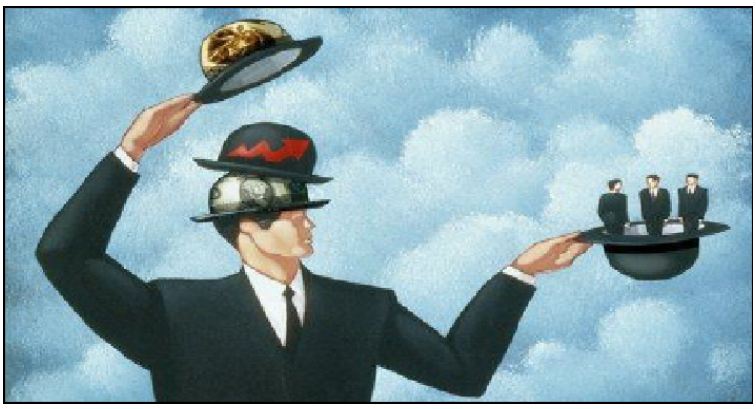
DECISION SUPPORT PRIMER
A BREAKDOWN OF INDUSTRY JARGON



a primer to help sort out the alphabet soup nomenclature
in decision support and business intelligence

Over the last decade, there has been a large proliferation of specialized software designed specifically to help organizations with their decision-making. These applications often gather data from a variety of sources and structure them in a manner where business understanding and decision-making is better and more agile. This science of decision support has spawned its own "legalese terminology" that can be confusing to the business user – this article is intended as layman's explanation of some of the most prevalent terms.

Data is obviously the most basic term used in the language of decision support. Even before the proliferation of these specialized systems, data was being recorded and stored in databases. There are various types of databases, but the important point is that databases are structured repositories of data that are often optimized for storage and/or retrieval of the data (depending on which is more important for a particular task).

Many software applications are associated with databases – depending on the systems architecture, databases are often independent of an application and often used by multiple applications. Some of the largest databases are associated with transaction-based applications designed to support typical day-to-day operations (e.g., logging customer contacts, recording retail sales).

Although these systems tend to be some of the largest in service today, they were also among some of the earliest developed corporate systems. As these transaction-based applications proliferated, they continued to record daily business activity – and the associated databases thus grew tremendously.

GENERAL SERIES

People soon began to realize that these massive databases contained a rich repository of historical data. And that this data, if analyzed properly, could identify interesting business trends and make decisions-making more predictable. So, analysts began writing queries to answer specific business questions. However, this process was difficult because the analyst had to have technical knowledge and the large databases were not designed for this type of ad-hoc query. In many cases, various databases from different transaction-based applications had to be queried to answer a simple business question – this was a difficult, inaccurate and time-consuming process.

To solve this problem, data warehouses were created. A data warehouse is a centralized repository for an organization's data that is usually founded from multiple internal and external sources. This compilation, often optimized for analysis and processing of large amounts of data, ultimately provides a comprehensive and homogenized view of an organization's data. This makes it much easier to run queries over data that originally came from disparate sources.

Whereas a data warehouse combines enterprise-wide databases, data marts focus on a particular business subject. Since all data repositories tend to be complex constructions, there is often a need to clarify their structure through metadata – data describing the data in the data repository (e.g., source of data, when it was collected).

Often, as part of process of importing data, a considerable amount of data cleansing occurs to ensure that correct and meaningful data exists within the data warehouse. Data cleansing is part of a more holistic process of data management which involves data gathering (collection of data from various sources), data profiling (understanding of the data), data integrity (improving the data), data assimilation (combining similar data from multiple sources) and data augmentation (improving the value of the data). For many enterprises, an alarming amount of the data entered into their transaction-based applications is incorrect. Although data issues would ideally be addressed at the point of data entry, in many cases, this is just not practical and thus the data cleansing process often happens after-the-fact.

Whereas operational data is generally real-time, data within the data warehouse is usually historical (although the near future involves active data warehousing and near-real-time data). The data warehouse is usually updated with data from appropriate transaction-based systems on a pre-determined time schedule (quarterly, monthly, weekly or daily as needed). This near real-time updating aspect is adequate since the data is used primarily for reporting and analysis purposes. An organization can choose to develop the data warehouse over time by adding imported data as time and resources permit. Sometimes, the data in the warehouse is loaded using a structured process called ETL which stands for extract, transform and load – essentially, this data migration process loads data from one database into another primarily to form data warehouses and data marts.

