The Challenge of Global Data Synchronization
Table of Contents

Introduction .................................................................2

  Data Quality Within the Organization  ..................................4
  Emerging Data Quality Standards .......................................5
  Data Pools and Data Registries .......................................5
  “A Call to Arms” .........................................................5
  People ........................................................................6
  Processes ....................................................................6
  Technology ..................................................................6
  Identify the Data Owners ...............................................7
  Set Data Standards ......................................................7
  Determine Data Quality Metrics and KPIs .........................7
  Audit and Monitor Data Quality ......................................8

Technology: The Informatica Data Quality Solution ..............8

  Experience and Know-how ............................................8

Conclusion ..................................................................9

White Paper

The challenge of Global Data Synchronization
Introduction

All enterprises strive to improve their efficiencies and maximize the return on their resources—and as they do so, they learn that their processes and resources are only as good as the data that underlies them.

Enterprises and consulting experts alike now accept the need for common data standards within and across industries. They acknowledge that the scope and importance of proper **Data Quality Management (DQM)** mean that high-quality data must be achieved not only inside the organization but also across organizations. This means adopting cross-industry standards for product identification, classification, and attributes.

Today, significant advances in productivity and profitability are available to enterprises that agree common data standards with their peers, vendors, and customers; and enterprises that work to agreed data standards can now achieve first-mover advantage in current and future e-business initiatives. Soon, however—within the next decade—the adoption of such standards will cease to be advantageous and will become simply essential for any enterprise that aims to be efficient, competitive, and compliant with industry regulations.

Manufacturing companies in the consumer packaged goods (CPG) and fast-moving consumer goods (FMCG) sectors are at the forefront of initiatives to adopt common data quality standards under the banner of **Global Data Synchronization (GDS)**. The Global Commerce Initiative (GCI), a leading proponent of data standards, defines GDS as “The continuous harmonization of data attribute values between two or more different systems, with the end result being the data attribute values are the same in all of the systems.” Crucially, the GDS initiative includes plans for international data registries that will maintain standardized product data on behalf of standards-compliant enterprises.

Closely related to GDS is the EPCGlobal initiative, a standard for uniquely identifying every item in a supply chain with **Electronic Product Codes (EPC)**. These codes represent the evolution of the UPC bar-code technology introduced in the 1970s and will provide the underlying data for items identified by means of **Radio Frequency Identification (RFID)**.

The GDS and EPC/RFID initiatives are complementary and parallel; the former is creating a series of global data registries, and the latter is creating a system of unique product identification.

These endeavors will streamline the flow and maximize the value of data between CPG manufacturers and their customers—including leading retailers. In the next decade GDS will inevitably redefine the nature and importance of data quality management. These changes present challenges and opportunities for many industries and for the CPG sector in particular.

This White Paper describes the coming paradigm shift in DQM and explains how CPG enterprises can empower themselves to improve and maintain their critical master data in the era of global data standards.

It also outlines the solution provided by **Informatica Data Quality**, a unique data quality suite that empowers business information owners and data analysts to take control of enterprise-wide data quality initiatives. The product offers a single platform for tackling data quality at multiple points across the organization while maintaining centralized control and management of processes; it manages data quality in line with internal and external standards while also monitoring key performance indicators.
The Challenge: Data in the Global Village

Multinational players are leading the way in responding to the changing data demands of the global marketplace. They have realized that developing and maintaining data that is accurate, complete, and standardized in format is a vital task, but they also know that data management on this level is no longer enough. Data must be stored and shared in a standard, reliable format across all enterprises that use it.

By treating data as a strategic asset, rather than taking it for granted, enterprises can reap significant and immediate benefits, including:

- Reductions in the cost of serving customers
- Fewer supply chain problems
- Improved customer service
- The ability to reduce inventory levels and free up working capital
- Better execution of promotions with trading partners
- Streamlined product introductions

These challenges are not faced by larger players alone. The traditional model of multinationals as giant manufacturing organizations no longer fits. As Peter Drucker, the renowned expert on organizations and management, has stated, “Most multinationals are not big. Rather, they are mostly small- to medium-sized enterprises.” Drucker also states that “The essence of the new world economy is that it is, above all, an economy of information and truly a global economy.”

