
TOTAL DATA QUALITY™

A Business Solution for Today and Tomorrow



TRILLIUM SOFTWARE
a division of Harte-Hanks

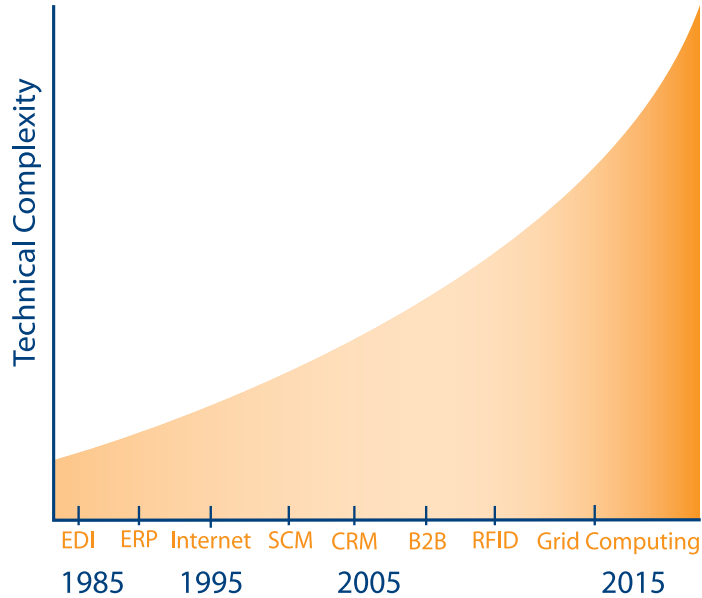
The Need for a Unified Solution

“The main job of technology is to get out of the way,” says David Gelernter, Yale computer scientist and prominent proponent of software that simplifies work by handling complex problems in a unified way.¹ Well-designed software should cut the time it takes to manage the complexities of information so people can more easily get their work done. But all too often technologies—diverse technologies, platforms, systems, applications—make work complex, creating barriers to reliable, unified information. As a result, businesses find that the latest software, if it works at all, fails to deliver the expected results. Where there should be simplicity and certainty, there is complexity and doubt.

The search for a unified solution to these complex technology problems is hardly new. During the nineties companies spent \$25B per year just on basic hardware, software, and services for ERP systems, which promised a single monolithic solution to the tangle of back-office systems and data. But ERPs hardly simplified work, users mistrusted the systems, and over 60% were eventually considered failures.² Meanwhile the Internet was fundamentally re-defining the information landscape into a decentralized and open system. Single monolithic systems, businesses have learned, can never provide a unified solution. Underlying the complexities of the systems, and forming the basis for a different kind of unity, is the data itself.

Total Data Quality resolves the technological complexity that enterprises face today in a way that unifies data regardless of platforms, systems, and business applications. Organizations need access to data residing in multiple business processes and systems—for summary reports, synchronized information about supply chains and customers, analysis of market trends and demands, and a host of other broad business initiatives. And all these activities require up-to-date, accurate consolidated data. Total Data Quality enables organizations to deliver such high-quality data, data that is fit for various business purposes, both simultaneously and over time.

Data accessible across systems has become the key to unifying and synchronizing enterprise



As technologies have become more complex, companies found that no single monolithic system could give them a unified solution across diverse platforms, systems, and applications. Today companies are focusing on the data and using Total Data Quality to create a unified enterprise solution.

More errors, less trust

Data quality becomes critical when organizations combine and repurpose data from multiple systems.

As organizations start to combine and repurpose information, basic data quality issues—the consistency and accuracy of it, how current and complete it is—become critical business problems. Unfortunately users are only seeing more errors that are having a bigger and more far-reaching impact on operations:

- 60% of all invoices contain errors³
- 30% of all brokerage transactions fail because of bad data
- 30% of manufacturing orders are incorrectly executed or delayed because of inaccurate information
- Forecasting error rates in retail industries average 35%
- An estimated 30% of customer data is inaccurate

All of this makes it harder to trust the data or any processes, reports, or analysis based on that data. It is hardly surprising that 75% of the world’s leading companies have reported significant costs resulting from poor data quality and only one-third express confidence in their data.⁴

Even as the pressure mounts from corporate executives and their boards to ensure that core information is reliable and complies with increasingly tough regulations, the data itself keeps changing and the data requirements of business users change as well:

Core business information—Product data, inventory data, order data, customer data, business partner data changes constantly. In some industries 35% of SKUs change annually, as businesses add new products, retire others, and expand their offerings, making the task of synchronizing product and inventory data a challenge in itself.⁵

Changing data needs for business users—Basic business information must change often to meet new business needs and competitive challenges. Repurposing data—modifying the information included in records, assimilating additional information from external sources, and/or restructuring data records—to meet new business imperatives has become a perennial activity.

