Business-Oriented Data Governance for Effective Master Data Management

Prepared by:

David Loshin
President,
Knowledge Integrity, Inc.
(301) 754-6350
loshin@knowledge-integrity.com

Sponsored by:
Semarchy
450 Townsend St, Suite 100
San Francisco, CA 94107
www.semarchy.com

The Data Governance Imperative

The explosive growth of the information economy has motivated a large number of organizations to rethink the way they ingest, manage, consume, and transform data into actionable information and intelligence. Increased oversight is rapidly becoming more pertinent as data is increasingly promoted as the emerging source of value. This is due to a number of key factors:

- Traditional enterprise data sources were intimately tied to the output of transaction processing. However, today's analysts demand access to a wider variety of organizational data beyond histories of sales transactions to include byproducts of operational processes (such as manufacturing and production, product quality, logistics, web logs) as well as alternate sources of corporate history including archived emails and documents, messages.
- Data volumes are growing exponentially.
- There is a growing desire to take advantage of different types of data artifacts beyond traditional structured data formats.
- Internal data sets are increasingly blended with externally provided data sources, including open data sets (such as those produced by government agencies), historical archives of documents, as well as streaming social media data from a variety of sources, with much of the content being unstructured.

To use the potentially massive volumes of data for productive purposes for business value improvement there is a need to **ingest, capture, prepare, manage,** and **analyze** that data in a way that ensures predictable and acceptable levels of data fidelity and usability. And while data management practitioners have long understood the value of data quality management, the C-level managers are also rapidly recognizing **data governance** as a "must-have" for the enterprise. It is now rare to see any big data, data quality, or master data management (MDM) initiative that is not combined with the institution of a data governance program.

The conventional wisdom on data governance proposes hierarchies, operating models, and processes for data policy definition and implementation. Unfortunately, poorly-designed and minimally-planned data governance processes are fundamentally ineffective because they are viewed as bureaucratic and overwhelming. This is especially true when they are imposed by fiat, take a long time, and don't result in any short-term improvement in information value.

Yet proper data governance is a critical success factor for master data management. In this paper, we examine the motivations for coupling data governance with master data management and consider how to evolve data policies and processes to position master data management for success.

Motivating Master Data Management in the Modern Environment

The technology fabric of most organizations is a quilt patched together from siloed systems originally designed within individual business functions in isolation to meet specific operational needs. Yet it has become clear that end-to-end business processes (such as "order-to-cash") will cross business functions with application touch points requiring pervasive accessibility to data about a range of critical master data business entities such as customers, vendors, suppliers, sites, employees, etc.

In essence, an organization's interactions with these key players revolve around intersecting lifecycles. Successfully executing the processes and lifecycle demand high-quality access to a composite view of each customer or vendor that is not only of high quality, but is also synchronized across the stages of those lifecycles.

Master data management (MDM) methods and technologies meet these needs by presenting a unified view of data associated with uniquely-identified entities. And in the modern environment, data about these entities may appear in many shapes and forms, necessitating data governance for more introspection and oversight to ensure proper data management.

Why the Conventional Approach to Data Governance is Deficient

The objectives of data governance include defining data policies, instituting the proper procedures to ensure that organizational data sets are suitably usable for their intended purposes, and ensuring that those procedures are followed to ensure policy compliance. However, the conventional approach to data governance centers more on the organizational structure for instituting data policies and defining processes, and less on practical methods for socializing compliance with those policies. The organizational structure encompasses the creation of a data governance council and the designation of data stewards, with the intent of defining and agreeing to corporate data policies. To complement the roles of the data stewards, the data governance council may create additional individual roles that report into the organizational hierarchy, including data quality analysts, data standards analysts, and metadata analysts.

However, the ways that this approach is typically implemented pose challenges to its effectiveness in improving the quality and usability of corporate information. First, data governance is typically organized as a practice defined outside of any specific business context. The perception of data governance as a "shared service" with no business backbone behind it will severely limit its acceptance across the organization. Second, the practice is often launched before any success criteria are defined. This prevents the data governance champions from being able to effectively articulate governance priorities, implying to the rest of the organization that data governance is not much more than an unnecessary bureaucracy.

Redefining Data Governance...

Under these circumstances, the outcome is bound to be unacceptable when there are difficulties in defining, agreeing to, operationalizing, and enforcing data policies. Yet standards and compliance with data policies are critical when managing the composite views of key master data domains like **customer** and **product**. If the conventional approach to data governance is inadequate to meet the needs of the enterprise, the concept of data governance must be redefined from the data consumer's perspective.

