Tamr

Agile Data Mastering

A Modern, Agile Approach to Solving
the Data Mastering Challenge

When it comes to data quality and availability, there is a gap between expectations and reality.

Expectation:

Reliable, integrated data is readily available in an easy-to-digest format that allows you to harness key business insights.

Reality:

Data comes from a wide array of sources that aren't easily accessible and are often times plagued with inaccuracies. Critical questions are difficult (or sometimes impossible) to answer because of the high level of effort required to find, assemble, prep and analyze data.

The problem isn't the technology—it's the approach.

Traditional approaches to data management don't acknowledge that data environments and end user requirements are constantly changing.

The traditional process of data mastering is slow. It can take months or even years to complete mastering projects using deterministic rules. And it places an unnecessary burden on data scientists: **Most data scientists spend up to 80% of their time** doing tedious, time-consuming data prep work before they can begin actually analyzing data.

Big Data brings new challenges with it.

The three V's of Big Data:

Big Volume:

Too much data to manage

Big Velocity:

Data is coming at you too fast, and you can't keep up

Big Variety:

Data is coming at you from too many places and from too many sources, causing a data integration problem

This is where Agile Data Mastering comes in.

Agile Data Mastering (ADM) connects people, processes, and tools—treating data unification as an iterative process. ADM engages stakeholders early and often, enabling teams to easily correct issues, accommodate emergent requirements, and quickly react to changing data.

The Pillars of ADM

Agile Mindset:

Data needs are constantly changing, which is why the top-down approach no longer applies. Enterprises should imagine their teams as software developers building data applications. The focus should be on **building operational** and **analytical applications** that solve major business problems, not on building reports.

Agile Skill Set:

Stakeholders from the entire data supply chain need to be included in the teams who build data applications (those who own the source systems, those who consume data in their analytical and operational applications, and those who provide the former for the latter). This might mean a reorganization of the data team for many businesses.

Agile Toolset:

The tools you are using should support rapid, iterative workflows that are interoperable with emerging technologies. Well designed tools allow for easy integration across all of your platforms so that you can unify and master your data through machine learning. The best tools offload lower-value work of data modeling and wrangling to algorithms and free up your people to build great analytical and operational apps with your data.

Key features of ADM tools

Machine learning and automation:

Instead of relying solely on deterministic rules, ADM tools also leverage probabilistic machine learning models to handle the heavy lifting of identifying relationships within your data.

Human and machine collaboration:

Subject matter experts are used to train and validate machine learning models, so the accuracy of the tool improves over time. This means that, inevitably, the amount of time a human needs to spend performing a specific task will decrease over time.

ADM drives key benefits for enterprises, enabling them to:

- **Scale easily:** The combination of machine learning and human expertise enables organizations to easily and quickly integrate datasets from multiple sources—so that you can scale as needed without compromising accuracy.

- **Improve time to value:** By leveraging machine learning, ADM tools can deliver results in days, not months or years.

- **Improve data science efforts:** By radically reducing the amount of time data science teams spend on data prep, they can be freed up to focus on more specialized, high value work.

- **Solve the 'too hard' problems:** When the 'cost to know' is greatly reduced, the projects that organizations have deferred for years because of the anticipated high cost and risk can finally be addressed.

- **Respond to the unexpected:** Digital transformation will never be straightforward. Unanticipated questions and challenges will inevitably arise, but an ADM capability provides the capacity to respond effectively when the unexpected happens.

At Tamr, we leverage human-guided machine learning to help organizations unify and prepare data across myriad silos.

By embracing an agile approach to data mastering, our solutions are able to deliver the flexibility, speed and scalability that enterprises need to keep up with the pace and challenges that come with Big Data. And because we are an API-first vendor, organizations can seamlessly integrate Tamr solutions within their existing technology stack to build a truly best-of-breed data pipeline.

ADM in action

Major Pharmaceutical Company:

A major pharmaceutical company leveraged Tamr's solution to make it easier to access and use data for exploratory analysis and decision-making about new medicines. The company had millions of different data elements to rationalize, so instead of taking a traditional master data management approach, which would have taken too much time and effort, the DataOps team turned to machine learning. The company was able to scale across 8 key business units to deliver visibility into over $380 million in cost-saving opportunities in 18 months.

The company was able to scale across 8 key business units to deliver visibility into over $380 million in cost-saving opportunities in 18 months.
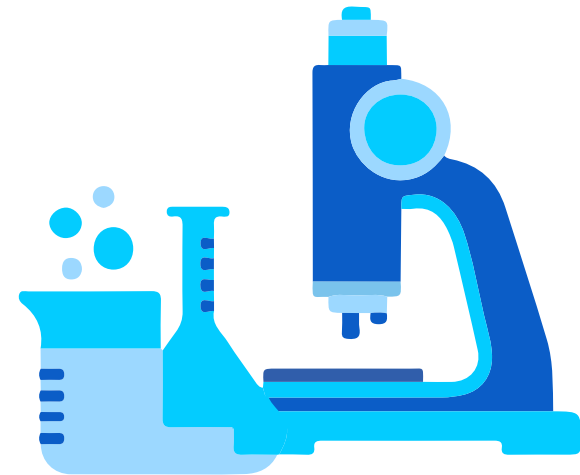
Multinational Media and Information Company:

A multinational media and information company faced challenges maintaining critical, accurate data. They had outgrown their manual curation processes, and looked to Tamr to provide a better solution for continuously connecting its data on millions of organizations with more than 5.4 million records. Using Tamr's solution, the company was able to rapidly master organizational entities for use in offerings— reducing manual resources required by 40% and eliminating several months of work.

Using Tamr's solution, the company was able to rapidly master organizational entities for use in offerings— reducing manual resources required by 40% and eliminating several months of work.

Next Steps

Stay focused on what you know will bring your enterprise the results they need to move into a modern data management model.

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 leading data mastering company to accelerate data-driven business outcomes. Industry leaders like: Toyota, Societe Generale, GE, and Thomson Reuters trust Tamr to manage their enterprise data as an asset. Tamr's unique approach of using human-guided machine learning algorithms to accelerate data mastering projects lets the world's largest organizations enhance their data operations, rapidly activate latent data, and increase the velocity of business outcomes through data-driven 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