

---

---

## Data Warehouse : Intelligent Management Decision Support

Shweta Pandya

Bhaumik Shroff

### Abstract

*Advances in computer and networking technology have led to the introduction of very powerful hardware and software platforms that can collect, manage and distribute large amount of pertinent data. Data Warehousing is one of the most interesting and dynamic among the new technological transitions. It is a repository of subjectively selected and adapted operational data, which can successfully answer any ad hoc, complex, statistical or analytical queries and it provides means for implementing effective decision support system by utilizing data scattered all over the organization. The aim of the paper is to generate interest and understanding among the library and information professionals regarding the concept of data warehousing. Major characteristics, goal, structure, application and advantages of data warehousing are also described in detail.*

**Keywords :** Data Warehouse, Information Retrieval

### 0. Introduction

“You have no choice but to operate in a world shaped by globalization and the information revolution. There are two options: adapt or die.” (Andy Grove, Chairman, Intel). The last few years have seen a growing recognition of information as a key tool for decision making. Those who successfully gather, analyze, understand and act upon the information are among the winners in this ‘information age’. Decision-making is progressive resolution of uncertainty and is a key to a purposeful behaviour. Decisions are required to be taken quickly and correctly using all the available data. As the data size increases continuously, the speedy processing of this data is a prime requirement, which helps to comprehend the meaning of the data. Competition also adds pressure on this situation and here the need for data warehousing arises, in terms of the ability to organize, maintain and analyze larger data in few seconds in the manner and depth as required. A data warehouse is a repository of subjectively selected and adapted operational data, which can successfully answer any ad hoc, complex, statistical or analytical queries. A data warehouse is becoming more of a necessity than an accessory for a progressive, competitive and focused organization. It provides the right foundation for building decision support and executive information system tools that are often built to measure and provide a feel for how well an organization is progressing toward its goals.

### 1. Data Warehouse

In 1993, the father of data warehousing, Bill Inmon gave the definition of data warehouse: “A data warehouse is a subject oriented, integrated, time-variant, nonvolatile collection of data in support of management decisions.”<sup>1</sup>

The data warehouse is oriented around the major subjects of the organization. The data-driven, subject orientation is in contrast to the more classical process/functional orientation of applications around which most older operational systems are organized. The data warehouse focuses on data modeling and database design exclusively. The data in the data warehouse is integrated. The integration in the data warehouse is shown in many ways – in consistent naming conventions, in consistent measurement of variables, in consistent encoding structures, in consistent physical attributes of data and so on. All data

in the data warehouse is accurate for some moment of time. This basic characteristic of data in the warehouse is very different from data found in the operational environment. Every key structure in the data warehouse contains implicitly or explicitly an element of time, such as day, week, month etc. Once correctly recorded with time the data cannot be updated in a data warehouse unlike the operational systems.

According to the definition, a data warehouse differs from an operational database in four major ways.

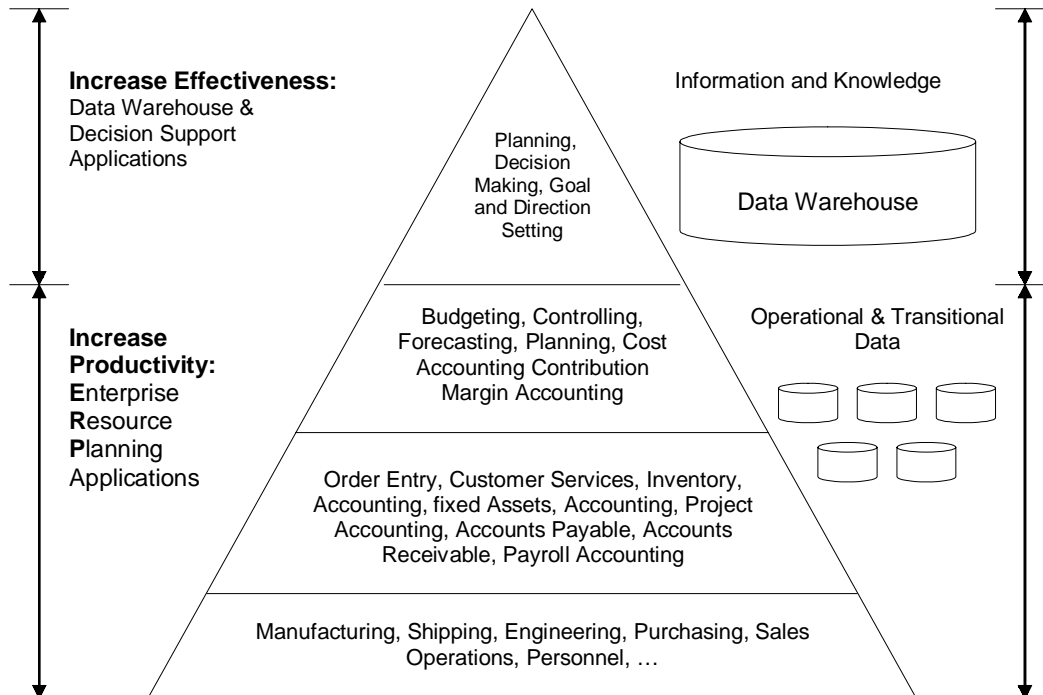
<b>Data Warehouse</b>	<b>Operational Database</b>
subject oriented	application oriented
Integrated	multiple diverse sources
time-variant	real-time
Nonvolatile	Updateable

An operational database is designed primarily to support day to day operations where as a data warehouse is deigned to support strategic decision making. Data Warehousing is an approach to the building of computer systems whose primary functions are the storage, cataloging, retrieval and analysis of data. It is a set of hardware and software components that can be used to better analyze the massive amounts of data that organizations are accumulating to make better management decisions.

