PERSONALIZED DATA SERVICE FOR MASTER DATA MANAGEMENT

Jinchul Han, Min-goo Lee
Prompt, Inc.110 Yangjae-Dong, Seocho-Gu Seoul, Korea

Jonghoon Chun
Department of Computer Engineering Myongji University Yongin, Kyunggi-Do, Korea

Keywords: Master Data Management, Personalized Data Service, Composite Master.

Abstract: Recently, the needs for enterprise-wide Master Data Management(MDM) have been increased drastically. MDM is the framework of processes and technologies aimed at integrating and managing the quality of disparate enterprise master data to provide consistent information across diverse users and application systems. Integrated master data environment may provide consistent, reliable, and accurate master data, yet it may not be suitable for individual data needs. Utilizing view definition of commercial DBMSs is one way to provide personalized data service, but it lacks flexibility in interoperability with other services provided by MDM systems. In this paper, we propose a new model to make the data service personalized, flexible, and efficient enough to use by physically combining underlying master data in order to meet individual data requirements.

1 INTRODUCTION

As different departments and branches have different data needs, they tend to build and use their own proprietary systems. The systems and master data used by them are suitable to their own needs thus normally maintained separately. From the enterprise-wide point of view, segregation of master data and the application systems incur duplication of master data, which will usually result in master data inconsistency problem. Master data residing in dispersed locations possibly with multiple copies obviously makes the master data management task a lot more complex, and almost impossible to streamline and execute enterprise-wide business processes and analyses. As a solution to it, MDM has long been recognized as a means to guarantee so-called "single version of truth" across an enterprise.

MDM is the framework of processes and technologies aimed at integrating and managing the quality of disparate enterprise master data, which almost always gets used by other users and systems as one and only source of reference. The main objective of MDM is to provide consistent information across diverse users and application systems by supporting sustainable framework for data standardization environment.

Data integration with MDM can provide consistent, reliable, and accurate master data to multiple data consumers within as well as outside enterprise. However, one may not be satisfied with contents and formats of master data if only a single static version of master data is enforced to be used across the enterprise. One can imagine such cases easily as people at sales departments may have different master data views and needs from people at R&D departments.

Such requirements may be fulfilled if one can implement an MDM environment in which different combination of master data are dynamically composed and feed to the individual users and systems as needed. We call that a "personalized data service," and believe that it is an indispensible feature that an intelligent MDM solution must provide in modern IT environment.

This paper presents a new model to support efficient yet flexible management and reference framework for personalized master data service. We identify requirements for personalization and devise technologies needed for it.

Han J., Lee M. and Chun J. (2009).
PERSONALIZED DATA SERVICE FOR MASTER DATA MANAGEMENT.
In *Proceedings of the International Conference on e-Business*, pages 287-293
DOI: 10.5220/0002227102870293

This paper has four sections. Section 2 introduces rudimentary concepts of master data and its management. Also in section 2, requirements for personalization are identified and contemplated. Section 3 defines methods and technologies needed for personalized data services. Section 4 concludes with summary of the paper and presents further studies.

2 MASTER DATA MANAGEMENT

2.1 Master Data

Master data is information that acts as a key to the operation of business and is the primary focus of the Information Technology (IT) discipline of MDM. This key business information may include data about products, customers, financials, materials, suppliers, employees, etc. which often are non-transactional in nature. Master data is usually needed by multiple functional groups and maintained in various systems in dispersed locations within an enterprise. Master data may not be maintained in a single central repository; therefore, the possibility exists for duplicated, inconsistent, and inaccurate master data. Thus Master Data is that persistent, non-transactional data that defines core business entities for which there is an agreed view across the organization.

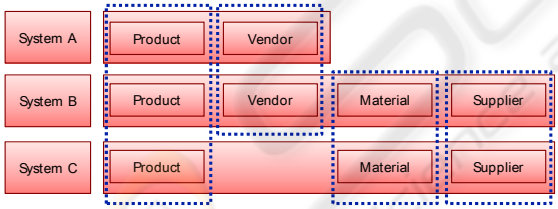

Figure 1: Master data used by multiple systems.

Figure 1 depicts a conventional system setup where each system manipulates its own master data. As seen in the figure, product, vendor, material, and supplier masters are all redundantly maintained in multiple systems, and there is no guarantee of these copies of masters would agree on each other. This is a very typical situation in modern IT system environment. In some cases, system setups like this take months to just generate a report on enterprise – wide spend analysis. Inconsistent, unreliable and bad data results in bad quality report although it comes in at a high cost, which in turn leads to a bad decision making.

