Mastering Big Data: The Next Big Leap for Master Data Management

By leveraging big data, organizations can provide a true 360-degree view of critical master data entities; meanwhile, master data can help convert information gleaned from big data sources into actionable insight.

Master data management (MDM) is critical to the success of many big data analytics projects. Why? Because it encompasses people, processes and technologies that ensure the accuracy, uniformity and uniqueness of data entities shared by multiple business units within an organization. (See our white paper “Innovations in MDM Implementation: Success via a Boxed Approach” for additional insights.) Big data, on the other hand, refers to the large quantities of human- or machine-generated bits and bytes that are generated at high speeds from multiple sources, including not only posts and comments on social media but also images, videos, PDFs, RFID data, point of sale data, call center comments, data generated by enterprise systems and more. A big data analytics strategy involves collecting this data in order to create meaningful information.

All this data, in various shapes, sizes and forms, represents valuable information that needs to be standardized and stored in a meaningful way within a central repository accessible by all business units. MDM provides a trusted, central point of storage and access to this information. This whitepaper describes how master data management and big data analysis need to work hand-in-hand to provide a 360-degree view of an organization’s critical data entities. (While this paper is largely focused on social data and customer MDM, there are many other applications of big data not explored here.)

The Relationship Between Big Data and MDM

MDM provides a meaningful context to big data and a panoramic view of critical data entities, such as product, customer, employee, agent, location, etc. (see Figure 1, next page). With all this information stored in a central trusted repository, all business units have access to a single version of the truth, on demand. In the absence of such a master data hub, records are typically stored in silos, such as loyalty systems, CRM systems, POS systems, Web portals and big data repositories. Such systems usually have their own version of the truth, causing headaches that ripple across the organization, from substandard customer service and bad business intelligence, through impaired operational efficiency (see Figure 2, next page).

While MDM is crucial to the success of many big data projects, it is important to note that not all big data initiatives require an MDM solution. MDM is needed only for big data initiatives that involve
master data entities, such as customer, employee, agent, product, location, etc.

The Scope of Big Data and MDM
When data is extracted from most big data sources, it contains two parts. The first part is entity attribute-level data, or data that can be used to identify an entity. The other is transaction-level data that is generated by each entity. For instance, if the big data source is Facebook, the two types of data include:

- **Attribute-level data**: Facebook ID, phone number, name, address, etc.
- **Transaction-level data**: Conversations, comments, likes, etc.

The big data extraction layer will pass the attribute-level data to the MDM system and the transaction-level data to the big data hub, to be analyzed by a big data analysis tool. After analysis, the brand advocate score, purchase intent, etc. will be passed to the MDM system to create the 360-degree view of the customer.

To help define the scope of big data, consider the following:

- MDM will act as a repository of the following information types obtained from big data sources:
  - Attribute-level data.
  - Information derived from transaction-level data.

With or Without MDM

The big data extraction layer will pass the attribute-level data to the MDM system and the transaction-level data to the big data hub, to be analyzed by a big data analysis tool. After analysis, the brand advocate score, purchase intent, etc. will be passed to the MDM system to create the 360-degree view of the customer.

To help define the scope of big data, consider the following:

- MDM will act as a repository of the following information types obtained from big data sources:
  - Attribute-level data.
  - Information derived from transaction-level data.
• MDM will cleanse, standardize and de-duplicate this entity-level data and maintain audit trails and hierarchies.
• MDM will not store entity-level transactional data, such as conversations, transactions, etc.
• MDM will not analyze big data.

In the absence of an MDM system, both attribute-level and transaction-level data can be stored in big data hubs. But such data stores are meant for heterogeneous, unstructured or semi-structured data that is intended for analytics purposes; hence, they are not optimal for managing shared enterprise master data entities. Additionally:

• Data models in big data hubs are optimized for analytics, while MDM data models are optimized for entity resolution.
• A big data hub stores all information gathered from a big data source; however, these records are neither unique nor cleansed and rectified. Hence, finding an authoritative golden copy can be difficult and time-consuming.
• The sheer volume of data in the big data hub will not allow a quick search for entity-level attributes for customer service or decision-making purposes.
• Big data hubs are built for a specific purpose—big data analysis. Hence, strong entitlement and hierarchy management are not included.

MDM for Big Data Analytics

To handle the demands of big data, next-generation MDM systems must successfully handle the volume, speed and formats across all data types. Required fundamental changes to MDM include:

• A data model to store big data-specific attributes, such as social media IDs, opt-ins, brand advocate scores, social network hierarchies, product preferences, purchase intent, churn possibility, etc. As previously mentioned, some of these attributes will be fed into the MDM system from big data analytics systems; others will be fed directly from the source.
• Improved MDM integration architectures to handle the speed at which this data is both added and updated.
• Advanced match and merge processes, especially for social analytics, which is among the most widely applied use cases in big data. The biggest challenge with social analytics is successfully identifying the social profiles of an organization’s existing customers. This is especially true since multiple results can be returned when searching for common names. Also, many people use aliases rather than real names in their social profiles, making it even more difficult to identify these profiles.
• Advanced data governance policies involve rules around not only data usage, ownership, etc. but also around big data-specific data governance and data stewardship. MDM
needs to enforce policies and procedures that ensure:

➢ External data does not affect the integrity of the internal data.
➢ Data is relevant for the purpose for which it is mined.
➢ Escalation paths exist for data conflicts.
➢ Privacy policies are enforced.

When Big Data and MDM Combine to Achieve a Business Result

**Airlines**

With enormous customer acquisition costs, customer centricity is the key to profitability in this industry, which makes airline carriers ideal candidates for big data programs.