## **2. Goal of the Data Warehouse**

Any organization that has some computerized record keeping systems and is interested in deducting or drawing logical conclusions from their voluminous, granular and detailed information pool should consider building an organization level data warehouse application. These organizations will then be capable of improving their insights into the trends in their operations and eventually increase the accuracy of their forecasts and plans. The effectiveness of the data warehouse application intensifies especially when the operational data resides in distributed, non-homogenous systems and replace manual data gathering and reconciliation procedures. Therefore, data warehousing can be useful to not only commercial sectors but also to sectors like government, health care, insurance, manufacturing, finance, distribution, education and so on.

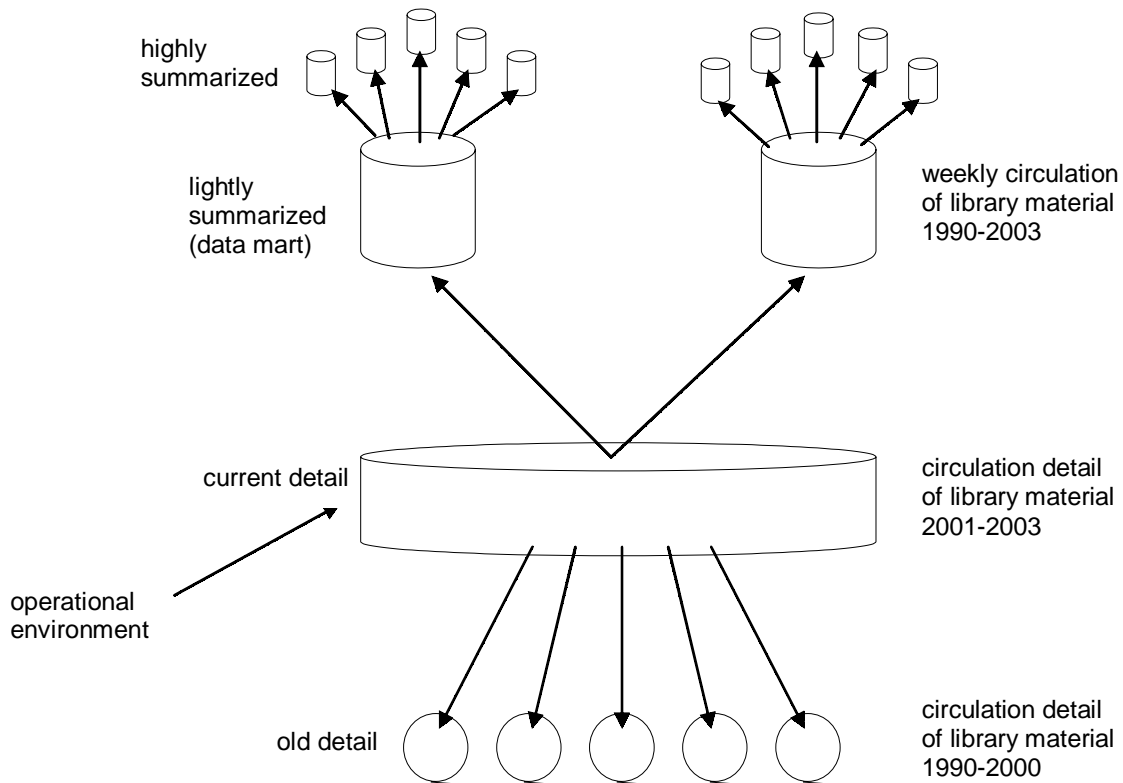
The goal of a data warehouse application in an organization is to increase the effectiveness of the decision-making and direction setting process. The following figure<sup>2</sup> gives an idea about the goal of the data warehousing.



**Figure 1: Goal of the Data Warehouse**

### 3. Structure of the Data Warehouse

There are different levels of detail in the data warehouse. They are an old detail, a current detail, a lightly summarized data and a highly summarized data. Data flows into the data warehouse from the operational environment. Once the data ages, it passes from current detail to older detail. As the data is summarized, it passes from current detail to lightly summarized data, then from lightly summarized data to highly summarized data. The structure of data warehouse is presented below in figure 2.<sup>3</sup>



**Figure 2: The Structure of Data Warehouse**

#### 4. Data Warehousing Applications

The applications served by data warehousing can be classified in one of three main categories.

