

ERP Glossary

ABC Classification – A sorting of the items in an inventory in decreasing order of annual value or other criteria. This array is then split into three classes, called A, B, and C. Class A contains the items with the highest annual value and receives the most attention. The medium Class B receives less attention and Class C, which contains the low value items, is controlled routinely. The ABC principle is that effort saved through relaxed controls on low-value items will be applied to reduce inventories of high-value items.

Abstract (Data) – Combine several items into one, summarize, perform computation and take out a sub-fact from one item to generate a new attribute and other actions that change data.

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

Accuracy – Accuracy is an important factor in assessing the success of data mining. When applied to data, accuracy refers to the rate of correct values in the data. When applied to models, accuracy refers to the degree of fit between the model and the data. This measures how error-free the model's predictions are. Since accuracy does not include cost information, it is possible for a less accurate model to be more cost-effective.

Action Message – An ERP output that identifies the need for and the type of action to be taken to correct a current or a potential problem. Examples of action messages are release order, re-schedule out and cancel.

Action Plan – A document used to guide the implementation of business process improvements. It contains task assignments, schedules, resource allocations, assignments and evaluation criteria.

Activation function – A function used by a node in a neural net to transform input data from any domain of values into a finite range of values. The original idea was to approximate the way neurons fired, and the activation function took on the value 0 until the input became large and the value jumped to 1. The discontinuity of this 0-or-1 function caused mathematical problems and sigmoid-shaped functions (e.g., the logistic function) are now used.

Activity – A named process, function or task that occurs over-time and has recognizable results. Activities use up assigned resources to produce products and services. Activities combine to form business processes.

Activity Accounting – The collection of financial and operational performance information about significant activities of an enterprise.

Activity Analysis – The breakdown of an enterprise into manageable segments for detailed analysis regarding cost and performance.

Activity Dependence – An activity intermeshed with other activities in such a manner that the first (i.e., dependent) activity cannot be executed until one or more outputs of other activities within the process have been received.

Activity Diagram – A graphic presentation of all or part of an IDEF0 activity model. Three possible types include node tree, context diagram and decomposition diagram.

Activity Measure – A performance value assigned to an activity's primary output.

Activity Model – A graphical representation of a business process, which exhibits the activities that make up the business process to any desired level of detail. An activity model reveals the interactions between activities in terms of inputs and outputs while showing the controls placed on each activity and the types of resources assigned to each activity.

- Activity Model (As-Is)** – An activity model that portrays how a business process is currently structured. It is used to establish a baseline for subsequent business process improvement actions or programs.
- Activity Model (To-Be)** – An activity model that results from a business process re-designed action or program. The 'to-be' model shows, how the business process will function after the improvement action is implemented.
- Activity Modeling** – A graphical representation of a business process that exhibits the activities and their interdependence that makes up the business process to any desired level of detail. An activity model reveals the interactions between activities in terms of inputs and outputs while showing the controls placed on each activity and the types of resources assigned to each activity.
- Activity, Non-Value-Added** – Any activity that provides a negative return on the investment or allocation of resources to that activity. Within broad limits, the enterprise benefits by allocating lesser resources to non-value-added activities.
- Activity, Value-Added** – Any activity that contributes directly to the performance of a mission and cannot be eliminated without impairing the mission.
- Activity Based Costing** – This is a form of cost accounting that focuses on the costs of performing specific functions (processes, activities, tasks, etc.) rather than, on the costs of organizational units. Activity based costing generates more accurate cost and performance information related to specific products and services than is available to managers through traditional cost accounting approaches.
- Activity Based Management** – A system of management that seeks to optimize the value-added activities performed by the enterprise while at the same time minimizing or eliminating the non-value-added activities, resulting in overall improvements in the effectiveness and the efficiency of the enterprise in serving its customers.
- Ad Hoc Query Processing** – The process of extracting and reporting information from a database through the issuance of a structured query. Programmers usually write queries using special languages that are associated with database management systems. Most relational database managers use a variant of 4GL (4th generation language originally developed by IBM). An example of an ad hoc query might be, "how many customers called the UK between the hours of 6–8 am?" There are several packages available that make the construction of queries more user-friendly than writing language constructs. These usually employ some sort of graphic/ visualization front-end.
- ADAPT** – Architecture Design, Analysis and Planning Tool. This tool provides a complete picture of an information system infrastructure including computers, communication links and devices, applications and databases. ADAPT is based on COTS software and can be also used as a management tool to support inventory control etc.
- ADC** – Automated Data Collection. Technologies that automate data collection at the source such as barcode, biometrics, machine vision, magnetic stripe, OCR (optical card readers), voice recognition, smart cards, and RFID (radio frequency identification).
- 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.
- Advanced Planning System (APS)** – A decision support tool that employs enhanced mathematical and statistical capabilities, a powerful simulation capability, and/ or other advanced techniques to help provide superior plans and schedules.
- Agent** – A program to perform standard functions autonomously. Typical uses are for automated data loading and exception reporting. Not to be confused with data mining.

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

Aggregate tables – Tables (usually held in a relational database) that hold pre-computed totals in a hierarchical multi-dimensional structure.

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

Allocation – In MRP, an allocated item is one for which a picking order has been released to the stockroom but not yet sent out of the stockroom. It is an un-cashed stockroom requisition.

Analytical Model – A structure and process for analyzing a dataset. For example, a decision tree is a model for the classification of a dataset.

Analytical Tools – This is an umbrella phrase used to connote software that employs some sort of mathematical algorithm(s) to analyze the data contained in the warehouse. Data mining, OLAP, ROLAP and other terms are used to designate types of these tools with different functionality. In practice however, a given analytical tool may provide more than one type of analysis procedure and may also encompass some middleware functionality. Such tools are difficult to classify.

Anomalous Data – Data that result from errors (for example, data entry keying errors) or that represent unusual events. Anomalous data should be examined carefully because it may carry important information.

Antecedent – When an association between two variables is defined, the first item (or left-hand side) is called the antecedent. For example, in the relationship, "when a prospector buys a pick, he buys a shovel 14% of the time," "buys a pick" is the antecedent.

Anticipated Delay Report – A report, normally issued by both manufacturing and purchasing to the master scheduling or material planning functions, regarding jobs or purchase orders that will not be completed on-time, explaining why not and telling when they will be completed. This is an essential ingredient of a closed-loop system.

API – Application programming interface: An interface that is used by one application program to communicate with programs of other systems. ERP vendors provide APIs for integrating other applications with their ERP systems.

APL – A programming language. The first multi-dimensional computer language and the title of a book by Ken Iverson, published in 1962. APL is still in limited use but is rarely used for new systems.

Architecture – An organized framework consisting of principles, rules, conventions and standards which serve to guide development and construction activities such that all components of the intended structure will work together to satisfy the ultimate objective of the structure.

Artificial Neural Networks – Non-linear predictive models that learn through training and resemble biological neural networks in structure.

ASCII – American standard code for information interchange. An eight-bit code for character representation; includes seven bits plus parity.

As-Is Model – A model that represents the current state of the organization modeled, without any specific improvements included. Can be either an activity model or rule model.

Assemble-to-Order – A process where the final products are finished to customers' configurations out of standard components. Many personal computers are produced and sold on an assemble-to-order basis.

Associations – An association algorithm creates rules that describe how often events have occurred together. For example, "when prospectors buy picks, they also buy shovels 14% of the time." Such relationships are typically expressed with a confidence interval.

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.

- Attribute** – A property or characteristic of an entity. An attribute has a name and a value. Attributes are used to identify and distinguish between entities and to provide descriptions of entities. Attributes are named with singular, generic nouns.
- 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.
- Automatic Rescheduling** – The process that allows the computer to automatically change due dates on scheduled receipts when it detects that due dates and required dates are out of phase. Automatic re-scheduling is usually not a good idea.
- Available-to-Promise (ATP)** – The uncommitted portion of inventory and/ or future production. This figure is frequently calculated from the master schedule and is used as the primary tool for order promising. See capable-to-promise.
- Back Propagation** – A training method used to calculate the weights in a neural net from the data.
- Back Scheduling** – This is a technique for calculating operations start and due dates. The schedule is calculated starting with the due date for the order and working backward to determine the required completion dates for each operation. This technique is used primarily in job shops.
- Back Flush** – The deduction from inventory of the components used in production by exploding the bill of materials by the count of parent items produced.
- Backlog** – All the customer orders that have been received but not yet shipped, irrespective of when they are specified for shipment.
- Base Tables** – The normalized data structures maintained in the targeted warehousing database. Also known as the detail data.
- Baseline** – A standard for comparisons. A baseline is a reference position for measuring progress in-process improvement. The baseline is usually used to differentiate between a current and a future representation.
- Batch (Transport)** – Data can be processed for quality or distributed either in a set of many transactions: "batch" mode, or as individual transactions.
- Benchmarking** – A method of measuring processes against those of recognized leaders. It helps establish priorities and targets leading to process improvement. It is undertaken by identifying processes to benchmark and their key characteristics; determining who to benchmark; collecting and analyzing data from direct contact, surveys, interviews, technical journals and advertisements; determining the "best of class" from each benchmark item identified; and evaluating the process in terms of the benchmarks set and the improvement goals.
- Best Practice** – A way or method of accomplishing a business function or process that is considered to be superior to all other known methods.
- 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.
- Bill of Activity (BOA)** – A listing of activities involved in the production of a certain output (cost object), together with the number of times each component activity is performed. When these factors are extended by the cost of each component activity, the result is the total unit cost of the final cost object.
- Bill of Material (BOM)** – A listing of all the sub-assemblies, intermediates, parts and raw materials that go into a parent item, showing the quantity of each component required.
- Binning** – A data preparation activity that converts continuous data to discrete data by replacing a value from a continuous range with a bin identifier, where each bin represents a range of values. For example, age could be converted to bins such as 20 or under, 21-40, 41-65 and over 65.
- Bolt-On** – A software application that performs specific tasks and that interfaces with an ERP system.

Bootstrapping – Training data sets that are created by re-sampling with replacement from the original training set, so data records may occur more than once. In other words, this method treats a sample as if it were the entire population. Usually, final estimates are obtained by taking the average of the estimates from each of the bootstrap test sets.

BPR – Business process reengineering is a radical improvement approach that critically examines, re-thinks and re-designs mission product and service processes within a political environment. It achieves dramatic mission performance gains from multiple customer and stakeholder perspectives. It is a key part of a process management approach for optimal performance that continually evaluates, adjusts or removes processes.

BPR Methodology – The BPR methodology is a structured sequence of activities that constitutes the typical BPR project. The typical BPR methodology develops an enterprise level model; identifies scope performance measure, opportunities and constraints; defines the current process and measures cost; benchmarks, analyzes and defines improvement; eliminates no-value-added activities; defines improved process (including measurement, cost and simulation); prototypes and field tests; prepares business case analysis; and implements the planned improvement.

BPR Principles – The principles of BPR form the foundation for achieving dramatic mission performance gains from multiple customer and stakeholder perspectives. These principles include: top management must be supportive of and engaged in reengineering efforts to remove barriers and drive success; an organization's culture must be receptive to reengineering goals and principles; major improvements and savings are realized by focusing on the business from a process rather than a functional perspective; processes should be selected for reengineering based on a clear notion of customer needs, anticipated benefits, and potential for success; process owners should manage reengineering projects with teams that are cross-functional, maintain a proper scope, focus on customer metrics, and enforce implementation timelines.

BPR Professionals – The individuals necessary to a business process reengineering effort. Five distinct skill areas have been identified as being essential to the task. They are: functional experts, strategic planners, data and process modelers, cost analysts, and functional activity coordinators (FAC).

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

Browser – A software application that is used to allow users to read an HTML document in a user-friendly readable format. Well-known browsers are Microsoft Internet Explorer, Mosaic and Netscape.

Bucketless System – An MRP, DRP or other time-phased system in which data are processed, stored and displayed using dated records rather than defined time periods or buckets.

