# Product Glossaries

Enter product terms you want to describe on this page.

**A  B  C  D  E  F  G  H  I  J  K  L  M  N  O  P  Q  R  S  T  U  V  W  X  Y  Z**

**A**

**Aggregate**

To collect data from multiple sources and merge it into one source. Aggregated data is also referred to as summarized data.

**Aggregate Table**

Aggregate tables contain pre-aggregated measures built from the fact table. It is registered in Mondrian's schema, so that Mondrian can choose whether to use the aggregate table rather than the fact table, if applicable for a particular query.

**Aggregation**

The process of merging multiple data values into one value. For example, sales data collected daily can then be aggregated to the week level, the week data could be aggregated to the month level, and so on. The data can then be referred to as aggregate data. Aggregate data and summary data are the same.

**Analysis Model**

Pentaho uses two data models, a Multidimensional Data Model and a Relational Data Model. See Pentaho Data Modelsfor more information.

**Attribute**

A property or field of an object in the directory.

**Audit Repository**

See Server Repositories.

**Analyzer Report**

A report created within Pentaho User Console or Pentaho Data Integration using Pentaho Analyzer.

**Authority, Role, or Group**

In the Business Analysis Server, these three terms are synonymous. A role is a string that is associated with a user. A role is said to be granted to a user. A user is said to belong to or be a member of a role. The same role can be granted to multiple users and users can be granted zero or more roles. The BA Server uses roles to make authorization decisions.

**Authentication**

The process of confirming that the user requesting access is the user that they claim to be. This is often done by presenting a user identifier (a username) paired with a secret known only to that user (a password), but can sometimes involve certificates or other means of establishing identity. In this documentation, authentication is synonymous with login.

**Authorization**

The process of deciding if the authenticated user is allowed to access the information or functionality he is requesting. A software system can protect itself at multiple levels. In the Pentaho BI Platform, pages in the Web-based user interface can be protected. In addition, objects within the Pentaho solution repository, such as folders and action sequences, can be protected using access control lists (ACLs).

**B**

**Business Analysis Platform**

The Pentaho Business Analysis Platform (BA Platform) is the core architecture and foundation of the Pentaho Open Business Analysis Suite. The BA Platform is composed of the libraries and compiled code that provide execution framework and services associated with logging, auditing, security, scheduling, ETL, Web Services, attribute repository, and rules engine. See Business Analysis Server and Product.

**Business Analysis Server**

The Business Analysis Server (BA Server) consists of the Pentaho Business Analysis Platform and the libraries that deliver end user analysis capabilities. The server runs inside a  J2EE-compliant Application Server such as Apache, JBOSS AS, IBM WebSphere, WebLogic, and Oracle AS.  The BA Server referred to in this document is your customized Pre-Configured Installation. See Product.

**C**

**Cube**

A cube is a collection of dimensions and measures, centered on a fact table. See Mondrian Schema.

**D**

**Data Definition Language**

Data Definition Language (DLL) is originally a subset of SQL. This language defines data structures, including rows, columns, tables, indexes, and database specifics, such as file locations. DDL SQL statements are more a part of a database management system and have large differences between SQL implementations. See http://en.wikipedia.org/wiki/Data\_Definition\_Language.

**Data Governance**

A control that ensures that the data entry by an operations team member or by an automated process meets precise standards, such as a business rule, a data definition and data integrity constraints in the data model. See http://en.wikipedia.org/wiki/Data\_governance.

**Data Mart**

A data mart is a database, or collection of databases, designed to help managers make strategic decisions about their business. Whereas a data warehouse combines databases across an entire enterprise, data marts are usually smaller and focus on a particular subject or department. Some data marts, called *dependent data marts,* are subsets of larger data warehouses. Specifically, a data mart is the access layer of a data warehouse environment that is used to get data out to the users. The data mart is a subset of the data warehouse that is usually oriented to a specific business line or team. In some deployments, each department or business unit is considered the owner of its data mart including all the hardware, software*,* and data. See http://en.wikipedia.org/wiki/Data\_mart.

**Data Model**

A data model in software engineering is an abstract model that documents and organizes the business data for communication between team members and is used as a plan for developing applications, specifically how data are stored and accessed.

**Data Store**

A data store is a data repository of a set of integrated objects. These objects are modeled using classes defined in database schemas. Data store includes not only data repositories like databases, it is a more general concept that includes flat files that can store data. See http://en.wikipedia.org/wiki/Data\_store.

**Data Stream**

A sequence of digitally encoded coherent signals (packets of **data** or **data** packets) used to transmit or receive information that is in the process of being transmitted. See http://en.wikipedia.org/wiki/Data\_stream.

