Chapter 2
Business Intelligence

Business intelligence (BI) systems are data-driven DSS, focusing on the manipulation of large volumes of company data in data warehouses. As other types of DSS, such as model-driven DSS, communication-driven DSS, and knowledge-driven DSS (Section 1.5.3), BI systems suffer from lack of cognitive decision support, despite their powerful data analysis functions. This chapter will discuss the basic BI concepts, architecture and some vendors' BI products. The drawbacks of today's BI are also analyzed in details.

2.1 What Is Business Intelligence?

The term business intelligence means different things in different domains. From technical point of view, BI refers to the process of extracting, transforming, managing and analyzing business data, in order to support decision making. This process is mainly based on large data sets, particularly data warehouse, with the mission of disseminating intelligence or knowledge across the whole organization, from strategic level to tactical and operational level. A typical BI process consists of five key stages (CSIRO 2003):

(1) Data Sourcing

A BI system is able to extract data from multiple data sources, representing different business units, such as marketing, production, human resource and finance. The extracted data must be cleaned, transformed and integrated for analysis.

(2) Data Analysis

In this stage, data is converted into information or knowledge through different data analysis techniques, such as reporting, modeling, visualization and data mining. The results of data analysis help managers to have a better understanding of the environment and make better decisions.

L. Niu, J. Lu, and G. Zhang: Cognition-Driven Dec. Supp. for Business Intel., SCI 238, pp. 19–29.
springerlink.com

(3) Situation Awareness

Situation awareness (SA) is a deep understanding of the current decision situation based on the results of data analysis. SA is a key prerequisite for decision making. BI systems should be able to aid decision makers to develop rich SA about their decision situations.

(4) Risk Assessment

Richer SA can help managers to make prediction about the future, identify threats and opportunities, and thus respond correspondingly. Today's businesses are operated in an increasingly complex environment. Business decision making is more likely to suffer risks from the external and internal environment. Thus, risk assessment is an important function of a BI system.

(5) Decision Support

The ultimate goal of BI is to help managers to make decision wisely, based on the current business data.

2.2 The Architecture of a Business Intelligence System

A typical BI system consists of four levels of components and a metadata management module (Codd et al. 1993; Inmon 2002). The generic architecture of traditional BI systems is shown in Figure 2.1. These different components cooperate with each other to facilitate the basic BI functions: extracting data from company operational systems, storing the extracted data in a center data warehouse, and retrieving stored data for various business analysis applications.

- Operational Systems Level

As the data sources of a BI system, business operational systems are mainly online transaction processing (OLTP) systems which support daily business operations. Typical OLTP systems are customer order processing systems, financial systems, and human resource management systems.

- Data Acquisition Level

This level is a data pre-process component including three phases: extracting, transforming, and loading (ETL). A company usually has different OLTP systems producing huge amounts of data. This data is first extracted from OLTP systems by the ETL process and then transformed according to a set of transformation rules. Transformed data is clean, unified, and aggregated and finally loaded into a central data warehouse. ETL is the most fundamental component of a BI system because the data quality of all other components mainly relies on the ETL process. In the design and development of ETL, data quality, system flexibility, and processing speed are the major concerns.

