# Attribute-Enriched Data Is Critical for Supply Chain Collaboration

**By Virgil Vincent, Weatherford International, and Richard Turner, Convergence Data Services**

Quality of data is essential if a company expects to realize significant value from its enterprise resource planning (ERP), product life-cycle management (PLM) and business intelligence (BI) tools.

Consider, for example, that reducing inventory (by identifying duplicate buy items), increasing purchasing power (by consolidating direct material spend across business units and product lines) and maximizing collaboration with suppliers and clients can only be truly realized by having rich, attribute-based data.

Unfortunately, the discipline with which many companies prepare data for migration from legacy applications to enterprise systems is limited - a problem that routinely compromises margins.

A particularly salient example is supply-chain management. With many corporations spending nearly 60 percent of their revenues on purchased products and services, supply-chain management has a fundamental impact on the bottom line.

*"A major obstacle to achieving improved efficiency in the supply chain is the inability to share and transfer meaningful information to the extended enterprise - particularly outside suppliers."*

However, improving efficiency in the supply chain is routinely constrained by organizations' inability to share and transfer meaningful information to the extended enterprise - particularly outside suppliers.

And the principal culprit is data corruption and inconsistencies (internal and external).

Typically there are several causes of redundant, erroneous, or nonstandard data:

- disparate systems that cannot communicate internally or externally;
- different commodity naming and numbering conventions;
- lack of a formal process for creating new parts; and
- lack of governance with respect to data standards.

Many companies have sought to overcome these problems by implementing enterprise systems. Unfortunately, many such efforts have been less than successful because of their failure to deal with unstructured information and the

lack of industry-standard taxonomies.

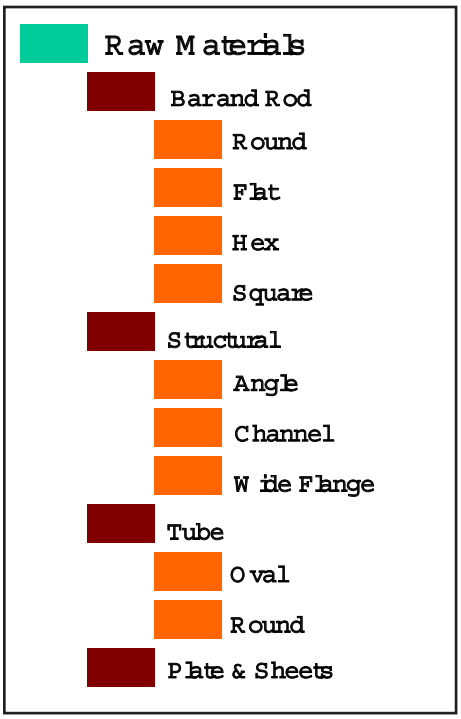
### *Using Enterprise Data Classification*

Enterprise data classification is a method of organizing data by grouping similar products into a common classification hierarchy.

As shown in Example 1, a company's classification structure often looks like a tree with branches, with a product grouping of attributes used to uniquely describe the products at the end of each branch. Products with common attributes thus can be found under the same category in the classification tree. Example 1 depicts the classification structure for raw materials commonly used in the oil and gas industry.

An enterprise classification structure contains the agreed-upon standards for describing all products. This begins with major product and service groupings, and bores down to defining acceptable attribute values for a specific group of products.

It usually is very easy to see the difference in data quality between companies that have done little with their data and companies that have applied




***Example 1: Classification tree structure for raw materials. Items are grouped based on identical attribute types.***

| OLD LEGACY DESCRIPTION | NEW ENTERPRISE DESCRIPTION |
|---|---|
| 03.250 02.440 41X 065-105Y TPN | Tube, 3.250 OD 2.440 ID 4130 SRL HF WS-440 |
| 4.750 X 3.890 80KSI | Tube, 4.750 OD 3.890 ID 4130 Q-T HF WS-410 |
| 1.875 x 1.502,1020 | Tube, 1.875 OD 1.502 ID 1020 SRL WS-107 |
| TUBE,BBL,CARB,2 1/4 X 2 3/4 | Tube, 2.750 OD 2.250 ID 1020 CARB WS-191 |
| TUBE, STL 3.09X2.29 | Tube, 3.090 OD 2.290 ID 1026 SRL CD WS-141 |
| 4 | Tu |
| 1. | Tu |
| 2.603 X 1.760,1015 | Tube, 2.603 OD 1.760 ID 1015/26 CD WS-143 |
| 4 X 3,4130,TB | Tube, 4.000 OD 3.000 ID 4130 ANL CD WS-451 |
| 1 X 0.625,304 STNLS STL | Tube, 1.000 OD 0.625 ID 304 ANL CD SS-201 |
| 1.500 X 1.125,304 | Tube, 1.500 OD 1.125 ID 304 ANL CD SS-201 |
| TBG,4130CDS,1.500X0.813X52-3/4 | Tube, 1.500 OD 0.813 ID 4130 NORM CD WS-451 |

