Managing Data in the Twenty-First Century

Mary Ann Robbert  
*Bentley College*, mrobbert@bentley.edu

Paul Leidig  
*Grand Valley State University*, leidigp@gvsu.edu

Matthew Merzbacher  
*Mills College*, matthew@mills.edu

Dian Rae Lopez  
*University of Minnesota*, lopezdr@mrs.umn.edu

Jane Fedorowicz  
*Bentley College*, jfedorowicz@bentley.edu

Follow this and additional works at: [http://aisel.aisnet.org/amcis2000](http://aisel.aisnet.org/amcis2000)
Managing Data in the Twenty-First Century

Mary Ann Robbert (Moderator), CIS Department, Bentley College, Waltham, MA 02154, mrobert@bentley.edu
Paul Leidig, Computer Science and Information Systems Department, Grand Valley State University, Allendale, MI, leidigp@gvsu.edu
Matthew Merzbacher, Computer Science Department, Mills College, Oakland, CA, matthew@mills.edu
Dian Rae Lopez, Department of Computer Science, University of Minnesota, Morris, MN, lopezdr@mrs.umn.edu
Jane Fedorowicz, Accounting Department, Bentley College, Waltham, MA, jfedorowicz@bentley.edu

In the fall of 1999, I did a survey of faculty members to determine the current status and content of database courses. The results from forty-five returns showed that twenty schools did not require a database course on the undergraduate or graduate level and three schools did not even offer a database course. The content and location of the database courses that were offered varied. The database course resided in the computer science or information systems department. The main topics covered were SQL, database design and modeling, and all aspects of the relational model. Not one survey response included data warehousing or data mining as a standard topic (some of the new database texts contain a chapter on data warehousing). Data warehousing was included in one course, if time permitted. These results led me to focus on two questions: Should data warehousing, data mining and other current data management topics be included in the curriculum? And where should these data management topics be included in the curriculum?

Data warehousing can be included in a database, statistics or accounting information systems course as well as an independent course. I offer data warehousing as a second semester graduate database elective. Next semester I am forming teams and holding some joint classes with the statistics class to enrich the data mining components. Each of the panelists has their opinion on how and where to cover the data warehousing and data mining concepts.

In this information age, data management must be understood. This understanding is achieved through education. This panel will examine the role of data management, including data warehousing and data mining and its place in the curriculum. The panelists represent different backgrounds and schools. Our objective is to determine if and where current data management techniques should taught in the curriculum.

Paul Leidig
Information is critical in today's competitive business environment. However, even as organizations are "drowning in data" it is becoming increasingly difficult to manage necessary information. At the same time, institutions of higher education are finding it more difficult to squeeze in a widening array of 'necessary' course content. Thus many programs in computer science or information systems must choose what content is required of their students, and more importantly, what realistically can be offered in a limited curriculum. Many programs do not require a course in database design and management, much less one specifically devoted to data warehouse and data mining. Often coverage of data warehousing and data mining within a database course(s) is superficial at best.

Our department requires all undergraduates to take a single database course, and strongly recommends one for our graduate students. At the same time, we have had an increasing amount of interest in advanced database course content, specifically data warehousing and data mining techniques. The approach we propose to satisfy this demand is to expand the existing courses, splitting the content and adding a new elective advanced database course that includes data warehouse and data mining.

Dian Rae Lopez
Today we are overwhelmed with information - from the World Wide Web to huge databases - in our libraries and in our personal computers. Like a little baby with new information pouring in from all sides, we must try to make sense of it all. We desperately need metadata - data to help us organize and understand such large amounts of data. Only then can we begin to analyze key sections of data and make informed decisions and/or future predictions.

Our future business management, information systems, and computer science graduates must be trained to manage mammoth amounts of data from many sources. They must understand the design of large data warehouses. They must design the computer software
(load managers, warehouse managers and query managers) needed to transform and aggregate data into a form suitable for extraction and analysis. An understanding of data warehousing is essential to the creation of flexible, dynamic data management systems capable of guaranteeing consistent data, fast query response time, and the extraction of information needed for decision making. We must prepare our students for this future. We must prepare our students for this future by teaching a data warehousing course or, if that is not possible, by adding a data warehousing component to an existing course.

Matthew Merzbacher
Small colleges face unique challenges in integrating new material, such as data warehousing and data mining, into the curriculum. Because so few courses are offered each year, the cost of adding an entirely new course represents a relatively large change in the overall curriculum. Thus, most small colleges are quite conservative in adding new material; they wait until a field is well established before considering it for inclusion.

Instead, new material is usually grafted onto an existing course for some time before it is spun off. In the case of data warehousing and data mining, the logical courses for adaptation are Databases and Artificial Intelligence, both upper-division electives. Unfortunately, this puts exciting material near the top of the curriculum where it is inaccessible to students across the college. As a remedy, data warehousing can be offered to a broader range of students in an introductory "applied concepts course" (the algorithmic details of data mining still require prerequisites and would remain in the database course). At Mills, we currently offer a similar course for non-majors in three-dimensional computer graphics. I envision a similar approach for data warehousing that would attract students from all disciplines.

Jane Fedorowicz
The Accounting Information Systems (AIS) major joins together the skill sets of two areas experiencing rapid growth and change. E-business, enterprise computing, paperless work processes, and many other technology-intensive innovations have created new challenges and opportunities for accountants who also have expertise in information systems. The AIS major comprises coursework in Computer Information Systems, Accounting, and specialized AIS courses to produce an information technology specialist with strong business knowledge.

Information quality has long been a focal point of accountants, both as providers of financial information and as auditors. The introduction of the data warehouse has expanded the accountant's responsibility for assuring information quality. Not only does the data warehouse contain the same kinds of data the accountant has audited in the past, but the data domain has expanded to cover a longer period of time, data that must be integrated from many disparate sources, and a large array of non-financial information must be monitored as well. The user base for a data warehouse is also different from that of the traditional financial information system.

The Advanced Accounting Information Systems course introduces the AIS student to the data warehouse concept, and focuses on developing metrics and measures with which to evaluate, cleanse and integrate data from disparate sources. Students are exposed to cleansing software (from Trillium). They work through a group exercise defining metrics for a proposed data warehousing application. They then produce a group project in which they develop metrics for assessing a new enterprise system for a project-based organization, and compare it with another system with which they are already familiar. These experiences give them a strong preparation for working in some of the new assurance service areas of professional service firms, or internally as an IT auditor of a company's data warehouse application.

References
Robertson, P. "Integrating Legacy Systems with Modern Corporate Application.," Communications of the ACM, (40,5), May 1997.