

# An Agile Approach to Creating Business Value with Master Data Management

---

Prepared by:

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

Sponsored by:

Semarchy  
450 Townsend St, Suite 100  
San Francisco, CA 94107  
[www.semarchy.com](http://www.semarchy.com)



## Enlightenment in the World of MDM

As data volumes grow and organizations become increasingly driven to seek advantage from their accumulated data sets, there is a corresponding need to integrate data sets containing entity data (such as customer, material, product, or vendor data) together to enable different business functions across the organization. Yet all entity-focused applications, whether they are operational (such as managing material inventories for production or managing customer account information), or analytical, (such as customer profiling, vendor management, recommendation engines, or product quality analytics) will ultimately depend on the ability to materialize a unified profile of each concept under scrutiny.

Master data management (MDM) tools are intended to help organizations link records about the same customer, product, or vendor within different data sets as a prerequisite to providing high quality entity data to both operational and analytical applications. MDM tools have evolved over time by combining algorithmic applications for linking records based on data value similarity with the modeling and management frameworks for capturing a unified representation of linked records for application consumption.

However, despite of the maturity of the concepts of master data management (MDM) tools and techniques over the past decade, many organizations still face many challenges in transitioning the technology from concept to productive use. Some organizations report success as they manage to aggregate source data sets into a single master repository, yet many still struggle to monetize that effort when attempting to support key business initiatives. In essence, one can differentiate successful programs from unsuccessful ones by examining the degree to which the organization internalizes these key themes:

- **The myth of the “golden record”:** The idea of producing of a “single source of truth” or a unique “golden record” assembled using values pulled out of records from data sets of varying levels of trust can blind the stakeholders from understanding how a unified composite view of customer, product, or any other master domain addresses business requirements.
- **“Integration-oriented” business rules:** When data practitioners only focus their attention on the integration of source data sets, they often misunderstand that different business users may have different requirements for consuming master data. Applying a single set of survivorship rules (that select values to populate a single master record) defined by the IT department will result in a sanitized record that may not satisfy any of the business users’ needs.
- **Accountability for business rules:** Since the business data consumers are the ones that need to see the composite views of entity data, the job of defining and interpreting what is stored in master data records should not be relegated to the Information Technology team. Instead, those rules should be defined and managed by the business data consumer.

This paper explores how the operational processes employed to create and populate a master data repository often deemphasize the needs of the data consumers, leading to impediments in productionalizing the master data management environment and difficulty in its integration within enterprise business functions. In turn, the paper discusses an alternative path that adapts aspects of the agile development approach to focus on master data consumption and how that shapes the materialization of composite master views. Finally, the paper suggests some of the key features than an agile MDM tool must provide as guidance for considering vendors and products. This agile approach to designing and implementing an MDM program will speed time-to-value, reduce the complexity of integration and deployment, and promote adoption among an expanding cadre of MDM consumers across the enterprise.

### Challenges in Productionalizing MDM

The conventional wisdom has long suggested a “data integration centric” approach to designing, developing, and implementing a master data environment. That approach typically begins with data set accumulation, then profiling and analysis to identify the set of data elements to be represented in the master model, followed by the creation of a master model that incorporates many of the available data elements into a very wide master data profile representation. The IT developer team will then define a set of survivorship rules that choose data element values from the matched records to create a master record to be stored in a single repository, often with limited focus on business requirements, data quality, or synchronization.

This approach enables the creation and population of a master data environment, but at the same time it can lead to challenges in moving that environment into production, for a number of reasons including:

- **Access and interoperability:** Business consumers will want to integrate their reporting and analyses with the master records but may find it difficult for their existing applications to access the records in the master data repository.
- **Source data integrity:** Business consumers will want to know how their master records were populated, as the fidelity of the original sources may impact their decision-making. This introduces a need for maintaining data lineage.
- **Model evolution:** Because the initial version of master models are derived through a process of consolidation of sources, each new consumer may require model modifications based on the interpretation of their business needs. The result is a need for continual and reactive redefinition of the master models.
- **Publication latency:** New requests for access to the master data environment require an integration task to provide access, yet there is often a lag in being able to materialize master records within a short time frame when the business users really need the data.
- **Ingestion latency:** Another challenge is being able to ingest new data sources and incorporate their records into the master data environment in a reasonable time

frame. The root cause of these challenges is straightforward: it is establishing the *creation and population* of a repository as the primary objective instead of meeting the business requirements for *accessibility* to a composite view of shared master data. More to the point: implementing and populating an MDM repository in the absence of business engagement and participation makes it a “technology solution” approach that is limiting and unable to adequately address business problems within a reasonable time frame.

### Adapting MDM to the Business Opportunities

Understand that unless these challenges are addressed, one may risk being pulled into a vicious cycle: A prospective user of master data is engaged, evaluates the existing master model, determines that it doesn't immediately meet their needs. The MDM developers embark on a project to tweak the data to make it usable while the prospective user finds a temporary alternative. However, as the development cycle lags, the user's temporary alternative morphs into a permanent solution that bypasses the master data environment. Development stops until the next prospective user comes around, and the cycle begins again.