2.2 Data Integration

MDM is the framework of processes and technologies aimed at integrating and managing the quality of disparate enterprise master data, which almost always gets used by other users and systems as one and only source of reference. The main objective of MDM is to provide consistent information across diverse users and application systems by supporting sustainable framework for data standardization environment.

Data integration with MDM can provide consistent, reliable, and accurate master data to multiple data consumers within as well as outside enterprise. Traditional master data integration processes are as follows.

- ✓ Analyze an individual system's master data
- ✓ Define an enterprise master data
- ✓ Define an enterprise standardization of classification hierarchy, attributes and UOM(unit of measure)
- ✓ Master data cleansing conforming to enterprise-wide data standard
- ✓ Master data integration
- ✓ Continuous master data quality management program running
- ✓ Provide personalized master data services to individual users and systems

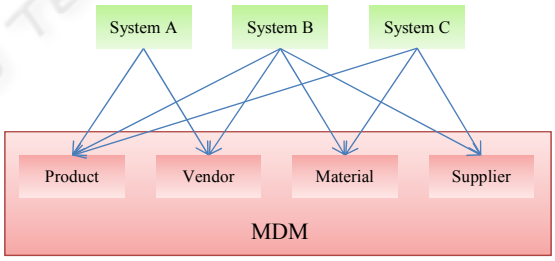

Figure 2: Data Integration.

Figure 2 shows an integrated master data management framework where individual systems can refer to. Essentially, MDM supports to share consistent master data throughout the enterprise by integrating and standardizing master data which, before, were maintained independently by individual systems.

Figure 3 is an architectural diagram of eCliX^{TM}MDM which is the commercial implementation of our master data implementation solution. eCliX^{TM}MDM consists of three layers. The interface layer is for master data interface to users and systems, the business layer is for main

functionalities of master data management including management of master data and metadata, and their quality management, and the storage layer is the repository for storing and maintaining master data and metadata.

Master data management module in the business layer performs registration/modification/deletion of master data. Metadata management module is mainly for manipulation of metadata such as classification hierarchy, attribute metadata, UOM values, and information for audit trailing. Data quality management module maintains quality of master data by utilizing the information provided by metadata. For example, if attribute metadata enumerates standard values for a particular field of master data, such information can be used to check the validity of new entry.

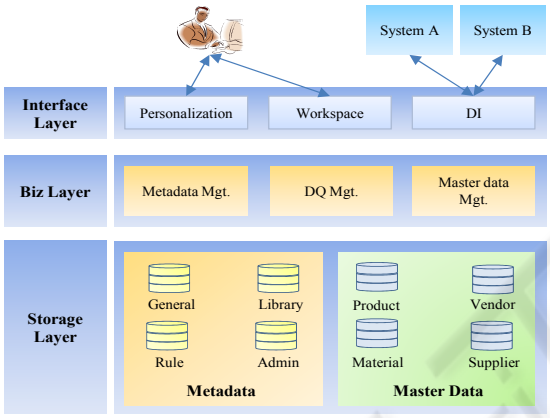

Figure 3: Master Data Management Architecture.

The first step toward enterprise-wide integrated master data management environment is to figure out master data from the entire set of data being used by various systems of the enterprise. Master data usually has distinguished characters that can be differentiated from transactional data. Table 1 shows a set of master data, namely product, vendor, material, supplier masters and their attributes maintained by individual systems. Example attributes shown in the table are very typical ones in real-life situations, as one needs careful consideration whether to include which attributes as master data attributes or not.

Generally master data are non-transient and works as a single reference point for various business processes. Any other data, that has tendency to change a lot as time passes by, are considered as transactional data. For example, 'Name' for vendor master does not change often, whereas 'Price' would change a lot. Thus 'Price' would not be included as a part of master data attributes whereas 'Name' would be. Obviously this type of design decision is purely subjective and it is up to the subject matter expert to provide sufficient information to let the designer make rational decision making.

From the implementation point of view, any data that is included in the MDM repository are master data, any others are transactional data.

Table 1: Redundant Master Data Attributes.

	Product	Vendor	Material	Supplier
System A	Name Price Date	Name Rank Price Quantity		
System B	Name Date	Name	Name Quantity	Name
System C	Name		Name Price Quantity	Name Rank

Separation of master from transactional data from table 1 results in the following.

Table 2: Enterprise Master Data.