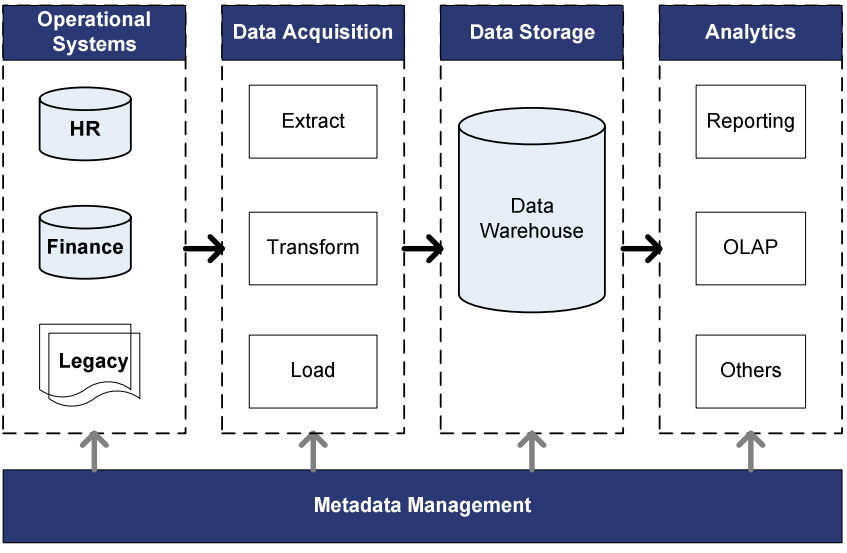

Fig. 2.1 The General Architecture of Current Business Intelligence Systems

- Data Storage Level

The data processed by the ETL component is stored in a data warehouse which is mainly implemented using traditional relational database management systems (RDBMS). A RDBMS is designed to support transaction processing. By contrast, a data warehouse is a subject oriented, time-variant, non-volatile and integrated data store (Inmon 1993)[1]. Data from company OLTP systems are extracted, transformed, and loaded into the data warehouse based on pre-defined schemas. The star schema and snowflake schema are the most popular data warehouse schemas. No matter what kind of schema on which a data warehouse is designed, the data warehouse always includes two basic types of tables: fact tables and dimension tables.

- Analytics Level

Based on the data warehouse, various kinds of analytical applications are developed, which represent the last level: Analytics. BI systems support two basic types of analytical functions: reporting and online analytical processing (OLAP). The reporting function provides managers with different business reports, such as sales reports, product reports and human resource reports. Reports are generated by executing queries into the data warehouse (DW). The DW queries are mainly pre-defined query sentences programmed by the DW developers. Thus, the reports generated by BI systems usually have static formats and contain fixed types of data.

[1] This is the original definition of data warehouses. Today's data warehouse systems can be company wide in scope and the data can be updated over time, for instance, as real-time BI systems.

The most promising BI analytics is OLAP. OLAP allows managers to efficiently browse their business data from different analysis dimensions through slicing, dicing and drilling operations at will (Codd et al. 1993). An analysis dimension is a perspective through which data is presented, e.g., product type, sales location, time and customer. Compared to the reporting function, OLAP supports ad hoc data analysis, i.e. managers have full control of the data by selecting different analysis dimensions of interest to them. OLAP is based on multidimensional data models (known as the snowflake and star schema).

In addition to reporting and OLAP, there are many other types of analytical applications which can be built on the basis of a DW system, such as data mining, executive dashboards, customer relationship management, and business performance management. Technically, these applications are not necessarily built on a data warehouse. However integrating them with DW systems has become a common practice in many practical BI systems.

- Metadata Management

Metadata are special data about other data such as data sources, data warehouse storage, business rules, access authorizations, and how different data is extracted and transformed. Metadata is crucial for producing accurate, consistent information and system maintenance. It affects the entire process of designing, developing, testing, deploying and using BI systems (Caserta 2004; Inmon 2002).

2.3 Analytics of Business Intelligence

The analytics are the core part of a BI system. Evelson and colleagues (2008) summarized BI analytics into eight categories based on a lab-based evaluation of popular BI tools on the market.

(1) Production/operational reporting for pixel-perfect mass report distribution

No matter how much BI self-service end users request, good old-fashioned report development tools, mainly used by professional programmers, remain at the heart of a BI product line. While these tools may also be used to analyze data and produce visual dashboards, they are primarily used for mass distribution of very sophisticated reports like customer statements. Requirements for these products often include pixel-perfect positioning of data and graphics, a scripting language equal in power to a full programming language, and the ability to handle complex headers, footers, nested subtotals, and multiple report bands on a single page.

(2) Ad hoc query tools provide a quick answer to a business question

When report formatting or distribution is not a requirement, and an information management professional just needs a quick answer to a business question like, “How many units of a product were sold yesterday across all stores and outlets?” or, “What were my total sales in 2007 in North America?” simple ad hoc query tools with an intuitive point-and-click user interface (UI) are the way to go.