The global imperative—Organizations are not naturally equipped to handle the endless string of problems that global data presents. Local linguistic, cultural, postal, and electronic formatting are beyond the capabilities of even the most multinational of companies. Without the capability to automatically handle a new customer from Sweden, a new supplier from China, or a new acquisition from Brazil, companies will falter.

Seamlessly handling product information from China or Chile has become a global imperative.

The Total Data Quality Solution

To meet these complex, changing, global data requirements, business users need a unified solution like Total Data Quality. With Total Data Quality, organizations obtain reliable, accurate data to use in various circumstances simultaneously today and to re-use tomorrow for different purposes. Total Data Quality enables companies to deliver information that is fit for multiple business purposes, however, wherever, and whenever an organization needs that information.

Total Data Quality provides a comprehensive strategy as well as the tactics for solving data quality problems. And it does it in a way that builds a foundation for meeting tomorrow's information challenges with the same confidence, efficiency, and expertise that was available to solve today's problems. Core business data is multifaceted and dynamic. It is not easily or naturally

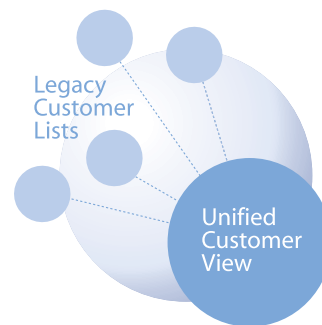
segmented by business unit or system. And it is most profitably leveraged when it is accessible to multiple users for diverse purposes. For these reasons, Total Data Quality is a solution that must evolve over time, span technological generations, and cross all geographic borders.

The Total Data Quality Process

It can start with something as simple as a list of customers. A company usually wants to improve its data quality to meet a specific and pressing need. But while the need may seem simple enough—a complete and accurate view of all customers throughout the enterprise—obtaining that view turns out to be complex. Inevitably, the first data quality initiative leads to a better understanding of the business information available and the possibilities for connecting it to other information. Organizations begin to see new opportunities and cost savings based on high-quality data. And they continuously expand the number of business users and processes that can tap into that information at each step along the way.

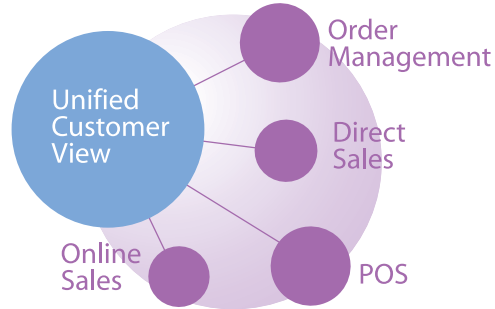
Here is one typical scenario: In order to market and cross-sell to its customer base, for example, a company may collect customer data from various business units and various transactional as well as database systems, including direct sales, online sales, customer service, accounting, and order management.

Total Data Quality enables companies to deliver information that is fit for multiple business purposes, however, wherever, and whenever the information is needed.



But each of these business functions has a different set of information about customers since each function has a different relationship with customers. Some records have information others lack; some have information that is more up-to-date than others.

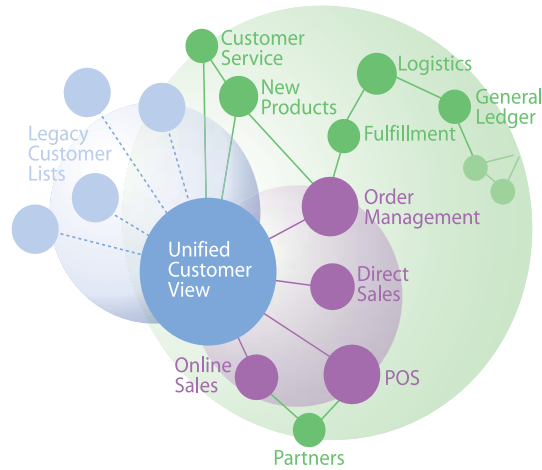
Building an accurate, complete, and standardized view of customers is no small task. It is not uncommon to have hundreds of diverse records that need to be reconciled and consolidated. Once the data quality process is complete, however, the company has a much clearer view of just who their customers are, what they've bought before, what they are likely to need next, and how to contact them.