**Data Warehouse**

A data warehouse (DW or DWH) is a database used for reporting and analysis. The data stored in the warehouse are uploaded from the operational systems, such as marketing, sales etc.The data may pass through an operational data store for additional operations before they are used in the DW for reporting. See http://en.wikipedia.org/wiki/Fact\_table.

**E**

**End-User**

An end-user is a person who consumes the results of analysis or reporting in dashboards. They are not technically savvy and do light work to personalize dashboards, analysis, reports, or the mobile interface.

**End-User Capability**

In Pentaho Business Analysis, end-user capabilities include reporting, analysis, dashboards, data mining, and personalizing the mobile interface. See Product.

**F**

**Fact Table**

In data warehousing, a fact table consists of the measurements, metrics, or facts of a business process. It is often located at the center of a star schema or a snowflake schema, surrounded by dimension tables. See http://en.wikipedia.org/wiki/Fact\_table.

**G**

**Group**

See Authority

**H**

**I**

**J**

**Job Entries**

An item in the Pentaho Data Integration job scheduler. A task to be completed as part of a scheduled job.

**K**

**L**

**LDAP User DN**

Used with Lightweight Directory Access Protocol (LDAP) authentication, the distinguished name (DN) consists of one or more strings identifying the user's assigned attributes in the LDAP backend server and a user password. See http://en.wikipedia.org/wiki/Lightweight\_Directory\_Access\_Protocol.

**M**

**Manager**

A user with read access to relevant objects in the directory. If you're familiar with the JDBC API, a manager is analogous to a user name given along with a URL and password in a [DriverManager.getConnection (url, user, password) |http://java.sun.com/j2se/1.4.2/docs/api/java/sql/DriverManager.html#getConnection%28java.lang.String,%20java.lang.String,%20java.lang.String%29] call.

**Message Bundle**

A messages sub-package that contains locale-specific translations is called a message bundle.

**Metadata Model**

See Pentaho Data Models.

**Microsoft Active Directory**

MSAD is created by Microsoft for Windows domain networks. It is included in most Windows Server operating systems.

An MSAD controller authenticates and authorizes all users and computers in a Windows domain type network—assigning and enforces security policies for all computers and installing or updating software.

**Mondrian Schema**

Defines a multi-dimensional database. A Mondrian schema contains a logical model, consisting of cubes, hierarchies, and members, and a mapping of this model onto a physical model. The logical model consists of the constructs used to write queries in the MDX language: cubes, dimensions, hierarchies, levels, and members. The physical model is the source of the data presented through the logical model. It is typically a star schema, which is a set of tables in a relational databases.See http://type-exit.org/adventures-with-open-source-bi/2010/07/a-basic-mondrian-cube-introducing-the-star-schema/,

**N**

**O**

**P**

**Persona**

Personas are archetypal users of applications, websites, and content that represent the needs of larger groups of users, in terms of their goals and personal characteristics. They act as stand-ins for real users and help guide decisions about design,functionality, voice, and content delivery among other things.Personas are used by the Pentaho Documentation Team to craft and deliver content for a particular audience. See Pentaho Personas for a complete list.

**Pre-Configured Installation**

The Pre-Configured Installation (PCI) is a ready-to-use sample deployment of the Pentaho Buiness Analysis Enterprise Edition that can be customized quickly and easily. The PCI deployment includes the following components: JBoss Application Server, JBoss Portal V2.0, sample JSPs that demonstrate platform component usage, sample data, sample reports and Business Analysis processes, users and roles used in samples. The PCI can be modified to work with MySQL, Postgres or Oracle for the RDBMS repository. See Pentaho Product. Is this the Graphical Installer??? Paula

**Pentaho License Key**

As of 2012, Pentaho provides six licenses for its products. Users enter these licenses in the Pentaho Enterprise Console.

* Pentaho BI Platform Enterprise Edition
* Pentaho Reporting Enterprise Edition
* Pentaho Dashboard Designer
* Pentaho Analysis Enterprise Edition
* Pentaho PDI Enterprise Edition
* Pentaho Mobile

The names that include *Enterprise Edition* have an open-source counterpart. Those without are not available as open-source.

**Pentaho Product**

Pentaho licenses Pentaho Business Analysis Enterprise Edition to its customers. This is a suite of data connection, transformation, loading, analysis, reporting, and visualizing components and plugins. Although Pentaho prefers to sell the entire suite, the components and plugins are often bundled in various combinations to meet customer needs. See Product Components and Plugins for a complete list.

**Pentaho Server**

The Pentaho Server is composed of Business Analytics components and Data Integration components in a unified server.

**Pentaho Solution Repository**

This is the RDBMS that stores the artifacts created by Pentaho servers and design tools, including reports, dashboards, charts, data models, properties files, and alike. See server repositories.

