

SearchDataManagement.com

Cleaning up the product data quality mess

Poor product [data quality](#) is a widespread and costly problem -- especially for large manufacturers, retailers and others who deal with huge volumes of products and product components, according to IT experts and technology professionals.

And though it may sting a bit at first, spending the time and money necessary to improve product data quality is a highly worthwhile investment, says Andy Hayler, president and CEO of Information Difference, a London-based analyst firm that specializes in [master data management](#) (MDM).

Hayler, who has worked on several product [data quality projects](#) in the past, reports that 20% - 30% of typical product data master files consist of duplicate entries and flat-out erroneous information. That's a major issue, he explained, because bad product information leads directly to problems like expensive shipping mistakes and supply chain hold-ups, wasted time, lost up-sell opportunities, unsatisfied customers and, perhaps worst of all, angry company executives.

"You need to clean up the [product data quality] situation that you have now," Hayler said, "and then put in place rules and processes that will prevent the situation [from] recurring."

For more on product data quality

Learn more [data quality techniques and best practices](#)

Read about more challenges associated with [product data quality improvements](#)

See how one [product data quality tool](#) works

One firm's approach to product data quality

High quality product data is a necessity for MDM initiatives, but the reverse isn't necessarily true. Many organizations undertake product data quality projects in the absence of a larger [MDM plan](#).

Other firms, like Premier Inc., a Charlotte, N.C.-based firm that specializes in helping healthcare providers reduce [supply chain](#) costs, choose to create a product data master using MDM techniques, but save the larger MDM plans for later.

"We deal with healthcare products, so we basically created a master data management solution of our own to [de-duplicate] and help clean data," Palmer explained. "We then store that [data] in an [Oracle database](#)."

The next step for Premier is to evaluate MDM platform vendors that can help the company move forward with multi-entity MDM, an approach to master data that lets users concurrently manage multiple master data domains -- such as customers, accounts and products -- across all business processes. Palmer said the list of "entities" that Premier is looking to manage via multi-entity MDM includes hospitals, healthcare providers and suppliers.

"Having a good, coherent strategy to address [master data](#) is something that is on the front burner of things that we're looking at," he said.

Responding to product data quality concerns with an all-encompassing MDM initiative clearly isn't the right move for every company. But it is worth considering, Hayler said, because even a scaled-down, product-only MDM initiative can be very beneficial in terms of helping organizations eliminate error-prone and duplication-friendly product data silos.

"If you are trying to get your product data sorted out systemically, then it may be sensible to look at a more top-down approach [that involves creating] a [master data hub](#) which acts as the master source for your enterprise," Hayler said. The data hub can then serve product master data to business applications and business users, thus eliminating the need to duplicate data and store it in multiple, disparate systems.

"Our research shows that the average large company has nine systems maintaining product master data, which is eight more than is ideal," Hayler said.

When launching a product data quality project, try not to faint

Companies getting started on a plan to improve product data quality had better get ready to experience information chaos, according to Richard Murnane, manager of the enterprise data operations group at iJET, an Annapolis, Md.-based enterprise risk management and crisis response firm.

Murnane -- who has learned a great deal about data quality issues during his time at iJET and previously, at a large electronics manufacturer -- said the best way to start improving product data quality is to "look under the covers" and profile the specific information in question. Then prepare to either laugh or cry.

"You're going to see lots of duplication and lots of records that you no longer need. You'll have products in there that you haven't sold in years," Murnane said. "You'll have products in there that don't make any sense. If you're supposed to be selling widgets, why do you have washers and bolts in the product list?"

After the initial shock wanes, the next step is to kick off any necessary data cleansing processes. Typically, Murnane said, this involves merging duplicate product files -- which may actually have slight differences between them -- into new master files.

The process of merging files requires what are known as survivorship rules for each column in the datasets. For example, a person could create a rule which states that when merging two files that are nearly identical, the date on the newer file should overwrite the date on the older file.

"These are the rules that I will put in place so that [only] one of those columns persists at the end of the day," Murnane explained.

Selecting a product-focused data quality tools vendor

Companies seeking to fix product data should use extra caution when shopping for [data quality tools](#), because most tend to focus on handling customer data, such as names and addresses, Hayler said.

"[Customer data] is very different from product data," Hayler explained. "Customer data is well structured and has just a few attributes. Product data is frequently unstructured and can contain hundreds of attributes."

Organizations should avoid the tendency to hastily throw technology at a problem, proceed slowly in the purchasing process and take the time to confirm all software vendor claims. Failure to take these

steps, Hayler warned, could easily result in the purchase of data quality tools that promise everything but fail to deliver on product data quality.

There are some niche vendors who offer a strong focus on product master data, Hayler added. They include [Silver Creek, which was acquired by Oracle last January](#), Datactics Ltd. and InQuera.

Additionally, several broadly-focused [MDM platform vendors](#) offer a solid understanding of the subtleties of product data, Hayler said. They include GXS Inc., Hybris GmbH, Heiler Software AG, Stibo A/S, TIBCO Software Inc. and bigger players like Oracle, IBM and SAP.

All Rights Reserved, [Copyright 2005 - 2010](#), TechTarget | [Read our Privacy Statement](#)