Tamr

Getting Started with

DATA PRODUCTS

A CDO's Guide

The data product strategy checklist

We've all heard the phrase **"every company is a data company."** And while true, many companies are failing to derive value from their data.

There are many reasons why this is the case. Some companies struggle to **integrate data across siloed systems[1]**, making it difficult for them to gain a holistic view of their customers or suppliers. Others are dealing with dirty, incomplete, out-of-date data that leads to faulty decision-making. While still others are trying to solve for all of the above.

We get it. Deriving value from data is hard work. But companies that are consistently providing the clean, curated, continuously-updated data their business needs know a secret: they treat data as a product.

"Large Fortune 500 companies used to think of themselves as product companies, whether they be a CPG or a car company. But now they're realizing [that products] are becoming commoditized. It's difficult to compete just on product. And it's difficult to compete just on brand. As a result, they're becoming data companies. They are all trying to build unique data assets that they can leverage to build out their businesses and grow."

Justin DeBrabant, SVP of Product at ActionIQ

Treating data like a product means implementing a data product strategy that brings structure to the ownership, processes, and technology needed to ensure their organization has clean, curated, continuously-updated data for downstream consumption. Data product strategies define key objectives and metrics, such as increasing competitiveness by **improving the customer experience**[2] or creating product differentiation, and deliver value by helping companies to drive growth, save money, and reduce risks.

Organizations with data product strategies also employ **data product managers**[3] who design, build, and manage the cross-functional development of a data platform, or a suite of specific data tools, in order to serve multiple internal and/or external data consumers.

Companies implement data product strategies through the design and use of a data product. But what, exactly, is a data product?
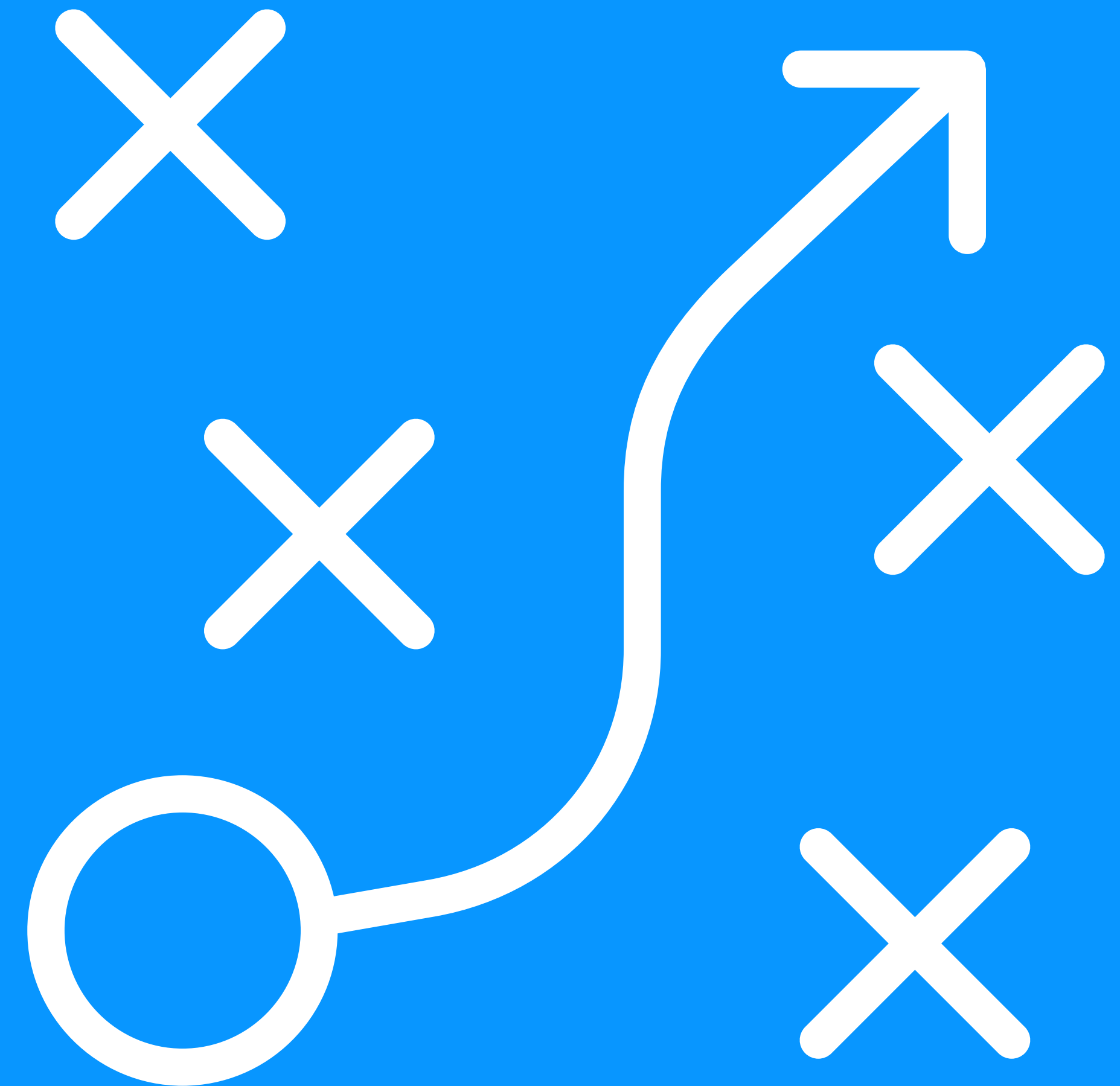
Data Product, Defined

