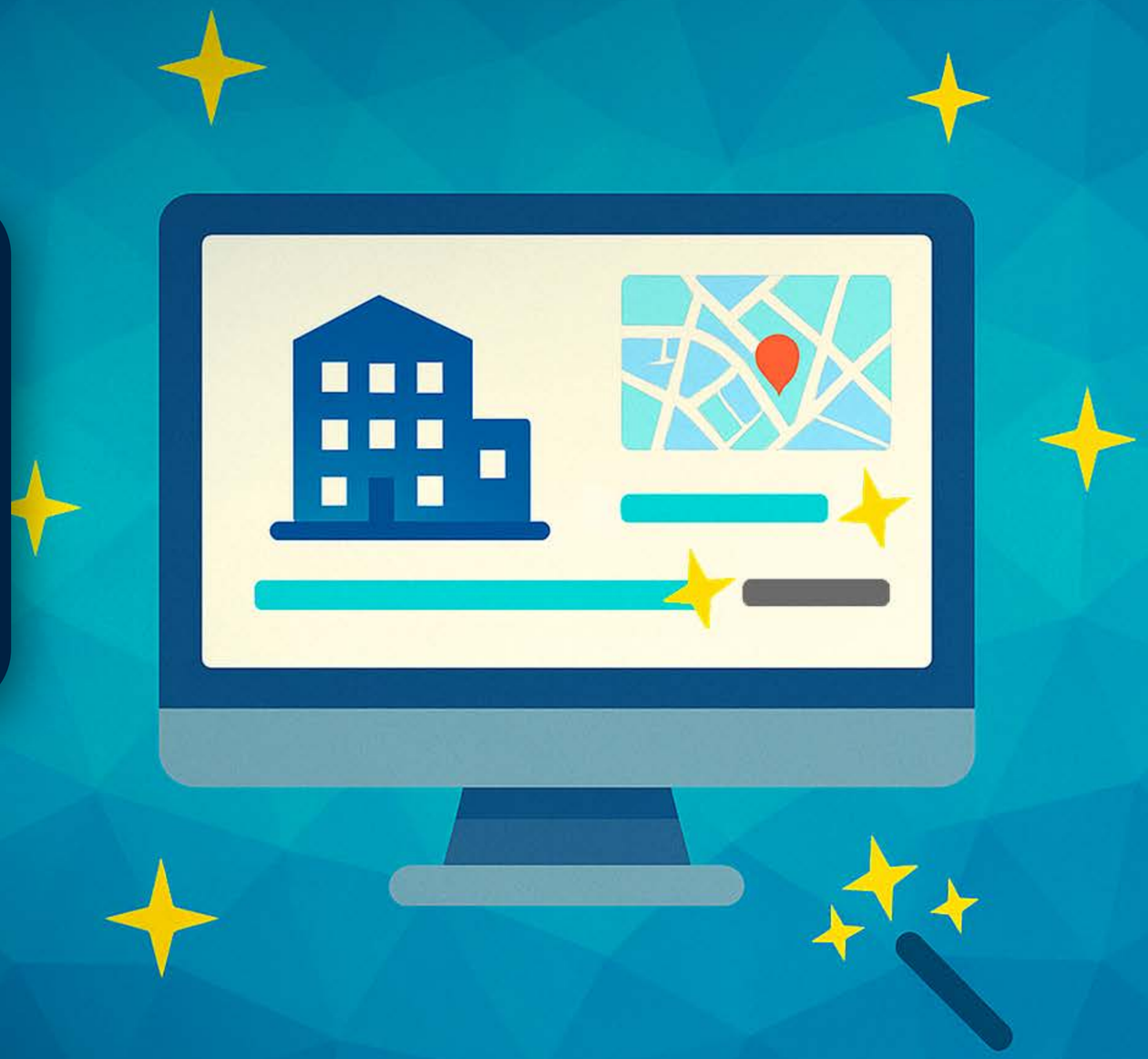
Guide to Data Enrichment

What It Is, Why It Matters, and How It Enhances the Power of Your Data



Tamr

Over the past 13 years, **the amount of data generated has grown 74x**. And it's predicted to grow even more, surpassing 180 zettabytes by the end of 2025. But it's not just the data's size that has expanded; its complexity has increased, too.