	Master Data	Transactional Data
Product	Name, Date, Price	
Vendor	Name, Rank	Price, Quantity
Material	Name	Price, Quantity
Supplier	Name, Rank	

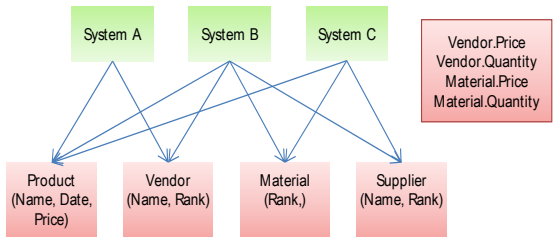

Figure 4: Enterprise Master Data.

Table 2 and figure 4 together shows a set of transactional data separated from master data. Vendor.Price, Vendor.Quantity, Material.Price, Material.Quantity are transactional data which needs to be maintained independently from master data. However, individual systems need these extra transactional data to carry out their business processes. Flexible integration capability is in need not by system integration but by data integration. We call that a "personalized data service."

In order to make personalized data service possible, the following must be considered.

- ✓ Ability to manage transactional data to later combine with master data on the fly
- ✓ Master data transformation based on classification mapping, attributes, and UOM
- ✓ Authentication, and authority control over users and systems that uses master data
- ✓ Adaptability to flexibly response to introduction of new requirements and systems

In this paper, we propose a new model to resolve aforementioned obstacles to provide flexible yet efficient personalized data service environment.

3 PERSONALIZATION

Established MDM system can provide consistent, reliable, and accurate master data to multiple data users across enterprise. However, one may need some fundamental transactional data as well as master data in order to carry out his/her business processes. For example, Vendor.Price, Vendor. Quantity, Material.Price, Material.Quantity of figure 4 are examples of fundamental transactional data, in this paper we call them additional properties of master data. These data may be transactional over all, yet they may well be very important master data to a single system. Consequently, we need to handle these transactional data differently from normal transactional data and need to derive an efficient method to integrate with enterprise-wide master data.

In this section, we propose a new framework to efficiently provide information contents in the form demanded by individual systems. In the framework, physical view generation to integrate master data with additional properties can easily be configured by the user of the system.

3.1 Composite Master

Once the initial load of standardized and integrated master data is complete, enterprise-wide master data quality management program can launch. The new program and system must not interfere with the existing individual systems which were already in operation. However, from the existing systems' point of view, additional functional requirements newly came up. One must be able to access the MDM system in order to look up and use the master data which before were locally maintained and used. The view definition supported by commercial DBMSs can be used to satisfy personalized data requirements. However, users who require master data, usually ask for additional information such as classification hierarchy as well as code information in a way that can be directly applied to their own business processes. Clearly simple view definition functionality of commercial relational DBMSs cannot support this type of requirements. To do so we propose new scheme based on the use of composite master. Composite master is essentially a personalized data view which are defined by underlying unite masters. The users and systems can dynamically create composite masters any time they need for any purposes without having to directly use bare SQL statement. By the system, any information that is inherently associated with composite master such as identification, attributes and classification information are automatically reflected on composite master.

Fig. 5 shows an example of the composite master schema. System A needed a personalized data view, which includes combination of information from product and vendor as well as some additional information. A composite master can be defined to include name, data and price from product unit maser, name and rank from vendor unit master, and price and quantity as additional properties.

Although composite masters are physically defined, actual values from unit masters are not replicated. Only their reference addresses of attributes and related information are physically maintained by composite master definition. This way, problems which may have been caused by data replication, are naturally prevented from being occurred. Additional attributes are defined and added to the composite master in the same way as other attributes are defined. Attributes are defined by separate module -- property manager -- which allows users of the MDM system to define new attributes independently from any tables.

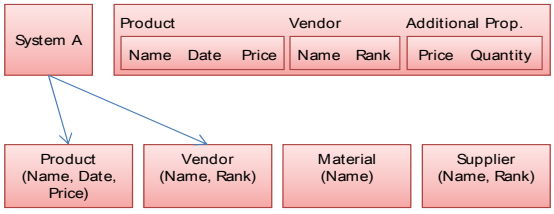

Figure 5: Example of Composite Master Schema.

3.2 Composite Manager

In order to establish personalized data service based on the use of composite master, the following requirements are identified.