Many organizations are unclear what, exactly, a data product is. Chances are, if you ask three different people to define what a data product is, you'll hear three different definitions. We're here to set the record straight.

A data product is a consumption-ready set of high-quality, trustworthy, and accessible data that people across an organization can use to solve business challenges. Organized by business entities and governed by domain, data products are the best version of data. They are comprehensive, clean, curated, continuously-updated data sets, aligned to key business entities, that both humans and machines can consume broadly and securely across an enterprise.

Said differently, data products make data tangible for the organization. A product is a familiar concept. It's something that the consumer - or data consumer, in this case - needs. It serves a purpose that they've defined and aligned it to an outcome they want to attain, and therefore they deem it useful. And because it's concrete, recognizable, easy to find, and easy to use, they are more likely to realize value from it.

"With a data product, you're not aligned so much to your organizational design. You're aligned to the operating outcome you're trying to achieve. We always have these giant org charts and everybody reports to their master. But products don't get made that way. Products get made by having the right people and the right skills at the right time to iterate, iterate, and iterate, and to make the product of use to the customer who asked for it."

Jennifer Agnes,
US Managing Director, Cynozure

Data product *[dey-tuh prod-uhkt],*
noun: *a consumption-ready set of high-quality, trustworthy, and accessible data that people across an organization can use to solve business challenges.*

Data Products In-Depth: Market Data Linkage at a Global Investment Firm

A leading investment firm that builds and invests in internet, software, and technology-enabled companies realized they needed better insights into market and company trends in order to outperform their competitors and, ultimately, attract more capital from outside investors. Their first step was to assemble a data products team to develop the infrastructure and processes to deliver on their objectives. The newly-formed team onboarded multiple next-generation applications, such as Snowflake, Fivetran, and dbt, but quickly realized they needed a scalable solution for data mastering to help them connect a growing list of alternative data sources.

The firm selected **Tamr's Market Data Linkage**[4] data product template because it enabled the firm to gain complete, accurate, and continuously-updated views of portfolio companies and potential investments with minimal effort from their scarce data engineering resources. The built-in data quality services and entity resolution models, built using Tamr's **patented**[5] machine learning-based technology, gave the team a high degree of confidence that the data product can scale as they acquire new, alternative data sources. And, it also gave the firm's partners confidence in the insights they use to invest.

Highly-scalable and SaaS-based, Tamr Market Date Linkage automated 99%+ of continuous entity linkage activity and provided reusability across new data sources and data products, enabling the firm to find better deals and attract more external capital.

What Makes a Good Data Product?

As is the case with many products, not all data products are created equal. But good data products embrace four principles.