To put this in perspective, imagine a typical marketing manager, whose job once involved juggling insights from just a few customer databases. Today, they will need to deal with data flooding in from dozens of sources: CRMs, social media, website analytics, third-party vendors—each with their own quirks and inconsistencies. More systems and more sources are capturing data, making it difficult for organizations to keep their data clean and consistent.

Making matters worse, much of an organization's data is incomplete. And incorrect, incomplete data leads to incorrect, incomplete insights, which, in turn, lead to bad decisions. In fact, according to **Gartner**, bad data is costing organizations over $12 million each year.

What's more troubling is that many organizations don't understand the full scope of their data quality problems. They've built self-service BI tools and dashboards—and connected their data to business-critical systems—before assessing, improving, and reviewing it with end users. As a result, they discover that while they've fixed the veneer, they've also exposed data that's garbage in disguise.

Incomplete, poor-quality data makes it nearly impossible for organizations to achieve a true **360-degree view** of the companies they work with (e.g., customers, prospects, suppliers, healthcare organizations, distributors, etc.) and the people they care about (e.g., consumers, contacts, healthcare providers, students, employees, etc.). To realize the full value of their data, organizations must prioritize improving its quality, an essential milestone on the MDM journey.

The MDM Journey

1. **Assess:** Know where you are…and where you want to go.
2. **Improve:** Make your data trustworthy by cleaning and enriching it.
3. **Review:** Put your data in front of end users to gather feedback and build trust.
4. **Operationalize:** Turn your data into a mission-critical asset by connecting it to key business systems.

Download our ebook to learn more about the MDM journey: what it is, where you are, and best practices for navigating each step.

Improving the quality of internal data begins with cleaning and enriching it. Yet, as organizations progress on their MDM journey, many have discovered that the best version of their data often resides outside of their company walls, living externally within sources such as ZoomInfo, Moody's, or **Dun & Bradstreet**. For example, while a company may trust the company name listed in their internal CRM record, they may prefer to rely on an external source such as Dun & Bradstreet for address information or DUNS number.

Think about it: Does the best, cleanest version of your personal data—your address or phone number, for example—live within the ERP or CRM systems of your service providers? Probably not.

That's why, to successfully improve their data quality, organizations must link their valuable, internal data with trustworthy external data provided by third parties, vendors, or public data sources. This is what data enrichment is all about, and it's an effective way to fill in the blanks in golden records and provide decision makers with data that is cleaner, more accurate, and more complete.

What is a Golden Record?

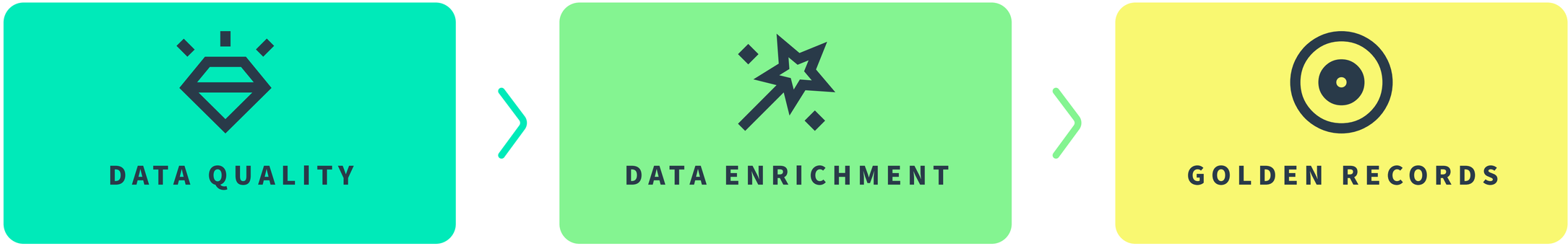
A **golden record** represents the pinnacle of data management: a single, authoritative, enriched, accurate version of a business entity's data across multiple data sources and datasets. This hard-to-achieve level of data integrity and comprehensiveness is now possible through an AI-native approach. By leveraging artificial intelligence (AI) and machine learning to create golden records, companies can achieve faster time to value, enabling them to sell more effectively, enhance marketing outcomes, uncover new opportunities, reduce operational costs, and make more informed decisions.