**Provider URL**

A URL usually specifying protocol (such as ldap:// or ldaps://), host name, port, and root DN. If you are familiar with the JDBC API, a provider URL is analogous to a URL given along with a user name and password in a [DriverManager.getConnection (url, user, password)|http://java.sun.com/j2se/1.4.2/docs/api/java/sql/DriverManager.html#getConnection%28java.lang.String,%20java.lang.String,%20java.lang.String%29] call.

**Q**

**R**

**Relational Online Analytic Processing**

Relational Online Analytic Processing (ROLAP) is an alternative to MOLAP (Multidimensional OLAP) technology. While both ROLAP and MOLAP analytic tools are designed to allow analysis of data through the use of a multidimensional data model, ROLAP differs significantly in that it does not require the pre-computation and storage of information. Instead, ROLAP tools access the data in a relational database and generate SQL queries to calculate information at the appropriate level when an end user requests it. With ROLAP, it is possible to create additional database tables or summary tables and aggregations to summarize the data at any desired combination of dimensions. See http://en.wikipedia.org/wiki/ROLAP.

**Role**

See Authority.

**Root DN**

The distinguished name (DN) of an object to which all search bases are relative.

**Run-time Repository**

See Server Repositories.

**S**

**Scheduling**

A tab on the User Console that enables a user to manage schedules (edit, create, delete, pause, etc).

**Security Back End**

A repository of usernames, passwords, and roles. The repository can be a flat file, an RDBMS accessed via JDBC, or a directory server accessed via LDAP.

**Security Data Access Object (DAO)**

A method of accessing the security back end. Examples of a security data access object are JDBC, Pentaho secruity (Hibernate-based), and LDAP. Both JDBC and Pentaho security data access objects talk to an RDMBS security back end, although they go about it in slightly different ways.

**Search Base**

A search base is appended to the root distinguished name (DN) to form a search base DN. An LDAP directory is hierarchical. Objects in the directory can have children and those children can have children, and so on. To search for relevant sub trees in the directory, a search base is necessary. The base indicates the DN of an object from which to start searching. Search bases are relative to the root DN.

**Search Filter**

A search filter is an expression that adheres to the rules specified in [RFC 2254|http://www.ietf.org/rfc/rfc2254.txt]. It is always enclosed in parentheses.

**Server Repositories**

The Business Analysis Server (BA Server) includes three embedded repositories that store the data necessary to define, execute, and audit a solution. These include: a solution repository, a runtime repository, and an audit repository. The solution repository contains the metadata that defines solutions. The run-time repository contains items of work managed by the workflow engine. The audit repository contains tracking and auditing information.

**Single Sign-On (SSO)**

Single Sign-On (SSO) is a property of access control of multiple related, but independent software systems. With this property a user logs in once and gains access to all systems without being prompted to log in again at each of them.

**Snowflake Schema**

A way of arranging tables in a relational database such that the entity relationship diagram resembles a snowflake in shape. At the center of the schema are fact tables which are connected to multiple dimension tables. Thus a snowflake simplifies to a star schema when relatively few dimensions are used. The star and snowflake schemas are most commonly found in data warehouses where the speed of data retrieval is more important than the speed of insertion. As such, these schemas are not normalized much, and are frequently left in third normal form or second normal form.

**Solution Engine**

The Business Analysis Server (BA Server) contains the engines and components for reporting, analysis, business rules, email, desktop notifications, and workflow. These components are integrated together so that they can used to solve a BA-related problem. In a solution, the behavior, inter-operation, and user interaction of each subsystem is defined by a collection of solution definition documents. These documents are XML-based and contain the definitions of business processes, definitions that execute as part of processes on-demand, or called by Web services. These activities include definitions for data sources, queries, report templates, delivery and notification rules, business rules, dashboards, analytic views.

**Solution Repository**

See Pentaho Solution Repository.

**Star Schema**

What is this? See Snowflake Schema and Mondrian Schema.

**Summarized Data**

See Aggregate.

**Summarization**

See Aggregation.

**Superuser**

A security user type. Specifically, the top-level user with security access to all software features. Some systems have names for superusers, such as *root* or *admin*.

**Step**

Used in Pentaho Data Integration, a step is a container for a set filters and commands that prepare data for analysis. Many steps comprize a PDI transformation. See **Transformation**.

**Superuser**

A user type for assigning security permissions. Superuser is the top-level user with security access to all software features, folders, and files. Some systems have names for superusers, such as *root* or *admin*.

**Supported Technologies**

The preferred name of the list of technologies and components that we officially support. In the past, this was known as the Support Matrix or Supported Components.

**T**

**Transformation**

In Pentaho Data Integration, a transformation is a way to automatically filter and manipulate data to prepare it for analysis. See **Step**.

**U**

**V**

**W**

**X**

**Y**

**Z**