... From the Consumer's Perspective

Business-oriented data governance ensures that data consumers have access to the data they need, the metrics of quality and usability of the data are aligned with defined business goals, and there are limited semantic conflicts about meaning and interpretation of data concepts. Because this redefinition augments the organizational structure with directives driving specification of policies that are directly related to business goals, there is a greater opportunity for socializing the benefits of compliance with defined data policies. In essence, business-oriented data governance highlights the importance of data policies by tying their observance to clear business benefits.

This approach integrates additional engagement with the data consumers to better identify what data sets the different business processes use and understand the scenarios in which definitions and specifications for data from different business processes may conflict. This provides the context for defining data policies that are aligned with the right business contexts to ensure that business policies are observed.

Examples

The specification of data policies and governance procedures is framed by business processes requiring high-quality data. Some examples include:

- **Product catalog management for revenue generation** – Item manufacturers often sell their products through third-parties, such as electrical supplies that are sold through eCommerce channels managed by big box hardware retailers. The eCommerce retailers will sell parts from different manufacturers, so the objective of each supplier is to provide part description data that is accurate and complete to ensure that when a customer visits the web site and performs a search of a product with particular characteristics, their items bubble up to the top of the result list. Ensuring high quality part data requires data governance policies for ensuring the quality of data associated with uniquely-identifiable products. In the end, it is still difficult to sell a high-quality product when it is accompanied by mediocre description data .
- **Regulatory compliance** – Safeguarding protected health information (PHI) is a key directive of the HIPAA Privacy Rule. This was reinforced by the HITECH Act, which states that any breach of a healthcare provider's or healthcare payer's environment

that affects more than 500 individuals must be reported to the Secretary of the Department of Health and Humans Services. These violations not only result in assessed fines, they also incur significant warranty costs to ensure those individuals whose data was exposed be protected from future identity theft or other types of fraud. Enterprises need data governance, data controls and ongoing validation of compliance. This can help support the ability to protect patient data and reduce the costs and negative impacts of breach notification.

Other business drivers include a broad range of performance improvement directives such as cost reductions, reduction and prevention of unscheduled downtime, increasing production quality, improving communications networks, reducing customer attrition, etc.

Implications of Business-Oriented Data Governance for Master Data Management

The goal of MDM is ensuring that the data representations are available that enable data consumers in different business contexts to have consistent accessibility to a shared composite view of uniquely identifiable entities. The practical implications from a data governance perspective suggest some key ideas to keep in mind:

- **Align with the business**: Identify corporate business objectives and determine what business policies govern meeting business goals.
- **Identify the business contexts**: Identify the different business contexts for consumption of master domain data and how information policies oversee consistent data use across the different processes.
- **Engage the data consumers**: Establish processes to engage business users to assess their information expectations in relation to business policies. Develop templates for transforming those expectations into business rules.
- **Maintain focus**: Concentrate on defining those data policies whose observance will contribute to complying with corporate objectives.
- **Plan for data architecture**: Isolate key master entity types necessary to meet business needs and specify usability criteria for their critical attributes.

Awareness of these ideas can inform the data governance team in devising procedures and processes for ensuring the usability of the master data assets. The set of processes for governing master data includes:

- **Semantic metadata inventory**: Assessing contextual definitions and cataloging data specifications for the business contexts from each user's perspective.
- **Harmonization**: Determining where there are conflicting business contexts and semantics and then harmonizing definitions, reference data, entity models
- **Distinction**: Differentiating between inherent characteristics associated with each entity (such as a product's weight) and its imposed attributes (such as a customer's purchasing profile).

- **Ontologies and taxonomies**: Devising contextual ontologies that distinguish between the entity and the roles that entity may play in the different business contexts.

These processes will improve your data architects' abilities to design master data models that will support the business objectives.

The Business-Oriented Approach Help to Overcome the Challenges of Deploying Data Governance

The business-oriented approach to data governance helps to overcome the challenges of deployment. Instead of a siloed data governance hierarchy defining policies outside of a business context, stewardship and data quality management activities are specifically aligned with the needs of business owners. In turn, when data governance is defined within specific business contexts, the corresponding data stewardship procedures can be directly integrated into business processes.

The alignment becomes more tightly-coupled when data governance priorities are driven by business priorities. Clear success criteria will be defined in terms of achieving business goals. To revisit one of the aforementioned examples, increased observation of data quality and completeness policies for part and product data should result in greater precision and accuracy in third-party retailer catalog "hits," which will reflect increased product sales for those highly-governed product data assets.