(3) OLAP tools, when business questions are more about "whys" than "whats"

While reporting and ad hoc query tools are typically used to answer the questions like "What happened?" and"When and where did it happen?", online analytical processing (OLAP) tools are used to answer the questions like "Why did it happen?" and also to perform "What if?" analysis. Otherwise known and "slicing and dicing" analysis (essentially a spreadsheet pivot table on steroids), OLAP tools allow a power user to see any facts (numerical, typically additive numbers, like transaction amounts and account balances) almost instantaneously regrouped, re-aggregated and resorted by any dimension (descriptive elements like time, region, organizational unit, or product line).

(4) Dashboards as an interactive, visual UI — not a reporting or analytical tool by itself

Dashboards should be used as a UI to operational or analytical information. Designed to deliver historical, current, and predictive information typically represented by key performance indicators (KPIs), dashboards use visual cues to focus user attention on important conditions, trends, and exceptions. The term *dashboard* is often used synonymously with *scorecard*, but Forrester defines a scorecard as just one type of a dashboard that links KPIs to goals, objectives, and strategies. Many scorecards follow a certain methodology, such as Balanced Scorecard, Six Sigma, Capability Maturity Models, and others. Other dashboard varieties include business activity monitoring (BAM) dashboards and visualizations of data/text mining operations.

(5) BAM will report on real-time data and process information streams

While a dashboard can be used as a graphical user interface (GUI) component, BAM also captures data and process events (e.g., number of credit applications processed today and number still pending in a queue), correlates and aggregates them into business metrics (e.g., ratios of processed, approved, and rejected applications per hour), and displays the real-time status of the metrics and trailing patterns.

(6) Predictive modeling answers questions about what's likely to happen next

Using various statistical models, these tools attempt to predict the likelihood of attaining certain metrics in the future, given various possible existing and future conditions. One typical predictive modeling class is called market basket analysis, which tries to predict the likelihood of a customer buying a certain product if and when he or she bought another product at a certain store at a certain season, date, and time, given certain economic conditions such as interest rates and price of gas.

(7) BI workspaces enable true end user self-service

While most BI environments attempt to address end user self-service requirements, they still impose many restrictions, such as fixed data models, an inability to add new dimensions on the fly, and sometimes restricted access to

production data. Forrester defines a BI workspace as a data exploration environment where a power user can analyze production, clean data with near complete freedom to modify data models, enrich data sets, and run the analysis whenever necessary, without much dependency on IT and production environment restrictions. Some examples of such workspaces are desktop-based multidimensional OLAP (MOLAP) cubes, in-memory data models, or BI software-as-aservice (SaaS).

(8) Guided BI search tools support free form ad hoc queries and analysis

While reporting, ad hoc queries, and OLAP tools work best when one knows the exact business question, they fall short when a user is looking for something that he or she is not quite sure of. A salesperson getting ready for an important client meeting may not know all of the information required to prepare for the meeting and may not be able to effectively construct the appropriate queries to pull the information he/she might need. What works much better is enabling this salesperson to simply enter a few keywords to find relevant customer dimensions in the database, then using a graphical interface to drill into the information he/she wants from a list of possibilities. This effectively solves one of the oldest dilemmas in BI: having to know exactly which questions to ask to get a meaningful answer.

2.4 Commercial Tools

2.4.1 SAS Business Intelligence

Website: *http://www.sas.com/technologies/bi/*

SAS BI offers a full breadth of SAS Analytics capabilities, including statistics, predictive analytics, data and text mining, forecasting, and optimization. These functions are integrated within the business context for better, faster decision making. SAS BI has two components: *Enterprise Business Intelligence* and *Business Visualization*.

The typical functions of SAS BI are as follows.

- Web and desktop reporting

SAS BI supports a wide variety of targeted, fit-to-task interfaces for report building, viewing and distribution for all levels of users across an organization.

- Portal and customizable dashboards

SAS BI provides users an easy-to-use, role-based Web portal, via which users can access aggregated information. It also includes a dashboard development environment, enabling users to create their own dashboards of different styles from virtually any data source.

