security Profile A list of role-based security specifications.

Service Level Agreement (SLA) A binding contract which formally specifies end-user expectation about the solution and tolerances. It is a collection of service level requirements that have been negotiated and mutually agreed upon by the information providers and the information consumers. The SLA has three attributes: STRUCTURE, PRECISION, AND FEASIBILITY. This agreement establishes expectations and impacts the design of the components of the data warehouse solution.

Sign-off Agreement with a client of the successful completion of a project, project phase, or deliverable.

Scalable:The attribute or capability of a database to significantly expand the number of records that it can manage. It also refers to hardware systems and their ability to be expanded or upgraded to increase their processing speed and handle larger volumes of data.

Snowflake Schema:A common form of dimensional model. In a snowflake schema, different hierarchies in a dimension can be extended into their own dimensional tables. Therefore, a dimension can have more than a single dimension table.

Source Module A physical program unit. An application system's repository of source code is controlled at the source module level.

Source System The internal or external computer system which provides the source data for the warehouse.

Stakeholder A person, group, or business unit that has a share or an interest in a particular activity or set of activities.

Standard A set of rules for ensuring quality. Usually standards are defined for products, deliverables or deliverable components and processes.

Structured Query Language (SQL) The ANSI internationally accepted standard for relational database systems, covering not only query but also data definition, manipulation, security, and some aspects of referential and entity integrity.

Subject Area An area of major interest or importance to the enterprise. It is centered on a major resource, product,

or activity. The subject areas provide reference information when conducting detailed requirements gathering.

Success Criteria The metrics and measurements established to determine whether the data warehouse solution has satisfied its objectives and met the requirements.

Summary Data The data that has been aggregated or transformed from the atomic level data. Summary data may reside in all of the database objects of the data warehouse.

Star Schema:A common form of dimensional model. In a star schema, each dimension is represented by a single dimension table.

Synchronization:The process by which the data in two or more separate database are synchronized so that the records contain the same information. If the fields and records are updated in one database the same fields and records are updated in the other

Synonym 1. A name assigned to a table or view that may then be used more conveniently for reference. 2. An alternate name for an entity.

System Test A project activity that tests an application system over its complete life-cycle, using scripts and associating scenario test specifications into chronological sequences.

T

Table A tabular view of data, on a relational database management system, defined by one or more columns of data and a primary key. A table populated by rows of data.

Tablespace A logical portion of a database used in allocating storage for table data and table indexes.

Target Database The data warehouse database object that is to store the source data once it is extracted, transformed and transported.

Technique A specific approach to performing a task. A methodical means of handling and communicating complex details.

U

Usability That quality of a system that makes it easy to learn, easy to use and encourages the user to regard the system as a positive help in getting the job done. User A person who uses a system to perform a business function.

User Preferences In many circumstances in computer systems there may be alternate ways a user can influence the

behavior of a utility, user interface, or other system process. Typically set by adjusting values in a set of user preferences; for example, in a program generator, preferences may be set for style, performance, user interface

behavior and code standards.

User Review A meeting at which some of the facilities of a system are demonstrated to and reviewed by the user. The objective of a user review is to elicit feedback on which to base future development and improvement of the facilities being reviewed.

Utility A program or system facility that performs a useful job for the users. It does not require the user to provide any interaction other than perhaps initially requesting the utility.

V

View A means of accessing a subset of data in a database.

Vision Session A technique used to identify opportunities, issues, and priorities for the client's data warehouse environment. These sessions are conducted during definition and establish the enterprise view of the warehouse as well as the prioritization of the increments.

Workshop 1. A meeting attended by users and developers to create a plan, specification or other documentation that can guide the developers in their development tasks. 2. A meeting designed to facilitate interaction and the exchange of information between individuals or groups
ata Naming Standards Working Group