Bucketed System – An MRP, DRP or other time-phased system in which data are accumulated into time periods or buckets. If the period of accumulation were to be one week, then the system would be said to have weekly buckets.

Bulk (Transport) – When data is moved in bulk, the entire database is refreshed periodically. The opposite strategy is to selectively refresh the database with data changes.

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 Case – A structured proposal for business process improvement that functions as a decision package for enterprise leadership. A business case includes an analysis of business process needs or problems, proposed solution, assumptions and constraints, alternatives, life cycle costs, benefits/ cost analysis, and investment risk analysis. In some government agencies a business case is called a functional economic analysis (FEA).

Business Consultant – A consulting services organization that assists in creating the functional specification of a DW/ BI system from the user's perspective. In a broader context, the business consultant helps design or architect the system from the business perspective.

- 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 Intelligence (BI)** – A phrase coined by (or at least popularized by) Gartner Group that covers any computerized process used to extract and/ or analyze business data.
- 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 Modernization** – The state of habitual practice of contemporary and proven superior business practices, typically achieved through elimination, simplification, consolidation, integration, and finally, automation of an organization's business processes and methods.
- Business Plan** – A statement of income projections, costs and profits usually accompanied by budgets and a projected balance sheet as well as a cash flow statement. It is usually stated in dollars. The business plan and the sales and operations plan, although normally stated in different units of measure, should be in agreement with each other.
- Business Process** – The formal or usual way in which business or an activity within the business is conducted. It is a collection of activities that work together to produce a defined set of products and services. All business processes in an enterprise exist to fulfill the mission of the enterprise. Business processes must be related in some way to mission objectives.
- Business Process Improvement (BPI)** – The betterment of an organization's business practices through the analysis of activities to reduce or eliminate non-value-added activities or costs, while at the same time maintaining or improving quality, productivity, timeliness, or other strategic or business purposes as evidenced by measures of performance. Also called functional process improvement.
- Business Process Re-design** – The action of analyzing 'as-is' activity models with the intent to construct a 'to-be' activity model that will yield potential improvements in the performance of the business process.
- Business Process Repository** – A shared database for storing, retrieving and interrelating business models.
- Business Rule Model (Data Model)** – A graphical representation of an organization's information and data assets expressed in terms of entities and relationships. Relationships are called business rules because they enable or constrain business actions. Rule models, like activity models, have 'as-is' and 'to-be' representations.
- Business Rules** – Specific English sentences that describe and specify the constraints the environment imposes on the data relationships. The parent entity name is the subject, the relationship name is the verb, and the child entity is the object. Another relationship in which the child entity name is the subject and the parent entity name is the object follows that sentence. Business rules are used with key-based and fully attributed rule models.
- 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.
- Buyer values** – The underlying values of customers that lead to buying decisions. Getting at these values requires digging more deeply than you will get through standard mail-in survey forms. For example, a car customer may say they want a bigger car. However, what they may really value is the security they feel riding in a bigger car.
- CAD/ CAM** – The integration of computer-aided design and computer-aided manufacturing to achieve automation from design through manufacturing.
- Calculated Member** – A calculated member is a member of a dimension whose value is determined from other members' values (e.g., by application of a mathematical or logical operation). Calculated members may be part of the OLAP server database or may have been specified by the user during an interactive session. A calculated member is any member that is not an input member.
- Capable-to-Promise** – An advanced form of available-to-promise (ATP). ATP looks at future production as specified by the master schedule. Capable-to-promise goes farther. It also looks at what could be produced,

out of available material and capacity, even though not formally scheduled. This capability is sometimes found in advanced planning systems (APS).

Capacity Requirements Planning – The process of determining how much labor and/or machine resources are required to accomplish the tasks of production and making plans to provide these resources. Open production orders, as well as planned orders in the MRP system, are input to CRP which translates these orders into hours of work by work center by time period. In earlier years the computer portion of CRP was called infinite loading, a misnomer. This technique is used primarily in job shops.

CART – Classification and Regression Trees. A decision tree technique used for classification of a data set. Provides a set of rules that you can apply to a new (unclassified) data set to predict which records will have a given outcome.

CASE – Computer-Aided Software Engineering.

Categorical Data – Categorical data fits into a small number of discrete categories (as opposed to continuous). Categorical data is either non-ordered (nominal) such as gender or city, or ordered (ordinal) such as high, medium, or low temperatures.

Category Management Application – Analysis of purchase trends and volume/revenue statistics to determine product and product category patterns. In the retail industry, results relate to issues such as store design and layout and new product introduction strategies.

Cell – A single data point that occurs at the intersection defined by selecting one member from each dimension in a multi-dimensional array. For example, if the dimensions are measures, time, product and geography, then the dimension members—sales, January 2007, chocolate bars and India—specify a precise intersection along all dimensions that uniquely identifies a single data cell, which contains the value of chocolate bar sales in India for the month of January 2007.

Cellular Manufacturing – A method of organizing production equipment, which locates dissimilar equipment together. The goal is to produce items from start to finish in one sequential flow as opposed to a traditional job shop (functional) arrangement, which requires moves and queues between each operation.

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. It is a style of computing where all the information systems are located and managed from a single physical location.

CHAID – Chi Square Automatic Interaction Detection. A decision tree technique used for classification of a dataset. Provides a set of rules that you can apply to a new (unclassified) dataset to predict which records will have a given outcome. Segments a dataset by using chi square tests to create multi-way splits. Preceded, and requires more data preparation than CART.

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.

Change Management – Change management is the balanced management of the resources (human and technical) associated with the change initiative. It is about people leading the change effort and those who are expected to implement the new strategies. It is concerned with the organizational culture and context in which change can occur and the management of the emotional connections essential for a successful transformation. A number of strategies involved in change management include education, training and communications.

Chi-Squared – A statistic that assesses how well a model fits the data. In data mining it is most commonly used to find homogeneous sub-sets for fitting categorical trees as in CHAID.

Churn Analysis Application – Profiling groups of customers in order to understand issues that impact attrition and customer retention.

Classic Data Warehouse Development – The process of building an enterprise business model, creates a system data model, defining and designing a data warehouse architecture, constructing the physical database and lastly, populating the warehouses database.

Classification – Refers to the data mining problem of attempting to predict the category of categorical data by building a model based on some predictor variables.

Classification – The process of dividing a data set into mutually exclusive groups such that the members of each group are as "close" as possible to one another and different groups are as "far" as possible from one another, where distance is measured with respect to specific variable(s) you are trying to predict. For example, a typical classification problem is to divide a database of companies into groups that are as homogeneous as possible with respect to creditworthiness variables with values "good" and "bad."

Classification Tree – A decision tree that places categorical variables into classes.

Cleaning (Cleansing) – Refers to a step in preparing data for a data mining activity. Obvious data errors are detected and corrected (e.g., improbable dates) and missing data is replaced.

Client – A single user computer which is connected via a network to, and works in conjunction with, one or more shared computers (called servers), with data storage and processing distributed between them in one of a number of ways. An OLAP client will not normally store data but will do some processing and most of the presentation work, with the server handling data storage and the rest of the processing. A client may be single or multi-tasking.

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 Architecture – A 'networked' environment where a smaller system such as a PC interacts with a larger, faster system. This allows the processing to be performed on the larger system that frees the user's PC. The larger system is able to connect and disconnect from the clients in order to more efficiently process the data.

Client/ Server Processing – A form of co-operative 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.

Closed-Loop MRP – The second step in the evolution of ERP. This is a set of business processes built around material requirements planning and also including the additional planning functions of production planning, master scheduling, and capacity requirements planning. Further, once the planning phase is complete and the plans have been accepted as realistic and attainable, the execution functions come into play. These include the plant floor control functions of input-output measurement, dispatching, plus anticipated delay reports from the plant and suppliers, supplier scheduling, and so forth. The term closed-loop implies that not only is each of these elements included in the overall system but also that there is feedback from the execution functions so that the planning can be kept valid at all times.

Clustering – The process of dividing a data set into mutually exclusive groups such that the members of each group are as "close" as possible to one another and different groups are as "far" as possible from one another, where distance is measured with respect to all available variables. Clustering algorithms find groups of items that are similar. For example, clustering could be used by an insurance company to group customers according to income, age, types of policies purchased and prior claims experience. It divides a data set so that records with similar content are in the same group and groups are as different as possible from each other. Since the categories are unspecified, this is sometimes referred to as unsupervised learning.

Collaborative Work Technology – This is the term used to describe electronic "groupware" products, which are focused on supporting the interaction and coordination of information, ideas and opinions within a work group. This may include planning, discussion, brainstorming, collaborative design or writing, prioritizing or just about anything that people can do together.

Collection – A set of data that resulted from a DBMS query.

Common Parts BOM – A type of planning bill which groups all common components for a product or family of products into one bill of material.

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

Consolidate – Multi-dimensional databases generally have hierarchies or formula-based relationships of data within each dimension. Consolidation involves computing all of these data relationships for one or more dimensions for example, adding up all departments to get total division data. While such relationships are normally summations, any type of computational relationship or formula might be defined.

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).

Context – The statement of the purpose, objectives and point of a modeling effort. It describes the part of the organization and its functions that the modeling team will concern itself with and by implication, what it will not be concerned with.

Continuous – Continuous data can have any value in an interval of real numbers. That is, the value does not have to be an integer. Continuous is the opposite of discrete or categorical.

Continuous Process Improvement – A policy that encourages, mandates, and/ or empowers employees to find ways to improve process and product performance measures on an ongoing basis.

Continuous Replenishment (CR) – Often called CRP for continuous replenishment process or program. The practice of partnering between distribution channel members that changes the traditional replenishment process from distributor generated purchase orders, based on economic order quantities to the replenishment of products based on actual and forecasted product demand.

Copy Management – A process that takes all or a snapshot of data from a source environment and copies that data to a new target environment.

Core Technologies Competencies – The methodologies and tools necessary to complete a BPR effort.

Cost – The price or imputed value of each resource assigned to an activity that is consumed in the process of producing the products and services of that activity.

Cost-Benefit Analysis – The analysis of the business benefit realized by the cost of expenditure on some resource, tool or application development.

Cost Center – A function in a business where the cost of producing a product or service is tracked and personnel are held accountable for performance.

Cost Driver – The factor(s) that cause work to be performed. Determination of negative cost drivers—those factors that cause non-value-added work to be performed—is the essence of activity based costing.

Cost Object – The single primary output of a certain process in an activity based costing analysis. A cost object may be a product, a process, a service, an event, etc.

Critical Business Issues – Enterprise-wide challenges ("pain" in solution selling terminology) being addressed at a strategic and executive level within an organization. Examples include increasing profitability, customer service, competitive edge, reducing costs, risks, etc.

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

Cross Validation – A method of estimating the accuracy of a classification or regression model. The data set is divided into several parts, with each part in turn used to test a model fitted to the remaining parts.

Cross-Dimensional Formula – Formulae with all operands within a dimension are common, even in non-OLAP systems: e.g. profit = sales – expense, might appear in a simple spreadsheet product. In an OLAP system, such a calculation rule would normally calculate profit for all combinations of the other dimensions in the cube (e.g. for all products, for all regions, for all time periods, etc.) using the respective revenue and expense data from those same dimensions. Part of the power of an OLAP system is the extensive multi-dimensional application of such a simply stated rule, which could be specified by the OLAP application

builder or created by the end-user in an interactive session. The true analytical power of an OLAP server however, is evidenced in its ability to evaluate formulae where there are members from more than one dimension.

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

Cube – See multi-dimensional array.

Cumulative Lead-Time – The longest time involved to accomplish the activity in question. For any item planned through MRP it is found by reviewing each bill of material path below the item, and whichever path adds up the greatest number defines cumulative lead-time. Cumulative lead-time is also called aggregate lead-time, stacked lead-time, composite lead-time, or critical path lead-time.

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

Customer – The recipient of an output product or service. May be internal or external to the organization.