1 **Discoverable and interoperable** using intuitive user and machine interfaces that support real-time operations

2 **Comprehensive and consistent** by aligning to a domain-specific, universal schema and enriching data with external data sources

3 **Clean, accurate, and reliable** with validation and cleaning based on global standards and version control for monitoring and provenance

4 **Curated and continuously-updated** using pre-built, machine learning-based matching models

Within these four principles are several key capabilities that define every good data product. Let's take a closer look.

KEY CAPABILITIES OF A GOOD DATA PRODUCT

Curated with feedback from end users

Data products drive usage of the data by people across the business. Using an intuitive user interface, data consumers across the broad organization can engage with data, evaluate the quality of data, give feedback on data, and provide expert stewardship, all of which create trust in the data. This interface also promotes discoverability, enabling people to find the data they need, when they need it.

Machine interfaces to support real-time operations

Machine interfaces (APIs) support real-time interaction with the data from operational applications. These APIs enable organizations to publish consumption-ready data to analytical endpoints, including changes in values, changes in attributes, and new records. And because data products are continuously-updated, they contribute to greater usage of data across the organization.

Aligned to a domain-specific, universal schema

When you align data products to a domain-specific, universal schema, you benefit by providing a common frame of reference for your messy source data that aligns to your organization's common understanding of the entity. For example, your organization may require that all customer records have a name and a social security number.

Enriched with external data sources

It's highly-likely that much of your source data is missing critical data attributes. Good data products use data enrichment through pre-integrated data sources to both fill in the gaps in your internal data as well as add additional data attributes that didn't exist in your source data, enabling you to improve analytics use cases and drive operational processes.

Validated and cleaned based on global standards

Creating consistency in your data is critical because consistency is what makes it easier for both machines and people to use it. Data products that utilize data quality services are able to automatically clean and normalize the data which, in turn, promotes higher usage.

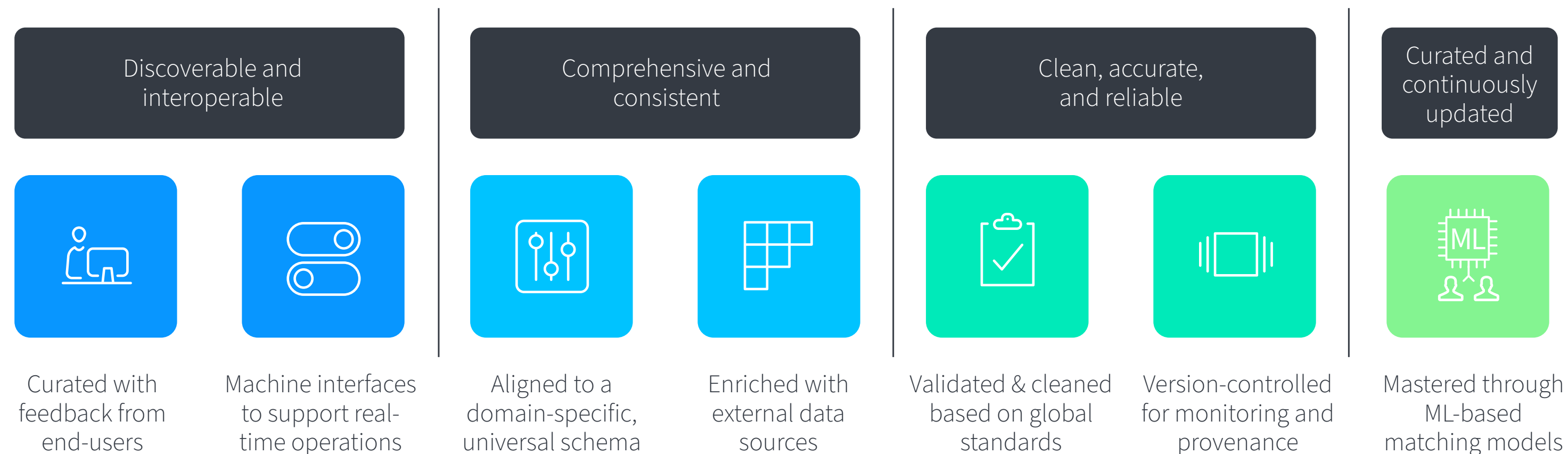## Version-controlled for monitoring and provenance

It's common practice for multiple versions of data to exist across your business. Using data version control, you'll be able to see data changes and provenance in your data products over time, allowing you to "roll back" data in the event of a disaster or other circumstance and to view data changes and improvements over time.

