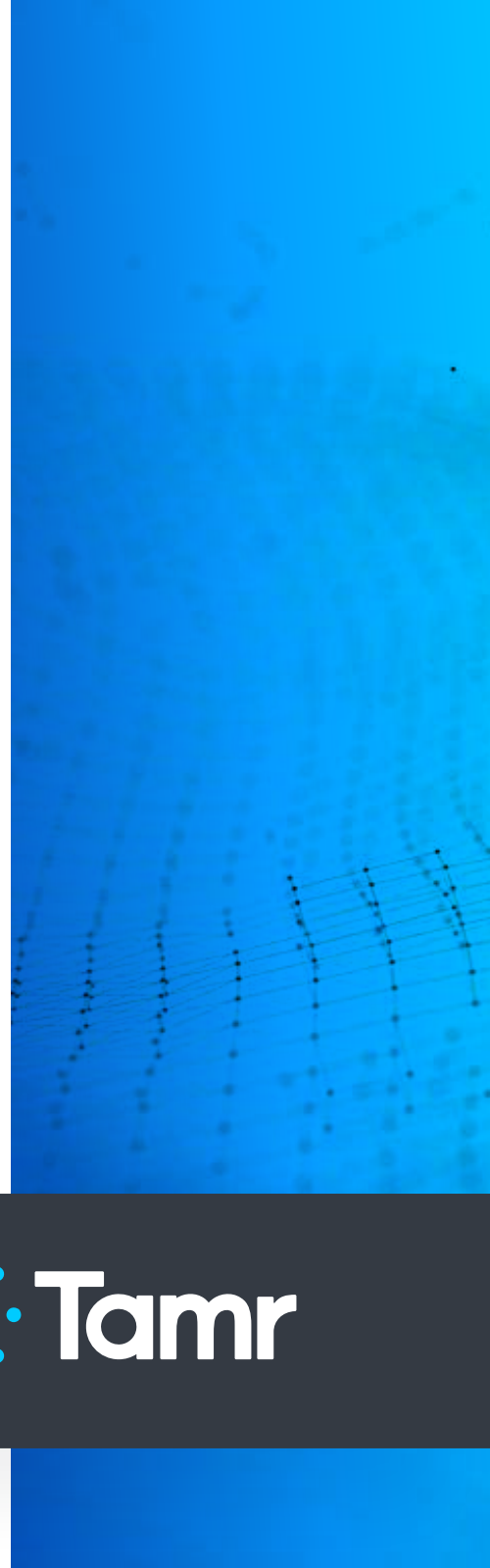
How to Modernize Your

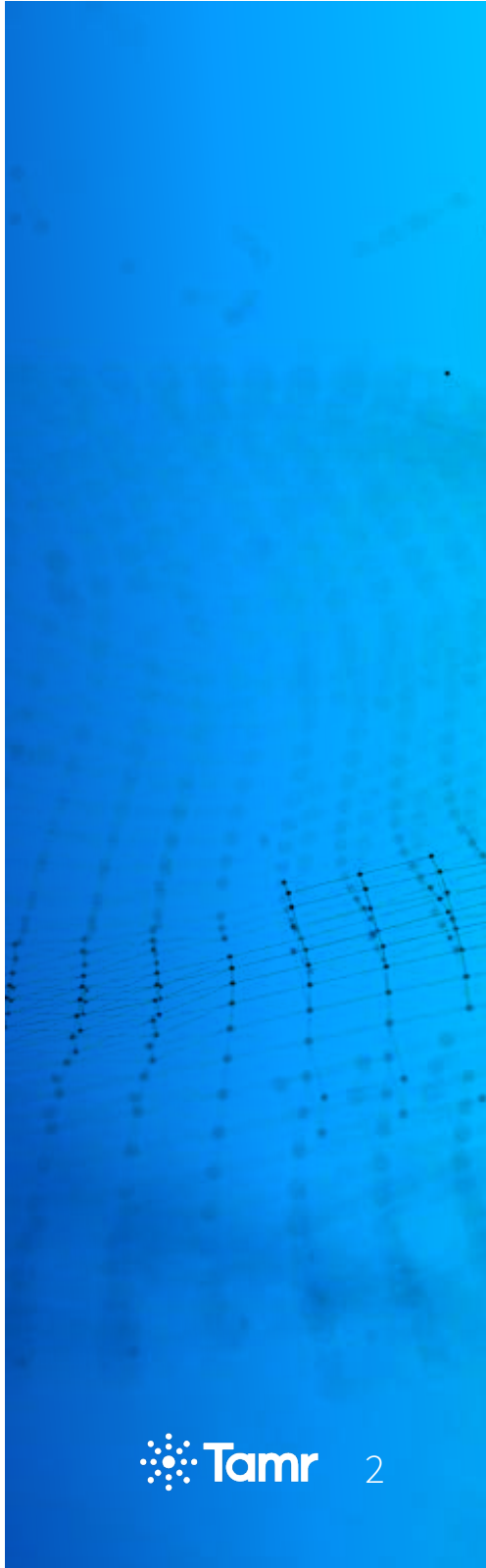
Reference Data Management Architecture

An evaluation guide for banks and financial institutions seeking an intelligent approach to managing and sharing reference data.



Tamr

Table of Contents



1

The Current State of Data Management

Financial services is a highly regulated and intensely competitive industry that requires data-driven insights to stay ahead. However, financial institutions are particularly susceptible to data management challenges as they are likely to operate in a siloed, decentralized manner using a variety of legacy systems across many regions. This siloed approach to data management restricts collaboration between departments and business units and prevents employees from making fast and effective, data-driven decisions. While this is not news, siloization is causing increased friction as banks continue to pursue transformation agendas and look to incorporate digital applications, whether home grown or external SaaS and Fintech offerings. Efforts to streamline payments, such as Straight-through processing (STP), rely on data quality.

So, how can financial institutions achieve an agile, collaborative, and integrated approach to data management? This guide will help your organization build a data architecture that eliminates silos and empowers employees to assess risk, meet compliance requirements, and make important decisions, quickly and effectively.

Why is Reference Data Management a Critical Process?

Data is a core asset for banks and financial institutions. The sheer number of external reference data sources has grown exponentially. The need to leverage these sources has also intensified.

Due to the rapid advancements in technology coupled with tighter regulations, your employees need access to a wealth of internal and external data at their fingertips. Reference data typically represents the majority of data used in financial transactions, providing key information on securities, counterparties, and corporate actions.). The number of reference data sources is growing every year. Gaining access to reference data is relatively easy. There's an increasing number of platform players in the market who attempt to offer an integrated view across third parties and expand the realm of data considered with increased provision of dynamic data. While this stand-alone data remains pure, challenges arise when this clean third-party data is combined with internal sources and a multitude of other reference data sources. Human error in transaction data entry and inconsistent internal classification makes reconciliation of data non-trivial work.