- ✓ Ability to dynamically define composite master
- ✓ Ability to define additional attributes for composite master from existing set of attributes as well as by defining new attributes on the fly
- ✓ Flexible composition and integration with underlying unit master data
- ✓ New master data registration process must be coordinated with the search/reference/validation process of the existing master data
- ✓ Attribute value and format transformation of mater data in the form required by individual systems and users
- ✓ Independent from changes made to underlying unit masters, but permanent deletion of master data may need some extra work

Authentication and authority control for data governance must be supported per individual unit master

3.2.1 Composite Master Definition

Composite manager is a module responsible to define and sustain composite masters

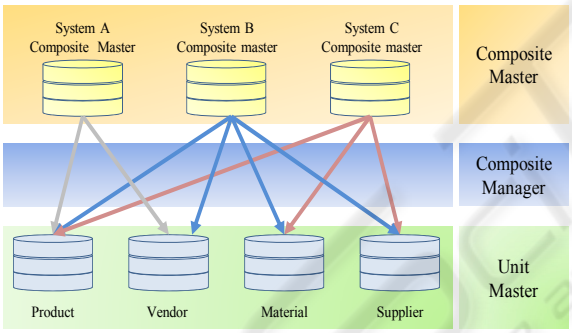

Figure 6: Composite Manager Data Model.

Figure 6 outlines the functionality of composite manager. Composite manager uses unit masters to define new composite masters. First, composite manager allows you to view and select predefined unit masters. All or some of the attributes of the selected unit masters may be used in definition of composite master. Additionally if there are some additional attributes needed while defining composite masters, one can create new attributes and use them as a part of composite master attributes. If there are some attributes previously defined by using property manager, one can select needed attributes from the existing list.

The main function of property manager is the creation of new attributes independently from any relations. Some of the attributes may be defined once and used by multiple unit and composite masters. Some of the attributes may only be defined and used as additional attributes for a single composite master. Creation of new attributes involves specification of constraints as well as some of physical aspects of it. For brevity, in this paper, we omit the rest of explanations of property manager.

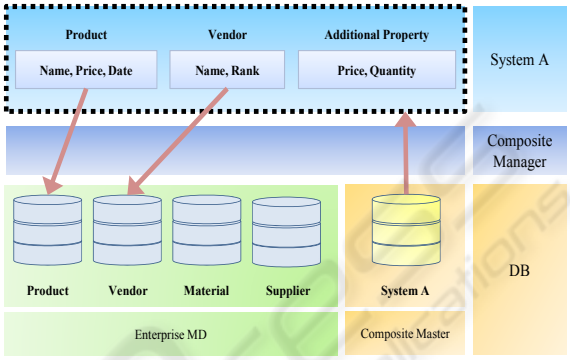

Figure 7: Definition of Composite Master.

Figure 7 shows the schema of composite master. A composite master consists of two parts – one for storing reference information of underlying unit masters, the other for storing additional attributes, and it is physically saved and maintained. While generating new composite masters, composite manager works in coordination with authentication and authority control module to block unauthorized access to composite and unit masters.

3.2.2 New Data Registration/Modification/ Deletion of Composite Master

For efficient and consistent master data management, standard procedure for new data registration to composite master is needed. A person who has access authority to composite master does not necessarily have access authority to unit masters. In other words, he/she only has the authority to register/modify/delete additional attributes of composite master, but he/she does not have the authority to register/modify/delete associated underlying unit masters.

The registration, modification, and deletion procedure for composite master are as follows.

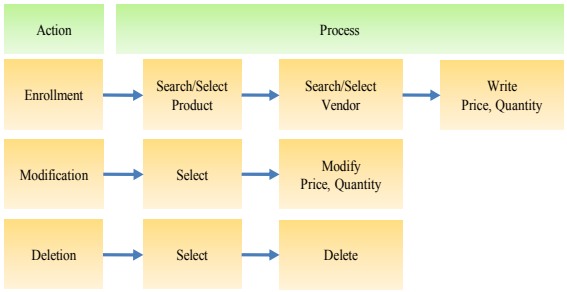

Figure 8: Registration, Modification, Deletion Procedure.

Utilizing the two layers of masters, data stewards can manage master data flexibly. A composite master only references locations of unit masters instead of maintaining copied values, it can sustain independently from any changes made to underlying unit masters. Efficient manipulation of unit master reference addresses must accompany to ensure the scalability.

3.2.3 Data Search in Composite Masters

In order for external users and systems to be able to retrieve needed data from a composite master, unit master search, transformation of the search results to the representation that the user/system needs, and dynamic composition with additional attributes are needed.