Mastered through machine learning-based matching models

Good data products leverage pre-built machine learning models for data mastering. Machine learning models do the heavy lifting when it comes to cleaning, curating, and enriching data, which frees up your human resources to focus on answering the most challenging data questions. This approach is a more effective use of your human resources' time as it leaves the yeoman's work to the machine.

Data Product Definition

A "data product" is a consumption-ready set of high-quality, trustworthy, and accessible data that people across an organization can use to solve business challenges.

| Discoverable and interoperable		Comprehensive and consistent		Clean, accurate, and reliable		Curated and continuously updated
Curated with feedback from end-users	Machine interfaces to support real-time operations	Aligned to a domain-specific, universal schema	Enriched with external data sources	Validated & cleaned based on global standards	Version-controlled for monitoring and provenance	Mastered through ML-based matching models

Recap

Simply put, data products make data tangible for your organization. They are consumption-ready sets of high-quality, trustworthy, accessible data that people across an organization can use to solve business challenges. And, they embody four key pillars and several critical capabilities in order to deliver the clean, accurate data needed by data consumers.

Getting Started with Data Products: A Data Product Strategy Checklist

Knowing where to start is the hardest part of implementing a new strategy. But when it comes to getting started with data products, there is a clear blueprint for CDOs and aspiring data leaders like you to follow. Below, we're sharing five steps to help you get started. Follow them, and you'll be on your way to implementing a successful data product strategy for your organization.
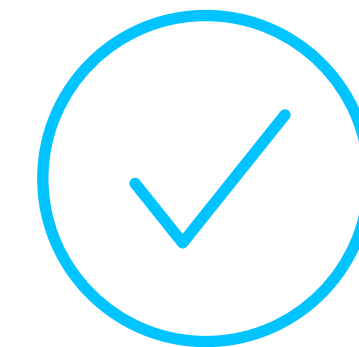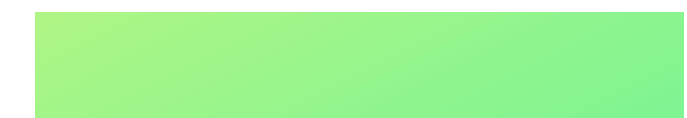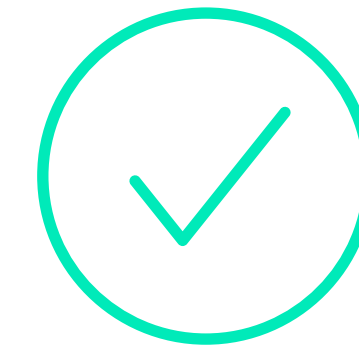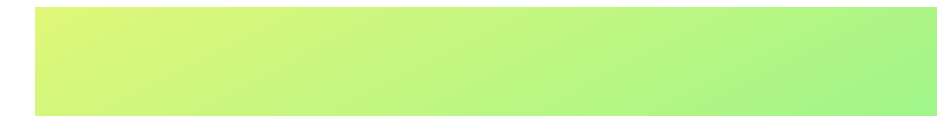
1

Know Your Why

Do you know why you are creating a data product strategy? If not, defining your objective is a good place to start. After all, if you don't know why you are creating a data product strategy, then it will be difficult to gain traction with it.

When starting out, it's best to start small. Identify a clear, specific objective that aligns to business priorities. For example, does your organization need to increase revenue, control costs, or manage risks? Knowing why you are creating a data product strategy – and what you are looking to achieve – will focus your effort and align

your team, the lines of business (LOBs), and your leaders. It's paramount because every decision you make from this point forward will aim to fulfill this objective. And it will help you gauge success as you implement your strategy.

