

REAL WORLD CASE 4

Emerson and Sanofi: Data Stewards Seek Data Conformity

A customer is a customer is a customer, right? Actually, it's not that simple. Just ask Emerson Process Management, an Emerson Electric Company unit in Austin that supplies process automation products. In 2000, the company attempted to build a data warehouse to store customer information from over 85 countries. The effort failed in large part because the structure of the warehouse couldn't accommodate the many variations on customers' names.

That's when the company hired Nancy Rybeck as data administrator. Rybeck is now leading a renewed data warehouse project that ensures not only the standardization of customer names but also the quality and accuracy of customer data, including postal addresses, shipping addresses, and province codes.

To accomplish this, Emerson has done something unusual: It has started to build a department with 6 to 10 full-time "data stewards" dedicated to establishing and maintaining the quality of data entered into the operational systems that feed the data warehouse.

Creating a data quality team requires gathering people with an unusual mix of business, technology, and diplomatic skills. At Emerson, data stewards review data and correct errors before it's put into the operational systems. They also research customer relationships, locations, and corporate hierarchies; train overseas workers to fix data in their native languages; and serve as the main contact with the data administrator and database architect for new requirements and bug fixes.

As the leader of the group, Rybeck plays a role that includes establishing and communicating data standards, ensuring data integrity is maintained during database conversions, and doing the logical design for the data warehouse tables.

The stewards have their work cut out for them. Bringing together customer records from the 75 business units yielded a 75 percent duplication rate, misspellings, and fields with incorrect or missing data.

"Most of the divisions would have sworn they had great processes and standards in place," Rybeck says. "But when you show them they entered the customer name 17 different ways, or someone had entered, 'Loading dock open 8:00-4:00' into the address field, they realize it's not as clean as they thought."

Although the data steward may report to IT—as is the case at Emerson and at pharmaceuticals company Sanofi-Synthelabo Inc.—it's not a job for someone steeped in technical knowledge. Yet it's not right for a businessperson who's a technophobe, either.

Seth Cohen is the first data quality control supervisor at Sanofi in New York. He was hired in 2003 to help design automated processes to ensure the data quality of the customer knowledge base that Sanofi was beginning to build.

Data stewards at Sanofi need to have business knowledge because they need to make frequent judgment calls, Cohen says. Indeed, judgment is a big part of the data steward's job—including the ability to determine where you don't need 100 percent perfection.

Cohen says that task is one of the biggest challenges of the job. "One-hundred percent accuracy is just not achievable," he says. "Some things you're just going to have to let go or you'd have a data warehouse with only 15 to 20 records."

Data stewards also need to be politically astute, diplomatic, and good at conflict resolution—in part because the environment isn't always friendly. When Cohen joined Sanofi, some questioned why he was there. In particular, IT didn't see why he was "causing them so many headaches and adding several extra steps to the process," he says.

There are many political traps, as well. Take the issue of defining "customer address." If data comes from a variety of sources, you're likely to get different types of coding schemes, some of which overlap.

People may also argue about how data should be produced, he says. Should field representatives enter it from their laptops? Or should it first be independently checked for quality? Should it be uploaded hourly or weekly?

Most of all, data stewards need to understand that data quality is a journey, not a destination. "It's not a one-shot deal—it's ongoing," Rybeck of Emerson says. "You can't quit after the first task."

Case Study Questions

1. Why is the role of a data steward considered to be innovative? Explain.
2. What are the business benefits associated with the data steward program at Emerson?
3. How does effective data resource management contribute to the strategic goals of an organization? Provide examples from Emerson and others.

Source: Adapted from Mary Brandel, "Data Stewards Seek Data Conformity," *Computerworld*, March 15, 2004. Copyright © 2004 by Computerworld, Inc., Framingham, MA 01701. All rights reserved.