Defining data policies in relation to a business context simplifies the agreement process while reducing the effort necessary for operationalization and enforcement. And when it is obvious that complying with data policies leads to improved business outcomes, instead of being seen as unnecessary bureaucracy, governance and stewardship oversight are recognized as providing value.

Considerations for Master Data Management

The business-oriented approach implies that the collection, integration, and presentation of shared master data to support the achievement of clearly-defined business directives requires thoughtful data policies throughout the design and development lifecycle. There are many MDM products in the market, but you must consider how those different products facilitate business-oriented data governance. Specific recommendations for developing and implementing master data management should ensure that your choice of MDM product be aligned with a proper business-oriented approach to data governance. Therefore, when considering master data management technologies and vendors, keep these ideas in mind:

- *Let the business drive your master model design*. While using data models bundled with MDM packages, remember that those models are designed to support a summarized perception of what the vendor thinks its customers need. Initial

designs may be motivated using vendor data models, but as your environment requires greater sophistication, avoid overreliance on vendor-supplied models. Instead, look for a product that lets you use your own business process architecture to motivate data architecture and model design for master entities.

- *Don't create new dependent data silos*. Rethink the concepts of "single source of truth" or "golden record." Because the interpretation of the different attributes within each business function may differ, you must be cautious when attempting to create a new "master" repository whose records are created from consolidating numerous sources into a presumed high-quality record. The objective of MDM is to provide access to composite information about uniquely-identified entities, so look for a product that provides different alternatives for materializing a master index.
- *Develop data models that are adaptable to different business function ontologies*. External business conditions are always morphing, exposing new opportunities to use existing master data. However, complex master data models that are tightly bound to specific meanings may constrain the agility that is demanded by the rapidly changing environment. It would be unfortunate if the limitations of the data model prevent you from taking advantage of emerging opportunities within a reasonable window. Look for a product that enables flexibility in model design that simultaneously ensures consistency in interpretation across business functions.
- *Evolve the attribution of shared data models as new applications are brought online*. Notwithstanding the need to support multiple business functions, attempting to develop master domains that are all things to all users will result in models that are never completed. Instead, prioritize your master data domain requirements based on the business drivers, take your best cut at designing the models and the master data services, and iteratively refine your environment as new applications are ready to integrate with the master data environment.
- *Focus on core master data services*. Governed master data requires that certain capabilities are provided, including
    - Identity resolution, or the ability to match sets of records together that represent the same real-world entity. This is critical to enable the creation of the master index that maps entity attributes to the composite view.
    - Federated access, allowing simultaneous query access to the records that are linked together in their original source. Whether you choose to create a segregated master repository or maintain a dynamic index, maintaining the data lineage and ensuring that quality business rules are applied are key governance directives that must be observed.
    - Composite views of shared data, to support different presentations of the master records depending on the data consumer expectations.
    - Application of business rules for presentation of master data that are defined by the data consumers based on their process requirements.

Conclusions

Adopting the business-oriented approach to data governance can streamline the introduction of data governance at a high level while simplifying the operational aspects of improving data quality, implementing master data management, and improving reporting and analytics, especially as data volumes grow and become more diverse. The business-oriented approach shares many of the same goals of conventional governance programs, but its focus is directed at creating and enhancing corporate value that can be derived from information.

Alignment with the business motivates collaboration and policy compliance while reducing resistance to adoption, and this helps ensure that the governance policies are effective and successful. The next step is to consider where impediments to successful data governance are caused by rote adoption of conventional data governance techniques and to tweak the existing structures and adjust the operating model with the business-oriented principles in mind.

About the Author

David Loshin, president of Knowledge Integrity, Inc, (www.knowledge-integrity.com), is a recognized thought leader and expert consultant in the areas of data quality, master data management, and business intelligence. David is a prolific author regarding BI best practices, with numerous article published at searchBusinessAnalytics.techtarget.com, and numerous books and papers on Big Data, Analytics, Business Intelligence, Data Warehousing, data governance, and data quality. Visit http://dataqualitybook.com for more of his insights. David can be reached at loshin@knowledge-integrity.com

About the Sponsor

Semarchy provides master data value in weeks not months, using an agile, value-driven approach to design and manage master and reference data. Our unique approach creates value for the business quickly and cost effectively, supports all domains within a single environment, adapts to evolving business requirements, and allows deployments in any environment with any style of MDM and RDM. Results can be trusted with complete visibility to all underlying data and lineage. Semarchy Convergence is a multi-domain, rules driven, fast to implement, and scalable enterprise solution. We are the Evolutionary MDM™ company.