Viewpoint

David Menninger
SVP &
Research Director

Time to Take Another Look at MDM

Master data management (MDM) as a discipline has been around for decades. And MDM software has been evolving over time to help organizations meet their changing data requirements. But if your understanding of MDM software is even a little bit dated, it's time to take another look. The early days of MDM were not even about MDM software, because applications dedicated to the purpose hadn't yet evolved. Let's unpack the history. "Master data" is the collective term for all the core reference data that organizations track, including customers, products, locations, employees and just about any other information that might be kept in a list. This information was first managed in the primary business software applications of earlier days, including enterprise resource planning (ERP), customer relationship management (CRM) and supply chain management (SCM) software. These applications were the repositories of the organization's "golden records," or master versions of the data.

Manual processes were used to ensure consistency of this information across systems and applications. If one company acquired another and they needed to merge software systems, reconciliation required days or weeks of manual work and batch jobs to load data from one system into another. This all may sound archaic today, and it became complicated enough that MDM software emerged to automate these processes. Automation made the work easier, the data more reliable, and offered the promise that organizations could spend less time on data management and more time focused on their core business activities.

However, it wasn't really as simple as that might sound. Despite the obvious importance of having real-time and accurate master data, our research also shows that fewer than one-quarter (22%) of organizations are very confident in the quality of their data. This is partly due to the fact that MDM is a complicated process. The various sources of master data need to be identified, extracted and loaded into a separate system. These systems helped identify conflicts in the master data, but it

Confidence in Data Quality
Results indicate room for improvement

■	Very confident
■	Confident
■	Somewhat confident
■	Not confident

Source: Ventana Research Data Preparation
Benchmark Research
© 2021 Ventana Research

wasn't always obvious how those conflicts should be resolved. And many early MDM systems were developed without the benefits of artificial intelligence or machine learning (AI/ML).

MDM software evolved primarily along two tracks: customer data integration and product information management. These domains were the lifeblood of key business processes. If organizations couldn't properly track who their customers were, they couldn't charge and collect the revenue associated with
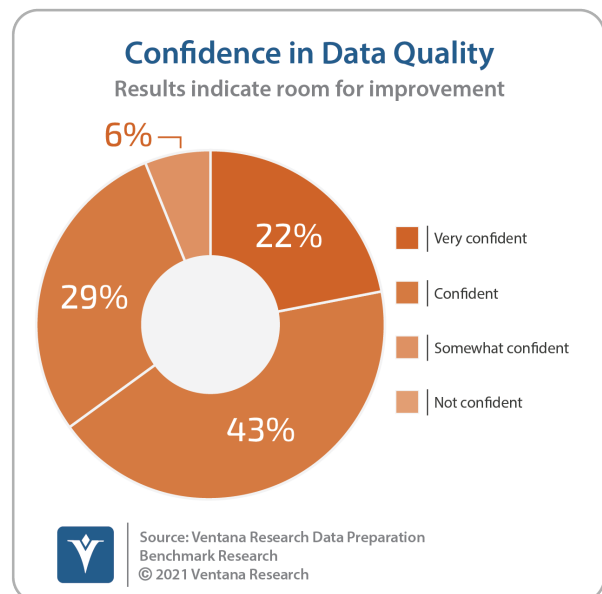
sales to those customers. Similarly, if organizations couldn't properly track the bills of materials for products, they couldn't produce, market and sell those products. As a result, these two areas drew the most attention of developers. Furthermore, driven, in part, by the advent of data warehousing, which allowed information to be brought together from across the organization and thus highlighted any inconsistencies, organizations eventually learned the value of consistent information across multiple domains. Today, any MDM system needs to incorporate such multi-domain capabilities.

MDM software, like most business applications, was initially developed as enterprise software that was installed on-premises. But enterprise applications have been moving to the cloud right along with an organization's data in general. Our research shows that already nearly one-quarter (24%) of organizations have the majority of their data in the cloud, and 86% expect the majority of their data to be there eventually. The move to the cloud has also been accompanied by a proliferation of cloud-based business applications. Now it's common for sales force automation, marketing automation, human resources and many other applications to be acquired as cloud-based services. So, today, MDM systems need to provide cloud-native capabilities.

Organizations' data patterns are changing as well. Data volumes are expanding and latency requirements are growing more stringent as organizations digitize their business processes. These days, all types of devices are streaming data into organizations, and nearly every detail is retained for analysis. Data warehouses have grown into data lakes to accommodate the volume. Many organizations now process billions of records on a regular basis, and they need to be able to process all that information quickly: for instance, our research shows that nearly one-half of organizations need to process streaming event data in seconds or sub-seconds. In this demanding business environment, it is nearly impossible to govern your data properly without MDM software that provides real-time responsiveness that scales with any volume of data.

MDM systems today need to include AI/ML capabilities to make it easier to manage master data. AI/ML can boost MDM effectiveness and efficiency by identifying and recommending resolutions to conflicts, deduplicating records with similar but not identical entries, and by identifying sensitive information. As data and analytics software have advanced, vendors have begun to incorporate these capabilities into their products and we expect that by 2022, more than one-half of all data processes will use AI/ML to boost the value that can be derived from the data.

Examine how your own organization's data requirements have evolved. As you seek to ensure consistent information between and across your various systems, consider the requirements for your MDM software. Consider whether you can master all the relevant domains for your organization without it. Consider whether you need cloud-native capabilities. Consider your real-time and scalability requirements. Consider the role that AI/ML can play in improving your MDM processes. And if you haven't considered master data management software in a while, it may be time for another look.

Dave Menninger – SVP and Research Director, Ventana Research

David Menninger is responsible for the overall direction of research on data and analytics technologies at Ventana Research. He covers major areas including artificial learning and machine learning, big data, business intelligence, collaboration, data science and information management along with the additional specific research categories including blockchain, data governance, data lakes, data preparation, embedded analytics, natural language processing (NLP) and IoT.