The first step in breaking this cycle is clarifying the success parameters for MDM so that it can support business opportunities. These success criteria include:

- Transitioning the organization's approach to managing shared master data to be business-focused to solve discrete business problems.
- Establishing the data and system architecture to rapidly provide business users with the needed data so they can quickly derive the benefits.
- Ensuring that functionality and operational system services are designed to support a variety of both operational and analytical use cases. Some examples include services for registering a new entity, searching for an entity, or linking entity records from multiple data sets.
- Enabling access to a variety of downstream data users who can consume the data within the context of their business process and based on their own needs.
- Supporting the mastering of reference domains shared across the enterprise so as to ensure consistency in interpretation.

### The Agile Approach: Leverage MDM Technologies to Solve Business Problems

Employing MDM tools to satisfy these directives means changing the way we think about “master data.” Instead of it being a technology “solution” meant to create some hypothetical single repository of golden records, start thinking about MDM as a suite of capabilities that provide business users with access to composite views of customer, product, asset, location, or any other master entity data to solve specific, well-defined business problems. And instead of focusing on a single master data domain at a time, realize that business processes typically touch multiple master domains and reference data sets, all of which can be considered from a business perspective and integrated.

For example, don’t try to justify the creation of a master customer model with the need for a “360-degree view of the customer.” Instead, identify where the absence of master customer profiles inhibits the effectiveness of ecommerce analytics in increasing sales volume by positioning the right products based on demographic and geographic parameters.

Likewise with justifying the creation of a single materials model. Instead, assess how the absence of unified information about products affects operational systems like configurations for just-in-time manufacturing that cannot accurately predict time and cost of production because there is inaccurate knowledge of raw material inventories.

This approach focuses on how the tools can be used to support corporate problem-solving. Some ideas would be to:

- Team business analysts and subject matter experts with their IT counterparts, which allows them to collaborate to rapidly solve each business challenge together.
- Have the IT teams work directly with the business users to understand the problems they are trying to address, evolve models that address those needs, and adapt the utilities to develop business solutions.
- Limit each development iteration to solving clearly specified business problems in a way that provides flexibility in moving onto the next problem.
- Identify commonly-used reference data sets, assess how inconsistent versions of these reference domains affects categorization for dimensional analysis, and create a standard set of master reference data domains.
- Be model-based, and allow for the use of virtualization tools to accommodate the production of a composite record crafted through the rules defined by the expert business data consumer.
- Importantly, design your models and services to be flexible, since your business customers’ needs may change over time and you need to be able to adapt quickly.

This approach provides value quickly, simplifies each iteration of deployment, and importantly creates credibility in monetizing the investment in the MDM technology. In turn that allows you to focus on flexibility and adaptability to eventually satisfy multiple data consumers using an evolved master environment.

At that point one can envision a best case scenario in which the MDM tools are flexible enough to allow adaptation of the master environment to accommodate different sets of survivorship rules for presentation of composite records that are customized to meet each downstream user's needs. The value of the agile approach is that adjustments to the master data environment do not require a complete reboot, but blend data stewardship and incremental development to continue to tackle and solve new business problems.

### Assessing MDM Technologies

Under the right circumstances, such as when you have the luxury of time for comprehensive requirements analysis and synthesis, your organization may be tolerant of using the conventional waterfall approach for design and development master data environment. However, as time windows contract for taking advantage of business opportunities, and the dependence on high-quality master data grows, your organization may benefit from transitioning to an approach that hyperfocuses on business uses for master data coupled with rapid iterative master data development.

If you are considering the agile approach, though, it is imperative that your MDM tool kit provides you with the flexibility to solve business problems but doesn't shoehorn you into years of design and development leading up to a "big bang" delivery of a technology solution that is still not ready to support specific business user needs. Instead, choose a vendor whose technology supports the agile approach, and allows you to establish an infrastructure that simplifies rapid development, continuous refinement, and adaptability.

Aside from the typical capabilities provided by a master data management tool such as identity resolution, data validation, survivorship directive, and data stewardship control methods, some of the key features that an MDM tool supporting agile development must provide include:

- **Model-driven design and development.** In a model-driven approach, the consumer's business needs are used by the designers to shape the representative model for key shared master entities.
- **Simplified stewardship:** Look for a tool with a simple, modern stewardship interface designed to maximize the effectiveness when used by a broad spectrum of Data Stewards. In the best scenario, the interface guides novices and provides training in stewardship practices.
- **Progressive and Transformative development.** As new business contexts for master data use emerge, the designers can incrementally add new rules, connect new data sources to the platform, make changes to the data model to add new entity attributes or even add new entities, or modify relationships.
- **Iterative extensibility.** There is no doubt that enterprise-wide master data requirements are going to change, so look for tools that simplify domain modification, introduction of new domains and corresponding relationships, or

## Knowledge Integrity Incorporated

### *Business Intelligence Solutions*

adding new consumer applications while maintaining semantic consistency across business functions.

- **Version control.** A corollary to extensibility is that there may be need to juggle multiple versions of the master environment as existing applications must remain in production while enhancements are added to support new business requirements.

These types of additional features not only provide the ability to integrate data and build the master repository, they also ensure that customized views of synchronized and semantically coherent composite master records can be rapidly produced and delivered to a variety of different business users in ways that are quickly adapted to new business opportunities.

## Knowledge Integrity Incorporated

*Business Intelligence Solutions*

### About the Author

David Loshin, president of Knowledge Integrity, Inc, ([www.knowledge-integrity.com](http://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 articles published at [searchBusinessAnalytics.techtarget.com](http://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](mailto: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.