

[The MIT Information Quality
Industry Symposium](#)

David Loshin

Knowledge Integrity



Title:

Data Profiling for Data Quality Assessment: Rapidly Building the Business Case for Data Quality Improvement

Abstract:

Poor data quality hampers business success by casting doubt on the data feeding both operational and analytical activities. The absence of metrics linking business impacts to data issues impedes the development of the appropriate business case for introducing data quality improvement. Because data issues contribute to operational inefficiencies, missed opportunities, and exposure to risk, there is a need for techniques to quickly identify high priority data issues and objectively evaluate the relationship between those issues and business productivity.

This presentation describes a process through which an organization can, using proper planning, tools, and expertise, identify high visibility data issues and characterize related business impacts. Simultaneously, opportunities for improvement can be identified, providing an ability to objectively review organizational data, determine whether the levels of data quality are sufficient to meet business expectations, and if not, evaluate the value proposition and feasibility of specific data quality improvements.

This rapid data quality assessment, performed in as few as three weeks, combines top-down and bottom-up approaches to evaluating specific data sets to both engage business data consumers and immediately identify opportunities for improvement. We also describe case studies from the insurance and financial services sectors.

Key topics include:

- Business impacts of poor data quality
- Using data profiling tools to find potential anomalies
- Synthesizing data profiling results in the context of business impacts
- Characterizing and prioritizing key data quality flaws
- Identifying immediate remediation tasks