INTERACTIVE SESSION: MANAGEMENT

AMERICAN WATER KEEPS DATA FLOWING

American Water, founded in 1886, is the largest public water utility in the United States. Headquartered in Voorhees, N.J., the company employs more than 7,000 dedicated professionals who provide drinking water, wastewater and other related services to approximately 16 million people in 35 states, as well as Ontario and Manitoba, Canada. Most of American Water's services support locally-managed utility subsidiaries that are regulated by the U.S. state in which each operates as well as the federal government. American Water also owns subsidiaries that manage municipal drinking water and wastewater systems under contract and others that supply businesses and residential communities with water management products and services.

Until recently, American water's systems and business processes were very localized, and many of these processes were manual. Over time, this information environment became increasingly difficult to manage. Many systems were not integrated, so that running any type of report that had to provide information about more than one region was a heavily manual process. Data had to be extracted from the systems supporting each region and then combined manually to create the desired output. When the company was preparing to hold an initial public offering of its stock in 2006, its software systems could not handle the required regulatory controls, so roughly 80 percent of this work had to be performed manually. It was close to a nightmare.

Management wanted to change the company from a decentralized group of independent regional businesses into a more centralized organization with standard company-wide business processes and enterprise-wide reporting. The first step toward achieving this goal was to implement an enterprise resource planning (ERP) system designed to replace disparate systems with a single integrated software platform. The company selected SAP as its ERP system vendor.

An important step of this project was to migrate the data from American Water's old systems to the new platform. The company's data resided in many different systems in various formats. Each regional business maintained some of its own data in its own systems, and a portion of these data were redundant and inconsistent. For example, there were duplicate pieces of materials master data because a material might be called one thing in the company's Missouri operation and another in its New Jersey business. These names had to be standardized so that the same name for a piece of data was used by every business unit. American Water's business users had to buy into this new company-wide view of data.

Data migration entails much more than just transferring data between old and new systems. Business users need to know that data are not just a responsibility of the information systems department: the business "owns" the data. It is business needs that determine the rules and standards for managing the data. Therefore, it is up to business users to inventory and review all the pieces of data in their systems to determine precisely which pieces of data from the old system will be used in the new system and which data do not need to be brought over. The data also need to be reviewed to make sure they are accurate and consistent and that redundant data are eliminated.

Most likely some type of data cleansing will be required. For example, American Water had data on more than 70,000 vendors in its vendor master data file. Andrew Clarkson, American Water's Business Intelligence Lead, asked business users to define an active vendor and to use that definition to identify which data to migrate. He also worked with various functional groups to standardize how to present address data.

One of the objectives of American Water's data management work was to support an enterprise-wide business intelligence program based on a single view of the business. An analytical system and data warehouse would be able to combine data from the SAP ERP System with data from other sources, including new customer information and enterprise asset management systems. That meant that American Water's business users had to do a lot of thinking about the kinds of reports they wanted. The company had originally planned to have the system provide 200 reports, but later reduced that number by half. Business users were trained to generate these reports and customize them. Most financial users initially tried to create their reports using Microsoft Excel spreadsheet software. Over time, however, they learned to do the same thing using SAP Business Objects Web Intelligence tools that came with the system. SAP Business Objects Web

Intelligence is a set of tools that enables business users to view, sort, and analyze business intelligence data. It includes tools for generating queries, reports and interactive dashboards.

At present, American Water is focusing on promoting the idea that data must be "clean" to be effective and has poured an incredible amount of effort into its data cleansing work—identifying incomplete, incorrect, inaccurate, and irrelevant pieces of data and then replacing, modifying, or deleting the "dirty" data. According to Clarkson, just as water treatment plants have measurements and meters to check water quality as its being treated, data management needs to ensure the quality of data at every step to make sure the final product will be genuinely useful for the company.

Sources: "SAP to Deliver Software Solution to American Water," www.sap.com, accessed January 31, 2014; David Hannon, "Clean Smooth-Flowing Data at American Water," SAP Insider Profiles, January–March 2013 and www.amwater.com, accessed February 2, 2014.

CASE STUDY QUESTIONS

1. Discuss the role of information policy, data administration, and efforts to ensure data quality in improving data management at American Water.
2. Describe roles played by information systems specialists and end users in American Water's systems transformation project.
3. Why was the participation of business users so important? If they didn't play this role, what would have happened?
4. How did implementing a data warehouse help American Water move toward a more centralized organization?
5. Give some examples of problems that would have occurred at American Water if its data were not "clean"?
6. How did American Water's data warehouse improve operations and management decision making?

Review Summary

1. *What are the problems of managing data resources in a traditional file environment?*

   Traditional file management techniques make it difficult for organizations to keep track of all of the pieces of data they use in a systematic way and to organize these data so that they can be easily accessed. Different functional areas and groups were allowed to develop their own files independently. Over time, this traditional file management environment creates problems such as data redundancy and inconsistency, program-data dependence, inflexibility, poor security, and lack of data sharing and availability. A database management system (DBMS) solves these problems with software that permits centralization of data and data management so that businesses have a single consistent source for all their data needs. Using a DBMS minimizes redundant and inconsistent files.

2. *What are the major capabilities of DBMS and why is a relational DBMS so powerful?*

   The principal capabilities of a DBMS includes a data definition capability, a data dictionary capability, and a data manipulation language. The data definition capability specifies the structure and content of the database. The data dictionary is an automated or manual file that stores information about the data in the database, including names, definitions, formats, and descriptions of data elements. The data manipulation language, such as SQL, is a specialized language for accessing and manipulating the data in the database.

   The relational database has been the primary method for organizing and maintaining data in information systems because it is so flexible and accessible. It organizes data in two-dimensional tables called relations with rows and columns. Each table contains data about an entity and its attributes. Each row represents a record and each column represents an attribute or field. Each table also contains a key field to uniquely identify each record for retrieval or manipulation. Relational database tables can be combined easily to deliver data required by users, provided that any two tables share a common