**Various naming conventions**

**Consistent descriptions across enterprise**

***Example 2: Legacy vs. Attribute Enhanced Descriptions***

enterprise classification mechanisms. Example 2 highlights these differences.

As shown in Example 2, there are significant differences between the old legacy product descriptions and the new enterprise product descriptions.

For example, the new descriptions depict a consistent order of information, as well as a consistent format involving the use of decimal places, use of abbreviations, use of the same noun to start a description and so forth.

The product descriptions on the left also are written haphazardly. Each description looks different, lacks useful information and has no standard format.

Basically, a lot more information can be obtained on the products on the right side, thus making it much easier to find what you are looking for. Descriptions on the right are made up of attribute values that can also be used to find a certain product.

Example 3 illustrates the value of attribute-based data in developing a consistent enterprise description.

Using an automated description generator with enterprise agreed templates, based on allowed attribute values, drives description conformity throughout a product category.

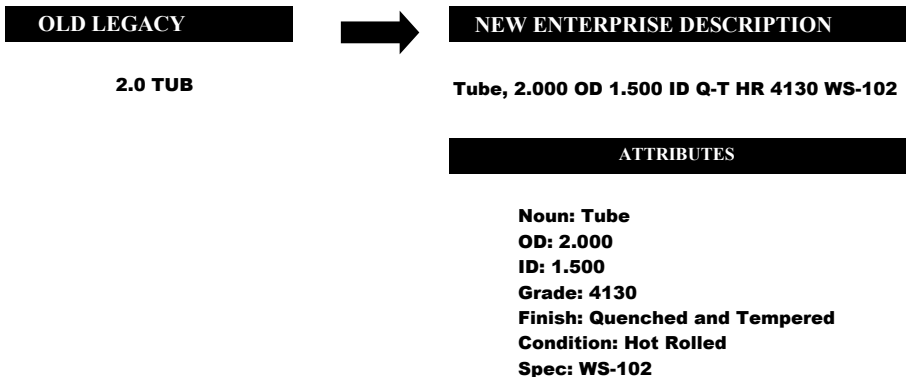
The process also serves as an effective validation tool in identifying any missing or incorrect attribute values.

### *Importance of Maintaining Data Standards*

Once a company has carefully converted its legacy data to an agreed-upon standard enterprise classification, the next step is to apply the same level of diligence to any new products that will be created going forward.

The focal point here is the new part creation process that typically is managed by today's product life-cycle management (PLM) systems.

Most PLM systems are not capable of policing the creation of new product classifications, which often contributes to a free-formatting approach to describing new products.




***Example 3: An example of enriched product data described by attributes***

Consequently, it is essential that companies introduce exceptional levels of order to their new-part-creation processes.

To keep up to date with new products and standards, companies also will need the ability to maintain and update their enterprise classification structure.

In other words, if a PLM system is a particular company's master data manager for its product data, then it will need to determine how to update and maintain the classification in its PLM system on an ongoing basis.

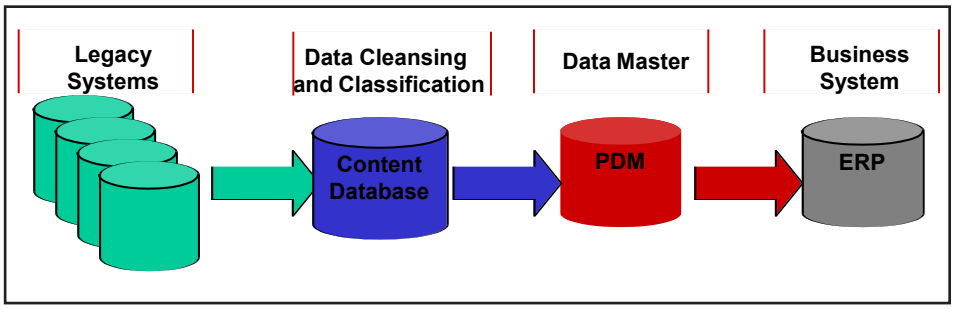
### Consolidating Duplicate Products

Once attributes have been defined for a company's key product groups, the process of identifying duplicate products becomes much easier.