- Microsoft Office integration

Microsoft Office can be integrated with SAS BI, bringing SAS capabilities in data access, reporting and analytics directly from Microsoft Office.

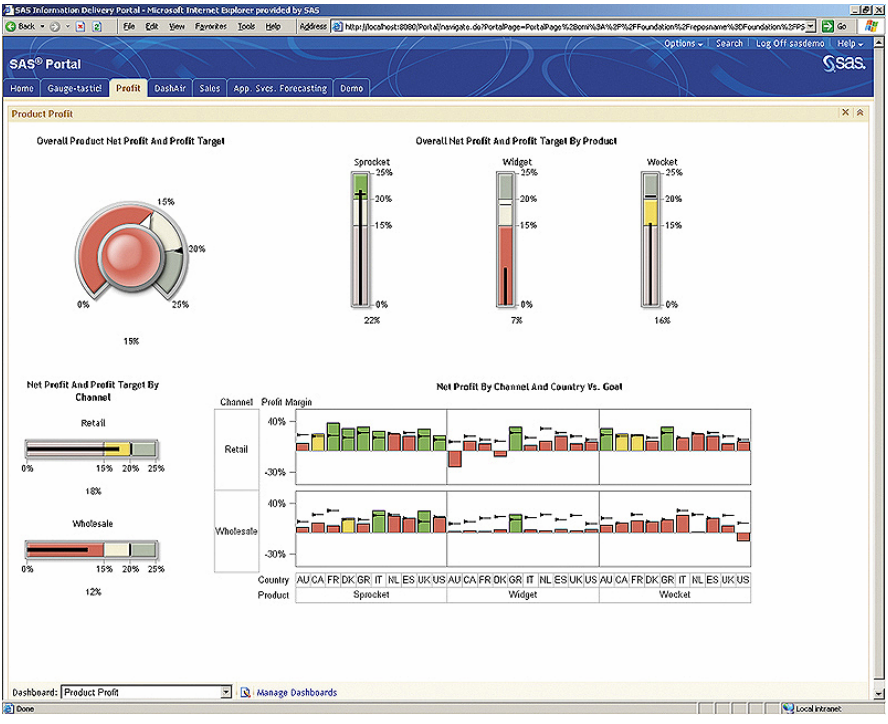

Fig. 2.2 A Screenshot of SAS BI

- Query and Analysis

SAS BI has a user-friendly interface to allow users to easily access and query data on their own without the help of IT staff.

- Interactive business visualization

SAS BI presents data in charts, graphs and geographic maps within multiple BI applications.

- OLAP storage and OLAP data exploration interface

Users can work on their Web browsers to create OLAP cubes and interact with SAS BI to view the multidimensional data from different business dimensions.

- Integrated analytics

Users can access sophisticated analyses directly from their BI interface for decision making.

- Guided analysis.

A dynamic Windows interface can guide users during model development. This function enables business analysts, statisticians and programmers to leverage SAS analytics and efficient processing across all enterprise platforms.

A screenshot of SAS BI is shown in Figure 2.2.

