Say Goodbye to Dirty Dashboards

Analytics Leaders Guide

Tamr

Duplicate records showing up in dashboards? Stale customer address information leading to errors? Missing attributes limiting the view of the customer? Here's a step-by-step guide on how to deliver comprehensive, clean, curated data for a single view of customers for accurate analytics.

Clean, holistic data is the foundation for good customer analytics. While the field of analytics continues to advance, and customers' expectations of cutting-edge digital experiences are higher than ever, the reality is that most organizations are held back by dirty data. The statement that '80% of a data scientist's time is spent on data cleaning and preparation' now feels like an age-old saying. It's time that changed. As companies modernize their data stack and continue the migration to the cloud, there is new potential to tackle the unsolved problem of the single view of the customer.

While the cloud ecosystem has transformed infrastructure management, attention can now turn to the usability of the data itself and its ability to feed analytics. The cost benefits of cloud migration are undoubtedly compelling - elastic compute has provided many with much-needed flexibility to scale up and down resources with the business's needs. However, the more interesting value proposition, and the fuel for digital transformation, comes from transforming the usability of the data itself and putting it to work to solve business questions. The aggregation of data has unleashed the

potential for advanced analytics as data silos are broken down. The next step is to master the data to ensure it's usable - clean, curated, and complete - before it reaches analytics tools and data scientists. Mastered, clean, accurate data is the cornerstone upon which digital customer experiences and advanced analytics are built.

Common customer data challenges to tackle for analytics include:

Inability to match records across data sources
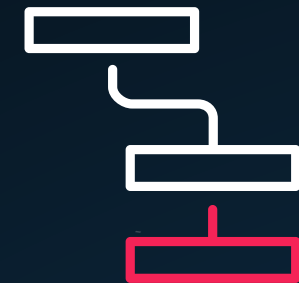
Errors in customer records

Stale data on key attributes such as address and telephone

No unique identifier to consistently refer to the customer

Lack of lack trust in the data from users

We'll provide a step-by-step overview of how to tackle these challenges through a next-generation approach to customer data mastering.

Customer Mastering for a Single View of Customers

Customer mastering enables true customer 360 - the creation of a unified, accurate, enriched view of customer data across systems and sources and a unique identifier so the customer can be consistently tracked. Mastered customer data can feed both operational and analytical systems, from CRMs, ERPs, and marketing automation tools to business intelligence tools and dashboards.

The goal is to ensure trustworthy, up-to-date, holistic information is driving decision-making for customers. On the right is an overview of how disparate records are transformed for a single, up-to-date view of a B2B customer.

1 IDENTIFY DUPLICATES

Apple
Cupertino CA
95014-0642
salesforce

Apple USA
1 Apple Park Way
salesforce

Apple Inc
Santa Monica
CA
HubSpot

Apple Inc
Cupertino

2 CREATE GOLDEN RECORDS

Apple Inc

1 Apple Park Way
Cupertino,
California, 95014-0642
Tamr ID 746342

3 ENRICH

Apple Inc

1 Apple Park Way
Cupertino,
California, 95014-0642
United States
Geocode: AFR93
(800) 692-7753
https://www.apple.com
Tamr ID 746342

4 UNDERSTAND HIERARCHIES

Apple Inc

1 Apple Park Way
Cupertino,
California, 95014-0642
United States
Geocode: AFR93
(800) 692-7753
https://www.apple.com
Tamr ID 746342