If it has done a good job of enforcing the use of standard attributes to describe each product group, then those attributes can be used to conduct consolidation searches (Example 4).

It is also important that - to enable true search results - a standard set of attribute values exists for each attribute.

Example 4 (depicting a large oil and gas OEM that uses raw materials across all of its business) shows the power of having a rich attribute set that helps identify duplicate raw materials.




*Example 5: Example of a Data Classification Architecture*

### Enterprise Data Classification Process

As we have noted, successfully extracting data from multiple legacy systems - and then passing the data on to new enterprise applications - demands a robust data-classification process.

In fact, a robust data-classification process can pay big dividends to an organization through its ability to pass a clean, rich data set that has been pre-validated to meet the requirements of new enterprise applications.

For this reason, it is vital that companies work to prepare their legacy data before loading the data into a new application architecture.

A robust classification process starts with extracting only the most pertinent information - not passing along information that should have been purged previously.

Once the optimal legacy data-set is determined, the information then should be loaded into a temporary classification content database.

Example 5 describes a classification architecture.

Prior to classifying a new batch of legacy data, the company should work to consolidate duplicates and map newly loaded parts to any existing, matching classified parts.

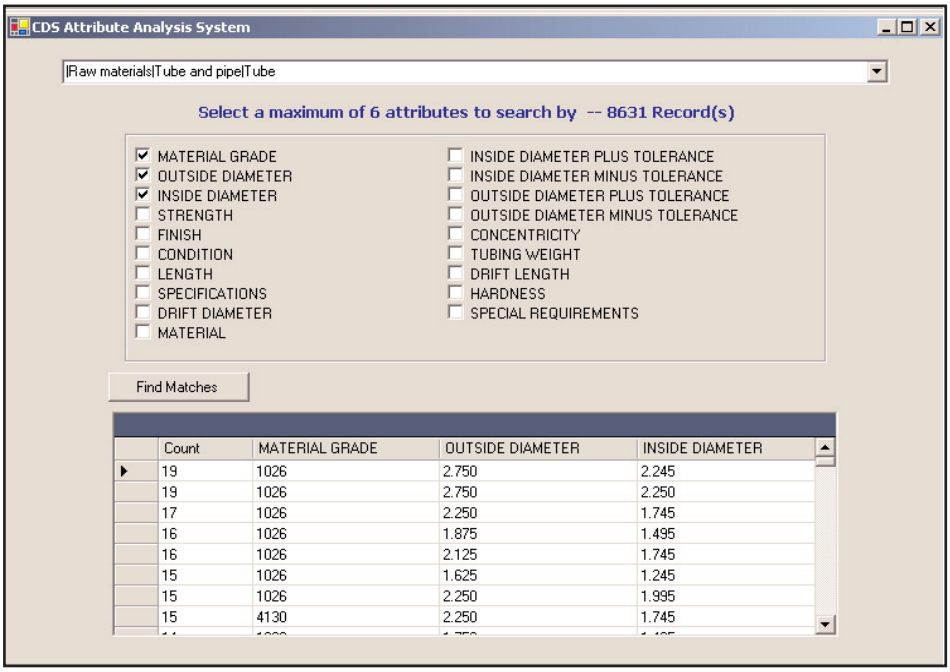
Data classification represents a good opportunity to consolidate potential duplicate parts prior to loading classified data into a new system.

The next step is to take the remaining legacy data-set and classify those products into the approved enterprise classification structure.

If an enterprise classification structure does not exist, then subject matter experts need to determine what that structure should be, including identifying the product categories and attributes used to describe these products.

Once a legacy data set has been classified, the data can be cleansed and populated with additional attribute information.

Example 6 describes some of the key activities that take place during the time that data are held in the classification content database.




*Example 4: Attribute Consolidation Tool*

### Data Cleansing and Classification

After populating data with attribute values, the classification and attribute values then will need to be validated against agreed-upon enterprise standards.

Once the classification cleansing and validation have been completed, various descriptions can be generated

# Quality Data is Key to Successful Strategic Sourcing, Spend Analysis, Collaboration

Strategic Sourcing. Suppliers are an increasingly valuable contributor to profitability. In fact, through innovation, continuous improvement and collaboration, many have become extensions of their customers' organizations.