What Is Data Enrichment?

Data enrichment is the process of enhancing existing internal datasets with information that is generated from additional, outside data sources. These sources could include data about organizations or people, for example, and be used for sales and marketing, risk management, and more.

Improving data quality involves validating and standardizing data to identify gaps and common attributes. Data enrichment takes the improvement process even further by using **unique IDs** to connect internal data with a third-party source to find additional or missing attributes to append to each record. Data quality and data enrichment build on each other to unlock greater value from your organization's internal data.

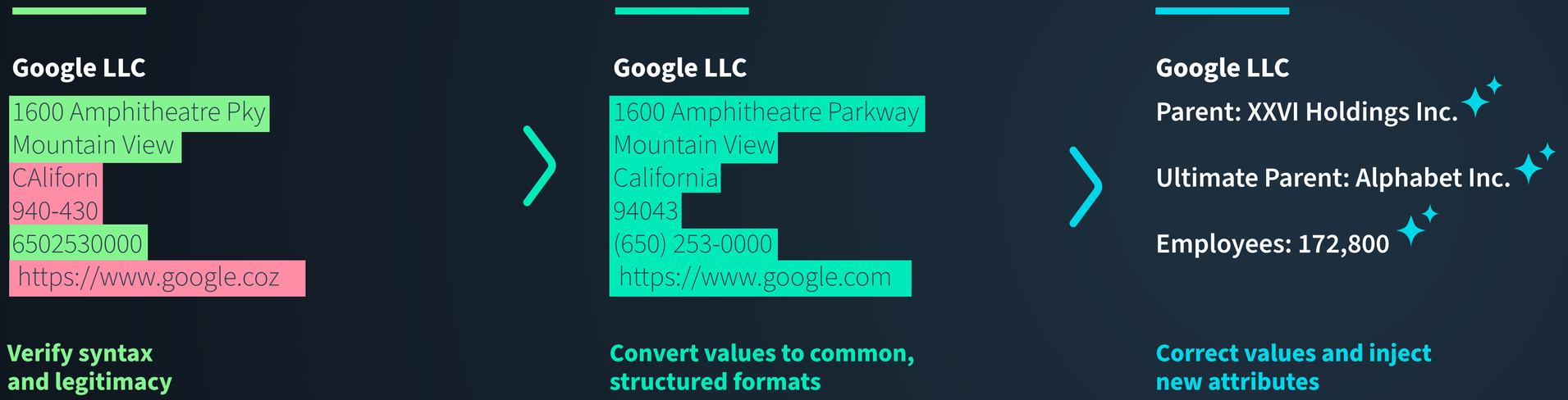
See the next page for an example of what each looks like using the corporate information for Google LLC.



Validate verifies data syntax and content legitimacy. For example, this step helps verify whether the email domain of Google LLC is legitimate, whether or not the address actually exists, and whether or not the phone number is a landline or mobile.

Standardize converts values to a common and structured format. For example, this step standardizes the phone number and the addresses into a standard format. This step also corrects spelling mistakes if they exist.

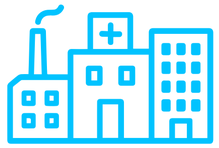
Enrich provides options so you can determine which value is correct and injects new attributes from third-party sources connected to your internal data. For example, this step tells us that Google LLC has an ultimate parent company that is Alphabet Inc. This step can also provide other attributes for analytics purposes, such as revenue or number of employees.



By improving data quality through validation, standardization, and enrichment, you can gain a full, holistic view of the entity in order to make the best business decisions.