Apple Pty Limited
366 Geor Street,
Sydney
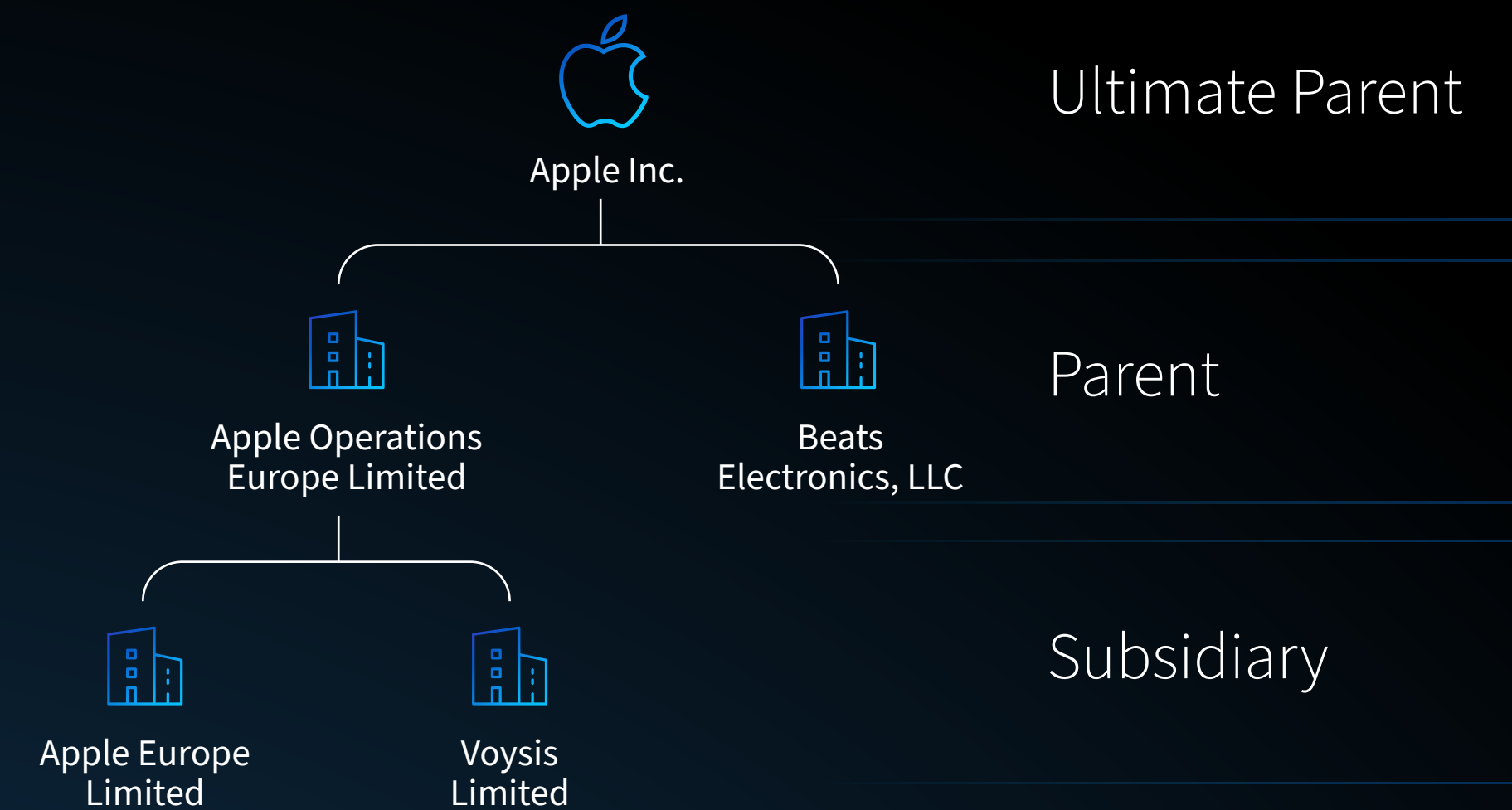**New South Wales 2000,**
Australia

Apple Retail Europe Ltd
Hollyhill Industrial Estate,
Hollyhill,
Ireland

Checklist for Clean Customer Data for Analytics in the Cloud

1   Understand the business use cases to accurately define the customer

2   Map schema to unify data sources

3   Incorporate third-party external data sources

4   Match records with a machine learning approach

5   Gain user feedback to ensure data is in a usable form for analytics

6   Enrich the data to ensure critical attributes are up-to-date

7   Create a unique ID to consistently identify customer records across systems

1 Understand the business use case to define the customer

Defining customers at the right level of granularity is complicated. For B2B companies, customers can be classified by site, country, region, contracted entity, legal entity, or corporate parent. While there is often variation across business units from sales to finance to marketing, it can also vary within departments - how sales classify customers for team coverage analysis might be different from how marketing classifies customers to calculate CAC. Aligning upfront how to define customers for a single view can save time later and better ensure the output is practical for use for analysis. A data matching solution needs to be robust to support business needs and potentially multiple definitions of customers, from sites and sales regions to corporate hierarchies, as various analytics questions arise. It's critical to link accounts at an overall hierarchy level that considers both contact and corporate level information.

