



WebSphere software

Product Information Management for e-commerce

*Divakaran Permalla
IBM WebSphere Product Center*

Contents

- 2 Introduction**
- 2 Product Information Management benefits to e-commerce**
- 3 Organization and logical groupings of items**
- 5 New item introduction and synchronization**
- 7 Frequency of data changes**
- 9 Content management**
- 10 Information exchange**
- 14 Keys to successful Product Information Management implementations**
- 15 Summary**

Introduction

Publishing enriched product data to new or existing e-commerce solutions presents unique challenges. This white paper presented by the IBM WebSphere® Product Center (WPC) team is a result of our experiences implementing Product Information Management (PIM) solutions for customers to meet these challenges better. The purpose of this white paper is to document various consumer requirements as they relate specifically to e-commerce. The goal is to provide readers with an understanding of the overall capabilities a PIM solution should include to meet those requirements.

Product Information Management benefits to e-commerce

While business-to-business (B2B) Web sites are primarily part of the supply chain management effort, business-to-consumer (B2C) Web sites focus on selling directly to consumers. For that reason, an e-commerce solution in a typical B2C environment provides features and functionalities to help appeal to end-consumers by enhancing their online shopping experience. At the same time, e-commerce solutions are usually integrated with transactional systems such as order management to ensure timely execution of the order placed on an e-commerce site.

A large amount of data from multiple sources is consolidated into these e-commerce solutions, which affects the quality of the end-consumer shopping experience. This information includes data classification by the type of items sold, organization of the product assortment for easy navigation and various assets that correspond to an item such as detailed descriptions, images, music, video or manuals.

Benefits of a PIM solution

Benefits of implementing a PIM solution to support e-commerce include:

- Implementing business processes around the introduction, enrichment and management of product information for data publication to an e-commerce solution.
- Consolidating data from existing repositories by feeding existing data sources into PIM.
- Increasing business-user involvement in managing data, leading to improved business processes and reduced dependencies on IT personnel.
- Going beyond publishing to e-commerce solutions by publishing data to a variety of external systems such as portals and print publishing solutions, eliminating the need to maintain product information in multiple applications.

The utilities that come with e-commerce applications such as bulk data import, manual user data maintenance and some approval processes provide tools to manage data. However, with the growing need to make dynamic changes on Web sites – to quickly add new products, to fix critical errors in product information such as price or to create online promotions, cross-sells or up-sells – these tools often fall short in the following areas.

- *Providing complex data validation to ensure accuracy throughout the management and enrichment of product information*
- *Managing multiple states of catalogs – those that are published or in draft stages*
- *Managing multiple hierarchies with varying category data*
- *Incorporating user-friendly data management tools to mimic the Web site structure*

Therefore, e-commerce is just one piece of the puzzle. A PIM solution can help manage product information for an entire organization, meeting many of the requirements a typical e-commerce solution cannot address alone.

Organization and logical groupings of items

Imagine visiting a new department store to shop for a black leather sofa from a popular brand. Unless you are aware of the physical layout of the store, it is going to take you some time to find the exact location where the item is displayed. The location of the sofa can be identified more quickly by following the signs on each floor/aisle, bringing you to the right area. Once you get to the exact location, you can narrow down your search to the specific brand, color and size of your choice.

Now that you have chosen the sofa, you might be interested in a chair. It is highly likely that the matching chairs are stored near the sofa. This type of association is classified into up-sell, cross-sell techniques. The store may study customer buying patterns and organize merchandise to improve the customer experience and also to increase store sales. The store may also wish to sell these products together such as in promotional bundling, which could have a combined price that is less than buying the sofa and chair at individual item prices.

Items are stored and organized in a similar way in most businesses, where the end consumer navigates to the specific item of choice to make a purchase. This applies to both physical and online stores. The better the store is organized, the better the customer experience, which adds to overall customer satisfaction.