3 Common Ways to Enrich Data

There are many ways companies can enrich their data. Below are three of the most widely used approaches:



Firmographic Data Enrichment

Companies use firmographic enrichment to update or add basic—yet critical—data elements to company records including name, location, industry, company size, number of locations, years in business, and ownership type. B2B companies often use firmographic data for sales and marketing, allowing them to better segment and target their customers.

Demographic Data Enrichment

Demographic data enrichment focuses on updating attributes of people, including consumers, patients, and students, for example. Using demographic data enrichment, companies can correct or complete fields such as age, gender, income, education, marital status, and ethnicity. Companies use demographic data enrichment to tailor campaign messaging, target relevant audiences, or better understand customer behavior.

Geographic Data Enrichment

Geographic data enrichment refers to the enrichment of data related to addresses, postal codes, coordinates, and geographic boundaries. Companies use geographic data to deliver location-based services as well as mapping and navigation functionality.

Why Do Businesses Use Data Enrichment?

Data enrichment helps companies create and realize business value in a number of ways, including use cases such as:

- **Customer analytics:** With data enrichment, you can better understand customer behaviors and target the right prospects. For example, by using firmographic data, employment growth, and technologies that companies use, you can identify ideal B2B prospects and their lookalikes for your products and personalize your offers to them.
- **Improved decision-making:** Data enrichment adds additional context from external sources to existing internal data sets, enabling organizations to make faster, smarter, and more confident decisions.
- **Duplicate reduction:** With data enrichment, you can preserve data integrity by preventing bad or duplicative data from entering your systems and corrupting your records.
- **Risk management:** Data enrichment enables you to identify key operational, financial, and reputational risks by identifying who you are dealing with. For example, you can avoid supply chain disruption by identifying the parent companies and seeing if they are sanctioned or in certain geographical locations that expose you to certain risks.
- **Regulatory compliance:** By enriching existing data with trustworthy external information, your organization can better meet regulatory requirements, reduce the risk of errors, and avoid costly penalties.

How to Overcome Obstacles to Data Enrichment

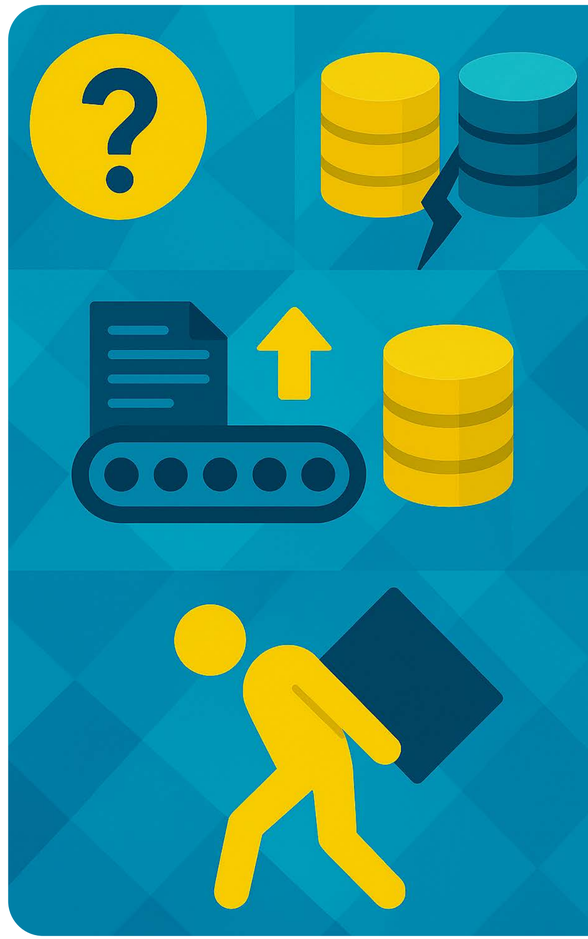
Enriching your internal data with data from external providers is a surefire way to improve its quality. However, many organizations encounter four primary obstacles as they look to realize value from external data.

