Introduction to the Minitrack on Data Warehousing

Barbara Wixom
McIntire School of Commerce, The University of Virginia
Charlottesville, VA 22904
bwixom@mindspring.com

Paul Gray
School of Information Science, Claremont Graduate University
Claremont, CA 91711
paul.gray@cgu.edu

Hugh J. Watson
Terry College of Business, The University of Georgia
Athens, GA 30602-6256
hwatson@arches.uga.edu

Data warehouse repositories and their applications offer great potential for understanding and managing customer relationships, uncovering important patterns and trends in corporate data, and supporting comprehensive performance measurement systems, such as Balanced Scorecards. The papers in this year’s data warehousing minitrack investigate how data warehousing can be used to support these kinds of initiatives and how data warehousing can be made more effective technically.

In “Method for Demand-Driven Information Requirements Analysis in Data Warehousing Projects,” Robert Winter and Bernard Strauch present a 13-step methodology for analyzing the requirements of data warehousing projects. Expert interviews were conducted and used to develop the methodology, and parts of the methodology were applied successfully in data warehouse projects.

Steven Hanes investigates approaches and tools used to collect, manage, and deliver institutional metrics to the senior command of the United States Marine Corps. The paper explores current performance management strategies and tool architectures, with emphasis on the balanced scorecard approach and the commercial, off-the-shelf software that supports it. The initial evaluation framework is presented.

The third study in this minitrack focuses on customer relationship management and describes a customer-oriented architecture for data warehousing. Hans-Georg Kemper and Phil-Lip Lee describe two kinds of customer relationship management initiatives – analytical and operational – in “The Customer-Centric Data Warehouse – An Architectural Approach to Meet the Challenges of Customer Orientation.” They suggest how the incorporation of an operational data store into a data warehouse environment helps to support combined analytical and operational needs.

As applications for data warehousing continue to grow, the amount of data that organizations must manipulate and store increases. The next two papers address some of the challenges with large volumes of warehoused data. The first paper, “Ad-Hoc Association Rule Mining Within the Data Warehouse” by Svetlozar Nestorov and Nenad Jukić, offers a new data mining framework that is tightly integrated with data warehousing technology. The framework analyzes star schema organized data; leverages the processing power of the data warehouse; queries data across the data warehouse; and incorporates varying levels of data aggregation.

A second way to manage large volumes of data is through virtual data marts, an approach presented by Leslie Hodge and Charles Milligan in “Managing Virtual Data Mart(s) With Metapointer Tables.” Hodge and Milligan describe a data management method that enables users to access relevant subsets of large data populations that reside in different data warehouses.

It is challenging for academics to stay abreast of changes to teach and research effectively in the data warehousing field. Teradata, a division of NCR, working closely with academics, recently introduced its Teradata University Network (TUN) initiative. TUN is a website that academics can use to find materials for data warehousing, DSS/BI, or database classes; to access knowledge and training materials from subject matter experts; and to connect with colleagues in the academic and business communities. In the final presentation in this minitrack, Hugh Watson, Director of the TUN Advisory Board, will describe the resources available on TUN and how academics can benefit from its offerings.