SPECIAL CONFERENCE ISSUE

OAUG™ oracle applications users group

insight

THE OFFICIAL PUBLICATION OF THE ORACLE APPLICATIONS USERS GROUP

SPRING 2012

Best Practices for Push & Pull Using Oracle Inventory Stock Locators

Introduction to Master Data and Master Data Management (MDM): Part 1

Utilizing Oracle Upgrade Advisor for R12

Networking: The Power Boost for Your COLLABORATE 12 Experience!

This article is the first of a two-part series. Part 1 provides an overview of what Master Data is and discusses the challenges associated with managing Master Data. Part 2 concludes the article with a look at the art and science of Master Data Management and shares insights into the latest innovations in process and technologies. Part 2 will appear in the summer 2012 *OAUG Insight* magazine.

Introduction to Master Data and Master Data Management (MDM)

(PART 1 OF 2)

By Mani Kumar Manda, Rhapsody Technologies, Inc.

As Master Data Management (MDM) applications mature to the level of suitability for enterprise deployment, many companies will want to competitively differentiate themselves with a strategic implementation of MDM solutions. Every day, **companies lose millions of dollars** due to making business decisions based on inaccurate Master Data. The ability to use **data that is free of duplications and errors** when making critical business decisions and conducting marketing campaigns is of inestimable value. This article defines Master Data, addresses the challenges of managing Master Data and shares insights into the latest innovations in process and technologies.

This article appeared in the spring 2012 issue of OAUG Insight *magazine, the official publication of the Oracle Applications Users Group (OAUG), and is reprinted with permission.*

Data Spectrum

Per the Merriam-Webster Dictionary, the word "data" is defined as "factual information (measurements or statistics) used as a basis for reasoning, discussion or calculation." The term "digital data" is used to refer to the capture, storage and movement of data using digital technologies by encoding it in numerical forms that can be easily transmitted or processed.

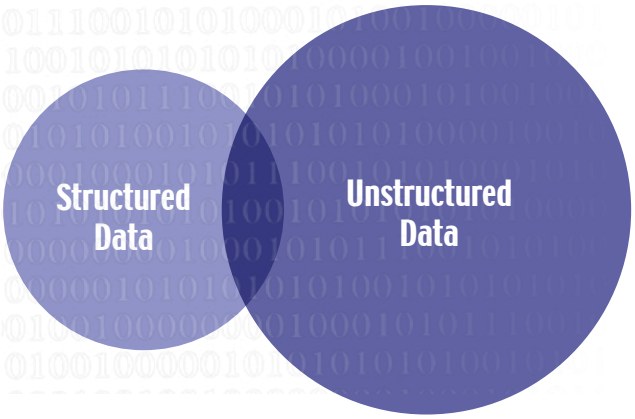
Digital data forms the foundation for modern day commerce and can be broadly classified into two categories: structured and unstructured data. For the purpose of this article, the terms "data" and "digital data" are used interchangeably.

Digital data is growing exponentially, with growth challenges being three dimensional: 1) more data is being created (amount of data), 2) data is being created faster and needs to be more quickly accessed (the velocity) and, lastly, 3) the types of data being captured is broadening (variety). The vast amounts of data being created in the digital media are unstructured and require technologies that are specific to processing and managing it. The further discussion about unstructured data is outside the scope of this article.

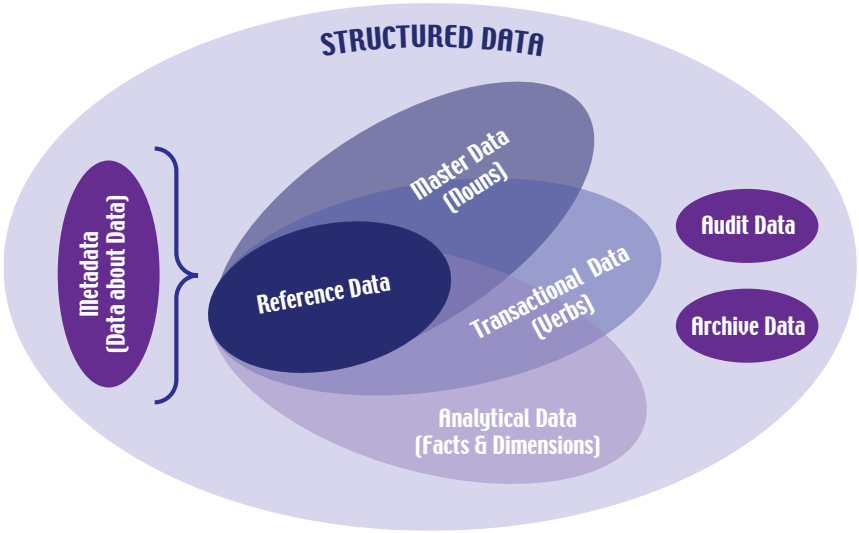
Structured data, which is addressed by MDM among other technology solutions, represents a lesser percentage of all data being created and is managed through technologies that are already well developed. This includes relational database technologies that can easily store, access and process such data.