If you are unsure where to begin, customer data is a good place to start. Building a customer data product enables you to increase visibility into your customer journey so you can improve the customer experience, drive cross/up-sell opportunities, improve targeting, and increase conversion rates.

why?

2

Assess Your Data, Your Organization, and Your Technology

Now that you've defined your objective, you need to determine if you have the right capabilities in place to implement a data product strategy. On the data side, it's important to understand:

• Where does the data live?
• Is it integrated across systems and departments?
• Is the data accurate and complete?
• How often is it updated?

Understanding the answers to these questions will help you determine not just the quality of your data, but also how much budget you need and the number and types of resources it will take to build a high-quality data product.

Next, review your team. You may also find that you need to add new skill sets to fill gaps in your organization. Many

organizations identify the **need for a data product manager**[6] to design, build, and manage the cross-functional development of a data platform and serve multiple internal and/or external customers. Identify resource gaps now, and make a plan to fill them.

It's also important that you have the supporting technology in place to deliver on your data product strategy. Your technology should enable you to easily **master your data**[7], **enrich**[8] it with external datasets, and integrate it across your systems and departments. It should blend pre-built, machine learning models with human feedback so that your data product delivers the best possible version of data. Without these capabilities, you'll struggle to deliver analytical insights or drive operational efficiencies.

Define Your Use Case

Now that you have an objective and have assessed the state of your data, your organization, and your technology, it's time to roll up your sleeves and dig in. It's likely that the leaders of your lines of business, e.g. marketing, R&D, or procurement, have a problem that is preventing them from reaching their respective objectives. And they may – or may not – realize that data can help. Defining a use case is the best way to create a vision for how a data product can help them solve a business problem and reach their objective. It brings clarity to the work you're doing, and makes it tangible for the lines of business leaders

To determine the requirements for your data product use case, you'll need to ask questions such as:

• Why do you need a data product?
• What are you going to do with the data product?
• Who is going to use it?
• Where will they consume it?
• What data does it need to include?

Your LOB leaders may not know all the answers, but that's where you come in. Connect the dots, identify the data sets, and show them how a data product can help them.

4

Secure Buy-in and Budget

At this point, you need to secure leadership buy-in and budget. Without it, your project will struggle to secure the funding, resources, and support needed to succeed. Share your defined use case and the roadmap for how you will deliver on it. Demonstrate not just how you will initially show value, but also how you will iterate on your data product to make it even more valuable.

It's also important to define how you will measure success. Work cross-functionally to build or strengthen partnerships and to define KPIs and success measures so that everyone is working towards the same goals and objectives. Leaders will ask you about measurement and how you define success, so be prepared to answer them.

5

Develop a MVDP: Minimum Viable Data Product

Once you've secured your buy-in and budget, you are ready to begin implementing your strategy. Priority number one: developing a MVDP: a minimum viable data product.

Just like in agile product development, it's important to start small, release quickly, iterate, and prove value. By focusing on a specific use case (see step #3), you can be nimble. Each time you release a data product, deliver a few more capabilities and a little more value. Not only will it help drive adoption of your data product, but it will also help you secure additional funding, more resources, or better supporting technology.