Customer Analysis – A customer analysis is the collection and dissemination of market intelligence about who the customers are and their needs. A customer analysis includes both quantitative data (demographics, satisfaction metrics, competitive ratings, etc.) and qualitative (customer profile, behavior patterns, focus group results, etc.). A customer analysis is a critical element of strategic planning, BPR and TQM.

Customer Information File (CIF) – A record designed to hold information about a single customer in a format suitable for DSS.

Customer Marketing Application – The process of acquiring, retaining and extending customers that places the customer at the center of the business model. Includes such things as market segmentation, profitability analysis and call center analysis.

Customer segmentation or profiling – The process of analyzing and categorizing a broad base of customers into specific segments based on customer needs, interests or buying trends in order to provide more targeted, personalized service. The ultimate segment is one customer.

Cycle Counting – A physical inventory taking technique where inventory is counted on a periodic schedule rather than once a year. For example, a cycle inventory count may be taken when an item reaches its re-order point, when new stock is received, or on a regular basis, usually more frequently for high-value fast-moving items and less frequently for low-value or slow-moving items. Most effective cycle counting systems require the counting of a certain number of items every working day.

Dampeners – A technique within MRP used to suppress the reporting of certain action messages created during the computer processing of MRP. Extensive use of dampeners is not recommended.

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 Administration – The application of a consistent set of disciplines and techniques to a definition, organization, operation, supervision and protection of data.

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 Cleansing – The process of ensuring that all values in a data set are consistent and correctly recorded. It is a method to check data for adherence to standards, internal consistency, referential integrity, valid domain and to replace/ repair incorrect data with correct data. For example, replacing an invalid pin code with a pin code derived from the state/ city information. Checks data quality and scrubs data by some combination of: look-up against valid data (e.g. a list of millions mailing addresses), look-up against domain values (e.g. a list

of valid pin codes), domain range checks (e.g. employees less than 15 or greater than 90 years old), consistency checks among table data, pattern analysis of exceptions, correlations and frequency distributions.

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 Content Quality – The accuracy and validity of the actual values of the data in contrast to issues of data structure and database design.

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. It is a catalog of all data elements, containing their names, structures and information about their usage and 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 Directory – A collection of definitions, rules and advisories of data, designed to be used as a guide or reference with the data warehouse. The directory includes definitions, examples, relations, functions and equivalents in other environments.

Data Element – The most elementary unit of data that can be identified and described in a dictionary or repository, which cannot be sub-divided.

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

Data Extraction, Cleansing and Transformation Process – The process by which data is extracted from an operational database, cleaned and then transformed into a format useful for a data warehouse based application.

Data Filter – Process to check data for adherence to standards, consistency, valid domain; then either clean or reject invalid data.

Data Format – Data items can exist in many formats such as text, integer and floating-point decimal. Data format refers to the form of the data in the database.

Data Integration – Agreement of data messages between senders and receivers, enforced by business rules.

Data Loading – The process of populating the data warehouse. DBMS specific load processes, DBMS insert processes, and independent fast load processes provide the data loading.

Data Mall – A data mall is a collection of data marts.

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 Mart (DM) – A data mart is a data warehouse that is restricted to dealing with a single subject or topic. The operational data that feeds a data mart generally comes from a single set or source of operational data.

Data Merging – Combining two or more data sets; values or structures. See abstract.

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 – An information extraction activity whose goal is to discover hidden facts contained in databases. Using a combination of machine learning, statistical analysis, modeling techniques and database

technology, data mining finds patterns and subtle relationships in data and infers rules that allow the prediction of future results. 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 mining is a process by which the computer looks for trends and patterns in the data and flags potentially significant information. An example of a data mining query might be, "what are the psychological factors associated with child abusers?" Typical applications include market segmentation, customer profiling, fraud detection, evaluation of retail promotions and credit risk analysis.

Data Mining Method – Procedures and algorithms designed to analyze the data in databases.

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 Model (Business Rule Model) – A graphical representation of an organization's information and data assets expressed in terms of entities and relationships. Relationships are also called business rules because they enable or constrain business actions. Data models, like activity models, have 'As-is' and 'To-be' representations.

Data Navigation – The process of viewing different dimensions, slices and levels of detail of a multi-dimensional database. See OLAP.

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 Repository – A specialized database containing information about data and data relationships. Used to provide a common resource of standard data elements and models.

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 – Involves 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 Visualization – The visual interpretation of complex relationships in multi-dimensional data.

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 facilitates forecasting and planning efforts. Data mining and data warehousing go hand-in-hand. Before you can effectively mine information, you have to put it all in one place—at least temporarily. Data warehousing involves integrating information

from different systems, functions and locations across an enterprise into a central database to allow more accurate analysis of customer needs, buying patterns and profitability and improved decision-making and marketing.

Data Warehouse (DW) – A repository for data organized in a format that is suitable for ad hoc query processing, data mining, OLAP and/ or other analytical applications. Data warehouses are built from operational databases. The operational data is "cleaned" and transformed in such a way that it is amenable to fast retrieval and efficient analysis. A single purpose data warehouse is sometimes referred to as a "data mart."

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 Engines – Refers to 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 – Delivering 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 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 – 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 – A collection of related data, organized to serve one or more independent applications, stored with security, privacy and integrity controls.

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 – Describes a class of systems or solutions that reveal meaningful trends and patterns within an enterprise's raw data, allowing people to make ad hoc inquiries for information they need to make more accurate decisions. Data mining and data warehousing are examples of decision support applications.

Decision Support System (DSS) – DSS is an umbrella expression that encompasses ad hoc query, data mining, OLAP/ ROLAP, vertical applications and, in the minds of at least some, the data warehouse as well. DSS appears to be falling into disuse in some circles and is being replaced with business intelligence (BI).

Decision Tree – A tree-shaped structure that represents a set of decisions. These decisions generate rules for the classification of a data set. See CART and CHAID.

Decomposition – The process of breaking down an activity into lower level component activities.

Deduction – Deduction infers information that is a logical consequence of the data.

Degree of Fit – A measure of how closely the model fits the training data. A common measure is r-square.

- Delta Update** – Refers to only the data that was updated between the last extraction or snapshot process and the current execution of the extraction or snapshot.
- Demand** – A need for a particular product or component. The demand could come from a variety of sources (i.e., customer order, forecast, interplant, branch warehouse, service part), or to manufacture the next higher level.
- Demand Management** – The function of recognizing and managing all of the demands for products to ensure that the master scheduling function is aware of them. It encompasses the activities of forecasting, order entry, order promising, branch warehouse requirements, interplant requirements, interplant orders and service parts requirements.
- Demonstrated Capacity** – Capacity calculated from actual performance data, usually number of items produced times standard hours per item plus the standard setup time for each job.
- De-Normalized Data Store** – A data store that does not comply with one or more of several normal forms. See normalization.
- Dense** – The majority or significant minority (at least ten percent), of potential data cells actually occupied in a multi-dimensional structure. A multi-dimensional database is dense if a relatively high percentage of the possible combinations of its dimension members contain data values. This is the opposite of sparse.
- Dependent Demand** – Demand is considered dependent when it comes from production schedules for other items. These demands should be calculated, not forecasted. A given item may have both dependent and independent demand at any given time.
- Dependent Variable** – The dependent variables (outputs or responses) of a model are the variables predicted by the equation or rules of the model using the independent variables (inputs or predictors).
- Deployment** – After the model is trained and validated, it is used to analyze new data and make predictions. This use of the model is called deployment.
- Derived Data** – Derived data is produced by applying calculations to input data at the time the request for that data is made, i.e. the data has not been pre-computed, and stored on the database. The purpose of using derived data is to save storage space and calculation time, particularly for calculated data that may be infrequently called for or that is susceptible to a high degree of interactive personalization by the user. The trade-off is slower retrievals.
- Derived Members** – Derived members are members whose associated data is derived data.
- Desktop Applications** – Query and analysis tools that access the source database or data warehouse across a network using an appropriate database interface. It is an application that manages the human interface for data producers and information consumers.
- Desktop OLAP** – Low priced, simple OLAP tools that perform local multi-dimensional analysis and presentation of data downloaded to client machines from relational or multi-dimensional databases.
- Detail Member** – A detail member of a dimension is the lowest level number in its hierarchy.
- Dimension** – Each attribute of a case or occurrence in the data being mined. Stored as a field in a flat file record or a column of relational database table.
- Dimension** – In a flat or relational database, each field in a record represents a dimension. In a multi-dimensional database a dimension is a set of similar entities; for example, a multi-dimensional sales database might include the dimensions product, time and city.
- Direct-Deduct Inventory Transaction Processing** – A method of inventory bookkeeping which decreases the book (computer) inventory of an item as material is issued from stock, and increases the book inventory as material is received into stock. The key concept here is that the book record is updated together with the movement of material out of or into stock. As a result, the book record represents what is physically in stock.
- Discounted Cash Flow** – A method of performing an economic analysis that takes the time value of money into account. Used to remove interest rates and inflation factors from a calculation so that the results of analysis are comparable.

Discrete – A data item that has a finite set of values. Discrete is the opposite of continuous.

Discriminant Analysis – A statistical method based on maximum likelihood for determining boundaries that separate the data into categories.

Dispatch List – A listing of manufacturing orders in priority sequence according to the dispatching rules being used. The dispatch list is usually communicated to the manufacturing floor via hard copy or CRT display and contains detailed information on priority, location, quantity, and the capacity requirements of the manufacturing order by operation. Dispatch lists are normally generated daily or more frequently and oriented by work center.

Distributed Center (DC) – A facility stocking finished goods and/ or service items. A typical company for example, might have a manufacturing facility in Noida and distribution centers in Kochi, Mumbai, Bangalore, Hyderabad and Chennai. A DC serving a group of satellite warehouses is usually called a regional distribution center.

Distributed support services – Services offered by a vendor that help you maximize the availability and effectiveness of the desktop systems, printers, software applications, networking equipment and other desktop assets used throughout an organization.

Distribution Requirements Planning – The function of determining the needs to replenish inventory at distribution centers. A time-phased order point approach is used, where the planned orders at the branch warehouse level are exploded via MRP logic to become gross requirements on the supplying source. In the case of multi-level distribution networks, this explosion process can continue down through the various levels of master DC, factory warehouse and so on, and become input to the master schedule. Demand on the supplying source(s) is recognized as dependent and standard MRP logic applies.

Distribution Resource Planning (DRP) – The extension of distribution requirements planning into the planning of the key resources contained in a distribution system: warehouse space, manpower, money, trucks and freight cars, and so forth.

DLL – Dynamic Link Library: A library of core elements required by the Windows architecture, a DLL contains all the functions and definitions needed to communicate with a program at run time.

DOLAP – Various, Desktop OLAP or Database OLAP. The desktop OLAP variant is more commonly used.

Domain expertise – In-depth knowledge of specific industries—such as financial services, government, communications and transportation—offered by an information services firm.

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.

Drill-Down/Up – Drilling down or up is a specific analytical technique whereby the user navigates among levels of data ranging from the most summarized (up) to the most detailed (down). The drilling paths may be defined by the hierarchies within dimensions or other relationships that may be dynamic within or between dimensions. For example, when viewing sales data for South India, a drill-down operation in the region dimension would then display Andha Pradesh, Tamilnadu, Karnataka, Kerala, etc. A further drill-down on Tamilnadu might display Chennai, Coimbatore, Madurai, etc.

Driver – An activity or condition that has a direct influence on the operational performance or cost structure of other activities.

DSS – Decision Support System. This is application for analyzing large quantities of data and performing a wide variety of calculations and projections.

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.