Structured data is classified into the following categories (also shown in the illustration to the right):

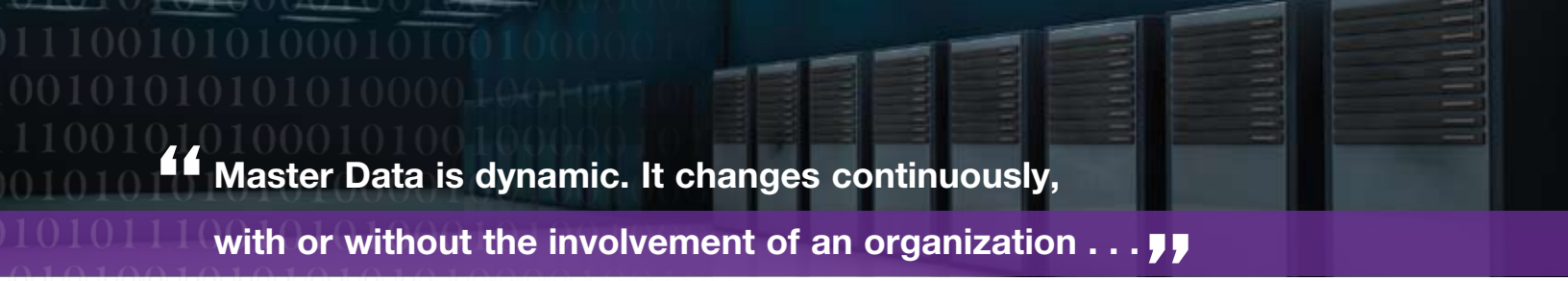
- **Reference Data**
- **Master Data**
- **Transactional Data**
- **Analytical Data**
- **Metadata**
- **Audit Data**
- **Archive Data**



Reference Data refers to the types of data that are fairly static in nature but do change over a long period of time. Reference data falls into one of four categories: a) data used as a classification scheme such as Standard Industry Classification (SIC) codes.; b) codes that are external to the business organization, such as country codes, currency codes, etc.; c) lookup codes such as "status code" using the values of active, inactive; and d) constant values such as tax rates. Reference data requires little effort to maintain, although the primary challenge in maintaining it is to ensure the accuracy of mapping across heterogeneous applications and incorporating it in all integrations in a timely fashion when new reference data is created or existing reference data is changed.



Reference Data Examples: Lookup Codes	Master Data Examples: Customer, Product
Transactional Data Examples: Orders, Invoices, Leads	Analytical Data Examples: Facts, Cubes, Dimensions



> "Master Data is dynamic. It changes continuously, with or without the involvement of an organization . . ."

Master Data is similar to Reference Data, but changes gradually over time, due to both external and internal factors to an organization. Master Data usually refers to nouns such as customers, employees, products, suppliers, etc. Organizations need to devote resources, both in people and technology, with the help of well-defined processes in order to keep this data at a consistent and accurate level for organizational users to consume it effectively. Master Data uses Reference Data.

Transactional Data, commonly associated with verbs and considered the "record of an event," refers to data that captures information about a specific event such as an order, an invoice, a lead, etc. Transactional Data uses both Master Data and Reference Data.

Analytical Data refers to data structures that are optimized for analytical purposes to infer knowledge based on massive amounts of data in shorter periods of time. Data for analytical purposes is stored into facts and dimensions, with data models following principles of star schema or snow flake schema to meet performance objectives. Analytical Data uses all three of the data types defined above: Transactional Data, Master Data and Reference Data.

Metadata refers to the data that describes other data.

Audit Data refers to data that describes changes made to existing data, such as who performed the change, when the change was made and in which system the change occurred. Audit data is captured and stored for the purpose of capturing data changes and to prevent or identify root causes of data changes when a data change impacts the business in some form and fashion.

Archive Data refers to the storage of all structured data that is presently not needed but is required to be accessible for the purpose of compliance with audit and regulatory reasons. Data is also archived in order to improve the efficiency in terms of performance of operational applications. How long data is archived varies greatly by industry and ranges anywhere from seven to more than 30 years for firms in the insurance industry.

Let's take a closer look at Master Data.

Master Data Defined

Master Data, previously defined as nouns, refers to the type of data that has a long life span and changes gradually over time. Master Data can be classified into four categories, as shown in the chart below. Each type of entity under these categories is known as a domain.

