The Essential Buyers Guide to Cloud MDM:

Four Criteria for Modernizing Master Data Management

Tamr

Every business is a data business. To power growth, lower risk, or optimize spending, data provides the foundation for accelerating business outcomes. Master Data Management, or MDM, is often at the heart of how companies organize, clean, and unify their data in order to coalesce large disparate sets of data into a single point of reference to power analytic insights. The demand for curated, mastered data is clear: A new report forecasts the MDM market will more than double in the coming years from **$11.3 billion in 2020 to $27.9 billion by 2025.** Coupled with the rapid movement of data to the cloud (companies spent **$107 billion for cloud computing** infrastructure services in 2019, up 37% from the previous year), the need for modern approaches to MDM—solutions that scale natively in the cloud and accelerate the ability to master millions of records across dozens of internal and external data sources—is driving buyers toward next-gen cloud data mastering.

As traditional solutions continue to fail the needs of the businesses that use them, what should companies be looking for in a modern MDM replacement? What are the critical components a modern MDM solution should possess? We will help answer these questions in this buyer's guide.

Key Competencies

In order to solve modern MDM challenges at scale, there are four key competencies a solution must have:

CLOUD-NATIVE TECHNOLOGIES

Organizations' rapid shift to the cloud is playing a major role in transforming data management. Cloud-native capabilities (technologies that leverage built-in elastic and ephemeral cloud and compute benefits of cloud technology) allow for a highly secure and scalable infrastructure that is able to add additional storage and compute power without adding to physical and hosting costs. With this built-in advantage, cloud-native solutions allow organizations to reduce the total cost of ownership of MDM projects, and enable data organizations to take advantage of ongoing product enhancements without needing to allocate additional resources to hardware, and system or software upgrades.

A MACHINE LEARNING-FIRST APPROACH TO MDM

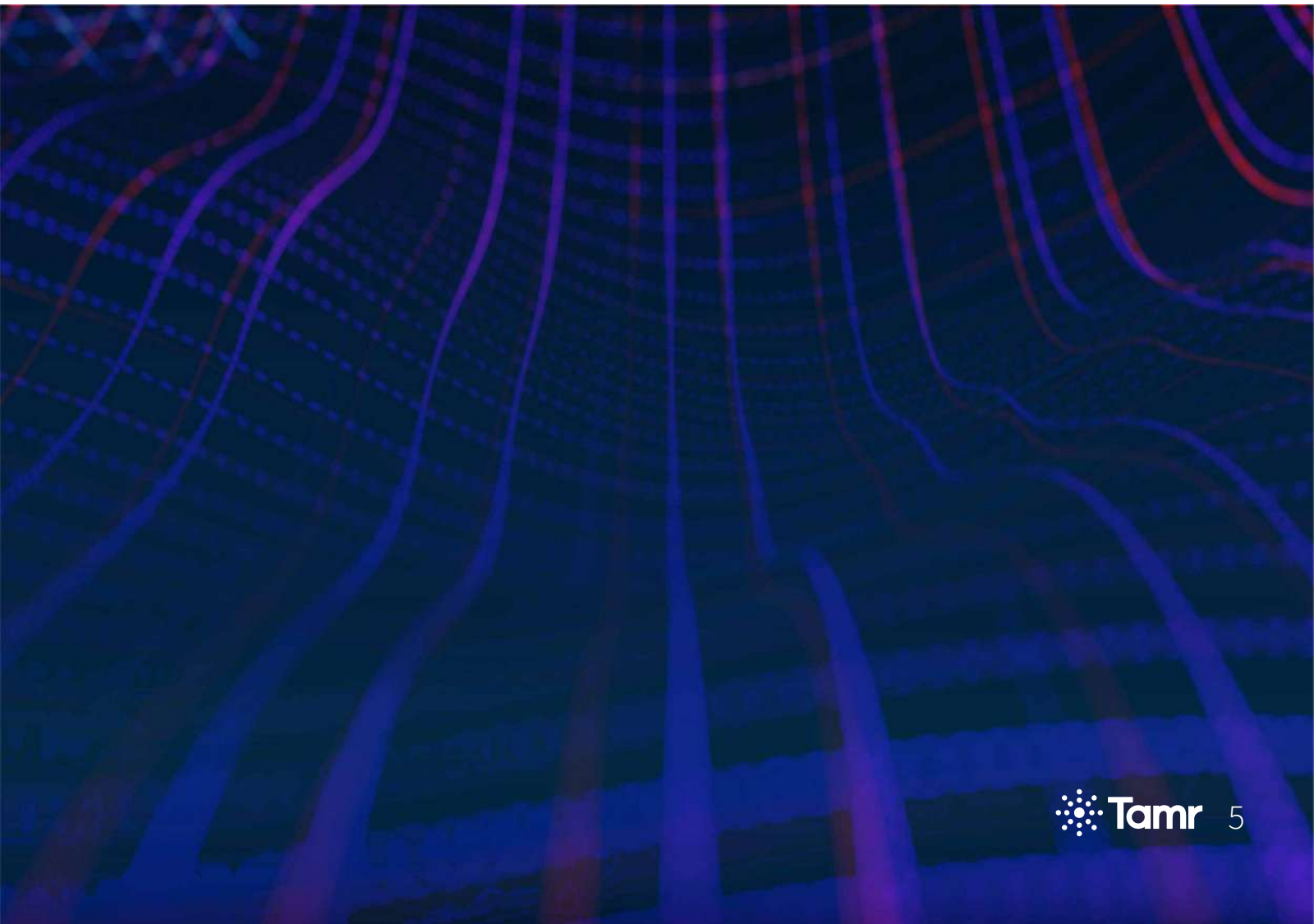
As you're evaluating MDM solutions, a primary consideration is whether it employs a traditional rules-based approach to mastering data or leverages machine learning. Traditional rules-based systems can be effective on a small scale, relying on human-built rules logic to generate master records. However, rules quickly fall apart when tasked with connecting and reconciling large amounts of highly variable data at scale. Rules, by their nature, are strict and inflexible. Machine learning, on the other hand, improves with more data; huge amounts of data (1M+ records across dozens of systems) provides more data points for the algorithms to identify patterns, matches and relationships, accelerating years of human effort down to days.

