Introduction to the Minitrack on Data Warehousing

Barbara Haley Wixom
McIntire School of Commerce
The University of Virginia
Charlottesville, VA 22903
bhaley@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

An estimated 95 percent of the Fortune 1000 companies either have a data warehouse in place or are planning to develop one [2]. The Palo Alto Management Group predicts that the data warehousing market will grow to a $113.5 billion market in 2002, including the sales of systems, software, services, and in-house expenditures [1]. These trends indicate that data warehousing continues to be an important topic in information systems as companies create data warehouses to transform business processes and create closer ties with their customers and other stakeholders. This minitrack was created to promote research in technical and managerial areas of data warehousing and to showcase efforts that can help companies improve the effectiveness of their warehouse initiatives. It intends to describe how data warehousing will continue to adjust to today's business environment in support of decision-making, customer relationship, and other applications.

Before a data warehouse can meet needs within the organization, data have to be extracted from data sources across the company and combined to create a “single version of the truth.” Unfortunately, this task can be quite arduous as project teams grapple with disparate data sources, large data volumes, and conflicting business rules. Jari Vanhanen, Kai Risku, and Pekka Kilponen present an open architecture for collecting and utilizing data from different sources in “Combining Data from Existing Company Data Sources: Architecture and Experiences.” They illustrate the effectiveness of their architecture with information about their experiences gained from an implementation within the industrial product development environment.

Steve McCown and Charles Milligan examine an innovative way to search through data once it is combined in “A Navigation System for Personalized Databases.” The authors recognize the importance of automating decision support systems to alleviate the burden on the decision maker, and they describe a method called “StarMap” by which decision makers can efficiently navigate and filter data and identify information content. Their vision is for decision makers to be able to augment subsets of data with metadata that allows efficient navigation and display; to use Web-based interaction tools for navigation support; and to share instances of personal “StarMaps.”

Another issue when creating a data warehouse involves the extent and way in which data are to be integrated. The amount of data centralization that occurs within the warehouse can impact the flexibility of the system to adapt to changing business needs after mergers, acquisitions, and other kinds of corporate restructuring occur. Ulrike Baumöl, Reinhard Jung, and Robert Winter conceptualize an interface that enables the integration of decentralized business units, while allowing for corporate restructuring. They describe a “management middleware” that is a read-write information junction that accepts input
from decentralized units and allows access to headquarters and other units.

Even if data is combined appropriately, the data warehouse will not create value unless users access the warehouse and use it to perform business tasks. One barrier to use is poor data quality, as described by Amit Rudra and Elimie Yeo in "Issues in User Perceptions of Data Quality and Satisfaction in Using Data Warehousing – An Australian Experience." Data quality can significantly affect organizational performance and frustrate warehouse users. Rudra and Yeo present findings from a survey conducted at a large organization in Australia that describe user perceptions of data quality and their satisfaction rating in using the data warehouse.

Outcomes from the use of the data warehouse are explored by Kathryn Brohman, Michael Parent, Michael Pierce, and Michael Wade in “The Business Intelligence Value Chain: The Process of Data Warehouse in Large Organizations: An Exploratory Study.” Fifteen semi-structured interviews in 6 organizations were conducted to explore how organizations use data warehouse applications to support decision-making and generate value for the business. The results of these interviews were used to formulate a model called the “Business Intelligence Value Chain.” The authors describe the details of how the model was created and the stages and relationships within it. They then present a case study of a large Canadian retail store to illustrate the way in which the model can be applied.

Customer relationship management (CRM) is a specific application of data warehouses that has received much attention recently. CRM software creates a comprehensive view of the relationship between a business and its customer CRM, and companies are beginning to incorporate CRM as an important piece of their business strategies. Ron Swift, Assistant Vice President of NCR Corporation will provide insight into some of these emerging trends of CRM in his presentation, "Customer Relationship Management in Data Warehousing." His experience with CRM applications provides him with unique exposure to cutting edge projects in practice.