DM Review
intelligent solutions

jonathan geiger

You Can be Sure if It's . . . Been Certified, Part 2

Last month's column introduced the concept of data certification and three levels of certification: the bronze level, which certifies correct processing by the data warehousing environment; silver level, which extends the certification to include all systems; and gold level, which extends the certification to include the business processes. Executives and business users can fully rely on gold-level data. This month we will explore the data quality certification oversight committee.

The major responsibilities of the data quality certification oversight committee are to determine the data that requires certification, establish and oversee certification processes, ensure that appropriate actions are taken to address quality issues and promote the data quality certification program. The oversight body needs to include people with the appropriate skills and the recognized authority to carry out these responsibilities. Members should include the chief information officer (who is ultimately responsible for managing the corporate asset of data), a representative from the legal or auditing department (who understands the regulatory, security and privacy requirements) and representatives from the major business groups.

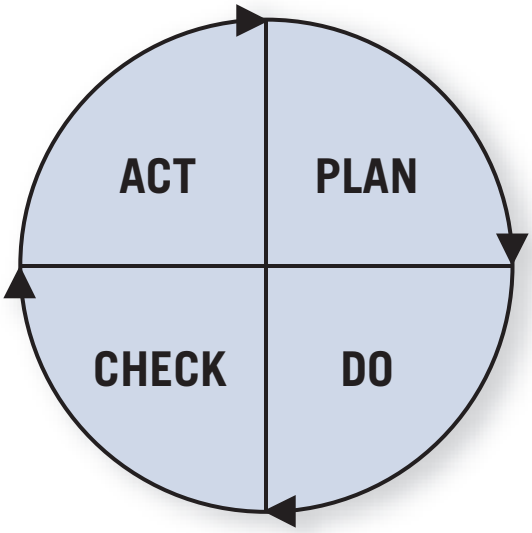


Figure 1: Continuous Improvement Cycle

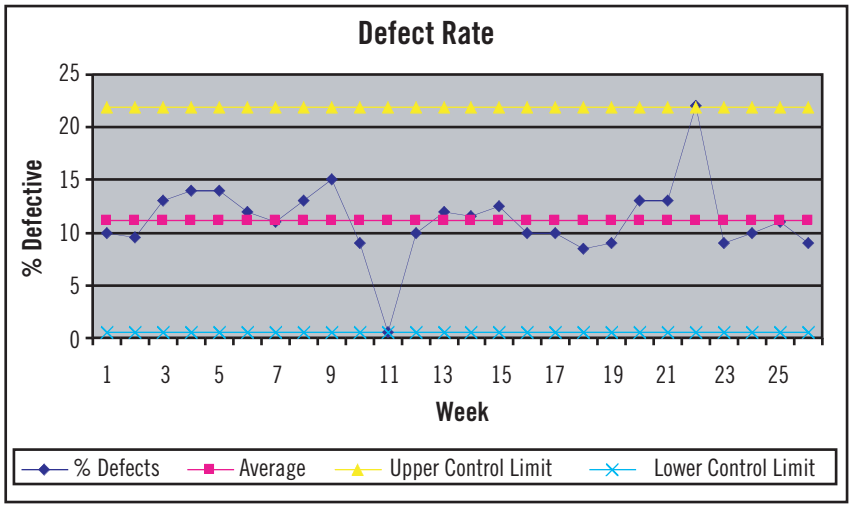
Each member of the certification body needs to understand the business priorities for data certification, the definition of quality, the continuous improvement cycle, the requirements of repeatable and measurable processes for achieving quality, the approaches for analyzing quality issues and the appropriate actions to be taken.

Business Priorities

The data certification process requires resources and a financial commitment, and not all of the data can be certified at once. The data quality certification oversight committee must determine which data is most in need of being certified. An objective assessment of the company's data quality can provide valuable input to the committee. The committee should apply the results of the assessment with specific requests, an understanding of the corporate priorities and regulatory requirements, and estimates for the certification process for each of the most significant data groups in determining the priorities.

Definition of Quality

Quality entails conformance to valid requirements. This definition requires two things: designation of the person(s) who can dictate the valid requirements and an



NOTE: Weeks 11 and 22 are "out of control," indicating that something unusual occurred and needs to be investigated.

Figure 2: Sample Control Chart

explicit specification of the requirements. These requirements should be included in the business meta data. If the data conforms to the requirement, it is quality data; if it does not, then it is not quality data. If a data stewardship committee exists, that committee generally has the responsibility to define the valid requirements. If no such committee exists, then as each set of data elements is tackled, the people to perform the role need to be designated.

Continuous Improvement Cycle

The continuous improvement process is based on the work of Dr. Walter A. Shewhart, who developed the continuous improvement cycle. The cycle recognizes the importance of planning a task, performing it, measuring the performance, taking appropriate corrective actions based on the measurements and then repeating the process. This continuous improvement cycle is shown in Figure 1.

The four components of the cycle must be understood by the data quality certification oversight committee. The cycle consists of *plan, do, check* and *act*.

Plan: Repeatable and Measurable Process. Companies that deploy effective total quality improvement programs establish and document repeatable processes for accomplishing work tasks and include appropriate metrics within these processes. This concept applies to the capture, management, maintenance, dissemination and disposal of data. Programming and data warehousing teams often receive their instructions for the repeatable processes in the form of programming or transformation specifications. The audit and control specifications for verifying successful compliance with the processes are sometimes missing. These must be included for data to receive the bronze or silver certification level. If the data is to receive the gold level, then the business processes must be defined along with requisite metrics.

Do: Process Execution. On a regular basis, the data is captured, processed and disseminated in accordance with the defined process. For a data warehouse, for example, this involves the regular execution of the data acquisition and the data delivery processes.

Check: Measure Results. The data quality certification oversight committee should ensure that the measurements are taken regularly and that these are tracked over time using appropriate techniques. Some measurements may be binary (e.g., Did a job complete or not?), while others may be quantitative (e.g., What percent of the records contain complete information?). The individual measurements are captured in the meta data. There are many techniques for tracking results over time, with control charts being a popular technique within Six Sigma programs (see Figure 2).

The results of the measurements should be published so that the business users of the data are aware of any quality issues that may exist.

Act: Take Appropriate Actions Based on the Measurement Results. The measurements provide information on whether or not the quality specifications are being met. If they are being met, no corrective actions are needed. If the measurements are not being met, the root cause for the nonconformance needs to be analyzed and appropriate corrective measurements pursued.

Next month, we will conclude this series with a description of the certification process.

Jonathan G. Geiger, executive vice president at Intelligent Solutions, Inc., has been involved in many Corporate Information Factory and customer relationship management projects. He has authored or coauthored numerous articles and three books, presents frequently at national and international conferences, and teaches several public seminars. Geiger can be reached at jgeiger@intelsols.com.

©October 2005 SourceMedia, Inc. and *DM Review*. All rights reserved. SourceMedia, One State Street Plaza, New York, N.Y. 10004 (800) 367-3989