KEEPING HUMANS IN THE LOOP

While machine learning is critical for scaling data mastering, it's equally important for MDM solutions to effectively engage business users and subject matter experts in the remediation process. By engaging those who know the data best, keeping humans in the loop not only trains the machine learning models more quickly, it also drives tighter alignment between the data and business outcomes that require curated data.

OPEN AND INTEROPERABLE ARCHITECTURE

Look for a solution with an open and interoperable architecture that allows businesses to pursue “best-in-breed” solutions for all their data needs. Today’s premier data organizations take a DataOps approach to their technology stacks, which means using the best tool for each specific need, instead of what’s easiest or readily available. While it may be convenient to pursue a vendor that offers everything under one roof, these single vendor solutions lack the depth to solve today’s complex and evolving data challenges. Look for solutions that play well with others and are complementary through RESTful APIs and robust integration capabilities



A Choice: Augment or Replace an Existing MDM

One of the big reasons people ultimately don't invest in new technology is because they believe there's a "sunk cost" associated with their incumbent technology and all the inherent infrastructure that's built around it. If the incumbent technology simply does not fit the business's objectives (e.g., there is no clear way to engage business users in IT-driven projects and many attempts have failed), the sooner you can architect a migration path the sooner you will save money and drive growth.

However you do not always need to start from scratch. Some modern MDMs can be used to effectively augment and improve the incumbent technologies. By choosing to replace individual pieces of the puzzle—for instance, adding machine learning to mastering but leaving the ETL system in place—instead of the entire thing, organizations are able to deliver results faster and increase buy-in across the company.

A tip: look for open architecture, APIs, and cloud-native offerings that provide flexibility integrating with legacy systems, new and future data pipelines.

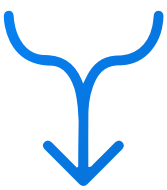
6 Modern MDM Use Cases

1. Spend Optimization

Use machine learning cloud-native MDM to consolidate and categorize spend with unprecedented granularity to uncover millions in indirect and direct spend savings.

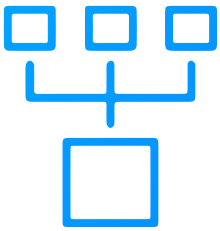
2. Customer 360

Cleanse and master customer data by consolidating internal data sources (e.g. CRM systems) and connecting structured and unstructured external data sources to create a comprehensive, up-to-date view of consumers and retailers.



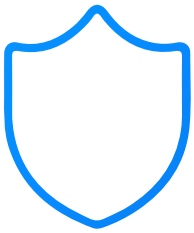
3. Merger & Acquisitions

Quickly rationalize and master data from new and legacy systems (e.g. CRM, ERP) from across business units and geographies to unify data schemas and generate common taxonomies, driving advantage of data and insights from newly acquired organizations.



4. Product Rationalization

Consolidate internal data sources (e.g., ERP and CRM systems) and connect to external data sources to offer a comprehensive, up-to-date view of market opportunities.



5. KYC / Anti-Money Laundering

Master critical data at scale to lower risk, improve regulatory operations, and enhance the customer experience. Match customer records from internal and external datasets (e.g., GLEIF, sanctions lists) and gain a comprehensive view of customer activity.

6. Analytic-Ready Datasets

Make informed decisions with cleansed, up-to-date, datasets that are curated and ready to use with analytics programs and data visualizers.