Information built for a marketing campaign may be re-used by product development for tracking buying patterns and by manufacturing for forecasting demand.

As the process continues to unfold, sales receives leads from the new marketing campaign. At this point, it makes sense for direct sales to adopt the same data quality processes within their systems so that they can easily share information with marketing and collaborate. Sales people contribute additional information. Over time, information from sales channels can also be added, including new online sales, retail points of sales, and business partners. These expansions mean connecting to other enterprise business processes and in real time over the Internet. A Total Data Quality solution is one that provides a foundation for such flexible, easy expansion and the various connections it requires.

The company now has extensive and up-to-date customer information, a valuable resource for many business units and functions to use for various purposes: New product development can repurpose this information for querying and trend analysis. It may also want track warranty problems by adding customer service information to its customer records. Manufacturing, on the other hand, may want to take the same information and aggregate it by product to understand demand fluctuations over the calendar year.



Total Data Quality in action: 1 + 1 = 3

Total Data Quality, however, is more than the sum of its information parts. As an enterprise-wide data quality initiative grows, more and more accurate, consistent, and actionable corporate information becomes available. Businesses can make it possible for individuals, groups, and systems to make better use of the data they have, developing a broader context and deeper understanding of the true nature of their business, their customers, and the road ahead.

Total Data Quality: The bottom line

One heavy manufacturer, which had grown by acquisitions throughout the world, first wanted to standardize its product and inventory codes for its European operations. It began its data quality program with a single division but soon extended the initiative, adding records country by country in five different languages. The resulting data quality program cleansed, standardized, and consolidated records on a regular basis from 70 databases, contained 2.8 million parts, and 900 million records. As a result, reports that used to take six months to compile are produced weekly from legacy databases within a matter of hours at the push of a button. Some 300 business decision makers gained access to up-to-date accurate information. They then expanded the program to include core business data from China, Australia, and North America.

Some 300 decision makers now easily access up-to-date consolidated reports that used to take six months to compile.

As the company realized how valuable consolidated product data was, it decided to expand the data quality process to include name and address data of its suppliers, and then its customers worldwide. Business units were able to realize significant operational efficiencies and cost savings as the use of this information expanded:

- Procurement reduced supplier contracts by 5% based on a unified view of 200,000 suppliers worldwide.
- Inventory synchronization cut procurement costs by 30% and increased on-time delivery.
- Invoicing streamlined its processing by eliminating marginal transactions and consolidating records.
- Manufacturing was able to automate business-to-business processes by cost-effectively appending multiple standard parts numbers and commodity codes.

Communicating with and about customers

A telecommunications giant, for example, began with its central marketing database, which held millions of customer records. The company soon expanded the data quality process to call centers and then to repair orders and resolution. Later it synchronized inventory data and then its customer data to achieve:

Synchronized data in supply chains can reduce procurement costs by 30% and increase on-time delivery.

- Call centers are more efficient: workers can now quickly access complete customer files and confirm information in real time.
- Repair services have access to unified customer records, including all repair requests.
- External data enrichment has sharpened target marketing and increased sales.
- Eliminating manual data analysis and data-entry checks has cut costs.
- Better logistics has improved technical service response on the road.
- The company has established ongoing enterprise-wide correcting and monitoring of existing customer records and of new data as it comes in over the Internet.

From a single database, the telecommunications company has now expanded data quality to hundreds of business processes across local and international services, including broadband and Internet.

What a Total Data Quality Solution Offers

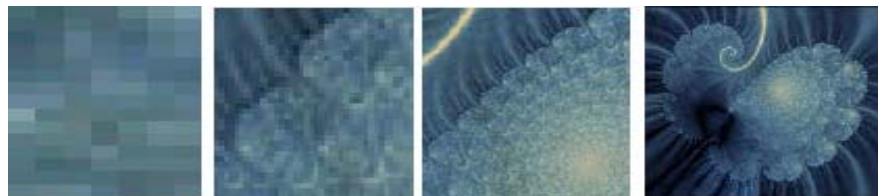
A Total Data Quality solution offers a comprehensive strategy and a full range of tactics for taking raw data from all kinds of systems, creating high-quality meaningful business information, and maintaining that information as corporate assets for distribution across the enterprise. It's all about building connections: connections between data sets to expand context and meaning, connections between systems to consolidate and update information stores, connections across operational units to share information, synchronize business processes, and find new opportunities.