MASTER DATA CATEGORIES & DOMAINS (OR ENTITIES)

PARTIES	THINGS	PLACES	CONCEPTS
Examples: • Customer • Supplier • Employee • Citizen • Partner • Bank	Examples: • Product • Part • SKU • Service • Asset	Examples: • Facility • Location • Geography	Examples: • Chart of Account • Contract • License • Policy

Parties refer to entities such as individuals and organizations that are involved as trading partners in conducting business. The Parties can be employees, customers, suppliers, citizens, etc. The management of customer data is known as Customer Data Integration (CDI).

Things refer to items such as products, parts, SKUs, etc., that are used in commerce. The management of product data is known as Product Information Management (PIM).

Places refer to data entities such as facilities, locations and structures that are tied to a physical address in a geographic space and the maintenance of all their associated information.

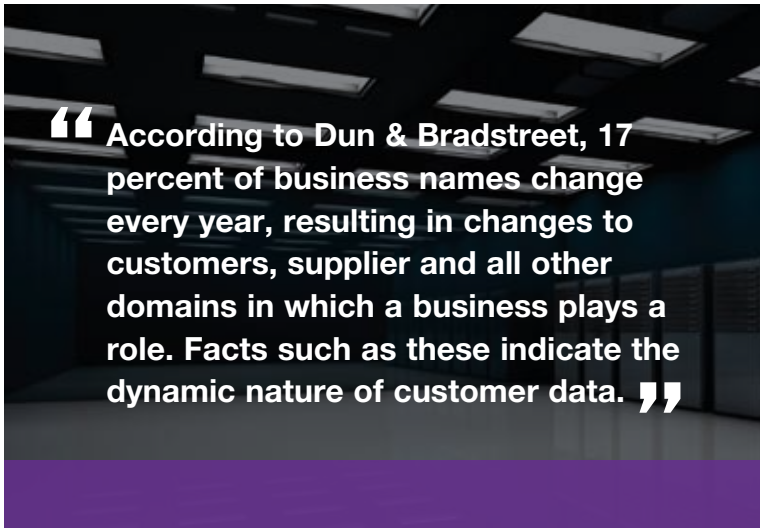
Concepts refer to data entities such as chart of accounts and contracts that are a means to group/classify to encapsulate pertinent data. For example, chart of accounts are used to group journal entries as per accounting principles. Contracts are used to capture a business relationship in contractual terms that govern the business interaction.

What Master Data domains are important to address and the relative criticality of each domain to an organization depends on the industry they are in and the types of products and/or services these organizations offer. For example, for an organization that builds and sells airplanes, given the number of parts and subassemblies used in each of these planes (often gets into million plus parts) and the revisions that are made to these planes and the subassemblies used over the years, the solutions to address both product Master Data and supplier data become more critical. Whereas, for distribution companies and organizations that operate in the B2C (consumer) space, the focus on customer data is more important, with some of them also needing solutions to address product Master Data.

The Challenge

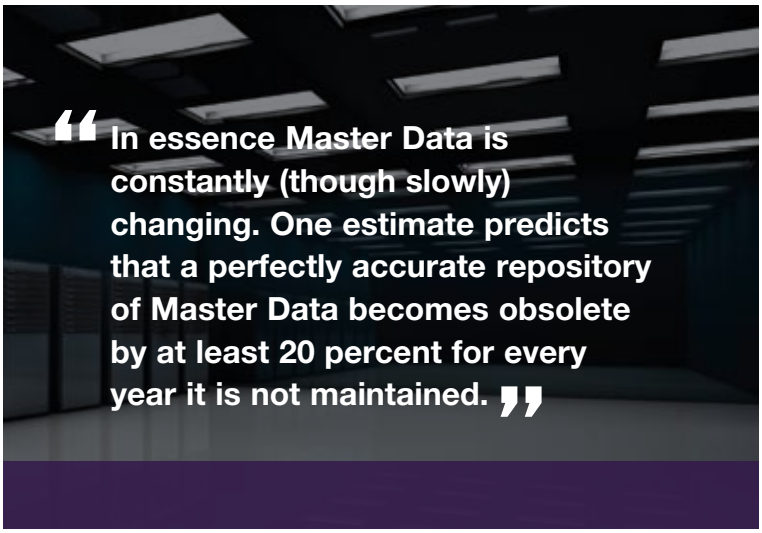
Master Data is dynamic. It changes continuously, with or without the involvement of an organization, and is universal across all domains of Master Data belonging to the four categories discussed.

According to Dun & Bradstreet, 17 percent of business names change every year, resulting in changes to customers, supplier and all other domains in which a business plays a role. Per the United States Postal Service (USPS), about 14 percent of the population moves yearly, resulting in address changes for suppliers, customers, employees, citizens, etc. Facts such as these indicate the dynamic nature of customer data. Most of these changes often are not communicated to organizations. Businesses are, therefore, left alone to devise their own solutions and approaches to obtain the most accurate data, sometimes using sources other than the customer itself. Many other challenges are tied to the management of party data.