What to Consider When Researching Cloud MDM

As you evaluate the features, benefits, and limitations of MDMs, here's a checklist of key things to consider.

Is the solution on-prem, hosted on the cloud, or cloud-native?
You want an MDM that makes it easy to access, master, and maintain your data wherever it is located. Consider solutions which offer a variety of hosting options and vendor flexibility to support your needs today, and in the future. Think about how often you will need to change workloads, add a new use case, or incorporate a new data source. If any of these scenarios are likely, it's best to focus on cloud-native solutions instead of on-premise. Cloud-native technologies are designed to automatically scale up and down resources on-demand, which in real terms means that the cost of scaling terabytes of data, at millions of records, stays linear instead of exponential. Regarding flexibility and future-proofing, a lot of companies now utilize a multi-cloud strategy to avoid being locked into a single cloud offering. Make sure your MDM can support this and is able to work seamlessly with data across different cloud instances.

Is machine learning at the "core" of the solution?

While many vendors have acquired or are building machine learning technology for their rules-based platform, you'll want to evaluate whether the machine learning technology is a feature, or truly at the core of the product. A critical part of scaling machine learning effectively is not just the algorithms and machine learning technology, but the ability to address the data remediation and edge case data reconciliation needs in a scalable way.

What are the upfront costs, and how much does it cost to maintain?

Software licencing costs for MDM are often just the start. MDM budgets of $10-20M per year are not uncommon at large multinationals. Traditional MDM solutions typically require high upfront costs as well as a lot of highly skilled - and expensive resources - to maintain. Human-built rules do not scale and cannot effectively master highly variable data. As Mike Stonebraker (Turing Award winner and Tamr co-founder) likes to point out, humans cannot manage more than 500 rules. So, in addition to the consulting fees common for a large scale implementation, many organizations will have a continual need for dozens of engineers to write, test, and maintain rules for the lifetime of its use. As you research pricing, it's critical to understand the resources you have, and the resources you'll need to add, in order to calculate the total lifetime costs of a particular MDM.

How does the MDM solution scale for millions of records?

Modern businesses have dozens to hundreds of siloed data sources generating information. Each system has its own taxonomy and way of organizing data. Add to this common data variances like language, units of measurement, and currencies, and you start to see the immense problem data variety presents when mastering data. If you are relying on rules, adding just one incremental source can mean adding thousands of new rules. Take a critical look at how a solution handles data variety at scale. How quickly can anomalies be identified and fixed? What are the time and capital costs of updating the system?

How long will it take to see results?

MDM implementations can take months, and in many cases years, to master enough data entities to power business outcomes. A common cause for delay is the interaction between IT and the business. MDMs that leverage machine learning are able to deliver faster time to value by relying on the algorithms to do the tedious work of matching and unifying disparate data. By leveraging machine learning instead of rules, some solutions can deliver analytics-ready datasets in days or weeks instead of months or years. Additionally, MDMs with an intuitive framework and UI for engaging SMEs will accelerate results.

Are MDM results explainable?

A benefit of a rules-based approach is its simplicity and transparency. It's fairly easy to understand and audit why a rules-based MDM system arrived at a particular result. Conversely, machine learning is often derided as a black box and knocked for its inability to "show it's work." This can be an especially big problem in regulated industries (like financial services and life sciences) which legally must be able to explain the logic behind decisions. As you investigate different machine learning technologies, it's good to ask whether the models are reviewable, can humans impart feedback and tune the model's results, and lastly, is the technology already being used by organizations with these strict regulatory requirements.

Final Thought

Businesses have spent three decades investing in people and technology in the pursuit of transforming their data into an asset. Yet old MDM systems persist and their shortcomings continue to throttle the pace of data-driven insights and innovation.

We know deterministic, rule-based approaches do not easily accommodate nor scale for the increasing flow of messy, diverse data coming from today's disparate data systems.

Tamr offers a better approach. By leaning into its machine-learning and interoperable, modular approach to the data mastering problem, Tamr solves many of the problems traditional MDM solutions still cannot overcome. With Tamr, businesses can accelerate critical analytical insights by reconciling internal and external data at scale.

Join us for an exclusive workshop to see first-hand how Tamr's modern cloud MDM approach can accelerate your data initiatives and turn your data into a competitive asset.
Join us. https://bit.ly/cloud-mdm

About Us

Tamr is the leading data mastering company, accelerating the business outcomes of the worlds' largest organizations by powering analytic insights, boosting operational efficiency, and enhancing data operations. Tamr's cloud-native solutions offer an effective alternative to traditional Master Data Management (MDM) tools, using machine learning to do the heavy lifting to consolidate, cleanse, and categorize data. Tamr is the foundation for modern DataOps at large organizations including Industry leaders like Toyota, Santander, and GSK. 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**