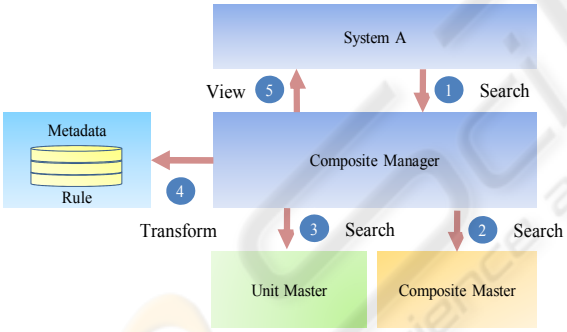

Figure 9: Composite Master Data Retrieval.

Figure 9 shows the step by step processes needed to retrieve master data from a composite master. The following scenario breaks down the typical procedures needed.

- ✓ System A submits master data retrieval requirements to composite manager (Step 1)
- ✓ Search statement submitted to execute on composite master(Step 2)
- ✓ Decompose search statements into several search statements of associated unit masters(Step 3)
- ✓ Search unit masters (Step 3 continued)
- ✓ Search for rules applicable to the representation format that system A requires (Step 4)
- ✓ Transform the result according to the rules found(Step 4)
- ✓ Dynamically combine transformed results and additional attributes(Step 5)
- ✓ Send final result to system A(Step 5)

Figure 9 merely shows the simplest case retrieval scenario. However, any information (e.g., classification hierarchy) that is related to the composite master can be retrieved effectively and efficiently in a similar manner. It is because such information is automatically inherited from the unit masters without having to have any additional operations to redundantly specify.

3.3 Functional Requirements

In order to define and maintain composite masters, the following functional and technological requirements are needed. These functions and technologies, in general, are not specific to our approach, but also necessary for implementation of master data management systems overall.

Table 3: Requirements Specification.

Technology	Functionality	Composite Mgr
CM Schema	Efficient definition of composite master	Composite Master Definition
Admin DB	Access authority & authentication definition and maintenance	Composite Master Definition
MD Searching	Keyword & class search for master data	Composite Master Register/Search
Rule DB	Storage & maintenance of transformation rule	Master Data Transformation
MD Monitor	Monitoring of master data change	Deletion Propagation

Table 3 summarizes main functions and technologies needed for composite master manipulation, and eCliX^{TM}MDM is our commercial product that implemented them.

4 CONCLUSIONS

In this paper, we investigated personalization issues in the context of enterprise-wide MDM. As for enablement of efficient personalized data service, we

proposed composite master and composite manager for management purposes. Composite masters are different from conventional view definition, in that it only maintains references to underlying physical masters. Composite masters can also be defined with additional attribues, and they provide extra flexibility to individual users and systems. New system deployment to MDM can no longer be any problem, because users of new system can define composite masters dynamically.

We believe that the concept of composite master is very important for any organization that might have some personalized data service requirements, and soon it will be realized as an indispensible part of master data definition.

Additionally, eCliX^{TM}MDM provides information retrieval capability to efficiently retrieve any information that is related to given master data. Quality management is another strong point of our prouct. Enrichment of metadata makes master data quality management process easy yet systematic. As business metadata plays an important role in MDM, enriched metadata repository and its management functionality provided by eCliX^{TM}MDM becomes the necessity for any good MDM practice.

Composite master is still implemented as a part of relational database schema physically realized. Therefore, irresponsible deletion or modification may complicate the constraints. Semi-automatic propagation of composite master deletion/ modification needs to be further studied.

ACKNOWLEDGEMENTS

This research was supported in part by the Ministry of Knowledge Economy, Korea, under the Industrial Technology Research and Development Project support program supervised by the Korea Institute of Industrial Technology Evaluation and Planning.

REFERENCES

Hyun-gu, Jo. 2007. Achieving Value through Master Data Management, Enterprise SOA Office, SAP Korea, 2007

IBM. 2008. Enterprise Master Data Management: Market Review & Forecast for 2008-12. An MDM Institute MarketPulse™ In-Depth Report

John, R. Andrew, W. 2008. The Seven Building Blocks of MDM: A Framework for Success. Gartner Master Data Management Summit.

John, R. 2008. MDM for Customer Data for Beginners. Gartner Master Data management Summit.

Sung-won, Kim. 2006. MDM Strategy for Next Generation Systems. SAMSUNG SDS Consulting Review, No.4.

A, Z. 2007. Multi-Entity Master Data Management: The 4 ᄉᆞ Generation of MDM. An MDM Institute MarketPulse™

Ted, F. 2008. Data Integration Technology and Architecture: Increasing the Value of Your Master Data. Gartner Master Data management Summit.