



IDEA GROUP PUBLISHING

701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-<http://www.idea-group.com>

ITB9392

Chapter XIV

Data Warehousing: The 3M Experience

Hugh J. Watson
University of Georgia, USA

Barbara H. Wixom
University of Virginia, USA

Dale L. Goodhue
University of Georgia, USA

ABSTRACT

Data warehouses are helping resolve a major problem that has plagued decision support applications over the years — a lack of good data. Top management at 3M realized that the company had to move from being product-centric to being customer savvy. In response, 3M built a terabyte data warehouse (global enterprise data warehouse) that provides thousands of 3M employees with real-time access to accurate, global, detailed information. The data warehouse underlies new Web-based customer services that are dynamically generated based on warehouse information. There are useful lessons that were learned at 3M during their years of developing the data warehouse.

INTRODUCTION

The information systems field has a long history of using computers to support decision making. The 1960s saw management reporting systems, the 1970s witnessed decision support systems (DSS) and the 1980s saw executive information systems (EIS). Throughout the 1990s and continuing on to today, there is great interest in online analytical processing (OLAP), data mining, customer relationship management (CRM) and other decision-support applications.

While there have been many successes with these applications, there have also been many failures. A frequent reason for failure has been the absence of a solid data infrastructure to support the applications. System developers simply could not access the data that was needed to support the applications. The data was often “locked up” in operational systems, either because of data ownership issues or the technical difficulties of accessing needed data. Even when the data could be accessed, it was typically “dirty.” Data was missing; dummy values were in fields (e.g., 999-99-9999 as a SSN); fields were used for multiple purposes; primary keys were reused; business rules were violated (e.g., a loan rate lower than the lowest rate) and so on. Then there was the problem of integrating the data from multiple source systems. Either there were nonunique identifiers (e.g., multiple account numbers) or the absence of an appropriate primary key. Developers knew how to build the applications, but, without a solid data infrastructure in place, they were doomed to failure.

A response to the data infrastructure problem emerged in the late 1980s (Inmon, 1992). A few leading-edge firms in the telecommunications, retail and financial services industries developed data warehouses — large repositories of data created to support decision making. These firms wanted to use their customer data as a basis for competing in the marketplace. By knowing their customer well, they hoped to maintain and enhance customer relationships, better meet customers’ needs and wants, and increase revenues and profits. Over the next decade, most large- and many medium-sized firms began data warehousing initiatives. At the turn of the century, data warehousing had become one of the top two or three strategic initiatives (along with e-commerce) in the information systems field (Eckerson, 1998).

Data warehouses are built because they promise to provide organizational benefits. They can generate cost saving through the consolidation of multiple, disparate decision-support platforms; provide time savings for information technology professionals and end users by making data easier to access and use; improve the quality of data that is available to decision makers; support the redesign of business processes and enable new business initiatives and strategies (Haley, Watson, & Goodhue, 1999). Data warehouses are challenging undertakings for both organizational and technical reasons (Crofts, 1998). On the organizational side, there must be executive sponsorship, management support and participation, effective handling of data ownership issues and user participation in determining data requirements. On the technical side, new hardware and software must be selected and implemented; data from source systems must be understood, extracted, transformed and loaded into the warehouse; appropriate data models must be developed and implemented and users must be trained and supported on data access tools. In the following section, the architecture for data warehousing is discussed.

In order to illustrate data warehousing in practice, a case study is presented. It describes 3M and the business and systems drivers for the global enterprise data warehouse (GEDW) that was developed. Next, the development of GEDW and its

13 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/data-warehousing-experience/27917

Related Content

Navigation Rules for Exploring Large Multidimensional Data Cubes

Navin Kumar, Aryya Gangopadhyay, George Karabatis, Sanjay Bapna and Zhiyuan Chen (2006). *International Journal of Data Warehousing and Mining* (pp. 27-48). www.irma-international.org/article/navigation-rules-exploring-large-multidimensional/1773

A Conformity Measure Using Background Knowledge for Association Rules: Application to Text Mining

Hacène Cherfi, Amedeo Napoli and Yannick Toussaint (2009). *Post-Mining of Association Rules: Techniques for Effective Knowledge Extraction* (pp. 100-115). www.irma-international.org/chapter/conformity-measure-using-background-knowledge/8439

A Lattice-Based Framework for Interactively and Incrementally Mining Web Traversal Patterns

Yue-Shi Lee (2009). *Handbook of Research on Text and Web Mining Technologies* (pp. 448-467). www.irma-international.org/chapter/lattice-based-framework-interactively-incrementally/21740

An Immune Systems Approach for Classifying Mobile Phone Usage

Hanny Yulius Limanto, Tay Joc Cing and Andrew Watkins (2007). *International Journal of Data Warehousing and Mining* (pp. 54-66). www.irma-international.org/article/immune-systems-approach-classifying-mobile/1784

Selecting Salient Features and Samples Simultaneously to Enhance Cross-Selling Model Performance

Dehong Qiu, Ye Wang and Qifeng Zhang (2010). *Strategic Advancements in Utilizing Data Mining and Warehousing Technologies: New Concepts and Developments* (pp. 329-337). www.irma-international.org/chapter/selecting-salient-features-samples-simultaneously/40415