Once the data warehouse is built, the organization can then use it to perform structured analyses. These analyses include data mining, the practice of extracting data from a data warehouse in order to analyze patterns and trends, and data modeling, the practice of analyzing an organization's data and identifying the relationships among the data. The resulting analysis engine is then used to create user-friendly reporting modules (which

GENERAL SERIES

sometimes includes ad-hoc analysis capabilities and what-if analysis tools) to provide information to help management make decisions about the future of the organization. Often, a data warehouse is so large and complex that it cannot provide efficient and effective access for certain applications. Online Analytical Processing (OLAP) refers to a category of software tools that provide analysis of multidimensional data (especially time trends) – often OLAP databases contained detailed views of a segment of the data (smaller than what is in the data warehouse). The OLAP server understands and is optimized for how data is organized and has special functions for expediting data analysis. Usually larger configurations (more data and more users) and special user application requirements need OLAP servers to sit between the OLAP database and user applications making data retrieval quicker and simpler.

The robustness of data warehouses, analysis tools and advanced information mapping techniques especially graphical data representations elevated the science of decision making into a whole new realm – and Business Intelligence was born. Business Intelligence (BI) systems were relatively easy to use and readily replaced the earlier,difficult-to-use, techno-centric systems called DecisionSupport Systems (DSS).

Business Intelligence, as its name suggests, is the amalgam of all that is necessary to turnmassive amounts of corporate data into useful information. Entire companies were formed selling holistic, stem-to-stern decision-making solutions to the Fortune 500 and beyond. These all-inclusive Business Intelligence (BI) solutions contain sophisticated data warehousing, analysis and reporting sub-systems pre-packaged for a cross-section of corporate functions such as finance, sales, marketing, and human resources.

Graphical representations of information (especially measurements) became the cornerstone of many Business Intelligence offerings. Many of these displays were used to provide information on organizational or departmental measurements – often interchangeably referred to as metrics or key performance indicators (KPI). Metrics and other key information about the enterprise were presented in scorecards meant to provide a ready assessment on organizational performance and dashboards meant to provide diagnosis and feedback. Scorecards and dashboards were meant to provide information regarding key business parameters enabling management to “steer the ship” quickly and in the right direction.

As Business Intelligence was maturing, the underlying data was being segmented into specialized functional categories such as finance, sales and marketing, and supply chain. Some BI solutions were designed specifically for a functional area while others ranged the spectrum of corporate functions. More detailed data along these business dimensions gave rise to specific decision support tools called Analytic Applications. These applications provide focused and detailed information along one functional dimension. They often include complex and detailed analysis and reporting capabilities often supplementing what is available in the core BI system.

The static, largely historical retrospective provided by BI tools although useful for some purposes fell short of meeting the needs of some companies. Organizations needed decision-making tools which were forward-looking as opposed to historical. Thus, over the last couple of years, we have witnessed the advent of the latest generation of decision support tools, Business Performance Management (BPM) systems. These systems are

GENERAL SERIES

forward-looking, process oriented and replace measuring with managing as the cornerstone parameter. Ultimately BPM (or alternatively called CPM) tools intend to provide a holistic profile of decision tools specifically catered to not only measuring but also improving performance.

In the area of performance improvement, vendors have pushed the envelope even further and have developed more specialized decision support systems. One of the problems that exist with most BI and BPM solutions is that they are not real time (and they rarely need to be). Most decisions, even important strategic ones, can be made based on data that is one week, one month, or even one quarter old (which is probably how often the underlying data warehouse accessed by BI or BPM tools is updated). There are, however, certain business situations which need to analyze more timely data (e.g., reports on the previous hour's web activity). Business Activity Monitoring (BAM) solutions provide near-real-time access to critical business information – the underlying data warehouses for the systems are intimately attached to transaction-based applications.

The current trend in decision support, BI and BPM is to provide "big elephant" solutions covering all the needs of the enterprise. While it is reasonable to think that one holistic, interconnected decision-support system can be "the be all, end all" solution, in practice, it often fails miserably. Often, customized analytic decision-support applications designed for specific purposes and custom-modified for a particular organization's idiosyncrasies deliver much higher ROI. The future of decision-support is likely to be smallish, stand-alone, business-specific analytic applications that communicate with one another and share an underlying data platform.

HIGHLIGHTS

The future of effective decision-support is likely to be smallish, stand-alone, business-specific analytic applications that communicate with one another and share an underlying data platform.