The challenge lies in creating a single source of truth for the millions of entities that comprise reference data and activating it to streamline financial transactions or support lending decisions. If you're searching for an intelligent approach to managing and sharing reference data, you've come to the right place.

Active management of reference data is critical given the extent of touch points across the organization. Ideally, data is validated and amended pre-transaction, whether it's internal traders or client information. When errors do occur, it's important to be able to reconcile the data as quickly as possible. Faulty reference data is a major source of operational risk. Financial institutions need to ensure the quality and integrity of data so that it becomes a reliable foundation on which multiple departments operate. Reference data use has multifaceted use cases across any financial institution- enhancing customer and client information, facilitating payment workflow and trade execution, and reducing risks by contributing to a complete audit trail. Inconsistent terminology is one key source of error, most frequent when silos exist among departments or acquisitions have led to large scale data integrations.

Before we dive into the remedy to eliminating data chaos when it comes to reference data management, let's explore the areas where financial institutions tend to stumble.

2

Three Reasons Reference Data Management Projects Fail

Reason #1 — Prioritizing Data 'Quantity' Over Data 'Quality'

I don't know what they want from me
It's like the more ~~money~~ 'data' *we come across*
The more problems we see

(Source: Christopher Wallace aka Notorious BIG)

Many financial institutions are drowning in data but starved of insights due to the poor quality of data. Often despite a data architecture plan, institutions fail to gain a truly unified view of their data across systems. Bringing external data sources into your systems will only benefit your organization if you can consolidate external data with internal data. Effective data management starts with clean, accurate internal data. If your data has been polluted by duplicated, stale, erroneous, and dirty data, layering reference data on top won't save your organization from drowning in a sea of unactionable, untrustworthy data. More data will create more problems unless your organization prioritizes data quality.