Apple Inc. — Ultimate Parent

Apple Operations Europe Limited — Parent
Beats Electronics, LLC — Parent

Apple Europe Limited — Subsidiary
Voysis Limited — Subsidiary

Apple — Account

Apple US — Country
Apple Ireland — Country

Apple Cupertino — Site
Apple Austin — Site
Apple Cork — Site

2 Map schema to unify data sources

The complexity of multi-channel sales and the growth of source systems for customer data have increased the need for sophisticated unification of data. While CRMs and ERPs continue to form the foundation for customer data, as the proliferation of data across the enterprise continues, the list of data sources is long and growing - from marketing automation software to online portals to adhoc CSV files. Those who have moved data to the cloud have largely tackled the data aggregation challenge. Data unification or integration is the next critical step to ensure the data can be put to use. Companies need a robust approach to mapping database schemas across a wide range of ever-changing sources and matching customer records within them to create a single view.

salesforce

HubSpot

Marketo

ORACLE

SAP

Microsoft Dynamics

eloqua

3 Incorporate third-party external data sources

The importance of integrated data pipelines extends to third-party data sources. There is no shortage of external data providers today and most companies use some form of third-party data to supplement their internal sources. Popular providers of external enrichment data include D&B, ZoomInfo, Companies House, and GLEIF. Unless external sources are integrated into pipelines, it can further fuel an unmanageable data swamp. For companies looking to get ahead, it's critical to ensure an integrated pipeline for unifying external and internal sources so that the data is in a usable form and a robust, 360-view of the customer is created.

zoominfo

dun & bradstreet

LEI GLOBAL LEGAL ENTITY IDENTIFIER FOUNDATION

4 Match records with a machine learning approach

To accurately match customer records at scale and carry out the mastering process efficiently, a machine learning approach is needed. Traditional approaches often rely on coding rules that quickly break down at scale and require high manual effort and monitoring to maintain. Human-guided machine learning is centered on providing feedback on machine-generated recommendations and has been shown to produce over 90% accuracy on record matching, as compared to 60-80% with rules-based systems. However, not all machine learning solutions are created equal. Machine learning engineering is one of the most sought-after fields in the industry and a build-your-own approach can quickly blow through budgets. Machine learning algorithms should be purpose-built for the data unification task. Tamr has 19 patents on topics from scalable Big Data Deduplication to Cluster-Centric Entity Resolution and data issue reporting

JP Morgan Chase or Chase Bank?

One of the most common, highly visible challenges is that customer names are often spelled differently:

JP Morgan Chase, **JPMorgan Chase**, **JP Morgan**, **JPM**, **Chase Bank**, **JPMorgan Chase & Co**.

It also requires using often imperfect information and judgment to decide if two customer entities are connected.