Total Data Quality is all about discovering and building connections: connections between data sets and between systems, expanding the context and enriching the meaning.

All types of data

The best data quality technologies offer a comprehensive way to process and maintain all types of core business data—product data, inventory data, sales data, manufacturing data, and contact data for customers, suppliers, agents, distributors, and other business partners. Most companies decide on a data quality solution when they are trying to solve one problem, such as consolidating customer lists or integrating product information. But business data is not naturally separated into such categories and the uses of data inevitably change and expand.

Businesses today frequently look to data quality software when they are trying to repurpose data from systems that organized information around products—the sales, manufacture, and servicing of products—to a more customer-oriented view. But sophisticated business intelligence (BI) tools—for reporting, analysis, predictive modeling, and monitoring—may require the information be organized differently and include some information available from another department. With a comprehensive data quality solution, businesses can use the high-quality data they have created in other systems for a wide variety of enterprise uses.



As a Total Data Quality initiative unfolds, users have access to more and more high-quality business information within a larger context. In the illustrations above, the detail on the left shows information viewed at the single system level. As that information is gradually understood within a larger context, users can see new connections and relationships that enrich the meaning of the data and expand its usefulness.

Full global capabilities

A comprehensive data quality solution must provide world-class capabilities, as an organization needs them. It should offer superior out-of-the-box solutions for every country in the world and be fully integrated so that it can produce completely consistent results and handle mixed international data sets. Businesses should be able to add geographic functionality as they need it for every country in the world.

Foundation for growth

Consistency of results is fundamental to good data quality practice, both at the basic level and at the enterprise level. A fully integrated data quality solution, built from the bottom up on widely accepted technology standards, is not just a “nice to have” feature for the IT department. Such a stable foundation guarantees consistent business results, regardless of the type of application, system, or platform environment used.

A stable foundation means that organizations can easily add data from additional operational units to their existing data process. It also means that the established data process can be exported to other environments with consistent results. A data quality system based on a solidly integrated technological framework truly “gets out of the way” and lets business users do the work that needs to be done.

Automated built-in intelligence

Data can be wrong in an endless variety of ways. Companies with massive volumes of data resolve these complex problems by using automated systems that can rapidly and intelligently interpret individual data fields one by one and create consistent data sets. The best data quality systems achieve intelligent interpretation through an open-ended set of rules based on advanced pattern-recognition techniques and accumulated best practices. Typically these systems should deliver high functionality out of the box for all business data.

Automated intelligence pertains to data profiling and discovery, too, so that large volumes of data are rapidly and fully analyzed and made available for both business and IT users to understand and discuss within a common workspace. Businesses should be able to write rules based on analysis that can be imported to data quality processes and re-analyzed when necessary.

Only automated intelligent systems provide rapid, consistent, re-usable processing that is fully auditable.

A Total Data Quality solution is structured so that the rules that govern data quality—both the built-in rules and any that might be added in-house—are external to the system and accessible to users. This makes them re-usable: other business groups can replicate them. It also makes them flexible: users can change, add, and finely tune the rules in a number of natural languages to meet each group's specific data needs.

The most comprehensive data quality solutions provide options for centralizing data standards for data governance, or allowing individual groups to modify those rules to suit their own business purposes. Businesses should also be able to track and audit all data quality processes to meet increasing demands for governance and compliance.

Access data from any source, send it to any destination

Being able to receive and distribute data using a wide variety of sources and systems is critical. Businesses must be able to leverage information across the enterprise, connecting across business functions—from sales to order processing to Web communications with customers and business partners to general ledger and many other functions. Total Data Quality provides a broad range of options for flexibly building high-quality corporate data assets over time.

Investing in Tomorrow: Corporate Data Assets

Organizations are beginning to understand that their data is a highly proprietary, unique, and valuable asset. It contains both a historical record and living picture of a company's entire business: customers, suppliers and trading partners, products and markets, operations and inventory, sales and profits. At a time when enterprise applications are being perceived as possible commodities, corporate data is assuming new prominence.⁶

Yet unprocessed, raw enterprise data, whether from legacy systems or streaming in real time, is very fragmented: It's hard to access, unreliable, difficult to use, impossible to trust. As much as 95% of companies' data goes unused according to industry experts, and upwards of 50% of it could be turned

Data can provide a living picture and a historical record of a company's entire business, an invaluable resource for competitive advantage.