- “Personal productivity applications” such as spreadsheets, statistical packages and graphics tools are useful for manipulating and presenting data on individual PCs. Though they are developed for a standalone environment, they address applications requiring only small volume of warehouse data.
- “Data query and reporting applications” deliver warehouse-wide data access through simple, list-oriented queries, and the generation of basic reports. These reports provide a view of historical data but do not address the need of organization for in-depth analysis and planning.
- “Planning and analysis applications” address such essential management requirements as budgeting, forecasting and customer profitability, financial consolidations and applications that use historical, projected and derived data.

---

---

## 5. Advantages of Data Warehousing

Data warehousing is being considered one of the most strategically significant developments in information processing in recent times. One of the reasons for this is that it is seen as part of the answer to information overload. Some of the advantages of data warehousing are described below.

### 5.1. Has a subject area orientation

Unlike classical operations systems that are organized around the applications of the organization, data is arranged in a data warehouse according to subjects.

### 5.2. Integrates data from multiple, diverse sources

Integration is the most important from all aspects of a data warehouse. Data is fed from multiple disparate sources into data warehouse. As the data is fed, it is converted, reformatted, resequenced, summarized and so forth. The result is that data once it resides in the data warehouse has a single physical organization image.

### 5.3. Allows for analysis of data over time

The operational database provides detailed current information, often with last year's data available for comparison, but analytical queries typically require much more. With a data warehouse, data snapshots taken at times that are significant to the decision making process make it possible to analyze trends over time.

### 5.4. Adds *ad hoc* reporting and enquiry

As pointed out in the definition, a data warehouse is designed to be accessible with end-user tools, and this allows *ad hoc* reporting and analysis by end-users.

### 5.5. Provides analysis capabilities to decision makers

Data warehouse provides dynamic data analysis, which provides an understanding of the changes occurring within an organization and may be used to identify appropriate solutions to specific problems or challenges, as they are uncovered and to facilitate the development of future strategies.

### 5.6. Relieves the development burden on IT

Data warehouse is useful to answer most of the users' queries according to their need, which relieve burden of development from IT staff. Users once trained for the use of appropriate desktop tool can find answers to their questions with the help of data warehouse.

### 5.7. Provides improved performance for complex analytical queries

A data warehouse provides improved performance for complex, analytical queries by denormalization and aggregation. Frequently used aggregates are often precomputed and materialized in views commonly known as summary tables. These materialized views provide fast access to integrated data, regardless of the original data sources.

### 5.8. Allows for a continuous planning process

Data warehouse provides powerful and integrated system, which helps in making planning a continuous process. Therefore, it becomes essential for quick and effective analysis and reporting.

### 5.9. Converts organization data into strategic information

Acquiring right information at the right time from the huge ocean of data is becoming difficult day by day. Data warehouse helps to convert the information overload in to meaningful strategic information. It is very useful to overcome the limitation of traditional tools by displaying different dimensions as required.

Thus, a well defined and implemented data warehouse can be used to understand trends and make better forecasting decisions, bring better products to market in a more timely manner and analyze daily information and make quick decisions that can significantly affect the organization's performance.

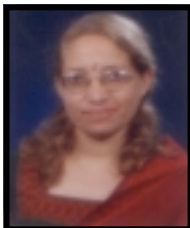
## 6. Conclusion

Data warehousing is not a new phenomenon. All large organizations already have data warehouse, but they are just not managing them. It is very helpful as it provides the means to change raw data into information for making effective decisions. Data in the data warehouse is preprocessed and presented such that it facilitates the cross functional monitoring and assessment of the overall direction of the organization. Thus, it is the hub for an intelligent management decision support. Successful implementation of a data warehouse requires a high-performance, scaleable combination of hardware and software, which can integrate easily with existing systems, so that users can use data warehouse to improve their decision-making. A data warehouse is incomplete until it provides the exploitation tools that enable end users to view, analyze and report on data in ways that support their decision-making. Data marts, data mining, data modeling and metadata are some other important concepts attached with data warehousing, the knowledge of which helps to a great extent in data warehouse implementation.

## 7. References

1. Inmon, W.H. Building the Data Warehouse. USA: John Wiley. p. 33.
2. Yazdani, Sima; and Wong, Shirley. Data Warehousing with Oracle. New Jersey: Prentice Hall, 1998. p. 4, 12.
3. Inmon, W.H. Building the Data Warehouse. 3<sup>rd</sup> edn. New Delhi: Wiley Dreamtech, 2002. p. 36.

### About Authors



**Ms. Shweta Pandya** is the Librarian of GLS Institute of Business Management and Computer Technology, Ahmedabad, India. She holds MLISc with 1<sup>st</sup> rank, CIC, DIM and qualified UGC-NET. She has got two best paper presentation awards and earlier worked with British Library Ahmedabad.

**E-mail : shwetapandya@yahoo.com**



**Mr. Bhaumik Shroff** is Lecturer at AES Institute of Computer Studies, Navrangpura, Ahmedabad, India and holds MCA. His research interests encompass Wireless Networks, Global System for Mobile (GSM), Short Message Service (SMS), Data Warehousing, Voice over IP, Wireless Application Protocol (WAP) etc. He won the first prize in paper contest conducted by Computer Society of India in 2001.

**E-mail : shroffbhaumik@yahoo.com**