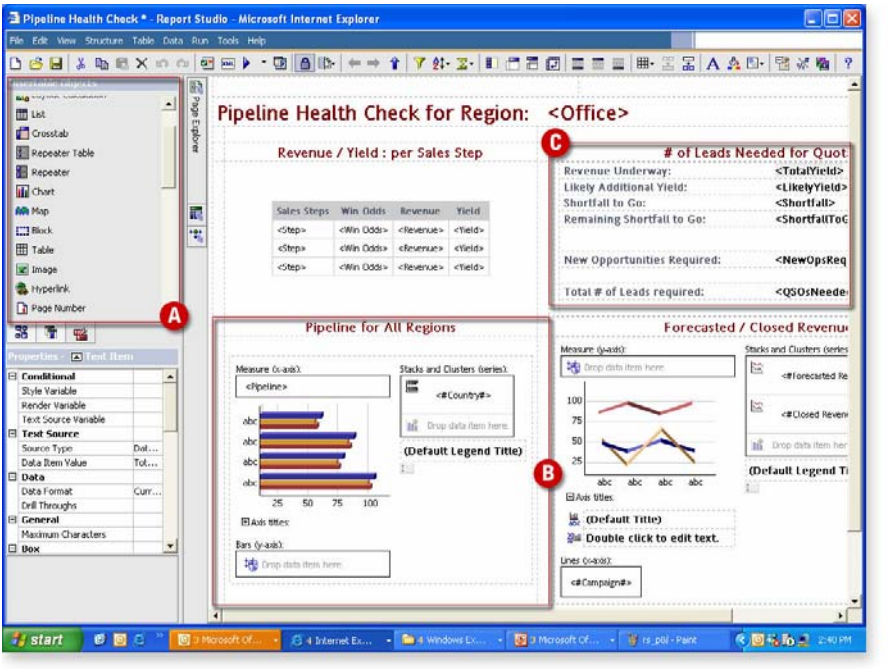

Fig. 2.3 IBM Cognos BI

2.4.2 IBM Cognos Business Intelligence

Website: **www.ibm.com**

IBM Cognos BI provides full breath of BI analytics, such as various styles of reporting, analysis, score carding, and dashboards. It has four major components:

- Analysis Studio

In Analysis Studio, users can explore their business data to find trends and comparisons that answer their business questions. The analysis in Cognos is no longer strictly against Cognos Power Cubes but also against relational data sources.

- Report Studio

Report Studio is the platform for power users and professional developers to create formatted reports that contains multiple charts or tabular data sets from multiple subject areas. Additional chart types, such as gauges and maps, that aren't available in Query Studio or Analysis Studio, are also supported in Report Studio. A screenshot of Report Studio is shown in Figure 2.3.

- Query Studio

Query Studio is used for ad hoc report authoring. Users can easily query any data sources (relational, multidimensional or planning data source) to create crosstabs, simple charts or detailed reports. Query Studio also provides formatted templates to give any report a standard corporate layout or logo.

- Metrics Studio

Metrics Studio is used for scorecarding. Users can monitor business performance through different parameters. The idea behind it is to put performance indicators next to the organization's key performance measures (red, orange, green status notation).

2.4.3 SAP BusinessObjects Business Intelligence

Website: www.sap.com

SAP BusinessObjects provides a full spectrum of BI functionality, ranging range from reporting, query, analysis, dashboards and visualization, to intuitive discovery and advanced predictive analytics capabilities, as well as data quality and extract, transform, and load functionality. The components and functions of SAP BusinessObjects are categorized as follows.

- Information Infrastructure

The information infrastructure of SAP BusinessObjects allows IT department to extend BI to any application or process in any environment. It provides following functions: Auditing, BI content search, BI widgets Central management console, Encyclopedia, InfoView, Java portal integration kit, Life-cycle management, Publishing, Process tracker, Query as a Web service, Microsoft Office SharePoint portal integration kit, Software development kits, and Universe designer.

- Reporting

SAP BusinessObjects contains a reporting tool: Crystal Reports providing users an interface to connect to virtually any data source, design and format interactive reports, and share them internally and externally. The final reports can be delivered via SAP BusinessObjects Enterprise, Crystal Reports Viewer, and Microsoft Office documents.

- Query, Reporting, and Analysis

There are two tools in SAP BusinessObjects, Desktop Intelligence and Web Intelligence, allowing users to perform ad hoc query, reporting and analysis, without having to understand complex database languages and underlying structures. Figure 2.4 is a screenshot of SAP BusinessObjects Web Intelligence.

- Dashboards and Visualization

SAP BusinessObjects has nine software tools to support information visualization, such as Xcelsius Enterprise, Dashboard Builder, and VizServer. This set of tools also provides professional developers a software development kit to customize advanced visualization interfaces.

- Advanced Analytics

SAP BusinessObjects Voyager is the OLAP interface through which users can explore multidimensional data. SAP BusinessObjects Predictive Workbench is a

software tool enabling users to conduct predictive analysis. SAP BusinessObjects Set Analysis is a software tool enabling users to conduct clustering analysis.

- Search and Navigation

With SAP BusinessObjects, business users can search BI contents from internal and external resources, such as structured databases, business intelligence (BI) systems, unstructured company and text content, search engines, and the Web.