- **Problem statement:** An airline wants to send targeted promotions to brand advocates and influencers to increase coupon redemption and positive word of mouth.

- **Solution:**
  ➢ Route a social media campaign through the airline’s loyalty portal to capture customers’ social media presence.
  ➢ Analyze big data to identify brand advocates and influencers among these customers and tag them appropriately in the MDM hub.
  ➢ A downstream campaign management system can readily obtain this data from the MDM hub to identify brand advocates among an organization’s customers and – based on the brand advocates’ preferences, opt-ins, etc. – send targeted promotions to them.

**Multibrand Retailers**

Due to cost pressures, severe competition and perishability of certain products, couponing is among the most commonly used marketing tactics in retail. Targeted couponing enabled by big data analytics would help retailers improve coupon redemption rates and increase profitability.

- **Problem statement:** A retailer wants to identify heavy users of certain products and send them targeted coupons to increase inventory clearance.

- **Solution:**
  ➢ Implement big data analysis software to analyze point of sale data, purchase history and data from social media (Polyvore, Pinterest, Facebook, etc.) and create a complete customer profile in the customer hub.
  ➢ Analyze data in the customer hub in real time to recommend products, coupons, etc. at the time of checkout or as part of marketing campaigns that send targeted coupons and promotions to these customers.

**Fashion Retailers**

With rapidly shifting styles, fashion retailers need to update their product assortment continuously. Garments that are in vogue one season can go out of style the next, which means retailers must clear their racks using discount pricing strategies. The longer that old merchandise sits in the warehouse, the more it loses value. So it’s imperative for retailers to sell their merchandise as quickly as possible.

- **Problem statement:** A retailer wants to build a complete profile of all customers so it can recommend ideal products and discount offers, at the right time and place, to improve customer service and profitability.

- **Solution:**
  ➢ Implement big data analysis software to analyze point of sale data, purchase history and data from social media (Polyvore, Pinterest, Facebook, etc.) and create a complete customer profile in the customer hub.
  ➢ Analyze data in the customer hub in real time to recommend products, coupons, etc. at the time of checkout or as part of marketing campaigns that send targeted coupons and promotions to these customers.

**Satellite TV and Cable Providers**

These providers can analyze customer channel-surfing behavior for many purposes.

- **Problem statement:** A satellite TV/cable provider wants to show targeted ads to customers to increase advertiser ROI.

- **Solution:**
  ➢ Analyze individual customer channel-surfing behavior to understand which customers are likely to watch which types of ads during which shows.
  ➢ Enrich customer data in the MDM hub based on this intelligence.
  ➢ Use downstream ad placement software to obtain channel-surfing information, along with viewer household information, opt-ins, etc. from the MDM hub on-demand, and based on that insight, target ads to viewers.
Multiproduct Manufacturing Companies

Such companies must manage a large number of suppliers and have a 360-degree view of them via a central repository. This is important not only for supplier selection purposes but also for regulatory compliance.

• **Problem statement:** A multiproduct manufacturing company seeks a 360-degree view of its global supplier base for quick and efficient supplier selection.

• **Solution:**
  > Analyze available information about each supplier via the media, the Web and from notes transcribed by procurement managers during transactions.
  > Enrich supplier data based on this analysis to build a 360-degree view of each supplier, including violations cited in the news, delay in past shipments, etc.

Automobile Manufacturers

Car makers tend to spend millions of dollars each year on market research for product line extensions or feature enhancements. Big data can be analyzed to understand what customers, partners and competitors think about a car and more.

• **Problem statement:** In an attempt to expand its business, a car manufacturer seeks to aggregate all feature-related commentary about its models and use these insights to shape future vehicles.

• **Solution:**
  > Analyze social media commentary, calls to the call center related to each model and data from automobile monitoring systems, if available.
  > Enrich the data about each car in the product information management system, including features to be added in the next model, so that all information about the vehicle resides in the same place.

Critical Move-Forward Considerations

While big data is already the next big thing in business-technology, and holds the key to achieving the next level of customer centricity, organizations that want to make sense of huge, proliferating pools of structured, unstructured and semi-structured data need to keep the following in mind:

• A big data strategy should be a companywide initiative, with complete buy-in from senior leaders.

• Big data will impact not only MDM but almost all enterprise systems in any size company (i.e., business intelligence systems, data warehouses, ERP, CRM, campaign management, etc.). Therefore, a proper IT integration strategy is crucial.

• It is important to establish a clear understanding of the expectations from big data analytics. The social Web holds terabytes of data for most brands; acquiring, storing and attempting to mine all of this data is not only impossible but also unnecessary.

• Finally, brands need to maintain a balance between targeted promotions that consumers find useful vs. targeted promotions that consumers may find intrusive or an encroachment on their privacy.

Final Thoughts

In this era of customer centricity, superior customer service and operational efficiency are two major differentiators that can set a company apart from its competitors. Big data analytics can help improve customer service, develop operational efficiencies and increase profitability. However, big data analytics cannot provide actionable insights unless the data and related information derived from big data sources can be traced back to the appropriate entity. Therefore, to get the full value from big data initiatives, MDM is not just an option but a critical necessity.

About the Author

Reshmi Nath is a Business Consultant within Cognizant’s Customer Solutions Practice, focused on master data management. She has seven-plus years of experience in the technology and consulting industries in the U.S. and India. Reshmi earned her M.B.A. in marketing and finance at Michigan State University and a bachelor’s degree in computer science at Bemidji State University. She can be reached at Reshmi.Nath@cognizant.com.