- Economic Analysis** – A formal method of comparing two or more alternative ways of accomplishing a set objective, given a set of assumptions and constraints and the costs and benefits of each alternative, such that the analysis will indicate the optimum choice.
- Efficient Consumer Response (ECR)** – A strategy in which the retailer, distributor and supplier trading partners work closely together to eliminate excess costs from the supply chain. This is a global movement to enhance the efficiency of product introductions, merchandising, promotions and replenishment.
- EIS** – Variously defined as Executive/Enterprise/Everyone's Information/Intelligence System/Service/Software. It is a category of applications and technologies for presenting and analyzing corporate and external data for management purposes. Extreme ease of use and fast performance is expected in such systems but analytical functionality is usually very limited.
- Electronic Data Interchange (EDI)** – The computer-to-computer exchange of information between separate organizations using specific protocols.
- End-User Data** – Data formatted for end-user query processing; data created by end-users; data provided by a data warehouse.
- Engineer-to-Order Product** – A product that requires engineering design, and bill of material and routing development before manufacturing can be completed. Such products typically require master scheduling of average or typical items or expected activities and capacities, with many individual components being identified only after preliminary design work is complete.
- Enterprise** – When used generically, an enterprise is defined as the aggregate of all functional elements participating in a business process improvement action regardless of the organizational structure housing those functional elements. A complete business consisting of functions, division, or other components used to accomplish specific objectives and defined goals.
- Enterprise Data** – Data that is defined for use across a corporate environment.
- Enterprise Data Model** – Refers to a single collection of data designed to serve the diverse needs of an enterprise. The opposing concept is that of a collection of smallish databases, each designed to support a limited requirement.
- Enterprise Data Repository** – A database containing "metadata" used to control data transformations for DW/BI systems. A leading exponent of the data repository concept is a software company called Platinum Technology.
- Enterprise Data Warehouse** – A single repository holding data from several operational sources that serves many different users, typically in different divisions or departments. An enterprise data warehouse for a large company might for example, contain data from several separate divisions and serve the needs of both those divisions and of corporate users wishing to analyze consolidated information.
- Enterprise Level** – The enterprise level provides the geographic, technological and managerial platform upon which all information systems development activity is based; it is the foundation that must support all that is built above it in the higher levels.
- Enterprise Model** – A high-level model of an enterprise's mission, function, process and information architecture used as a standard reference for constructing data and activity models and information systems.
- Enterprise Modeling** – The development of a common consistent view and understanding of data elements and their relationships across the enterprise.
- Enterprise Resource Planning (ERP)** – ERP predicts and balances demand and supply. It is an enterprise-wide set of forecasting, planning and scheduling tools, which links customers and suppliers into a complete supply chain, employs proven processes for decision-making and coordinates sales, marketing, operations, logistics, purchasing, finance, product development and human resources. Its goals include high-levels of customer service, productivity, cost reduction and inventory turnover, and it provides the foundation for effective supply chain management and e-commerce. It does this by developing plans and schedules so that the right resources—manpower, materials, machinery and money—are available in the right amount when

needed. ERP is a direct outgrowth and extension of MRP and as such, includes all of MRP II's capabilities. ERP is more powerful in that it applies a single set of resource planning tools across the entire enterprise, provides real-time integration of sales, operating and financial data, and connects resource planning approaches to the extended supply chain of customers and suppliers.

Enterprise server – A powerful, high-end system that combines the capabilities of mainframes—security, robustness, resiliency—with the cost advantages and flexibility of open systems. By virtue of their openness, enterprise servers can function as central information servers for multi-vendor environments.

Entity – A set of real-world objects (people, places, things, events, ideas) with characteristics in common and within the scope of a model. Each entity has a name that is a singular noun or noun phrase describing the object it represents. A box represents an entity and the entity name along with the entity number is placed above the box.

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.

Entropy – A way to measure variability other than the variance statistic. Some decision trees split the data into groups based on minimum entropy.

ERP – Enterprise Resource Planning: the current evolution of manufacturing resources planning (MRP & MRP II) systems. ERP is being positioned as the foundation and integration of enterprise-wide information systems. Such systems will link together all of a company's operations including human resources, financials, manufacturing and distribution as well as connect the organization to its customers and suppliers.

ETL – Extraction, Transformation and Loading. These are activities required to populate data warehouses and OLAP applications with clean, consistent, integrated and probably summarized data.

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

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.

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.

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.

Exploratory Analysis – Looking at data to discover relationships not previously detected. Exploratory analysis tools typically assist the user in creating tables and graphical displays.

Exploratory Data Analysis – The use of graphical and descriptive statistical techniques to learn about the structure of a dataset.

Extendibility – The ability to easily add new functionality to existing services without major software re-writes or without re-defining the basic architecture.

External Data – Data not collected by the organization, such as data available from a reference book, a government source or a proprietary database.

Extract – Selects data from various source system platforms.

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.

FASMI – Fast Analysis of Shared Multi-dimensional Information. This is the summary description of OLAP applications proposed by the authors of this report.

Fast Load – A technology that typically replaces a specific DBMS load function. A fast load technology obtains significantly faster load times by pre-processing data and bypassing data integrity checks and logging.

Feed-Forward – A neural net in which the signals only flow in one direction from the inputs to the outputs.

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 sub-set of information in a data warehouse.

Final Assembly Schedule (FAS) – Also referred to as the finishing schedule as it may include other operations than simply the final operation. For make-to-order products it is prepared after receipt of a customer order, is constrained by the availability of material and capacity, and it schedules the operations required to complete the product from the level where it is stocked (or master scheduled) to the end item level.

Financial Analysis and Reporting Application – Includes all types of analysis of all levels of operating financial data to determine financial efficiency and effectiveness within any functional organization.

Finite Loading – Conceptually, the term means putting no more work into a work center than it can be expected to execute. The specific term usually refers to a computer technique that involves automatic plant priority revision in order to level load operation-by-operation. Finite loading is also called finite scheduling.

Formula – A formula is a database object, which is a calculation, rule or other expression for manipulating the data within a multi-dimensional database. Formulae define relationships among members. OLAP database builders use formulae to provide great richness of content to the server database. End-users use them to model enterprise relationships and to personalize the data for greater visualization and insight.

Fraud Detection Application – The process of detecting patterns, trends or correlations in consumer or corporate behavior that might indicate that fraudulent activity is taking place; e.g., identifying potential or existing fraud through analysis and comparison of standard and aberrant behaviors.

Frequency – The timing characteristics of the data.

Function – A specific set of skills and resources that can be used to perform one or more activities that make up a process. Usually, several functions are associated with a single process.

Function Library – A body of ready-made, re-usable units of code for specific programming tasks. They can be implemented in an ERP program or called by external applications.

Functional Activity – A sub-division of a functional area.

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.

Functional Economic Analysis (FEA) – A technique for analyzing and evaluating alternative information system investments and management practices. It is also a document that contains a fully justified proposed improvement project with all supporting data, i.e. business case or decision package. FEA is also called a business case analysis.

Functional Management – A philosophy of management that organizes an enterprise by type of work performed. See also process management.

Functional Process – A sub-division of a functional activity.

Functional Process Improvement (FPI) – A structured approach by all or part of an enterprise to improve the value of its products and services while reducing resource requirements. Also referred to as business process improvement (BPI), business process re-design and business reengineering.

Fuzzy Logic – Fuzzy logic is applied to fuzzy sets where membership in a fuzzy set is a probability, not necessarily 0 or 1. Non-fuzzy logic manipulates outcomes that are either true or false. Fuzzy logic needs to be able to manipulate degrees of "maybe" in addition to true and false.

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

Genetic Algorithms – A computer-based method of generating and testing combinations of possible input parameters to find the optimal output. It uses processes based on natural evolution concepts such as genetic combination, mutation and natural selection.

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.

Group Technology – An engineering and manufacturing approach that identifies the sameness of parts, equipment or processes. It provides for rapid retrieval of existing designs and facilitates a cellular form of production equipment layout.

Groupware – Application for allowing a workgroup to share information, with updates by each user being made available to others, either by automated replication or concurrent sharing.

Hedge – In master scheduling, a quantity of stock used to protect against uncertainty in demand. The hedge is similar to safety stock, except that a hedge has the dimension of timing as well as amount. In purchasing, a purchase or sale transaction having as its purpose the elimination of the negative aspects of price fluctuations.

Heterogeneous Multi-Processing – A Unisys computing architecture for enterprise servers that integrates traditional mainframe software environments with Microsoft NT and UNIX applications on the same platform. This approach protects existing investments in software and technology, while allowing an organization to enjoy the benefits of client/ server computing.

Hierarchical Relationships – Any dimension's members may be organized based on parent-child relationships, typically where a parent member represents the consolidation of the members, which are its children. The result is a hierarchy and the parent-child relationships are hierarchical relationships.

Historical Database – A database that provides an historical perspective on the data.

HOLAP – Hybrid OLAP. A product that can provide multi-dimensional analysis simultaneously of data stored in a multi-dimensional database and in an RDBMS. Becoming a popular architecture for server OLAP.

Horizontal Dimension – See page display.

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.

Hypercube – An OLAP product that stores all data in a single cube, which has all the application dimensions applied to it.

I-CASE – Integrated Computer-Aided Software Engineering. A set of software design and development tools operating with an integrated shared repository to support the entire systems development life cycle.

ICOM – The acronym for the roles for data or material on an activity model. ICOMs are represented by arrows that interconnect activity boxes and they are named using a noun or noun phrase. INPUT (Data or material used to produce an output of an activity.) CONTROL (Data that constrain or regulate the activity. Controls regulate the transformation of inputs into outputs.) OUTPUT (Data or materials produced by or resulting from the activity. It must include the input data in some form.) MECHANISM (Resources—usually people, machines or systems—that provide energy to, or perform the activity.)

IM – Information Management.

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.

Implementation – Implementation refers to the actual installation of the change project's approved recommendations.

Implementation Consultant – A consulting services organization that provides the services required to build the DW /BI system including software development, selection of middleware, implementation of data models and related tasks.

Improvement Initiative – A set or package of planned improvements resulting from the analysis of baseline processes, inspection of strategic and business plans, and benchmarking results that if implemented, will result in-process improvement.

Improvement Opportunities – Refers to situations that can be changed to produce a more effective or more efficient process or product. Improvement opportunities may involve processes, business rules or both. Opportunities are often packaged together as an improvement initiative.

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.

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.

Independent Demand – Demand for an item is considered independent when unrelated to the demand for other items. Demand for finished goods and service parts are examples of independent demand.

Independent Variable – The independent variables (inputs or predictors) of a model are the variables used in the equation or rules of the model to predict the output (dependent) variable.

Index – An index is a link between one table and another; it allows for rapid access to the rows of a table based on the values of one or more columns in another table.

Indirect Costs – Costs common to a multiple set of cost objectives and not directly assignable to such objectives in a specific time period. Such costs are usually allocated by systematic and consistent techniques to products, processes, services or time periods.

Induction – A technique that infers generalizations from the information in the data.

Infinite Loading – See capacity requirements planning.

Info-Glut – Too much data!

Information – Data that has been processed in such a way that it can increase the knowledge of the person who receives it.

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

Information Engineering – The translation of certain types of process requirements into software programs.

Information Management – The treatment of information as an asset to enhance an organization's competitiveness and responsiveness. It results from effectively identifying, collecting, and analyzing information—and then directing it to the points of decision-making and customer service.

Information Management Company – A company that offers a full range of services, technologies and industry expertise needed to help organizations manage information for greater competitiveness and customer responsiveness.

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

Information Services – Services such as management consulting, systems integration and outsourcing designed to help an organization manage information and apply information technology to achieve core business goals.

Information System – An engineered arrangement of computers, communications facilities, software code and data designed to support a business process.

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 Technology (IT) – A package of equipment and/or systems related to data and/or communications that can be used as an enabler of process reengineering.

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

Input Members – Input members have values that are loaded directly from either manual entry or by accessing another computer-based data source, as opposed to being calculated from the raw data.

Input-Output Control – A technique for capacity control where actual output from a work center is compared with the planned output (as developed by CRP and approved by manufacturing). The input is also monitored to see if it corresponds with plans so that work centers will not be expected to generate output when jobs are not available to work on.

Integration – The process of taking discrete products or components and making them work together as if they were one product; or the same concept applied to processes.

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.

Interaction – Two independent variables interact when changes in the value of one change the effect on the dependent variable of the other.

Internal Data – Data collected by an organization such as operating and customer data.