It is increasingly clear that DQM now involves twin processes: reaching target levels of data quality within the organization, and agreeing on data standards across organizations. Put another way, if the data used by an enterprise is unreliable, then the synchronization of this data to a common standard will yield at best partial improvements in efficiency. A recent Capgemini report stated that “If the data quality is poor, initiatives like Global Data Synchronization (GDS) and Electronic Product Code/Radio Frequency Identification (EPC/RFID) will not fully deliver the promised benefits.”

Enterprises need to tackle data quality on both fronts—internally and externally, in accord with partners and peers—to stay efficient into the future. They must also future-proof their data management practices against the changes to be ushered in by the evolution of electronic product codes and RFID and changes in global standards as GDS matures. This means a re-appraisal of the role of data and information in the enterprise, and the implementation of a DQM solution that is readied for the globalized future.

Moreover, the time for action on these issues is now. There is a window of opportunity for data-dependent organizations to align their data management processes on an ahead-of-the-curve basis with the emerging data standards in their marketplace.

1 From Peter Drucker’s essay Trading Places (National Interest magazine, Spring 2005). Copyright © 2005 The National Interest. All rights reserved.
Data Quality Within the Organization

The data quality challenges faced by organizations large and small can be summarized as six key data performance indicators:

- **completeness**
- **conformity**
- **consistency**
- **duplication**
- **integrity**
- **accuracy**

- Data records are incomplete if their fields are blank or filled with null or default data values
- Data records that contain correct information that has been entered in a non-standard or illegible format have conformity problems
- When data values are not aligned to a common format or standard across systems, the dataset has consistency problems
- When the same information occurs in multiple forms or in multiple records, the dataset has duplication problems
- Integrity is concerned with recognizing all relevant information in a data record, for example, common product or household information
- When data is not accurate, it cannot be used

Considered in terms of an enterprise’s internal processes, these DQM issues can severely impact the efficiency of operations and business processes. When considered in the context of the emerging global data standards, they combine to present an urgent data quality challenge.

For CPG manufacturers, low-quality data already generates significant problems throughout the supply chain—affecting new product introductions, promotions, order taking, fulfillment, logistics, pricing, and invoicing. Further automation in this scenario will only accelerate the rate at which data-related errors occur. These errors directly affect the bottom line by introducing the need for data re-entry (a costly manual process) and through too much or too little inventory being carried, resulting in missed sales through out-of-stock situations and poor customer service.

Furthermore, implementing software systems such as product information management (PIM) systems to centrally manage product data also increases the need for data quality measurement, validation, and cleansing. For example, data that does not conform to the GDS or EPC format is de facto inaccurate and will not be accepted into a data pool.
Emerging Data Quality Standards

Many enterprises are unprepared for the emerging standards and have yet to take on board the importance of aligning their data to them. Whereas every CPG enterprise uses bar codes, few are ready for the next step—Electronic Product Codes and RFID.

RFID is a supply chain management initiatives gaining traction in the CPG/FMCG sector. The electronic product code (EPC) is fundamental to RFID and will also share standards with GDS: for example, Global Trade Identification Numbers (GTIN) and Global Location Numbers (GLN). These standards are being developed in a collaborative spirit by groups such as the Global Commerce Initiative, EPCGlobal, and GS1, the global standards body created by the merger of the European Article Numbering association (EAN) and the Uniform Code Council (UCC).

Data Pools and Data Registries

High-quality, industry-standardized data will permit a quantum leap both in supply chain efficiencies and cost savings—but only when the enterprise data is available to the enterprises that need it and is of high enough quality to share with trading partners. To meet this need, the GCI and GS1 have created the Global Data Synchronization Network (GDSN) and are developing global data pools that can store and disseminate CPG data.

“A Call to Arms”

For CPG enterprises, this is a call to arms: their critical item data, from multiple systems and multiple locations, must be brought to the required levels of accuracy and consistency in order to benefit from the promised efficiencies and cost savings.