into useful information.⁷ And data continues to stream into enterprises at hitherto unfathomable rates, tripling every two years by some estimations. Total Data Quality offers a way to harness the power of enterprise data for productive use. As business economists Hal Varian and Carl Shapiro observed, like it or not, all businesses are in the information business, and a major function going forward lies in “locating, filtering, and communicating what is useful” to the enterprise.⁸

Getting Started

An organization must be able to begin a Total Data Quality process anywhere, working with whatever business data sets are considered critical today. No one, however, can predict what their data needs are going to be tomorrow. Thus, businesses need to know that the time and money they invest in a data quality solution today won't be lost tomorrow because the solution was too limited or the technology not robust enough to evolve over time. That's why a Total Data Quality solution must be built to last.

¹ David Gelernter, *Machine Beauty*, (New York: Basic Books, 1998,) p. 50.

² “The Chaos Report,” Standish Group, 1994, http://www.standishgroup.com/sample_research/PDFpages/chaos1994.pdf.

³ http://www.technologyevaluation.com/Research/Research_Highlights/PIM/2004/07/research_notes/MI_PI_XBG_07_14_04_1.asp, Gartner Group, “Poor Data Quality Dooms Many IT Projects,” May 2004, http://www.fmi.org/supply/data_synchronization.pdf.

⁴ PriceWaterhouseCoopers, “Data Quality Management,” Global Data Management Survey, 2004.

⁵ http://www.fmi.org/supply/data_synchronization.pdf.

⁶ Nicholas Carr, *Does IT Matter?*, (Boston: Harvard Business School Press, 2004).

⁷ “Business Tools Get Smart,” *Internet World*, Feb. 2002.

⁸ Hal Varian and Carl Shapiro, *Information Rules*, (Boston: Harvard Business School Press, 1999), p. 12.

About Trillium Software

Trillium Software[®], a division of Harte-Hanks, has been selected by companies worldwide, both large and small, to improve their operational and analytic business decisions through accurate and timely information. Trillium Software offers an integrated suite of Total Data Quality software and services architected to turn raw, inaccurate data into usable, real-time information that supports business functions. The Trillium Software System[®] and Trillium Software Discovery are recognized as critical to the success of data warehouse, business intelligence, customer relationship management, enterprise resource planning, supply chain management, e-business, and other enterprise applications, and data integration, data migration, data stewardship, and data governance initiatives.

The **Trillium Software System[®]** provides cleansing, standardization, matching, enrichment, and linking functions for global customer data and operational business data, in any and all systems and applications. Regardless of data source or structure, the Trillium Software System ensures that data adheres to established standards and is adaptable to fit each organization's specific needs. Both single and double byte data are processed in local languages to provide a unique and centralized view of customers, products and services.

Trillium Software Discovery provides data investigation, discovery, and profiling in support of quality, migration, integration, stewardship, and governance initiatives. Fully integrated with the Trillium Software System, Trillium Software Discovery includes a broad range of features that expand upon basic profiling, including automated monitoring, business rule validation, and trend analysis.

Trillium Software Enrichment provides additional data enhancement services to complement, supplement, and amplify data available in-house. Choose from over 5000 third party data sources, and administrate enrichment through a single vendor.

Usage Notice

Permission to use this document is granted, provided that: (1) The copyright notice "©2005 by Trillium Software[®], a division of Harte-Hanks" appears in all copies, along with this permission notice. (2) Use of this document is only for informational and noncommercial or personal use and does not include copying or posting the document on any network computer or broadcasting the document through any medium. (3) The document is not modified from the original version.

It is illegal to reproduce, distribute, or broadcast this document in any context without express written permission from Trillium Software[®]. Use for any other purpose is expressly prohibited by law, and may result in severe civil and criminal penalties. Violators will be prosecuted to the maximum extent possible.

This document and related graphics might include technical inaccuracies or typographical errors and is subject to change at any time by Trillium Software. Trillium Software does not guarantee the suitability of the information contained in this document, which is provided "as is" without warranty of any kind. Trillium Software hereby disclaims all warranties and conditions with regard to this information, including warranties and conditions of merchantability, whether express, implied, or statutory, fitness for a particular purpose, title and noninfringement. In no event shall Trillium Software, and/or its respective suppliers be liable for any special, indirect, or consequential damages or any damages whatsoever resulting from loss of use, data, or profits, whether in an action of contract, negligence, or other tortious action, arising out of or in connection with the use or performance of information available from this white paper.