To support this type of experience while shopping online, a PIM solution supporting an e-commerce solution should be able to:

- Store the various sellable items (SKUs).
- Create logical groupings of items (products). In the previous scenario, the ultimate purchase is the SKU. However, sometimes the user first looks at the leather sofas from Brand XYZ and then narrows down to the specific size and color. In this case, the sofas from Brand XYZ form a logical group (see Figure 1).
- Create logical groups of existing logical groups. For example, bundled items can be a separate logical group. The additional flexibility needed includes the creation of logical groups comprising specific SKUs from other logical groups, logical groups comprising of other logical groups (sub-groups) or primary vs. secondary associations between logical groups and SKUs (see Figure 2).
- Provide the ability to create associations between logical group to logical group, logical group to SKU, SKU to SKU and SKU to logical group (see Figure 3).
- Create flexible views (hierarchies) and organization (categorization) of items with the ability to have items mapped to multiple views and to multiple locations (categories) within a view (see Figure 4), enabling the item being searched to be located in multiple ways.

Figure 1

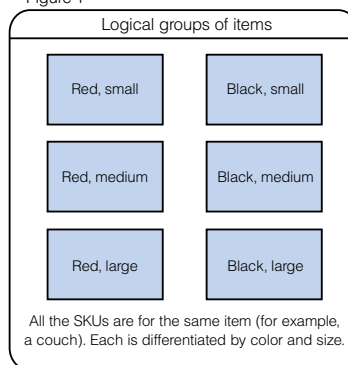


Figure 2

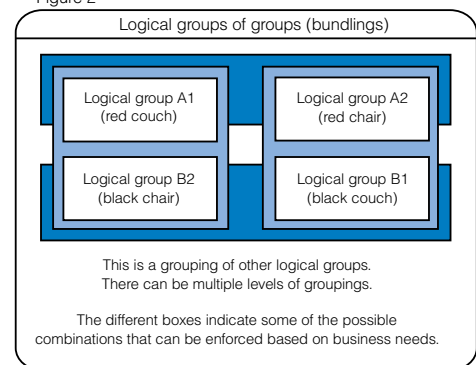


Figure 3

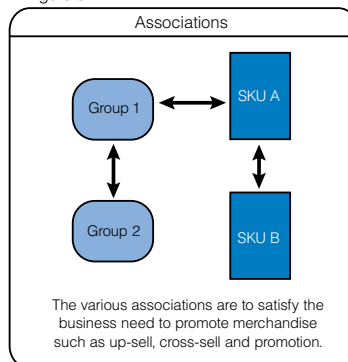
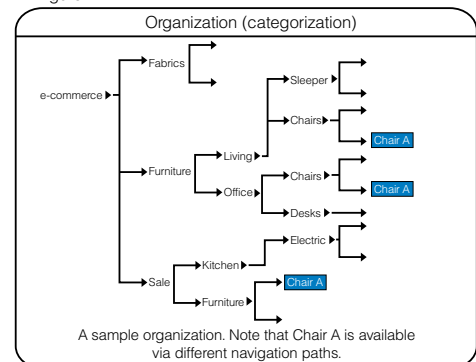


Figure 4



New item introduction and synchronization

Ways of introducing items to support an e-commerce solution

Existing SKUs can be introduced into a PIM system, and can trigger the logical group creation process. In such situations, the SKUs are known and uniquely identified. There will be situations where businesses start from the logical groups and add new SKUs to these groups. This requires new SKU creation with a unique identity. Rarely, businesses have existing SKUs maintained in a master repository and would also like to use PIM to introduce new ones. This is not a recommended practice, primarily because this defeats the purpose of a single master repository, which a PIM system would provide.