The nature of product (things) data is such that most products are uniquely identified by their characteristics (which often form the basis for generating item descriptions). Though a unique identifier (item number) makes it appear that there are no duplicate products in the database, the reality is often different. It is frequently discovered that two or more products

defined with unique identifiers indeed are the same products. And multiple products are often created by different users, with the characteristics represented in the description in a unique manner (descriptions created with representation of selected characteristics in different order and/or using varying abbreviations) making it difficult to identify these products as duplicates. This scenario makes it very difficult to identify duplicates since the very nature of identifying duplicates cannot be easily done using pattern matching technologies. Identifying duplicates in product data requires a technology different from pattern matching, and semantics-based matching technologies had great success in matching products for many industries.



Similar challenges exist for other domains of Master Data, including data in the realm of sites or concepts categories.

Another challenge with Master Data is that it is often scattered across the organization, housed in various applications that are used by a specific business line, business function, purpose or geography. It would take a tremendous effort to compile a full portrait of the Master Data of an instance, e.g., specific customer, specific supplier, specific product, specific site, etc. This task requires first matching similar data across a multitude of applications then comparing the data where similar information exists in multiple applications and identifying the correct value when conflicts exist. This is easier said than done. The reality is that changes made in one application often are not reflected in other IT applications. The task of operating the business, which is based largely on insights gained from information for which Master Data forms the basis, is error prone because Master Data is often not uniform across the organization.

In essence Master Data is constantly (though slowly) changing. One estimate predicts that a perfectly accurate repository of Master Data becomes obsolete by at least 20 percent for every year it is not maintained. Other measures peg the change in Master Data at much higher rates, even approaching two percent per month. Compounded, this monthly change becomes a change of 27 percent of data in one year, a change of 64 percent in two years and a whopping 104 percent change over three years.

What should organizations do to address these challenges?

Watch for the conclusion of this article in the summer 2012 issue of *OAUG Insight* magazine.

Mani Kumar Manda is founder of Rhapsody Technologies, Inc. and founder/coordinator for the OAUG Customer Data Management Special Interest Group (OAUG CDM SIG). Mani has been a consultant for his entire career to mid-market and large customers and is a recognized Expert and Speaker in the areas of MDM with numerous presentations at OAUG, Open World, MDM Summit and several local GEO and SIG groups. Mani has developed a proprietary RHYTHM Methodology that significantly increases the "success factor" for all MDM implementations. Mani can be reached at mmanda@rhaptech.com.

Acknowledgments: *Special thanks goes to Eugene Breger for reviewing and providing changes that enhanced the quality of this version of the article.*

> “The reality is that changes made in one application often are not reflected in other IT applications.”

REFERENCES

1. Customer Modeling: It’s your call between Rock and Hard Place, *http://www.slideshare.net/RhapsodyTechnologies/customer-modeling-its-your-call-between-rock-and-hard-place.*
2. Oracle CDH – the past (R11i), the present (R12), and the Future (Fusion MDM), *http://www.slideshare.net/RhapsodyTechnologies/oracle-cdh-the-past-11i-the-present-r12-and-the-future-fusion.*
3. Oracle TCA 101 (with audio track), *http://www.slideshare.net/RhapsodyTechnologies/oracle-tca-101.*
4. MDM – Oracle Site Hub 101, *http://www.slideshare.net/RhapsodyTechnologies/mdm-oracle-site-hub-101-1419759.*
5. Site/Location Hubs – A Hot Trend in Master Data Management (MDM), *http://www.slideshare.net/RhapsodyTechnologies/sitelocation-hubs-a-hot-trend-in-master-data-management-mdm.*
6. R12 features and functionality in TCA and CDH/EBS, *http://www.slideshare.net/RhapsodyTechnologies/r12-features-and-functionality-in-tca-and-cdhebs.*
7. Oracle Trading Community Best Practices Setting Up Customer and Prospect Data, Search Oracle’s support site with Note# 269124.1; or *https://support.oracle.com/cgi-bin/cr/getfile.cgi?p_attid=269124.1:2.*
8. To receive an Oracle CDH Poster, send an email to “OracleMDMPoster at RhapTech DOT com.”

Websites and Discussion Forums

OAUG CDM SIG	*http://groups.yahoo.com/group/cdmsig* *http://cdmsig.oaug.org*
Rhapsody Technologies, Inc.	*http://www.rhaptech.com* *http://www.rhaptech.com/resources.html*
Product Data Quality	*http://tinyurl.com/29xocwv*
Supplier Life Cycle Management (SLM) and Supplier Data Hub (SDH)	*http://tinyurl.com/2v2yyct*
Oracle Fusion Applications	*http://tinyurl.com/3232emw*