

Key Points

A number of important concepts and issues were covered during the seminar.

- A vast majority, between 50 and 80 percent, of data warehouse projects fail because participants don't fully understand the value of integrated data and aren't sure how to use the combined information.
- Data repositories can integrate disparate sources of data to evaluate the impact of a set of programs and services on a specific population. Such repositories must resolve issues about: the intersection of programs and services, business drivers, data, technology, and the policy environment. They are built on a foundation of leadership, sponsorship, partnerships, and common vision.
- Data warehouse projects can be realized through the process of building a bridge from the original idea to the actual system. This is done by: defining a service objective that meets real business needs, developing a business case that shows how the project will work and what benefits it will provide, establishing partnerships with all key stakeholders, encouraging open communication among all participants, building and testing a data warehouse prototype of relevant data, resolving problems, and being persistent.
- Data by themselves are a worthless luxury unless they are used. Thus a prototype system, which converts data into actionable information, must be developed so that users can appreciate the value of the data.
- To define relevant data, focus on what an organization must do and know rather than what would be nice to do and know. The goal is to provide information that will be actively used for identifying and solving problems.
- The development of a data warehouse should employ an evolutionary approach in which each phase builds on the preceding one so that users derive value by using the data from each phase. Start with the core data elements and build a prototype system with real data, not dummy data, so that users can get a real feel for the information when they "test-drive" the prototype. If the prototype does not demonstrate value to the users, the project should be buried before more time and money are wasted on building a data warehouse that will not be used.
- Data marts have three components: data feeding, storage, and use. First, operations and external data are extracted, transformed, and fed into the warehouse. The data itself, and the meta data that describes the data, are then stored in the system. Finally, online analytical processing tools and applications are employed to use the data.
- Some of the challenges encountered when developing data marts include: extraction of data from multiple sources, data quality, ease-of-use, user support, data gaps, performance, implementation time frame, bad or missing data, and changing requirements.
- When doing data warehouse projects, 80 percent of the effort is in the data.
- Technology is important, but management and policy issues must also be addressed early in the development process. Organizations need to know who will be involved in the project, what stakeholder groups will be affected by it, what the system will do, where the data will come from, and how it will be obtained. They also need to create business rules, standard data definitions, confidentiality and privacy policies, and other guidelines for data use.