The systems that can benefit include warehouse management systems, supply chain management systems, product information management systems, and enterprise resource planning systems. Enterprise master data needs to become a “single point of truth,” and global data formats empower the enterprise to create and maintain truthful data.
Meeting the Challenge: DQM Solutions

It is essential that CPG enterprises respond to the challenges outlined above.

The first step in doing so is to address **internal data quality**. This requires an intelligent DQM methodology and powerful DQM tools.

A proper DQM strategy addresses the key data performance indicators in a continually improving process of data monitoring, analysis, and enhancement. Changes in data storage structures such as migration to new systems can provide a catalyst for root-and-branch data audit and data cleansing. Conversely, where new data enters the system, the DQM solution must address the issue of real-time data quality management.

The next step is **global data quality**.

An enterprise does not need a global reach in order to benefit from global standards. As outlined above, adherence to cross-industry standards will become close to essential for enterprises of all sizes in an increasingly global marketplace.

Crucially, the most successful DQM strategies also accept that data quality is not an IT issue but a core business issue. This means that the response to global data alignment must be enterprise-wide.

Three main areas within the enterprise must be addressed to ensure that enterprise-wide data quality is achieved:

**People**

The enterprise must promote a culture of data ownership. People must be given responsibility for their data, and the enterprise must acknowledge and reward data ownership at all levels. Importantly, enterprises must empower the key individuals who understand how to resolve data inconsistencies in source systems—those close to the originating processes.

**Processes**

Many enterprises manage their product information through discrete and disconnected processes. However, data quality has cross-functional impact, affecting the supply chain, buying and merchandising, and sales and marketing. Therefore, cross-functional data quality processes must be defined.

**Technology**

In many enterprises, product information resides in multiple systems, in multiple formats, and with multiple rules and standards applied, depending on the part of the enterprise using it. Enterprises must review their data management technologies and replace them if necessary with solutions that can ensure the required data standards are met.

These three areas indicate that meeting the data challenge is not simply a matter of buying more and newer software: a sound data quality methodology is equally important.
People and Processes: Changes in the Enterprise

Several CPG enterprises are already accumulating valuable experience relating to data standardization and global data standards. They have realized that there are four main building blocks to achieving the required step-change in data quality:

Identify the Data Owners

There must be a major shift in attitude to data ownership across the enterprise. It must be viewed as a business priority. Senior functional managers across all business units must be given ownership of the critical master data within their remit: they must be responsible for its data quality and rewarded when data quality targets are met.

Set Data Standards

Enterprise-wide data standards must be agreed upon and publicized. These standards will be defined by the nature of the business and by the current quality characteristics of the datasets in question, as determined by data audit. Moreover, in all appropriate cases, the standards set should be aligned with external standards. These standards must govern all important data objects and elements, and they must encompass the key data quality performance characteristics: completeness, conformity and validity, consistency, accuracy, duplication, and integrity.

Determine Data Quality Metrics and KPIs

Every enterprise faces its own unique set of DQM challenges. However, data quality trends among CPG enterprises are known. From the GDS perspective, enterprises face challenges of data standardization—to effect standardization internally and in sync with the global standards applicable in their industries.

It must be noted that data alignment of this type cannot be achieved overnight. Capgemini estimates that the process of internal data alignment can take six to twelve months to complete for large organizations. Selecting the correct technologies to apply to the job can streamline this process, however, and part of the process is the identification of appropriate KPIs—the information areas in which the need for quality improvement is most urgent, and the areas in which high data quality is a prerequisite for quality improvements elsewhere. These KPIs and the attendant metrics (establishing measurable objectives for DQ improvements over set time frames) must be agreed right across the organization. (These exercises also help to grow the intellectual capital of the enterprise.)
Audit and Monitor Data Quality

In addition to a data audit at the start of the formal DQM process, the enterprise must implement systems and processes to monitor data quality on an ongoing basis—and, where appropriate, at the point of data entry. Regular auditing and monitoring measure progress toward quality targets during the formal data quality process and also measure compliance with internal and external standards once the targets have been achieved. It is also important to measure the progress of ongoing data cleansing initiatives and to identify any process “leaks” that allow low quality data into the organization. These processes must be managed centrally but controlled locally by the business users who understand the data and are close to the processes that create it.