There are different ways in which today's businesses introduce new items (SKUs/ logical groups). These include:

- *Maintaining a master repository of items (SKUs) in which a manual or automated process provides the introduction.*
- *Including approval processes at various stages of the item introduction, which again can be a mixture of manual and automated processes.*
- *Introducing various peripheral items such as products or bundles to address e-commerce needs. In some situations, this peripheral item creation is left to the application that needs this information such as pictures in an image management system.*
- *Including manual/semi-automated processes for managing obsolescence in the master repository and propagating changes in a timely manner to downstream systems.*

PIM solutions are designed to play the role of a master repository. In such cases, a PIM solution is required to include several features to enable item introduction including:

- *A one-time integration of the SKUs from existing legacy systems.*
- *The formation of a single repository, serving the needs of all systems needing item information.*
- *Manual or semi-automated mechanisms to introduce new items (SKUs/logical groups) including creating new SKUs with unique identifiers, creating logical groups with existing SKUs and/or new SKUs and maintaining unique SKU characteristics with a logical group.*
- *Mechanisms for associating various types of related content and data, such as images, manuals, audio or video files, with SKUs/logical groups.*
- *Approval mechanisms for the various stages of item introduction and maintenance processes.*
- *Management of obsolescence and propagation of timely changes to all destination systems.*

In cases where PIM plays the role of a data source for new items or item changes in e-commerce systems, PIM also needs to include the ability to:

- *Import SKUs from external sources, such as partners or suppliers, and include them in existing or new logical groupings.*
- *Associate any e-commerce related information to the imported SKUs, such as promotional pricing.*
- *Create new logical groups with existing SKUs and/or new SKUs.*
- *Include approval mechanisms for the logical group creation/changes.*
- *Maintain the uniqueness of the SKU characteristics within a logical group.*
- *Manage an active assortment of communication to the e-commerce system.*

As a best practice business process, the introduction of SKUs should all take place in one system to avoid unnecessary integration and synchronization.

Synchronization

Each organization has specific requirements when it comes to introducing a new item into a PIM system and all downstream systems. These requirements range from a simple process to a more complex one, which are defined by the nature of the business and, more importantly, the evolution of processes and related corporate policy of an organization. This in itself is an exhaustive topic, but the important point to consider is the need for synergy of this process between PIM and e-commerce related information.

The information used in a commerce system consists of two types at a high level. The first is static information. This is product information including attributes and related static assets (such as images, fact sheets or assembly instructions). The second type is dynamic information such as inventory and various transactional data including the price of an item that might change depending on the buyer viewing that item. PIM and content management solutions typically store static data, while ERP or other operational systems store dynamic (transactional) data.

For a product to be sellable on an e-commerce site, both static and dynamic information are essential. Due to the nature of this data, both types should be managed separately outside of e-commerce, which can lead to having separate hooks into the e-commerce solution.

Because these disparate systems exist in retail environments, an item tends to be introduced into each solution at different times. When a new item is introduced into a PIM solution and is then published to an e-commerce solution, this item technically cannot be sold until it is recognized by all other systems requiring this new product information. It is essential to synchronize the item introduction process with dependent systems while also providing information to e-commerce. There are various ways in which this can be achieved. The ideal way is to control the information flow into e-commerce until all the systems have the required information. This can be achieved by:

- *Introducing a message-driven integration between PIM and dynamic/transactional systems to initiate and acknowledge the item introduction. Enterprise application integration (EAI) tools can be useful here.*
- *Having a manual or automated approval process for new items to ensure the item is introduced to the required systems before publishing to e-commerce.*
- *Allowing the e-commerce system to handle the consolidation of all information flow and providing access to an item on the site only if this item is available in all the components.*

Frequency of data changes

With the explosion of Internet commerce, most companies now have an online presence with e-commerce capabilities for selling directly to end-consumers. For many, this online presence augments their main selling channel, which might be through physical stores or through catalogs and call centers. This has recently become quite common among companies in industries such as discrete manufacturing and consumer electronics, which traditionally sell through extensive networks of channel partners.

With increasing consumer demand for product information to assist in the evaluation process before purchasing, it became a requirement for end-consumers to access information or to even buy directly from the manufacturer. In the case of retail businesses, some entered the market strictly through the online channel while others have traditionally been retail outlets who have now started online e-commerce as another sales channel.