2

Reason #2 — Taking a 'Rules Only' Based Approach to Data Consolidation

Reference data management is a messy, inefficient, and labor-intensive process when handled by rules based systems. Exceptions are not as exceptional as they should be. When resolution involves manual handling, the lessons learned are often not captured in a systematic way. When organizations allocate labor resources to manually reconcile data, innovation within its employee base slows down. Many projects get shelved until resources are freed up as bandwidth is spent on time-intensive manual checks. Imagine having an innovative idea that could give your organization a competitive edge... If you're relying on humans to perform manual data reconciliation, your idea could take years to deliver or worse, risk being stalled because those resources are needed elsewhere. Tamr's machine learning solution can handle 100s of sources and integrate new sources as quickly as 3-4 days. Data challenges are the main reason innovative projects fail, automating data consolidation will significantly contribute to the success of your initiatives.

2

Reason #3 — Build Vs Buy. The Great Debate.

‘Machine Learning Applications? We Can Build Those…’

Empowering employees with the accurate and enriched data needed to make fast and effective lending decisions should be a top priority. Building an in-house machine learning solution to establish your data foundation? If you’re working with a small number of clean datasets, a home-grown solution will suffice. However, to truly become a data-driven organization, the need to consolidate a large volume of highly variable datasets—both internal and external—grows exponentially. And while some off-the-shelf ML libraries can solve discrete data consolidation problems for a chunk of the data w/in the sources, the long tail is critical.

Tamr’s data mastering technology is built to scale to millions of records across hundreds of sources, and can account for a broad set of business-critical different entities (e.g., customers, lenders, transactions and organizations). Another area often unresolved by homegrown solutions is the ability to engage data experts within the organization effectively. For example, who is going to review edge cases for ML matching or rules-based systems? How will you effectively scale the

2

operations of engaging other team members in the review process? A critical component of driving over 90% accuracy of data and accelerating data mastering projects to days and weeks instead of months and years, are intuitive feedback workflows. Tamr's data feedback workflows—yes or no questions built in-app to accelerate ML accuracy and speed to insights—are designed to bridge the gap between data challenges and business outcomes. These workflows enable data experts within the organization, or data curators, to speed up the process of consolidating large volumes of highly variable data to make it analytics-ready. The ability to "crowd-source" improves accuracy of outputs and promotes higher adoption of data as an asset.

Finally, as teams consider building in-house, it's important to consider the architecture and interoperability needed to truly streamline reference data management. Does the solution leverage best-of-breed scale-out and compute capabilities? Does the solution deploy in the cloud leveraging cloud-native capabilities to lower costs? Tamr can also augment existing MDM (Master Data Management) investments to improve speed and scalability while maintaining quality. With robust APIs and cloud-native and hybrid deployment capabilities,

2

Tamr can integrate up-to-date complete data seamlessly with new or existing pipelines and scale effectively as needed. When organizations choose a vendor who will act as a true partner – with data science expertise and a track record of delivering curated data quickly that is outcomes driven, the result is faster time to insights that truly grow the business.

3

What to Focus on Based on Your Organization's Business Objectives

Start by asking: What is the shared 'vision' for data at your organization?

Here's an example of three objectives your organization could achieve in a matter of months:

Increase efficiency by reducing extensive manual data work. Release technical talent from the burden of manual data preparation and free up time for high-value projects.

Enhance data quality by eliminating dirty, duplicate and de-centralized data and create a scalable model that can be leveraged for both front and back office tasks, from lending and trading to compliance and risk management.

Align key business and technical stakeholders around a vision that data is a powerful resource shared between departments which should be usedto make better decisions

4

Key Considerations When Evaluating New Solutions

Agility

Can my organization deploy this solution quickly to deliver quick wins?
Will this solution allow us to iterate over time as we learn and grow?

Cloud Capabilities

Is this solution cloud-native?
Does this solution have enough power to rapidly process a large volume of data from hundreds or even thousands of sources?

API Integrations & Documentation

Does this solution have robust APIs needed to build a scalable and flexible architecture?
Will my developers have access to comprehensive documentation needed to implement and maintain the solution?

4

Time to Value

Does this vendor have a solid track record of delivering successful projects on-time?

What does the onboarding process look like and what kind of training will be required to enable internal teams?

Overcoming Human/Behavioral Challenges

Employees within your organization will either embrace or fear change based on how they interpret this reference data modernization initiative. Can the vendor demonstrate how they have helped existing customers overcome behavioral challenges?