Furthermore, the enterprise should be willing and able to carry out ad hoc data quality audits, and the data quality tools in place should enable high-level and fine-grained audits of data across part or all of the enterprise.

Technology: The Informatica Data Quality Solution

Informatica is a world leader in the fields of data quality and information management. Informatica offers a unique combination of state-of-the-art software and an enterprise-level DQM methodology that builds human capital and enables the required step-changes in capability and attitude.

Crucially, Informatica’s solution—embodied in its Informatica Data Quality CPG application suite—is easy to use, flexible, and unobtrusive in the enterprise. Informatica Data Quality CPG does not add another layer of management in the enterprise, and it can deliver results to an executive desktop.

Experience and Know-how

Informatica has developed the CPG solution in conjunction with several global supply-side organizations to enable the CPG enterprise to generate a detailed picture of its data quality and to improve and standardize its data cost-effectively and quickly. With Informatica Data Quality CPG, manufacturing enterprises have access to a complete DQM solution—one which enables them to identify and resolve data quality and data consistency issues across the entire organization and for all critical data types.

The product empowers business owners to audit and monitor data quality across multiple applications and to ensure conformance with internal and external standards. Maintaining visibility of data quality in the key conformance areas, and easily enhancing data quality where non-conformance is identified, will be essential for enterprises that aim to take advantage of GDS and EPC/RFID: Informatica Data Quality CPG delivers these tools to the desktop via its data quality scorecard and dashboard features.

Informatica Data Quality CPG is a one-stop-shop for data quality analysis, standardization, de-duplication, and consolidation within the firewall of the CPG enterprise. It includes industry-specific reference content (including EAN/UCC product code reference content and GS1 product requirements), validation and consistency checking rules, and out-of-the-box data quality plans. (Plans are the rule-based solutions run by the product to analyze and enhance data.) The solution encompasses all critical enterprise data, including master data stored in enterprise resource planning, customer relationship management, warehouse management, business
intelligence, and product information management systems.

Informatica has built its CPG solution around its core data quality solution. Informatica Data Quality offers a single platform for tackling data quality at multiple points across the enterprise while enabling the control and management of data quality standards and KPIs from any desktop in the organization. A single solution provides all the functionality required for solving data quality problems, including profiling, matching, standardization, and consolidation. It also includes powerful reporting, scorecarding, and grading capabilities.

Informatica Data Quality integrates easily with key business processes and technologies such as Siebel, Oracle, and SAP.

• Informatica Data Quality SAP integration layer provides native two-way connectivity between the product and SAP’s mySAP Business Suite (including SAP ERP, Business Warehouse and CRM). Informatica Data Quality’s projects and processes can run ad hoc, in scheduled/batch mode, and in real time—like many other business systems, SAP can call Informatica Data Quality in real time to validate data as it is entered into the system.

Conclusion

Manufacturing enterprises and vendors of all types are approaching a crossroads in data management, and recognizing this crossroads is vital. Enterprises will have to choose between renovating and future-proofing their data systems or staying in the past and risking extinction.

A successful data quality management solution requires data quality applications that can monitor the state of an organization’s data according to key quality indicators; cleanse data where required; and neutralize problematic data at the system’s point of entry. For CPG enterprises, these attributes are magnified by the need not only to achieve data targets internally but also to align data with industry standards.

Informatica’s data quality audit, monitoring, and cleansing solution—Informatica Data Quality CPG—unlocks the potential inherent in enterprise data and meets the challenges involved in implementing GDS and similar initiatives. The solution includes software, industry domain professional services, relevant reference data content, and pre-built data quality plans. The centrally managed, distributed solution enables the enterprise to empower those local users who are close to the data to fix problems before they arise.

Informatica’s Data Quality CPG represents joined-up thinking for enterprises that wish to maximize their data’s potential and that also wish to implement DQM strategies that facilitate initiatives like GDS.