As well, make sure that you are supporting your lines of business partners so that they understand how they can use the data product in their everyday processes. Show them how they can access approved data products via their analytics tool. Get their feedback on the quality of the data. And use it to help make the next release of the data product even better.
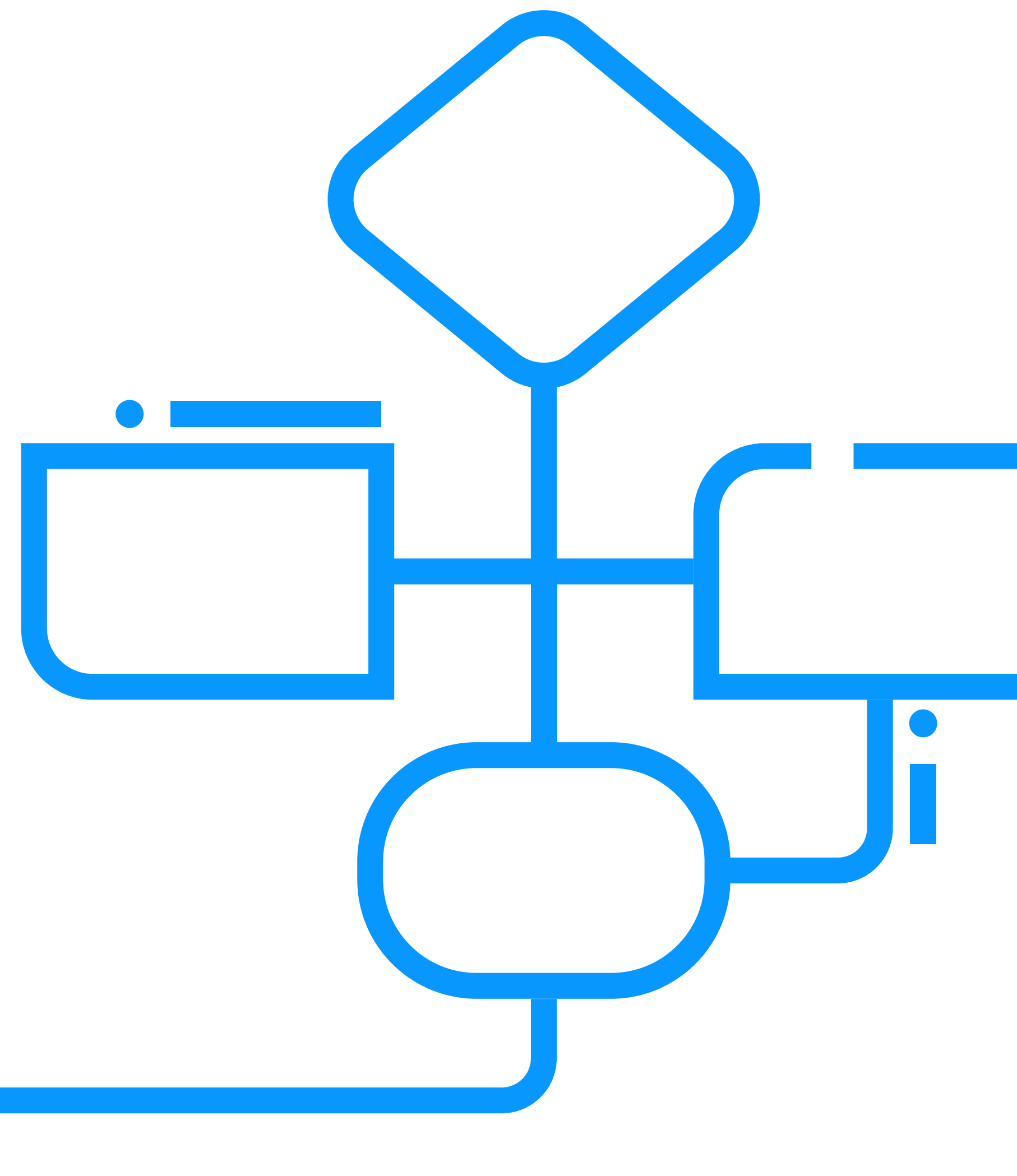
Tamr: the Platform for Data Products

Having the right technology to support your data product strategy is critical. Tamr provides integrated, turn-key data product templates that combine machine learning, a **low-code/no-code[9]** environment, and integrated data enrichment to streamline operations. With Tamr's data product templates, you can deliver a consumption-ready set of comprehensive, clean, curated, and continuously-updated data sets for key business entities.