Measuring Success & Aligning with Business Priorities

One of the most important questions you should ask yourself and your team is what business priorities does this initiative align with? Which business priorities does this initiative align with? Driving significant improvements in the quality of lending decisions, regulatory compliance, customer retention/acquisition, and spend reduction are examples of the priorities that Tamr customers often align with. Data quality, accessibility, time to insight, project delivery speed and cost savings are just a few of the metrics used by Tamr customers to measure value.

5

Customer Spotlight: Santander UK

Last year, Santander UK set out to implement a new credit lending system, nCino, which required a single view of each retail customer in order to accurately calculate financial exposure.

Before selecting Tamr as their preferred vendor for Data Mastering, customer records at Santander were stored in four different systems that weren't interoperable – a result of past acquisitions. The bank had thousands of duplicate customer records across different systems and no process in place to clean and maintain the data so that it could be used to drive business outcomes.

Creating a consolidated, coherent single view of each customer allowed Santander to:

1. Implement a new credit lending system, with reliable input data, and decrease lead cycle and credit decision times by more than 50%

5

2. Increase the reliability of risk and regulatory data by providing a consolidated customer view that can contribute to financial reporting, credit risk and regulatory reporting

3. Better understand credit risk profiles and customer propensity, as well as the input for analytics and visualization software

"Data mastered by Tamr underpins the entire digital journey and data architecture – if we didn't have a single customer view and weren't able to surface data from a myriad of product systems, including legacy and acquired systems, we wouldn't be able to operate in a digital format where we're putting that data in front of customers or automating processes and decision making"

Click here to watch the full webinar on-demand.

6

Three Key Takeaways

Modernizing your reference data management architecture can drastically transform how your organization uses data to deliver successful business outcomes.

When you're seeking a more intelligent approach to managing and sharing reference data here's the top three areas to focus on:

1. Quality > Quantity When it Comes to Your Organization's Data

Having a reliable solution that can integrate internal and external data sources at scale is critical. Relying on human-intensive processes or building solutions in house might not be the best use of your organization's resources.

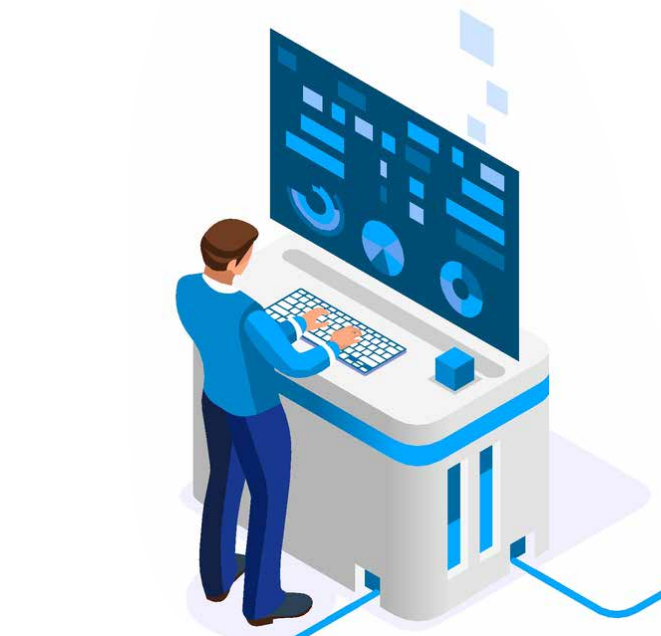
2. Focus on Clear Objectives and Align with Business Priorities

Establish the shared 'vision' for data at your organization. Identify short term results that can be delivered in a matter of months while building a long term strategy supported by robust solutions that will evolve with your business over time.

6

3. Don’t Overestimate What a Single Vendor or Piece of Software Can Do

Your organization should avoid relying on a single vendor at all costs. Successful digital transformation requires a thoughtfully designed ecosystem of loosely coupled ‘best of breed’ tools. Align your vendors with APIs and expectations that they MUST work together.



Next Steps

Do you aspire to a centralized reference data management model like those above? It takes a few steps to get there, but it's achievable. Connect with Tamr to get started.

To learn more about Tamr, please visit www.tamr.com or contact us to schedule a meeting and a demo.

SCHEDULE DEMO

Tamr

About Us

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 cofounding 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**