Interoperability – The capacity to integrate technology between or among different technical platforms. This form of integration is achieved through information engineering and translates process requirements into software programs.

Interplant Demand – Material to be shipped to another plant or division within the corporation. Although it is not a customer order, the master scheduling system handles it in a similar manner.

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

Inventory Turnover – The number of times that an inventory turns over during the year. One way to compute inventory turnover is to divide the average inventory level into the annual cost-of-sales. For example, if average inventory were three crore rupees and cost-of-sales were thirty crore, the inventory would be considered to turn ten times per year. Turnover can also be calculated on a forward-looking basis, using the forecast rather than historic sales data.

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.

Investment Justification – A functional economic analysis indicating that it is better to do a certain action than not do it. Investments may be compared and ranked by various criteria, including return on various categories of capital, risk-adjusted discounted cash flow, affordability, internal rate of return, etc.

IT – Information Technology. Sometimes used as a synonym for the computer professionals in an organization.

JIT – Just-in-Time. A policy calling for the delivery of materials, products or services at the time they are needed in an activity or process. Used to reduce inventory, wait time and spoilage.

- Job Shop** – A functional organization whose departments or work centers are organized around particular types of equipment or operation such as drilling, blending, spinning or assembly. Products move through departments by individual production orders.
- 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.
- Just-in-Time (JIT)** – In the broad sense, just-in-time is an approach to achieving excellence in manufacturing. In the narrow (and less correct) sense, JIT is considered by some as a production and logistics method designed to result in minimum inventory by having material arrive at each operation just in time to be used. JIT is used to reduce inventory, wait time and spoilage.
- Kanban** – A method for just-in-time production in which consuming (downstream) operations pull from feeding (upstream) operations. Feeding operations are authorized to produce only after receiving a Kanban card (or other trigger) from the consuming operation. Kanban in Japanese loosely translates to "card."
- Knowledge Discovery** – A phrase coined by (or at least popularized by) Gartner Group defined as the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories (e.g., data warehouses), using such technologies as pattern recognition, statistics and other mathematical techniques. Knowledge discovery is really the same thing as data mining.
- Knowledge Infrastructure** – A set of rules that controls the symbols that can be employed in sending and receiving messages.
- Knowledge Management** - An umbrella term that is used by some in the same context as business intelligence.
- LAN** – Local Area Network. High-speed connections between desktop PCs and server machines, including file servers, application servers, print servers and other services.
- Layer** – Nodes in a neural net are usually grouped into layers, with each layer described as input, output or hidden. There are as many input nodes as there are input (independent) variables and as many output nodes as there are output (dependent) variables. Typically, there are one or two hidden layers.
- Lead-Time** – A span of time required to perform an activity. In a logistics context, the activity in question is normally the procurement of materials and/ or products either from an outside supplier or from one's own manufacturing facility. The individual components of any given lead-time can include some or all of the following: order preparation time, queue time, move or transportation time, receiving and inspection time.
- Lean Manufacturing** – An approach to production that emphasizes the minimization of the amount of all the resources (including time) used in the various activities of the enterprise. It involves identifying and eliminating non-value-adding activities in design, production, supply chain management and dealing with the customers.
- Learning** – Training models (estimating their parameters) based on existing data.
- Least Squares** – The most common method of training (estimating) the weights (parameters) of a model by choosing the weights that minimize the sum of the squared deviation of the predicted values of the model from the observed values of the data.
- Legacy System** – Existing systems and technology in which an organization has considerable investment and which may be entrenched in the organization. Some systems have been in place for many years; some are considered old or inadequate technology; many are host based with terminal emulation. Customers may be trying to replace or merely update legacy systems.
- Life Cycle Management (LCM)** – A management process that governs a process or system from conception to final disposition.
- Linear Model** – An analytical model that assumes linear relationships in the coefficients of the variables being studied.

Linear Regression – A statistical technique used to find the best fitting linear relationship between a target (dependent) variable and its predictors (independent variables).

Load – The amount of scheduled work ahead of a manufacturing facility, usually expressed in terms of hours of work or units of production.

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 sub-sets of certain shared data tables and particular attributes therein, or sub-sets 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.

Logical Data Model – The actual implementation of a conceptual module in a database. It may take multiple logical data models to implement one conceptual data model.

Logical Systems Architect – A person or organization that designs the software for the DW/ BI application.

Logistic Regression – A linear regression that predicts the proportions of a categorical target variable, such as type of customer, in a population.

Logistics – In an industrial context, this term refers to the functions of obtaining and distributing material and product. Also the tracking and movement of raw materials to finished products and consumption throughout the supply chain.

Lot-for-Lot – An order quantity technique in MRP, which generates planned orders in quantities equal to the net requirements in each period.

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

Make-to-Order Product – The end item is finished after receipt of a customer order. Frequently, long lead-time components are planned prior to the order arriving in order to reduce the delivery time to the customer. Where options or other sub-assemblies are stocked prior to customer orders arriving, the term assemble-to-order is frequently used.

Make-to-Stock Product – The end item is shipped from finished goods off-the-shelf and therefore, is finished prior to a customer order arriving.

Manufacturing Resource Planning (MRP II) – The third step in the evolution of ERP. This is a method for the effective planning of the resources of a manufacturing company. It addresses operational planning in units, financial planning in dollars, and has a simulation capability to answer what if questions. MRP II is made up of a variety of functions, each linked together: business planning, sales and operations planning, demand management, master scheduling, material requirements planning, capacity requirements planning and the execution support systems for capacity and material. Output from these tools is integrated with financial reports such as the business plan, purchase commitment report, shipping budget, inventory projections in dollars, and so on. Manufacturing resource planning is a direct outgrowth and extension of closed-loop MRP.

Market Basket Analysis Application – The process of analyzing and correlating sales data to determine allied product purchasing patterns; e.g. if X is in the market basket, what else is likely to be in it?

Master Production Schedule (MPS) – See master schedule.

Master Schedule (MS) – The anticipated build schedule. The master scheduler maintains this schedule and in turn, it drives MRP. It represents what the company plans to produce expressed in specific configurations, quantities and dates. The master schedule must take into account customer orders and forecasts, backlog, availability of material, availability of capacity, management policy and goals.

Material Management (MM) – An organizational structure which groups the functions related to the complete cycle of material flow, from the purchase and internal control of production materials to the warehousing, shipping and distribution of the finished product.

Material Requirement Planning (MRP) – The first step in the evolution of ERP. This is a set of techniques, which uses bills of material, inventory data and the master schedule to calculate requirements for materials. It makes recommendations to release replenishment orders for material. Further, since it is time phased, it makes recommendations to re-schedule open orders when due dates and need dates are not in phase.

Maximum Likelihood – Another training or estimation method. The maximum likelihood estimate of a parameter is the value of a parameter that maximizes the probability that the data came from the population defined by the parameter.

Measurability – One of the characteristics that makes an output suitable for being named as the "primary output" of an activity. Other characteristics include homogeneity and consumption of resources in some direct ratio to the quantity of output.

Member Combination – A member combination is an exact description of a unique cell in a multi-dimensional array, consisting of a specific member selection in each dimension of the array.

Member Dimension – A dimension member is a discrete name or identifier used to identify a data item's position and description within a dimension. For example, January 1989 or 1Qtr 93 are typical examples of members of a time dimension. Wholesale, retail, etc. are typical examples of members of a distribution channel dimension.

MES – Manufacturing Execution Systems. These are systems that use network computing to automate production control and process automation. By downloading "recipes" and work schedules and uploading production results, MESs bridge the gap between business and plant-floor or process-control systems.

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 – 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.

Middleware – An umbrella term used to describe software that bridges various parts of a DW/ DSS system. For example, software that extracts, cleans or separates data.

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.

Migration System – An existing information system that has been officially designated to support standard processes and is intended to be the means of arriving at a target system or architecture (as in open systems architecture).

Mini Marts – A small sub-set of a data warehouse used by a small number of users. A mini mart is a very focused slice of a larger data warehouse.

Mini Cube – A sub-set of a hypercube, with fewer dimensions than the encompassing hypercube. The hypercube will consist of a collection of logically similar mini cubes.

MIP-O-Suction – A query that consumes a high percentage of CPU cycles.

Missing Data – Data values can be missing because they were not measured, not answered, were unknown or were lost. Data mining methods vary in the way they treat missing values. Typically, they ignore the missing values, or omit any records containing missing values, or replace missing values with the mode or mean, or infer missing values from existing values.

Mode – The most common value in a data set. If more than one value occurs the same number of times, the data is multi-modal.

Model – An important function of data mining is the production of a model. A model can be descriptive or predictive. A descriptive model helps in understanding underlying processes or behavior. For example, an association model describes consumer behavior. A predictive model is an equation or set of rules that makes it possible to predict an unseen or unmeasured value (the dependent variable or output) from other, known values (independent variables or input). The form of the equation or rules is suggested by mining data collected from the process under study. Some training or estimation technique is used to estimate the parameters of the equation or rules.

Modular BOM – A type of planning bill, which is arranged in product modules or options. Often used in companies where the product has many optional features (e.g., automobiles, computers).

MOLAP – Multi-dimensional [database] OLAP. We avoid the use of this term because all OLAPs are by definition, multi-dimensional and prefer the more explicit MDB.

MPP – Massively Parallel Processing. This is a computer configuration that is able to use hundreds or thousands of CPUs simultaneously. In MPP each node may be a single CPU or a collection of SMP CPUs. An MPP collection of SMP nodes is sometimes called an SMP cluster. Each node has its own copy of the operating system, memory and disk storage, and there is a data or process exchange mechanism so that each computer can work on a different part of a problem. Software must be written specifically to take advantage of this architecture.

MRP – Materials Requirements Planning. Phases in the development of computerized methods for planning the use of company resources, including scheduling raw materials, vendors, production equipment and processes.

MRP II – Manufacturing Resource Planning.

Multi Cube – An OLAP product that can store data in the form of a number of multi-dimensional structures, which together form an OLAP database. May use relational or multi-dimensional file storage.

Multi-Dimensional – Data structure with three or more independent dimensions.

Multi-Dimensional Analysis – The objective of multi-dimensional analysis is for end-users to gain insight into the meaning contained in databases. The multi-dimensional approach to analysis aligns the data content with the analyst's mental model, hence reducing confusion and lowering the incidence of erroneous interpretations. It also eases navigating the database, screening for a particular sub-set of data, asking for the data in a particular orientation and defining analytical calculations. Furthermore, because the data is physically stored in a multi-dimensional structure, the speed of these operations is many times faster and more consistent than is possible in other database structures. This combination of simplicity and speed is one of the key benefits of multi-dimensional analysis.

Multi-Dimensional Array – A group of data cells arranged by the dimensions of the data. For example, a spreadsheet exemplifies a two-dimensional array with the data cells arranged in rows and columns, each being a dimension. A three-dimensional array can be visualized as a cube with each dimension forming a side of the cube, including any slice parallel with that side. Higher dimensional arrays have no physical metaphor but they organize the data in the way users think of their enterprise. Typical enterprise dimensions are time, measures, products, geographical regions, sales channels, etc.

Multi-Dimensional Data Structure – See multi-dimensional array.

Multi-Dimensional Database – A database designed for on-line analytical processing. Structured as a multi-dimensional hypercube with one axis per dimension.