2.5 Limitations

Business intelligence is promising to turn 'data' into 'knowledge' and help managers survive data tsunami and eventually succeed in decision making. However, BI systems are essentially data-driven DSS. Current BI systems can only partially support managers' work (Singh et al. 2002). The emphasis of BI analytics is manipulation of large volumes of business data, rather than supporting managers' decision making from the cognitive perspective.

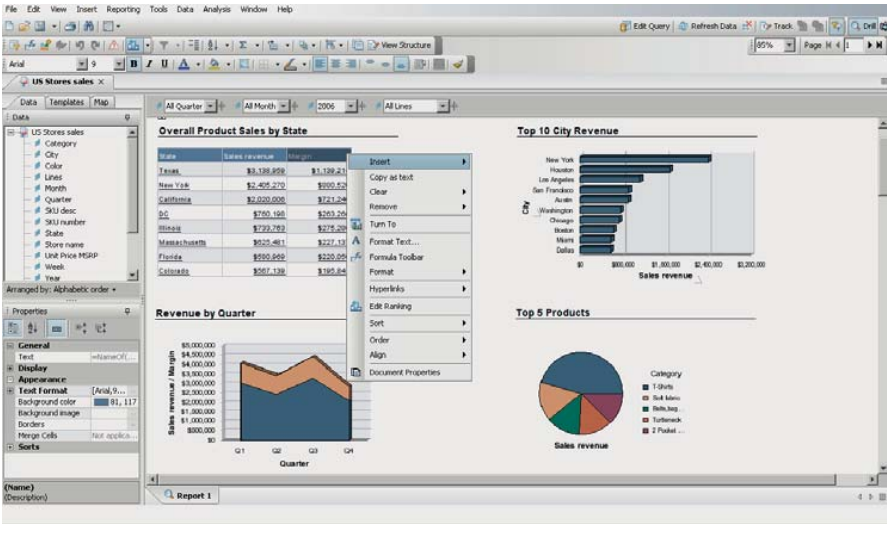

Fig. 2.4 SAP BusinessObjects Web Intelligence

A BI system is capable of providing managers with huge amounts of internal and external business data, but more data does not equal more valuable information (Endsley et al. 2003). On one hand, the reporting function is mainly pre-defined information representation. That is, business reports are generated in fixed types and formats by executing pre-defined queries into the data warehouse. The pre-defined reports are efficient and effective in reporting repetitive and structured business events, for example periodical (daily, weekly, monthly, quarterly and annual) product sales. However, pre-defined reports are not as flexible as many other ill-structured events require, for example, unpredictable marketing campaigns. On the other hand, OLAP-based ad hoc analysis gives

managers the full control of their data. Managers can easily obtain any data of their interest by selecting analysis objects and customizing analysis dimensions. Nevertheless, mangers often feel lost when confronted with a large body of business data concerning a decision problem (Resnick 2003).

A recent survey by Economist Intelligence Unit (2006) shows that 73 per cent of senior managers agreed that it is important to have less but more timely data to improve the quality and speed of decision making. This result corresponds to the research result by Sutcliffe and Weber (2003) about knowledge accuracy. Their research implies that having a lot of facts about a decision situation is less important than having a clear and consistent overview. Resnick (2003) criticizes the current executive dashboard design (a type of BI application) for emphasizing improvement on data analysis functionality while falling short of cognitive engineering consideration. More recently, an industrial report from InfoWorld Media Group shows that 'BI has a reputation for being a resource sink that delivers reports almost no one reads. It doesn't have to be that way. And you can no longer afford to let it be' (Gruman 2007, p. 22).

2.6 Summary

The identification of the advantages and disadvantages of traditional BI system motivated us in this research to seek a new way to better support managers' work, particularly for handling ill-structured decision problems. This book aims to provide models and techniques to facilitate cognitive decision support on the basis of BI platforms. Thus, it is necessary to look at the nature of cognitive decision support from the cognitive psychology perspective. In next chapter, relevant concepts and decision models in cognitive psychology will be reviewed.