Tamr's templated approach accelerates your time to value for key data products including:

- B2B Customer Mastering
- B2C Customer Mastering
- Healthcare Provider Mastering
- Market Data Linkage
- Patient Mastering
- Supplier Mastering
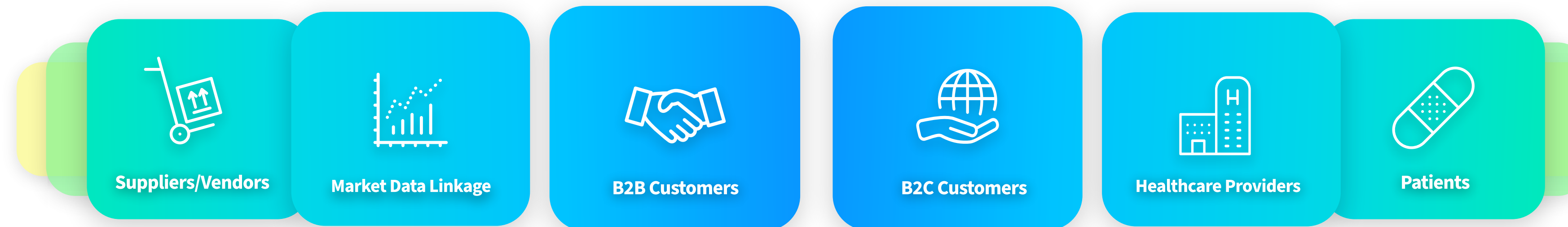
Our customers are using Tamr data product templates to drive efficiencies, save money, and increase opportunities and revenue. Let's take a look at a few examples.

| Suppliers/Vendors | Market Data Linkage | B2B Customers | B2C Customers | Healthcare Providers | Patients |

Data Products In-Depth: Supplier Mastering at Novartis

NOVARTIS

Novartis, one of the largest pharmaceutical companies in the world, has a purpose: to reimagine medicine to improve and extend people's lives.

But siloed, messy data was preventing Novartis from gaining quick and accurate access to supplier information. Millions of records across several SAP systems had different attribute names, a variety of formats, and inconsistent quality. And combining this data was a manual effort, which impacted efficiency and caused the company to incur unnecessary costs. Novartis leadership established a team to resolve these issues and deliver data products that can support a constantly-changing set of business and technical requirements.

Novartis is using **Tamr Supplier Mastering**[10], and as a result, data-driven insights are soaring across the business.

Data consumers are embracing data products and using them to reduce the time it takes to create sourcing strategies, identify changes in their supply base that require action, and, above all else, minimize the cost and risk associated with delivering critical medicines for patients.

With Tamr Supplier Mastering, Novartis was able to build a supplier data product and gain a holistic view of their suppliers across multiple source systems, giving them the up-to-date data they need to support current and future needs of the business. This data is enabling Novartis to continuously improve time to market for new products as well as improve research efficiency through access to critical lab data for researchers.

Data Products In-Depth: Customer Mastering at Analog Devices

Analog Devices is a global leader in the design and manufacturing of analog, mixed signal, and digital signal processing integrated circuits.

Analog Devices was looking to drive growth through cross-sell and upsell of existing customers. Data leaders quickly recognized that achieving this growth would require a new approach to their data infrastructure. They shifted to a data lake-based architecture which made it easier for people to get access to account and contact data, but it did not provide the single view of the customer needed to drive automated decision-making. As new leads came in, the team struggled to match them to existing accounts. In fact, only 36% of new leads matched existing companies. Slow and costly third-party vendors reduced prospect response time. And, managing a mix of customer languages became increasingly difficult as well.

Analog Device implemented **Tamr B2B Customer Mastering**[11] and experienced immediate results. In the first two weeks, the company saw a 50% increase in the match rate of sales leads to existing accounts. Lead response time improved from weeks to days, enabling the company to save hundreds of thousands of dollars.

By leveraging existing customer relationship information through Tamr B2B Customer Mastering, Analog Devices built a customer data product to gain a global view of their accounts to enable holistic contract management and inform negotiations with key accounts, resulting in more cross-sell and upsell opportunities. The data product approach not only delivered significant ROI in the short term, but also set the stage to accelerate future implementations of next-gen tools across lines of businesses, demonstrating the potential for even greater long-term value.

[11]

In Summary

More organizations are realizing that implementing a data product strategy through the design and use of data products is a surefire way for them to finally treat data as an asset and drive greater value from it. Data products elevate the value of data as an asset by making it discoverable and consumable for everyone across the organization.