- 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.
- Multi-Dimensional Query Language** – A computer language that allows one to specify which data to retrieve out of a cube. The user process for this type of query is usually called slicing and dicing. The result of a multi-dimensional query is a cell, a two-dimensional slice or a multi-dimensional sub-cube.
- Multi-Processor Computer** – A computer that includes multiple processors connected by a network. See parallel processing.
- Navigation** – Navigation is a term used to describe the processes employed by users to explore a cube interactively by drilling, rotating and screening, usually using a graphical OLAP client connected to an OLAP server.
- Net Change MRP** – A method of processing MRP on the computer whereby the material plan is continually retained in the computer. Whenever there is a change in requirement open order, or inventory status, bills of material, or the like, a partial re-calculation of requirements is made only for those parts affected by the change.
- Net Requirements** – In MRP, the net requirements for an item are derived as a result of netting gross requirements against inventory on-hand and the scheduled receipts. Net requirements lot sized and offset for lead-time, become planned orders.
- Network Integration** – Describes the services involved in helping an organization design, build and support local-area, wide-area, e-mail, and voice/data-networks that distribute information seamlessly to points of decision-making and service. These networks are typically multi-vendor, using premier hardware and software products from leading suppliers.
- Network/Traffic Pattern Analysis Application** – Normally associated with telecommunications, this application analyzes traffic patterns in order to discover facts about customer behavior, project future demand or to conduct market analysis to determine need for new services. Both, telcos and large private network operators can use the results to reduce network costs and/ or to improve network efficiency by analyzing such issues as capacity and maintenance.
- Neural Network** – This is a complex non-linear modeling technique based on a model of a human neuron. A neural net is used to predict outputs (dependent variables) from a set of inputs (independent variables) by taking linear combinations of the inputs and then making non-linear transformations of the linear combinations using an activation function. It can be shown theoretically that such combinations and transformations can approximate virtually any type of response function. Thus, neural nets use large numbers of parameters to approximate any model. Neural nets are often applied to predict future outcome based on prior experience. For example, a neural net application could be used to predict who will respond to a direct mailing.
- Non-Applicable Data** – Refers to missing values that would be logically impossible (e.g. pregnant males) or are obviously not relevant.
- Non-Value-Added Activity** – A process or activity that results in waste of resources or that can be eliminated or reduced without deterioration of service to customers. For example, a corrective process or activity performed because the function initially failed to comply with customer specifications is a non-value-added process or activity
- Non-Value-Added Cost** – The price or cost of any resource consumed by an activity that does not add value to the product or service. Non-value-added costs are generally the penalty for poor quality or poor decision-making actions in activities "upstream" of the activity that incurs the non-value-added cost.
- 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.
- OEM** – Original Equipment Manufacturer. A company that sells products (including software) under its own label that includes technology licensed from another vendor.

OLAP – On-Line Analytical Processing. Refers to a category of applications and technologies for collecting, managing, processing and presenting multi-dimensional data for analysis and management purposes.

OLAP Client – End-user applications that can request slices from OLAP servers and provide two-dimensional or multi-dimensional displays, user modifications, selections, ranking, calculations, etc., for visualization and navigation purposes. OLAP clients may be as simple as a spreadsheet program retrieving a slice for further work by a spreadsheet literate user or as high-functioned as a financial modeling or sales analysis application.

OLAP Product – A product capable of providing fast analysis of shared multi-dimensional information. Ad hoc analysis must be possible either within the product itself or in a closely linked product.

OLAP/ROLAP – On-Line Analytical Processing/ Relational On-Line Analytical Processing (OLAP/ROLAP) are applications that seek to verify complex hypotheses. An example of an OLAP query might be, "compare the costs of shipping to customers in the east to those in the west."

OLE – Object Linking and Embedding. This is a Microsoft Windows technology for presenting applications as objects within other applications and hence to extend the apparent functionality of the host (or client) application.

OLE DB for OLAP – Microsoft's OLAP API, effectively the new industry standard for OLAP connectivity. Used to link OLAP clients and servers using a multi-dimensional language, MDX.

OLTP – On-Line Transaction-Processing. Operational systems for collecting and managing the base data in an organization, such as sales order-processing, inventory, accounts payable, etc. Usually offer little or no analytical capabilities.

On-hand Balance – The quantity shown in the inventory records as being physically in stock.

Open Order – An active manufacturing order or purchase order.

Operational Data Store (ODS) – A database of operational data that is formatted as it is collected for use as a data warehouse. This is opposed to the more common process of creating a separate database out of existing operational databases designed expressly for use by the data warehouse. 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.

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 or OLTP Database – Operational data is the data collected from operations such as order-processing, accounting, manufacturing, marketing, etc. Most modern companies collect most of this data using a form of on-line transaction-processing (OLTP). Data generated by these systems is generally not in a format that makes for efficient query processing or analysis.

Optimization Criterion – A positive function of the difference between predictions and data estimates that are chosen so as to optimize the function or criterion. Least squares and maximum likelihood are examples.

Option – A choice or feature offered to customers for customizing the end product. In many companies, the term option means a mandatory choice (i.e., the customer must select from one of the available choices). For example, in ordering a new car the customer must specify an engine (option) but need not necessarily select an air-conditioner.

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.

Order Entry – The process of accepting and translating what a customer wants into terms used by the provider. This can be as simple as creating shipping documents for a finished goods product to a far more complicated series of activities including engineering effort for make-to-order products. A key element in the order promising process is customer order promising.

Order Promising – The process of making a delivery commitment (i.e. answering the question, "when can you ship?")

Order Quantity – The amount of an item to be ordered. Also called lot size.

Outlier – A data item whose value falls outside the bounds enclosing most of the other corresponding values in the sample. May indicate anomalous data. Should be examined carefully; may carry important information. Technically, outliers are data items that did not (or are thought not to have) come from the assumed population of data—for example, a non-numeric when you are expecting only numeric values. A more casual usage refers to data items that fall outside the boundaries that enclose most other data items in the data set.

Overfitting – A tendency of some modeling techniques to assign importance to random variations in the data, by declaring them important patterns.

Overlay – Data not collected by the organization such as data from a proprietary database that is combined with the organization's own data.

Owner – The object owner is a user or users who have authority over that object, where object is a table, view or attribute.

PAEI Skills – PAEI is the abbreviation for interpersonal or managerial skills, which are as follows:

- PRODUCER skills—the ability to get a job done
- ADMINISTRATIVE skills—the ability to organize and delegate
- ENTREPRENEURIAL skills—the ability to see new ideas, opportunities and methods
- INTEGRATIVE skills—the ability to bring diverse opinions together and achieve consensus

The four PAEI skills are not usually found in one individual but can be achieved through team formation. These skills are vital for the success of the team and must be considered when the team is formed.

Page Dimension – A page dimension is generally used to describe a dimension, which is not one of the two dimensions of the page being displayed but for which a member has been selected to define the specific page requested for display. All page dimensions must have a specific member chosen in order to define the appropriate page for display.

Page Display – The page display is the current orientation for viewing a multi-dimensional slice. The horizontal dimension(s) run across the display, defining the column dimension(s). The vertical dimension(s) run down the display, defining the contents of the row dimension(s). The page dimension member selections define which page is currently displayed. A page is much like a spreadsheet and may, in fact have been delivered to a spreadsheet product where the user can further modify each cell.

Parallel Processing – The coordinated use of multiple processors to perform computational tasks. Parallel processing can occur on a multi-processor computer or on a network of workstations or PCs.

Parallel Processing System – A system that strings together dozens or even hundreds of microprocessors that work in parallel to attack computing tasks. Parallel systems are typically "scalable", meaning that you can increase the performance of the system linearly as you add processors.

Pattern – Analysts and statisticians spend much of their time looking for patterns in data. A pattern can be a relationship between two variables. Data mining techniques include automatic pattern discovery that makes it possible to detect complicated non-linear relationships in data. Patterns are not the same as causality.

PDA – Personal Digital Assistant such as a Palm, Psion or Windows CE computer. Can now often be linked to mobile phones or the Internet for sending and receiving e-mails or faxes.

Pegging – In MRP pegging shows, for a given item, the details of the sources of its gross requirements and/or allocations. Pegging can be thought of as live where-used information.

Performance Measure – An indicator that can be used to evaluate quality, cost or cycle time characteristics of an activity or process usually against a target or standard value. It is an established, consistent way to measure the rate of change within an organization.

Period Order Quantity – An order quantity technique in which the order quantity will be equal to the net requirements for a given number of periods (days or weeks) into the future. Also called day's supply, week's supply or fixed period.

Picking – The process of issuing components to the production floor on a job-by-job basis.

Picking List – A document used to pick manufacturing orders, listing the components and quantities required.

Planner/ Buyer – See supplier scheduler.

Planning BOM – An artificial grouping of items in bill of material format, used to facilitate master scheduling and/or material planning. A modular bill of material is one type of planning bill.

Plant Floor (Shop Floor) Control – A system for utilizing data from the plant floor as well as data processing files to maintain and communicate status information on shop orders (manufacturing orders) and work centers. The major sub-functions of shop floor control are assigning priority of each shop order, maintaining work-in-process quantity information, conveying shop order status information, providing actual input and output data for capacity control purposes, providing quantity by location by shop order for work-in-process inventory and accounting purposes, providing measurements of efficiency, utilization, and productivity of manpower and machines.

Platform – A combination of hardware and system software.

Plug-in – Locally stored, helper programs that are used to augment browser capabilities. Require different versions for different platforms and possibly for different browsers, but currently capable of delivering better human factors and more functionality than Java applets. Occasionally used in OLAP Web products.

Portability – The ability of an ERP system to run various operating systems, databases and networks without requiring any major adjustments or sacrificing any functionality.

Post-Deduct Inventory Transaction-Processing – A method of inventory bookkeeping where the book (computer) inventory of components is reduced only after completion of production of their upper level parent. This approach has the disadvantage of a built-in differential between the book record and what is physically in stock. This is also called back flush.

Pre-Calculated/Pre-Consolidated Data – Pre-calculated data is data in output member cells that are computed prior to, and in anticipation of ad hoc requests. Pre-calculation usually results in faster response to queries at the expense of storage. Data that is not pre-calculated must be calculated at query time.

Precision – The precision of an estimate of a parameter in a model is a measure of how variable the estimate would be over other similar data sets. A very precise estimate would be one that did not vary much over different data sets. Precision does not measure accuracy. Accuracy is a measure of how close the estimate is to the real value of the parameter. Accuracy is measured according to the average distance over different data sets of the estimate from the real value. Estimates can be accurate but not precise, or precise but not accurate. A precise but inaccurate estimate is usually biased, with the bias equal to the average distance from the real value of the parameter.

Predictability – Some data mining vendors use predictability of associations or sequences to mean the same as confidence.

Predictive Model – A structure and process for predicting the values of specified variables in a dataset.

Pre-Requisites – Related knowledge and skills, which are required in order to perform, work effectively. A competency is defined as "core" if it is important to performance of a core technology. The BPR team members will be trained in BPR principles, tools and techniques. They must bring the core competencies to the BPR effort. Obviously, if the team member has prior training or work experience that include some or all of the core technologies little or no training would be necessary.

Present Value – The current value of a future series of cash flows, given a discount factor or interest value. Used to evaluate the alternative investments.

Prevalence – The measure of how often the collection of items in an association occurs together as a percentage of all the transactions. For example, "in 2% of the purchases at the hardware store, both a pick and a shovel were bought."