1. **Knowledge of data availability:** It's difficult for data teams to gain a basic understanding of what data is available. Why? Because the ever-expanding volume and variety of external data are expanding quickly, making it difficult to keep up with the pace of change in the data sources. In addition, vendors only share a subset of the data prior to purchasing, which may be polished—and misleading.

2. **Fragmented data sources:** Data teams often struggle with the numerous data sources available through a multitude of channels and providers. And the sources often come with heterogeneous formats and schemas to clean up as well.

3. **Operationalizing external data:** In order to efficiently use external data in operations and analytics, organizations may need to conduct a major upgrade to their existing data environment or workflow.

4. **Resource burden:** Adding new sources can strain internal data teams, requiring them to allocate extra time and effort to integrate, validate, clean, and maintain incoming records.



To overcome these obstacles, we believe there are five key questions you must ask—and answer—to realize the most value possible from external data.

1. What data attributes do you need—and how do you identify them?
2. How do you assess the sources you need?
3. How do you integrate the third-party data with your own internal data?
4. How do you get value—and how do you measure it?
5. How do you make the process continuous to ensure you're always gaining value?

Let's take a look at each one in detail.

1. What attributes do you need—and how do you identify them?

When it comes to enriching data, the best place to start is with the business question. It's important that you understand and assess which questions your business leaders ask the most—and where they most often get the wrong answer.

These questions form your hypothesis, helping you to determine which questions you need the data to answer. Refine and validate these questions with business users to ensure you are addressing the right ones. That way, you can identify and source the data that delivers the insights your business stakeholders need to make informed, confident decisions.

PRO TIP

Too often, companies make the mistake of talking with popular third-party providers and assuming that once they have the data, the insights will come. This is simply not true. Instead, enter these discussions with your hypothesis on which questions you need to answer. Then, ensure that the provider has the data you need to answer them. It's never a good idea to select a provider and assume you can define your use cases later.

2. How do you assess the sources you need?

Many third-party data providers talk about their quality and coverage rates. And while important, what's even more critical is the ability to link sources and attributes together.

Ask your external data providers what level of detail they can provide. And then make sure they are providing what you need. It's a smart move to conduct a proof-of-concept or match test to ensure the vendor you're working with is a strong provider of the data you need.

PRO TIP

When evaluating external data providers, don't fall into the trap of relying exclusively on the big players. After all, covering millions of companies doesn't guarantee that they will deliver the specific data your organization requires. In fact, many companies discover that niche players provide more value, especially for data sets that are smaller and more unique to your business.

3. How do you integrate the third-party data with your own internal data?

Integration of internal and external data is a critical step. And to integrate your data successfully, you need an automated process for verifying and matching data and workflows that allow you to easily clean up unmatched records.

You also need stewardship over the data so that you get the enriched records to the point that you can master them. It's a process that requires both humans and machines, working together.

PRO TIP

Avoid relying too heavily on your third-party provider's match services. Why? Because if you use multiple data providers, each one with its own match service, you may end up with different results from each. Instead, use independent verified match capabilities and ensure that your team becomes an expert at the discipline of enriching data. In the long run, having this ability, along with the right data enrichment tools and technologies, is more cost-effective for the business.