Regardless of industry, not only do companies need a strong market presence, but they also need to be up-to-speed and adaptable to market changes in order to be competitive. As a result of market awareness, data accuracy and speed-to-market of updated information on commerce sites are essential and key to the business. Companies have implemented various techniques to ensure that the latest data appears on their e-commerce sites in order to meet the demands of the consumer. But these techniques can vary based on the nature of the business. For example, a hardware or auto parts retailer may not have items that change often when compared to apparel or a specialty store, where the nature of business demands changes at a higher frequency.

These data changes may include introducing new items, changes in characteristics of items such as modified descriptions, new image associations or changes in up-sell or cross-sell relationships, and changes due to discontinued items.

Business factors influencing these changes can include seasonal data or data driven by dates or events. Some retail businesses such as apparel or specialty stores are driven by seasons. These retailers make changes to the majority of their assortments to correspond to a season in which the products are offered. In some cases, these changes can result in the re-introduction of items from prior-year seasons. But in most of the cases, a large number of new items are introduced. Date or event-driven changes, on the other hand, can be related to special events, promotions or discontinued products.

In order for a PIM solution to support these needs, the following capabilities are required.

- *The ability to introduce a new assortment of products quickly*
- *A mechanism to publish data in full and incremental modes to capture frequent changes and to reflect those changes on the e-commerce site*
- *The capability to capture and publish attribute-level state changes of an item based on effective dates including all types of assets such as images, files and prices*
- *Features to incorporate major assortment changes for businesses driven by seasonal factors while providing archival capabilities to enable reintroduction of existing items*

Most of these features are available in PIM products. The areas that need additional work include attribute-level state changes based on effective dates and enablement of major assortment changes without duplication of data.

Content management

Web sites that include e-commerce usually contain an extensive amount of information. This includes images, banners, manuals, articles, video and a host of other information in addition to the pure product information for the purposes of e-commerce management. Information that is not tied to products can be stored and managed by content management applications.

Content management and PIM solutions can jointly manage e-commerce data that falls into two areas. Pure content management is an area that includes the storage and management of unstructured, or binary, information such as images, manuals and video. This type of information cannot be broken down but it can be referenced in PIM while being stored and managed elsewhere by a content management system. Structured content management, on the other hand, includes the management of information such as an article, recipe or FAQ. This information can be structured by providing a defining set of attributes where the information is captured. This makes it easy for a data entry user to introduce such new information by identifying the common elements and breaking them down into manageable attributes. An article, for example has a header, introduction, multiple paragraphs and footer. Each element can be managed using a separate attribute.

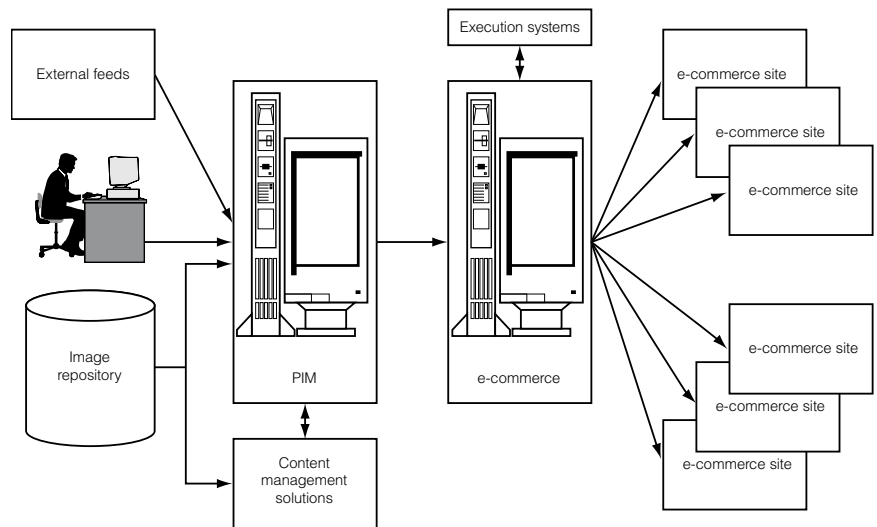
The PIM solution can integrate to an external content management system to provide various associations between existing content. For example, if an image managed by a content management system is associated with a product managed in a PIM system, the PIM solution will provide a mechanism to associate this image as a product attribute and is responsible for managing the product information and all such associations. However, the external content management system is responsible for managing and storing pure content-related data.

