



The CDO's Guide to Data Mastering at Scale

Introduction

As a new Chief Data Officer (CDO), you recognize that the way enterprises have been managing their data through traditional Master Data Management (MDM) is a dated process. And while MDM might have worked well for the type and scale of data mastering challenges that existed 15 years ago when the solution came about, today's data challenges require a different approach.

Individual business units (IBUs) within a company often still **utilize traditional means of data mastering**—which leads to large data silos that can be difficult to overcome. Now, utilizing those traditional approaches on the vast amounts and variety of data that enterprises are accumulating is slow, labor intensive, and extremely costly.

Traditional Approaches Create Traditional Results

The fact is, traditional approaches to data mastering produce traditional results—which is to say, limited. The velocity and variety of data has outgrown the old approaches we used, limiting a corporation's ability to quickly and cost-effectively analyze data. Let's explore this.

Exploring the limitations of traditional MDM

Master Data Management defines and manages an enterprise's critical data to provide a single point of reference. This single truth allows you to accurately answer key questions about any metric or KPI. This has been invaluable for companies seeking a single source of truth about an entity that other sources can reference further down the stream. There's a lineage to the records and flexibility in how the records are created. Absent of errors, there's no duplicates or unmatched data—creating accurate views of each entity.

The problem is that MDM requires a human-intensive process to deliver rules-based truths. This is a complex way of saying that it's not scalable, and moreover, it's dependent on constant manual reviews of exceptions. This means that there will be a large portion (and ever growing portion) of data that remains unmastered, and the only solution will be for enterprises to pay a premium in resources.

Exploring the limitations of ETL

Extract, Transform & Load (ETL) creates a global schema up front. Certainly, it's effective at moving data and performing masterings with simple rules, but it can't be scaled because it takes too much time. Enterprises with a huge amount of data simply can't consider ETL to be a viable option. Additionally, it isn't designed to create a single point of truth, so it's missing that consistency across consumption points.

These traditional approaches can be for small problems without real time requirements. But to solve for large numbers of complex records with real time requirements that enterprises like yours experience, you have to evolve a data mastering approach scale. As a new CDO, it's crucial that you guide your company in rethinking data integration and mastering to tilt the odds of success more heavily in your favor. Ultimately, you need to implement an agile data mastering approach that utilizes machine learning aka MDM 2.0.

The Future of Data Management: Agile Data Mastering

The software development industry has been employing agile approaches for years, and you can use those same practices in data management. Agile Data Mastering (ADM) connects people, processes and tools together to treat data unification as an iterative

process. Simply, ADM brings humans and machines together—business experts train and validate machine learning models, so the tool's accuracy improves. This means the smarter the model becomes, the less human interaction is required. This collaboration can drive key benefits for enterprises and change the way businesses use and manage their data. Here are just a handful of the benefits of ADM:

- 1. It's scalable:** Human expertise combined with machine learning allows companies to integrate datasets from a variety of sources, allowing scale without sacrificing accuracy.
- 2. It's faster:** ADM tools can deliver results in days that traditional methods might have needed months or even years to get.
- 3. It allows for team optimization:** When experts aren't spending endless days on data prep, they can focus on more specialized efforts.
- 4. It opens new opportunities:** With cost-to-know reduced, projects shelved for "high costs and risk" can finally get the attention they deserve.
- 5. It promotes flexibility:** ADM allows teams to respond to the unexpected, faster and more effectively than ever before.

Companies like Tamr utilize Agile Data Mastering to drive analytic outcomes and solve data problems every day. This human-guided, machine learning modern data management model is crucial for enterprises because it unifies data up to 10 times faster than traditional methods at up to 90% lower cost.

Building a modern data mastering, machine-learning model

You know what agile data mastering can do to advance traditional MDM. Now, you need to know how to do it. There are a few essentials required:

- 1. Machine learning:** Traditional MDM required manual analysis of systems, coding of data, and definition of rules. Modern data mastering utilizes machine learning models that do these things in a fraction of both the time and the cost.
- 2. Expert input:** Humans will train the machine learning models in order to ensure accuracy. Instead of juggling thousands of rules, now, human experts can simply define a handful of rules that turn the machine learning models into experts.
- 3. Transparency:** Visibility is key in machine learning management. This helps the models ensure accuracy and compliance with regulatory audits.
- 4. Low latency matching:** Data ops teams will integrate MDM systems back into ops systems as a system of record. These operational systems will then need sub-second access to the MDM in order to suggest mastered data to users or update master data at the source. This ensures source data remains better synchronized across the entire enterprise.
- 5. Continuous innovation:** Future data mastering opportunities will come to market as machine learning models, including data error detection and correction and model management. By remaining open to these options, enterprises can ensure they are on the forefront of data mastering and don't fall behind the curve.

These requirements will help ensure you can achieve the full promise of modern master data management at scale for your organization.

Start Building Your Model for Data Mastering at Scale

As a new CDO, your success is based on your ability to hit the ground running. Here are our top 3 recommendations for your first few months on the job:

- 1. Communicate with peers and stakeholders:** Implementing new data processes is as much about people and process as it is about the technology you are using, and it often requires cultural change. Creating a cultural shift is going to be one of your biggest challenges to bridge the gap between operational agility and corporate control. This means opening the lines of communication with business stakeholders and their enterprise data architecture teams is going to be necessary to your success.
- 2. Establish benchmarks:** Map out a plan prior to starting your work, so you can measure your progress, the company's progress, and so on. It pays to get analytical about the analytic outcomes your team is driving. Three areas that we recommend focusing on are:
 - Errors: How accurate is your data?
 - Deployment speed: How quickly is your team able to get results or reports in you customers' hands?
 - Productivity: Is your team able to effectively work together?

Leverage the data from traditional MDM to show the power and potential (and real-time) results of your agile data mastering. Creating value and increasing revenue are important goals, and creating a map will help you reach them more clearly and know when to pivot if something isn't working.

- 3. Seek out your CDO networks and agile data mastering experts:** The old adage of not reinventing the wheel exists for a reason—other people have been in your shoes with this monumental task ahead of them. Find your network and related associations and ask for insights from other professionals and peer forums. Don't be afraid to **talk to data mastering experts**.

Next Steps

Stay focused on what you know will bring your enterprise the results they need to move into a modern data management model: human-guided machine learning within a data ecosystem that drives value and increases revenue.

To learn more about how Tamr can help you create an analytics-driven organization using Agile Data Mastering, [schedule a demo today](#).

[Schedule a Demo](#)



About Tamr

Tamr is the enterprise-scale data unification company trusted by industry leaders like Toyota, Societe Generale, GE, Thomson Reuters, and GSK. The company's patented software platform uses machine learning supplemented by human expertise to unify and prepare data across myriad silos to deliver previously unavailable business-changing insights. With a co-founding team led by Andy Palmer (founding CEO of Vertica) and Mike Stonebraker (Turing Award winner) and backed by investors including NEA and Google Ventures, Tamr is transforming how companies get value from their data.

To find out more, visit tamr.com