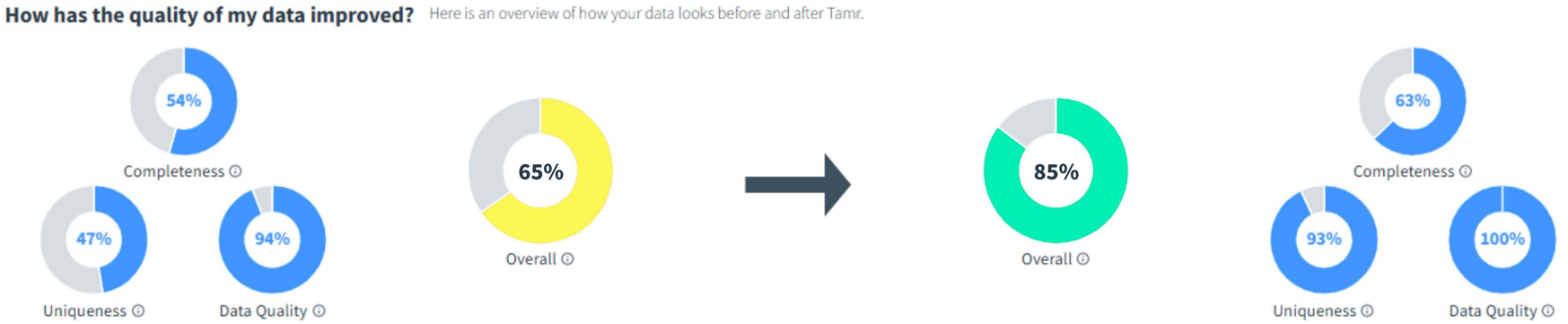
4. How do you get value from enriched data—and how do you measure it?

When you make an investment in data enrichment, it's important to show the value to your organization. By correctly defining the attributes you need, as we suggested in question #1, then you should be well on your way to gaining value from your investment. Measure how well the external data provider delivers on these attributes and the improvement in your data sources as a result.

Internal feedback is another way to assess value. Ask your business users for feedback—and ask them to be honest. That way, you can adjust your strategy if an external source isn't delivering its promised value

PRO TIP

Two overlooked measures are time savings and quality of insights. Many people in your organization are likely using web searches to enrich their data by looking up addresses or phone numbers on a company's website one by one. But this method of enrichment (obviously) doesn't scale. Using third-party providers will enable you to enrich the data at scale, saving business users time and enabling you to gain an aggregate understanding of the enrichment process. Further, ask your users for feedback on the quality of the insights the enriched data provides to them. That way, you can assess the performance of various third-party providers and determine the ones that work best for your business.



5. How do you make the process continuous to ensure you're always gaining value from your data enrichment efforts?

To ensure that you are continuously gaining value from your data enrichment efforts, there are two steps you must take.

First, set up a pipeline for continuous enrichment. Make it an ongoing process instead of a point-in-time effort. That way, you are continuously improving the quality of your data and, in turn, your analytics and operational use cases as well.

Second, be mindful of when your internal data begins to drift. Over time, new questions will arise that your business users must answer. And these questions could require enrichment from new external data sources.

PRO TIP

Because data is constantly evolving and changing, so, too, must your data enrichment strategies and providers. Keep your finger on the pulse of the business, and you'll always stay one step ahead of your business users' needs

Simplifying and Improving the Data Enrichment Process

As companies continue to spend significant amounts of time and money to derive business value from their disparate data sources, they need AI-powered data enrichment capabilities that make their data cleaner and more complete. And the best data enrichment capabilities are ones that are built into an AI-native master data management (MDM) solution.

Tamr's AI-native MDM solution provides seamless data enrichment capabilities that enable companies to tap into third-party data as a key part of their data mastering process. Companies can validate and standardize common attributes, use verified match capabilities to link internal records to a vast corpus of third-party firmographic data, and take advantage of pre-built integrations with leading data providers, enabling them to provide users with trustworthy golden records that support both analytical and operational use cases.

"As we find new data about a provider, even from external sources, Tamr enables us to now bring that data together through the mastering process and begin to enrich all of our divisions' data."

Sr. Manager of Engineering

Read the case study to learn more about this leading physician staffing company's MDM journey.

Tamr

Tamr provides the only AI-native Master Data Management (MDM) solution that delivers real-time master data for every dashboard, application, and person in your business. Tamr accelerates the discovery, enrichment, and maintenance of golden records, enabling informed decision-making, improved revenue growth, and better customer experiences.

Tamr's patented, AI-centric approach – with human refinement and oversight – delivers value in days or weeks, not months or years like traditional rules-based MDM and DIY solutions. And with intuitive Customer 360 pages, your business can improve data accessibility across the organization and leverage the best, most accurate data to support analytical and operational use cases in real time.

Learn more at **tamr.com**