For structured content that can be broken down into manageable attributes, it is beneficial to have the capability of defining these elements in a PIM solution with features to modify such data. Only structured content that can be broken down into manageable attributes should be managed in PIM, however. This will make the data maintenance process easier. The presentation of this content should use a template approach so that the PIM solution only provides the data for the elements managed in PIM that are applicable to the template.

Information exchange

Ensuring complete information exchange

The elements of an entity in PIM include the data stored in the system in addition to the other components stored externally. There are various ways for this data to be delivered. It can be in the form of user entry, imported information from external systems or messages.



This information is accumulated and enriched in PIM before publishing to external systems. If the information related to an entity in PIM is incomplete, it cannot be published to external systems. At the same time, it is essential for a PIM system to ensure that all the various components of an element are published as a single unit to the external systems. This is to ensure complete data integrity. For example, if an entity in PIM has a list of attributes with textual information, associated images (stored in content management systems) and linkages to unstructured content, this information all needs to be provided to e-commerce for the entity to be complete. This requires coordination with all the contributing systems used by e-commerce.

This type of information exchange falls into:

- *Complete configuration, assortment information and all the associated attributes.*
- *Referenced information such as images, files, unstructured content or any other externally stored information.*

For completeness of information in e-commerce, and to enable execution of orders, the required information flow from transactional systems managed outside of PIM is also essential.

Transmission of data

The data transmission approach depends primarily on the capabilities of the PIM system, the e-commerce system and the integration framework. The PIM system should be able to provide a complete assortment publish when necessary, while the e-commerce systems should be capable of accepting and loading the volume data. Sometimes only a subset of information or only changed information is needed. In these instances, the PIM system provides a delta feed and the e-commerce system loads the data. This is specifically true for businesses that have many data changes due to frequent assortment changes, seasonal factors or promotions, and where complete assortment publishing would become impractical.

Various transmission modes can be used for the data transfer. This could be a simple file transfer or a more robust and reliable mechanism such as EAI. The risk with file transfers is its reliability and its recoverability from transmission errors. EAI tools, on the other hand, are created to provide a more reliable mechanism and are the preferred approach for such transmissions for the following reasons.

- *Significantly more robust than File Transfer Protocol (FTP)*
- *Transformation capabilities without low level programming*
- *Guaranteed delivery*
- *Automatic restart and recovery*
- *Auditing and journaling capabilities, which simplify maintenance*

Implementing EAI requires some amount of planning for the overall usage and integration with various other applications within an enterprise. Consumers lacking this infrastructure can use FTP to quickly enable this data exchange. However, a long-term approach of moving to EAI should be put in place for the best long term results.

Preview capability

Successful e-commerce sites include ease of navigation, appealing presentation and the availability of accurate information. While navigation is highly dependent on the design of the site, the presentation and information are dependent on the data provided by the PIM and content management solutions. In such cases, there will be multiple iterations of the data sent to e-commerce to make sure the site looks good and provides accurate information. These multiple iterations can lead to numerous incremental updates.

A preview mechanism can help enable quick verification of the PIM data to ensure it matches e-commerce needs. There are different ways in which a preview can be done including:

- *The data entry user explicitly logging into an e-commerce environment and previewing on the system.*
- *Providing a preview capability within PIM systems by mimicking the e-commerce presentation.*
- *Integrating PIM systems with e-commerce to invoke the relevant e-commerce screens and present these screens with the user-entered data.*

Although the first option is the simplest mechanism, it reduces user productivity and increases the number of iterations of data push, directly impacting the time taken to validate data.