For this reason, most extended enterprises must work to ensure the delivery of real-time information to suppliers.

They also must acknowledge that customer-supplier relationships of this type have characteristics that are different from other types of supply chain activities; for example, "total supply cost" is more important that purchase price only.

All in all, these are relationships that demand information and technology sharing to achieve common standards and cost reductions.

Work flows and supply chain information will support common standards such as XML for interoperability among critical suppliers.

Spend Analysis. Consolidating purchasing power - driven by repeatable spend analysis - is a major component of strategic sourcing.

Purchasing at the enterprise level, rather than at the single-facility level can have significant impact on spend reduction.

In many ways, maximizing buying leverage across divisions and product lines is made possible by data with a rich set of attributes, enterprise descriptions, reconciled units of measure and accurate supplier information.

Collaboration. A collaborative supply chain efficiently integrates suppliers, manufacturers, warehouses and stores, so that products are produced and distributed so as to minimize total system costs and meet customer requirements.

Enabling supply-chain technologies will provide more efficient communication - with decisions based on real-time information - without cause to worry about data discrepancies.

The collaborative flow of data must be: 1) attribute based; 2) integrated into enterprise systems; and 3)web-based to maximize flexibility.

Supplier benefits associated with an attribute-based, collaborative process include:

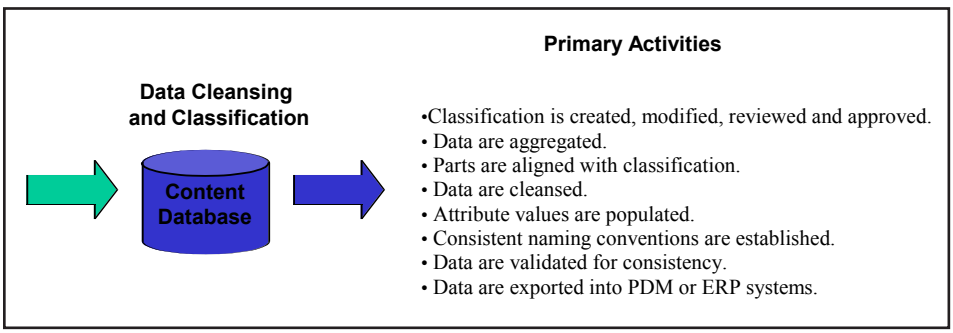
- an easier way of doing business;
- fewer errors;
- better inventory management and forecasting accuracy;
- reduced data maintenance, e.g., not having to map part-numbering systems;
- improved ability to engage in e-commerce;
- value-added service that differentiates a supplier from its competitors;
- unit-cost savings and lower operating costs; and
- opportunity to manage all of a customer's data, thereby simplifying change control and maintaining industry standards versus local standards.

from the existing attribute set.

The next step of the classification process then is to run final validations on both the classification structure and the new data set to ensure that the information will be accepted by the new enterprise applications.

Typically, the data is classified in batches, as described in the previous example. The size of each batch can vary widely, depending on how many disparate legacy databases exist across the enterprise.

The initial condition of the legacy data and the desired future state will dictate the number of resources needed to help classify and cleanse the data for each batch.




***Example 6: Typical activities involved in the classification process***

## *KEY POINTS*

The mission of this brief paper has been to emphasize the importance of consistent, high-quality data, and to identify the key activities associated with introducing that data into a new enterprise application environment.

The key points of this mission are summarized below.

- Data must be classified, enriched with standard attributes and developed with a consistent naming convention to achieve maximum benefit from enterprise systems.
- Identifying duplicate items using attribute searches can help reduce inventory levels

and enhance purchasing power by enabling consolidation across business units and product lines.

- More effective collaboration with suppliers - as well as syndication to clients - can only be truly realized through rich attribute-based data.
- Focusing on the supply chain (where the majority of the spend occurs on direct materials) provides the most cost-benefit opportunities.
- Enterprise data classification is a means of developing a common language - an attribute-based system where information can readily be shared with partners.
- Participation by an industry forum of OEMs and suppliers is key to moving toward a single taxonomy and reducing the need for suppliers to develop multiple "tailor-made" systems for each client.
- Increasing efficiency in the supply chain by automating supply-side processes will become even more important as companies exhaust internal opportunities to reduce costs.

*Editor's Note: A related article on PLM systems appears in the May 2003 issue of Upstream CIO on page 1.*