"At its core, every business is a data business. Which is another way of saying every business should have data products and think about managing their product – which they might think is software or retail or healthcare. But it's not. It is, in fact, the data. And they should manage that asset like a product."

Anthony Deighton,
Chief Product Officer, Tamr, Inc.

Tamr is the data product platform, providing all the capabilities CDOs need to deliver data products. Designed for consumption and ready for use, Tamr data product templates deliver high-quality, trustworthy, and accessible data that people across an organization can use to solve business challenges. Comprehensive, clean, curated, and continuously-updated, data products deliver data sets that humans and machines can consume broadly and securely across an enterprise.
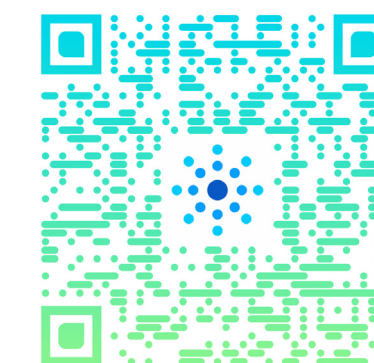
Organized by business entity and governed by domain, data products are, simply put, the best version of data.

To learn more, visit **tamr.com**.
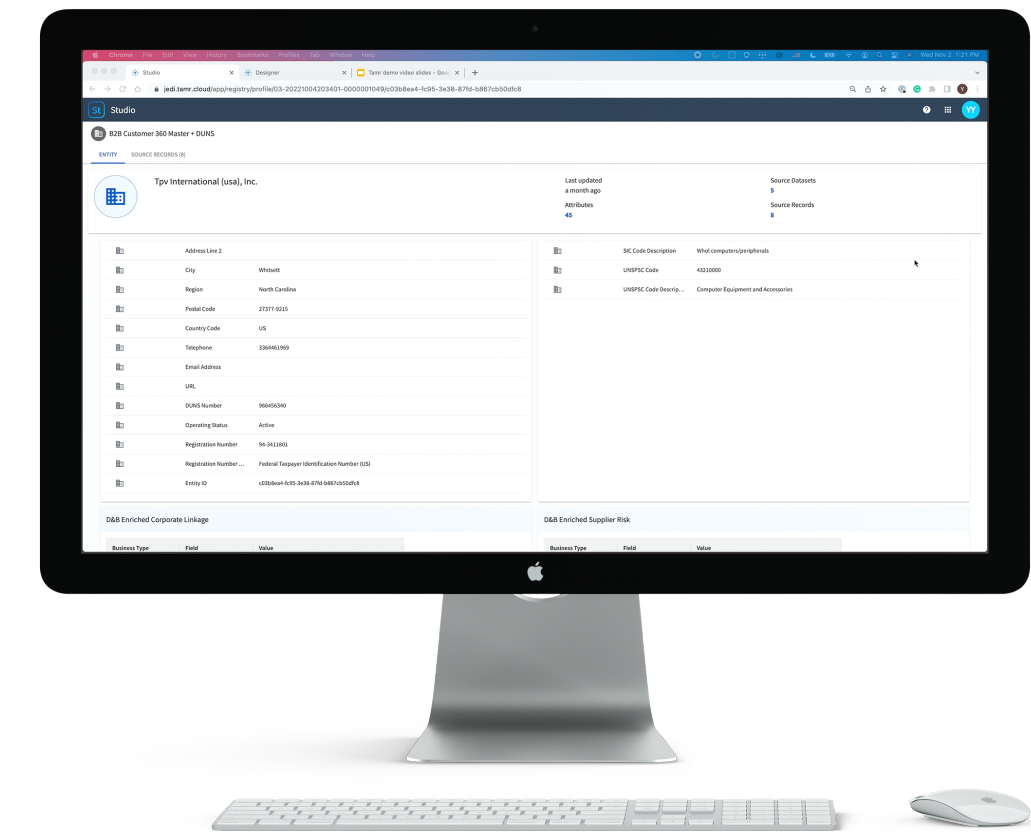
Data Products In-Depth: The Future of Data Products

As the leading platform for developing and using data products, Tamr sets the foundation for executing a successful data product strategy.

Watch the demo to see how it works.