The second option gives the user a look similar to the e-commerce page. However, the preview is not the same, which can lead to last-minute surprises on the site. There is also more development effort involved in PIM constantly keeping up with e-commerce screen changes. And there are challenges in implementing some specific functionality tied to the actual e-commerce environment such as presenting other data residing on e-commerce when PIM does not have visibility into this data.

The third option is preferred, even though it involves some technical challenges. With this option, the PIM system invokes an e-commerce application from a PIM application by passing the entity information. This gives users the ability to perform a preview on a real e-commerce system from PIM as data is entered. This way, users can quickly update incorrect information. This significantly reduces physical data transfers, increases user productivity and reduces the time it takes to introduce products onto the site.

Enabling a simplified integration requires PIM to extract and push a single entity (such as a product) along with all its relevant data. Similarly, e-commerce needs a mechanism to invoke the system with this data resulting in a presentation of the relevant page with the provided data. The end result is to have the ability to preview changes in product content on a Web site – to check for accuracy, for example, before committing those changes to production.

Keys to successful Product Information Management implementations

There are some challenges to overcome in order to implement a PIM solution successfully. This includes:

- *Providing state management of information.*
- *Catering to various commerce application structures while defining logical groupings of products.*
- *Classifying between pure content and structured content.*
- *Ensuring complete information exchange.*
- *Overcoming the technical limitations of implementing preview capabilities with various e-commerce systems.*

First, a mechanism should be introduced that can provide state management of information. This type of information management is based on a date, seasonality or other business-driven factor. The challenge in implementing this mechanism successfully is to be able to track attribute-level state changes based on an event and triggering mechanisms. It can potentially be an expensive operation from an overall data maintenance and systems performance perspective. This expense can be minimized by defining a uniform state management logic that satisfies various business needs, and moving this logic to a low level code.

In implementing a mechanism for state management of information, it is also important to avoid duplication of data while creating new seasons using existing seasonal data. The ideal scenario is to leverage data reuse capabilities without duplicating data when there is less than 50 percent of data changing between seasons.

It is also important to provide a way to support various commerce application structures while defining logical groupings of products. The challenge in doing this is to find a way to represent the various e-commerce business models in the PIM solution. Integration can also be complicated as the data representation of each e-commerce system can be very different. This can be minimized by incorporating flexible groupings and associations of various entities managed in PIM.

Classifying between pure content and structured content can be difficult as well, unless all binary information can be stored in content management systems. Any information that can be structurally broken down into manageable attributes can be stored in a PIM solution. The references from PIM solutions to content management solutions are still needed, however.

Because e-commerce systems are integrated with various other systems, it is critical that the information provided by PIM meets the needs of all systems in order to ensure complete information exchange. To achieve this, you need to identify all data points into e-commerce as part of the discovery phase and take these into account while designing the integration. This includes systems that have no information in the PIM solution.

Finally, there can be technical limitations in implementing a preview that incorporates various e-commerce systems. To overcome these limitations, the e-commerce system can provide mechanisms to enable the preview capability. It can then be used by the PIM solution to provide integration.

Summary

Implementing a PIM solution for e-commerce does have its challenges but the rewards can be experienced in the end product: a retail site with well-managed data providing a consistent user experience as well as an enterprise-wide solution for managing product information for other purposes. A PIM solution can help manage product information by improving business processes, consolidating existing information, reducing IT personnel dependency and eliminating the need to maintain product information in multiple applications. And used together with a content management solution, a PIM solution can help make complex e-commerce information more manageable.

For more information about using an IBM WebSphere solution pairing Product Information Management with a WebSphere Commerce investment, visit ibm.com/software/websphere



© Copyright IBM Corporation 2006
IBM Corporation
Software Group
Route 100
Somers, NY 10589
U.S.A.

Produced in the United States of America
02-06
All Rights Reserved

IBM, the IBM logo, and WebSphere are trademarks of International Business Machines Corporation in the United States, other countries or both.

Other company, product and service names may be trademarks or service marks of others.

References in this publication to IBM products and services do not imply that IBM intends to make them available in all countries in which IBM operates.