- Process** – A collection of activities that together produce a usable product or service by applying resources from one or more functional areas.
- Process Action Team (PAT)** – A group of "hands-on" people assembled as part of a total quality management/total quality leadership (TQM/ TQL) project to solve a specific operational problem.
- Process Analysis** – Process analysis is a combination of graphics and narrative symbols and rules designed to capture the processes and structure of information in an organization.
- Process Management** – A philosophy of management that organizes an enterprise by the series of activities, that combine to produce related types of goods and services for internal or external customers. See functional management.
- Process Organizations** – A process organization is an enterprise viewed from a "process" perspective. A process organization re-designs its processes first, and then determines the optimum organization form needed to make the process work best. The goal of a process organization is to create a high-performance workplace, a high-quality work environment noted for excellence in efficiency, effectiveness and customer satisfaction. With a focus on process, it is very common to see process organizations managing interdisciplinary work teams instead of specialized units more commonly seen in traditional organization charts.
- 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.
- Product Structure** – See bill of material.
- Product/ Performance Analysis Application** – Analyses to define how well a product or group of products is performing in the marketplace and how well it compares to or is positioned against competitive products. Can include analyzing customer response data, warranty exercise information and sales trends.
- Production Data** – Source data, which is subject to change. It is a data capture system often on a corporation's mainframe.
- Program Mission** – A statement of purpose that clearly specifies top management realistic expectations. It delineates the program's objectives, scope and viewpoint. It defines the program schedule and budget.
- Project Management** – Project management is the ability to define, schedule and assign project activities; record project issues; monitor progress and report changes in activity accomplishment and issue resolution; and maintain and control changes to designs, plans and issue lists.
- Promotion/ Marketing Campaign Analysis Application** – Analyze the effectiveness of promotional campaigns by examining purchase patterns, revenue derivation and other parameters.
- 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.
- Prospective Data Analysis** – Data analysis that predicts future trends, behaviors or events based on historical data.
- Pruning** – Eliminating lower level splits or entire sub-trees in a decision tree. This term is also used to describe algorithms that adjust the topology of a neural net by removing (i.e., pruning) hidden nodes.
- Pull System** – Usually refers to how material is moved on the plant floor. Pull indicates that material moves to the next operation only as needed by that next operation. See Kanban.
- Pump** – A data pump extracts data from several mainframe and client-server platforms, performs some filtering and transformation, and distributes and loads to another database(s). Usually the term pump is used rather than "replicator" to connote its applicability in a cross-platform environment.
- Push System** – Usually refers to how material is moved on the plant floor. Push indicates that material moves to the next operation automatically upon completion of the prior operation.
- Q&R** – Query and Reporting tool. Normally used for list-oriented reporting from relational databases.

Quality - The degree of excellence possessed by a product, service or other output of a business activity or business process (traditional definition). The total quality management definition of quality is conformance to the customer's requirements.

Quality Application – A manufacturing related application that analyzes manufacturing and/or sales data to determine how to improve the manufacturing process. Typically helps standardize production data so that the same quality products can be manufactured in any factory location.

Quality Assurance – The process of ensuring a correct result.

Quality Function Deployment (QFD) – A requirements identification analysis, flow down and tracking technique. It focuses on quality and communication to translate customer needs into product and process design specifics. Also known as the "house of quality."

Query – A (usually) complex select statement for decision support.

Query Governor – A facility that terminates a database query when it has exceeded a pre-defined 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 database. 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.

Queue – In manufacturing, the jobs at a given work center waiting to be processed. As queues increase, so do average lead-times and work-in-process (WIP) inventories.

Queue Time – The amount of time a job waits at a work center before work is performed on it. Queue time is one element of total manufacturing lead-time. There are direct increases to manufacturing lead-time as a result of increases in queue time.

Quick-Slice – A method of implementing most of the ERP functions into a small slice of the business, typically one product or product line, in a very short time.

RAID – Redundant Array of Inexpensive Disks. This is a technology for the efficient parallel storage of data for high-performance computer systems.

Reach-Through – Reach-through is a means of extending the data accessible to the end-user beyond that which is stored in the OLAP server. A reach-through is performed when the OLAP server recognizes that it needs additional data and automatically queries and retrieves the data from a data warehouse or OLTP system.

Real-Time – The immediate availability of data to an information system as a transaction or event occurs.

Re-Design – Business process re-design. The transformation of a business process to achieve significant levels of improvement in one or more performance measures relating to fitness for purpose, quality, cycle time and cost by using the techniques of streamlining and removing non-value-added activities and costs. Re-design projects typically take about six months to complete.

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.

Reengineering – This refers to business process reengineering. The radical transformation of a business process to achieve orders of magnitude improvement in one or more performance measures relating to fitness-for-purpose, quality, cycle time and cost; usually requiring the application of technology enablers. Reengineering projects typically take a minimum of two years to complete.

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.

Regeneration MRP – A method of processing MRP on the computer whereby the master schedule is totally exploded down through all bills of material to maintain valid priorities. New requirements and planned orders are completely regenerated at that time. See net change MRP.

Relationship – The representation of the association between two different real-world objects. A connection relationship has cardinality and may be either specific or non-specific.

Repetitive Manufacturing – Production of discrete units; planned and executed via schedule, usually at relatively high-speeds and volumes. Material tends to move in a sequential flow.

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 also be implemented across multiple hardware platforms and database management systems (DBMS).

Required Activity – An activity that is either primary or secondary but cannot be eliminated because of law, regulation or policy. (May nevertheless be non-value-added).

Re-Scheduling Assumption – A fundamental piece of MRP logic, which assumes that existing open orders can be re-scheduled in nearer time periods more easily than new orders can be released and completed. As a result, planned order receipts are not created until all scheduled receipts have been applied to cover gross requirements.

Resource Planning Applications – Refers to packaged software that uses databases and DBMSs for applications in functional areas such as finance, manufacturing and human resources. Also known as enterprise resource planning applications or ERP, these applications typically run on top of a DBMS provided by another vendor. Vendors of ERP software sometimes claim, that their packages are actually specialized DW/ BI applications and can obviate the need for a separate DW/ BI system.

Resource Requirements Planning – See rough-cut capacity planning.

Resources – Enterprise assets that are assigned to activities and consumed (used up) in the process of producing an output product or service. Examples of resources are labor hours, funds, machine hours, materials, contract labor.

Re-Substitution Error – The estimate of error based on the differences between the predicted values of a trained model and the observed values in the training set.

Retrospective Data Analysis – Data analysis that provides insights into trends, behaviors or events that have already occurred.

Reverse engineering – Reverse engineering derives a consistent set of metadata from several potential source systems metadata.

Risk Adjusted Discounted Cash Flow (RADCF) – A future cash flow applied to a discount rate with a range identified within which the cash flow is likely to fall to determine its present value using a randomly generated value within that range for projected cash flow. When applied, this makes each value more realistic by simulating the effect of risk on the process.

Risk Management Application – The process of discovering and analyzing risk factors and providing decision-makers with risk/ return models. Portfolio management, credit issuance and risk determination to support insurance actuarial determinations are typical uses. Also included are predictive models associated with mutual funds and derivatives pricing.

ROLAP – Relational OLAP. This is a product that provides multi-dimensional analysis of data, aggregates and metadata stored in an RDBMS. The multi-dimensional processing may be done within the RDBMS, a mid-tier server or the client. A 'merchant' ROLAP is one from an independent vendor, which can work with any standard RDBMS.

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

Rotate – To change the dimensional orientation of a report or page display. For example, rotating may consist of swapping the rows and columns, or moving one of the row dimensions into the column dimension, or swapping an off-spreadsheet dimension with one of the dimensions in the page display (either to become one of the new rows or columns), etc. A specific example of the first case would be taking a report that has time across (the columns) and products down (the rows) and rotating it into a report that has product across and time down. An example of the second case would be to change a report, which has measures and products down and time across into a report with measures down and time over products across. An example of the third case would be taking a report that has time across and product down and changing it into a report that has time across and geography down.

Rough-Cut Capacity Planning – The process of converting the production plan (from sales and operations planning) and/ or the master schedule into capacity needs for key resources: manpower, machinery, warehouse space, suppliers' capabilities and, in some cases, money. Product load profiles are often used to accomplish this. The purpose of rough-cut capacity planning is to evaluate the plan prior to attempting to implement it.

Routing – Information detailing the manufacture of a particular item. It includes the operations to be performed, their sequence, the various work centers to be involved, and the standards for set-up and run times. In some companies the routing also includes information on tooling, operator skill levels, inspection operations, testing requirements, and so forth.

RPC – Remote Procedure Call.

R-Squared – A number between 0 and 1 that measures how well a model fits its training data. One is a perfect fit; however, zero implies the model has no predictive ability. It is computed as the co-variance between the predicted and observed values divided by the standard deviations of the predicted and observed values.

Rule Induction – The extraction of useful if-then rules from data based on statistical significance.

Rule Modeling Technique – A technique used to support the discovery and documentation of data elements and their relationships in an organization. It provides the procedures and rules that guide the development of the rule model.

Safety (Buffer) Stock – A quantity of stock planned to be available to protect against fluctuations in demand and/ or supply.

Safety Time – A technique whereby material is planned to arrive ahead of the requirement date. This difference between the requirement date and the planned in-stock date is safety time.

Sales & Operations Planning (S&OP) – A business process that helps companies keep demand and supply in balance. It does that by focusing on aggregate volumes—product families and groups—so that mix issues (individual products and customer orders) can be handled more readily. It occurs on a monthly cycle and displays information in both units and rupees. S&OP is cross-functional, involving general management, the sales and marketing department(s), operations, finance and product development. It occurs at multiple levels within the company up to and including the executive in charge of the business unit. S&OP links the company's strategic plans and business plan to its detailed processes the order entry, master scheduling, plant scheduling and purchasing tools used to run the business on a week-to-week, day-to-day and hour-to-hour basis. Used properly, S&OP enables the company's managers to view the business holistically and provides them with a window into the future.

Sales Force Analysis Application – Involves measuring the effectiveness of policies and management controls on the effectiveness of a sales force. It is the analysis of techniques, methods and practices against revenue production to depict models of successful and unsuccessful behaviors.

- Sales Plan** – The overall level of sales expected to be achieved. Usually stated as a monthly volume of sales for a product family (group of products, items, options, features, etc.). It needs to be expressed in units identical to the production plan (as well as rupees) for planning purposes. It should represent the sales and marketing department manager's commitment to take all reasonable steps necessary to make the sales forecast accurately represent actual customer orders received.
- Same Name/ Same Meaning Test** – This implies that if two entities have the same name, they must represent the same type of real-world object. If two entities have different names, they must represent different types of real-world objects. This test also applies to attributes, activities and ICOMs.
- Sampling** – Creating a sub-set of data from the whole. Random sampling attempts to represent the whole by choosing the sample through a random mechanism.
- 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.
- Scheduled Receipts** – Within MRP, open production orders and open purchase orders are considered as scheduled receipts on their due date and will be treated as part of available inventory during the netting process for the time period in question. Normally, the computer does not automatically alter scheduled receipt dates and/or quantities. Further, scheduled receipts are not exploded into requirements for components, as MRP logic assumes that all components required for the manufacture of the item in question have either been allocated or issued to the plant floor.
- Schema** – The logical and physical definition of data elements, physical characteristics and inter-relationships.
- Scope** – The breadth and depth of a subject area in an organization or function, which will be analyzed in the modeling effort.
- Scoping** – Restricting the view of database objects to a specified sub-set. Further operations, such as update or retrieve, will affect only the cells in the specified sub-set. For example, scoping allows users to retrieve or update only the sales data values for the first quarter in the east region, if that is the only data they wish to receive.
- SDK** – Software Development Kit. Comprises a set of programs that allows software developers to create products to run on a particular platform or to work with an API.
- Segment-of-One Marketing** – The ability to provide personalized products, services or promotions to individual customers. This requires an in-depth knowledge of the customer's individual needs or interests, often gained through sophisticated customer segmentation or profiling.
- Selection** – A selection is a process whereby a criterion is evaluated against the data or members of a dimension in order to restrict the set of data retrieved. Examples of selections include the top ten salespersons by revenue, data from the east region only and all products with margins greater than 20 percent.
- Semantic Mapping** – The mapping of the meaning of a piece of data.
- Semantics** – The unique definitions that describe what is being portrayed by the symbols in a model. Refers to the content of a model rather than its form.
- Sensitivity Analysis** – Implies varying the parameters of a model to assess the change in its output.
- Sequence Discovery** – The same as association, except that the time sequence of events is also considered. For example, "twenty percent of the people who buy a VCR buy a camcorder within four months."
- Service Parts** – Parts used for the repair and/or maintenance of a product. Also called repair parts, spares.
- Shop Floor Control** – See plant floor control.
- Significance** – A probability measure of how strongly the data support a certain result (usually of a statistical test). If the significance of a result is said to be .05, it means that there is only a .05 probability that the result could have happened by chance alone. Very low significance (less than .05) is usually taken as evidence that the data mining model should be accepted since events with very low probability seldom occur. So, if the

estimate of a parameter in a model showed a significance of .01 that would be evidence that the parameter must be in the model.

Simulation – An analytical technique to prove a business practice concept by acting out or by creating the imitative representation of the proposed concept. The ability to test a real world environment by running transactions or business case situations against a process design to determine if the process design functions as expected. The outcome of simulation is measurement results providing operational performance timing, resource consumption costs and identification of transaction bottlenecks. This information is used for calculation of benefits and expected return on investment for reengineering. Within ERP, utilizing operational data to perform what-if evaluations of alternative plans, to answer the question: "can we do it?" If yes, the simulation can then be run in financial mode to help answer the question, "do we really want to?"

Simultaneous Equation – A set of calculations with circular interdependencies. Most OLAP products cannot resolve these automatically.

Skills – The ability to perform a task of function to an agreed-upon-criterion.

SKU – Stock-Keeping Unit. A term used by retailers to identify the lowest level of product detail. Such level of detail is often not included in OLAP applications.

Slice – A slice is a sub-set of a multi-dimensional array corresponding to a single value for one or more members of the dimensions not in the sub-set. For example, if the member actuals is selected from the scenario dimension, then the sub-cube of all the remaining dimensions is the slice that is specified. The data omitted from this slice would be any data associated with the non-selected members of the scenario dimension, for example budget, variance, forecast, etc. From an end-user perspective, the term slice most often refers to a two-dimensional page selected from the cube.

Slice and Dice – A term used to describe a complex data analysis function provided by MDBMS tools. It is a user-initiated process of navigating by calling for page displays interactively, through the specification of slices via rotations and drill-down/ up.

SMP – Symmetrical Multi-Processing. This is a computer hardware architecture, which distributes the computing load over a small number of identical processors, which share memory. Is common in Unix systems, but also supported by Windows NT and OS/2.

Snowflake Schema – A variant of the star schema with normalized dimension tables.

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

Sparse – A multi-dimensional data set is sparse if a relatively high percentage of the possible combinations (intersections) of the members from the data set's dimensions contain missing data. The total possible number of intersections can be computed by multiplying together the number of members in each dimension. Data sets containing one percent, .01 percent, or even smaller percentages of the possible data exist and are quite common.

Standardize – A collection of numeric data is standardized by subtracting a measure of central location (such as the mean or median) and by dividing by some measure of spread (such as the standard deviation, interquartile range or range). This yields data with a similarly shaped histogram with values centered on 0. It is sometimes useful to do this with inputs into neural nets and also inputs into other regression models. (Also see normalize)

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

Strategic Planning – Strategic planning is the top management decision process that focuses on the long-range direction of the enterprise and establishes the means by which that direction is reached. It includes the definition of missions and objectives—how the enterprise sees its purpose and where it wants to go. Strategic planning provides the basic direction and focus of the organization, the so-called big picture. Some of the organization's basic strategic decisions might relate to questions such as: What business are we in? What

business should we be in, now and in the future? What should be the geographical scope of operations? What are our research and development goals? How should products be sourced? Where are we the weakest? Strongest?

Structural Modeling Language (SML) – A set of defined terms and syntax used to textually define IDEF1X models in a format that can be processed both by a computer and manually.

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.

Summary tables – Often used in RDBMSs to store pre-aggregated information, rather than holding it in the same table as the base data. Used to improve responsiveness.

Supervised Learning – The collection of techniques where analysis uses a well-defined (known) dependent variable. All regression and classification techniques are supervised.

Supplier Scheduler – This is a person whose main job is working with suppliers regarding what is needed and when. Supplier schedulers are in direct contact with both MRP and the suppliers. They do the material planning for the items under their control, communicate the resultant schedules to their assigned suppliers, do follow-ups, resolve problems, and so forth. The supplier schedulers are normally organized by commodity, as are the buyers. By using the supplier scheduler approach the buyers are freed from day-to-day order placement and expediting and therefore have the time to do cost reduction, negotiation, supplier selection, alternate sourcing, and the like.

Supplier Scheduling – A purchasing approach, which provides suppliers with schedules rather than individual hard-copy purchase orders. Normally a supplier scheduling process will include a business agreement (contract) for each supplier, a daily or weekly schedule for each supplier extending for some time into the future and individuals called supplier schedulers. Also required is a formal priority planning system that works well, because it is essential in this arrangement to provide the suppliers with consistently valid due dates. Some form of supplier scheduling is essential for just-in-time purchasing.

Supply Chain – The processes from the initial acquisition of raw materials to the ultimate consumption of the finished product linking across supplier-user companies. Includes the functions inside and outside a company that enable the value chain to make products and provide services to the customer.

Supply Chain Management – The use of information technology to give automated intelligence to a network of vendors, suppliers, manufacturers, distributors, retailers and a host of other trading partners. The goal is for each player in the supply chain to conduct business with the latest and best information from everyone else in the chain, guiding supply and demand into a more perfect balance. Effective management of the supply chain enables a company to move product from the point of origin to that of consumption in the least amount of time at the smallest cost.

Support – The measure of how often the collection of items in an association, occur together as a percentage of all the transactions. For example, "in 2% of the purchases at the hardware store, both a pick and a shovel were bought."

Systems Integration – The services involved in designing and building complex information systems that integrate different makes and classes of systems, software and networking protocols to improve the flow and management of information. As with network integration, these systems typically include systems and components from multiple vendors.

Systems Integration Engineering – Implies the application of a consistent set of techniques and disciplines to the combination of software and hardware components and sub-systems in such a way that they function as one system.

Systems Integrators – Companies that connect computers, instrumentation and equipment to share data or applications with other components in the same or other functional areas. Systems integrators often provide consulting services and hardware sales as well.

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.

Target Activity Model – A graphic representation, which exhibits the activities of the business process, as you would like them to be upon completion of a business process improvement initiative. Also referred to as the "to-be" activity model.

Target Database – The database in which data will be loaded or inserted.

Team Roles – Modeling is a team effort. Each team member must be assigned one or more roles to ensure that the team meets its objectives. Some of the roles require a full-time effort; they are called the core team roles. Other roles require part-time or as-needed effort; they are referred to as the extended team roles.

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.

Technical Infrastructure Engineering – The application of a consistent set of techniques to the operation of the basic facilities, equipment and installations required for the functioning of information systems.

Technical Integration – The process of making technical requirements consistent between and among the internal architecture structures of an organization.

Technical Support Consultant – A consulting services organization that provides hardware/ software/ networking support, ensuring that the DW/ BI system runs properly.

Time Box – A fixed, specific period of time during which specific tasks must be completed.

Time Fence – A point in time where various restrictions or changes in operating procedures take place. For example, changes to the master schedule can be accomplished easily beyond the cumulative lead-time; whereas, changes inside the cumulative lead-time become increasingly more difficult. Time fences can be used to define these points.

Time Series – A series of measurements taken at consecutive points in time. Data mining products, which handle time series, incorporate time-related operators such as moving average. Also see windowing.

Time Series Analysis – The analysis of a sequence of measurements made at specified time intervals. Time is usually the dominating dimension of the data.

Time Series Model – A model that forecasts future values of a time series based on past values. The model form and training of the model usually take into consideration the correlation between values as a function of their separation in time.

To-Be Model – Models that are the result of applying improvement opportunities to the current (as-is) business environment (see also 'as-is' model).

Tool Encyclopedias – Encyclopedias, repositories or dictionaries used by application development tools. Includes the non-definable "repository" used by a tool.

Total Quality Management (TQM) – TQM is both a philosophy and a set of guiding principles that represent the foundation of a continuously improving organization. TQM is a strategic, integrated management system for achieving customer satisfaction. It involves all managers and employees and uses quantitative methods to continuously improve an organization's processes. At the foundation of TQM are three principles: focus on achieving customer satisfactions, seek continuous improvements and fully involve the entire workforce. Achieving these principles requires the establishment of a cultural shift within an organization aimed at making the new culture more participative.

Training – Another term for estimating a model's parameters based on the data set at hand.

Training Data – A data set used to estimate or train a model.

Transformation – A re-expression of the data such as aggregating it, normalizing it, changing its unit of measure, or taking the logarithm of each data item.

Transformers – Rules applied to change data.

Triggering Data – Data that selects and loads data on a scheduled basis.

Two-Level Master Schedule – A master scheduling approach for make-to-order products where an end product type is master scheduled along with selected key options, features, attachments, and common parts.

Unit Cost – The total costs in resource and material to produce one instance of a product or service.

Unit of Work Consolidation – The process of consolidating multiple updates to a single row image into a single update.

Unsupervised Learning – This term refers to the collection of techniques where groupings of the data are defined without the use of a dependent variable. Cluster analysis is an example.

Update – Not allowed in a data warehouse.

Validation – The process of testing the models with a data set different from the training data set.

Value-Added Activity – An activity in a process that adds value to an output product or service that is, the activity merits the cost of the resources it consumes in production. It contributes to producing a designated product or service that meets customer's requirements and that the customer is willing to pay for.

Value-Added Cost – Resources consumed in the performance of a value-added activity.

Value-Added Reseller (VAR) – A person or company that purchases the products of other companies, adds software, services of special hardware components of its own design and manufacture and then sells the whole package to a customer.

Values – Collective judgments of worth desired and shared with the organization. Also the specific facts represented by data.

Variable Cost – A cost element that varies directly with the amount of product or service produced by an activity or cost. Variable costs go to zero if the activity stops. See also fixed cost.

Variance – The most commonly used statistical measure of dispersion. The first step is to square the deviations of a data item from its average value. Then the average of the squared deviations is calculated to obtain an overall measure of variability.

Vendor Performance Analysis Application – Analyzing data that show how well a vendor is meeting its requirements from the standpoints of quality, cost and supply perspectives. Related to the supply chain management application.

Vendor Scheduler – See supplier scheduler.

Vendor Scheduling – See supplier scheduling.

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

Vertical Applications – Vertical applications are industry-specific or function-specific applications that have one or more DSS elements either incorporated into them or are front-ends to analysis tools provided by other vendors. Examples might be fraud detection in the banking industry or risk analysis in the insurance industry.

Vertical Dimension – See page display.

VGA – Video Graphic Array. This is an IBM PC display standard with 640x480 pixels. Introduced with the original PS/2, it has largely been superseded by higher resolution screens, usually referred to as Super VGA and X[V]GA, typically with 800x600 or 1024x768 pixels. With larger monitors, sizes of 1280x1024 and 1600x1200 are becoming more common. Objects of fixed pixel size designed for VGA screens, look very small and may be hard to read when displayed on higher resolution monitors.

Virtual Memory – Apparently extended memory on a computer, consisting partly of real memory (RAM) and partly of disk space. Is a technique to handle programs and applications that are too large to fit into real memory. Can degrade performance if used too heavily.

Visualization – Visualization tools graphically display data to facilitate better understanding of its meaning. Graphical capabilities range from simple scatter plots to complex multi-dimensional representations.

VITAL – An acronym for Virtually Integrated Technical Architecture Life cycle.

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.

WAN – Wide Area Network. Usually, two or more geographically dispersed LANs connected by lower speed links. Can cause problems with client/server applications that transmit large quantities of data between servers and clients.

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.

Windowing – Used when training a model with time series data. A window is the period of time used for each training case. For example, if we have weekly stock price data that covers fifty weeks and we set the window to five weeks, then the first training case uses weeks one through five and compares its prediction to week six. The second case uses weeks two through six to predict week seven, and so on.

WMS – Warehouse Management Systems: It is a software that integrates mechanical and human activities with an information system to effectively manage warehouse business processes and direct warehouse activities. These systems automate receiving, put away, picking and shipping in warehouses, and can prompt workers to do inventory cycle counts. Most support radio frequency communications, allowing real-time data transfer between the system and warehouse personnel.

Work-in-Process (WIP) – Product in various stages of completion, including raw material that has been released for initial processing and completely processed material awaiting final inspection and acceptance as finished product or shipment to a customer.

Yield Analysis Application – Addresses issues such as product line profitability, customer retention and churn analysis through targeting and modeling of customer segments and products. Used by some vendors to mean the same thing as the churn analysis and/ or product performance analysis applications.