5 Gain user feedback to ensure data is in a usable form for analytics

Humans should be involved to guide the machine, and build trust in the data through 'bottom-up' feedback. Analytics leaders can't compromise on this point if they want usable data results. The traditional approach to cleaning and unifying data of 'top-down' rules typically ignores the context of how data is consumed. Users should not only be involved in the design stage but also review the model results before the data is fed back to a data warehouse or system. As every analytics leader knows, trust in the data is key. Gaining efficient input and timely review allows users to take ownership of the data. By involving human input at critical points in the data pipeline, and leveraging the feedback efficiently through machine learning, overall manual effort can be reduced by 80%.
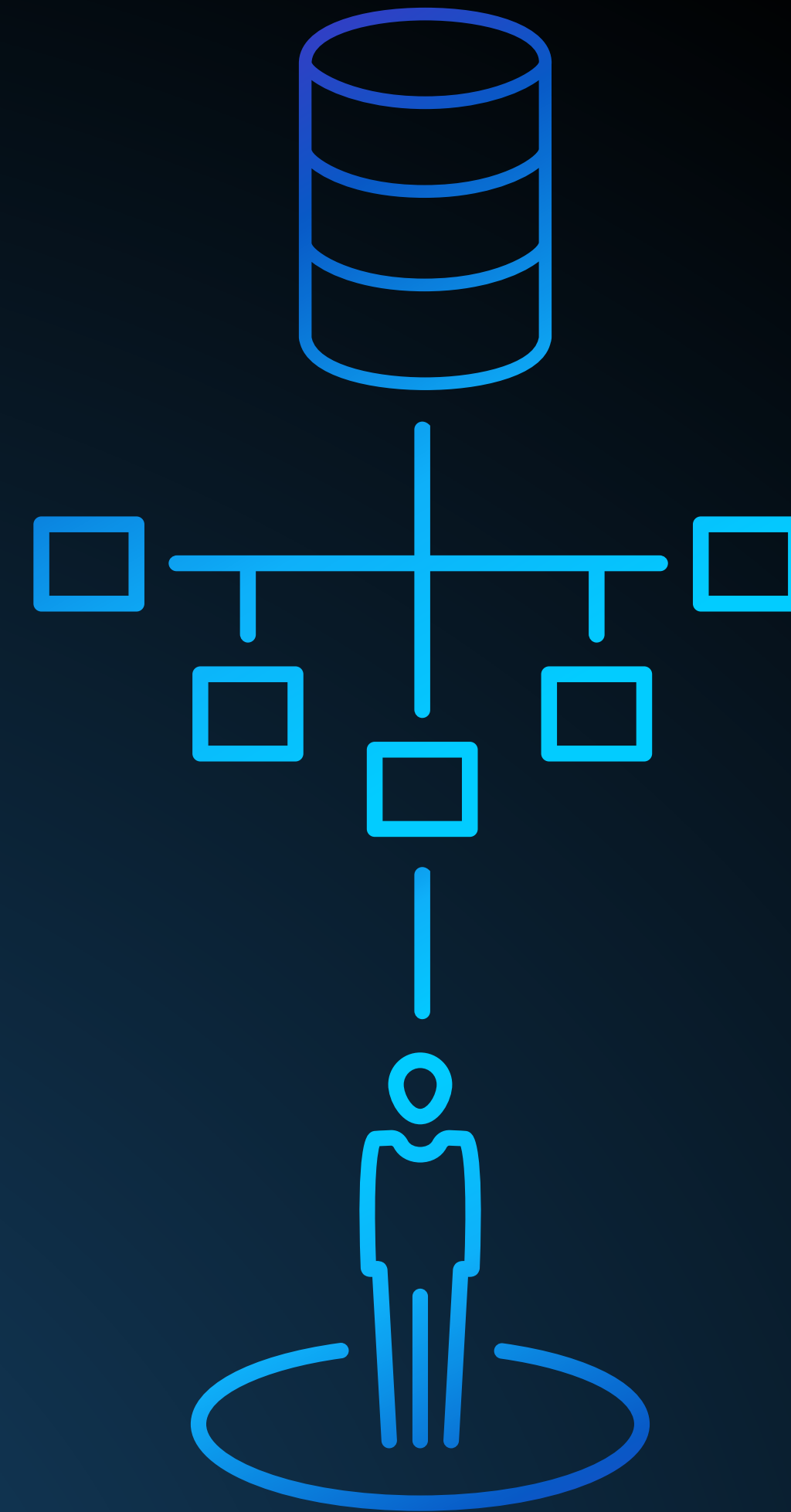
6 Enrich the data to ensure critical attributes are up-to-date

Most enterprises continue to struggle with data quality and consistency for customers. Typical errors include misspelled organization names, wrong contact information, and unintentional duplicate records within databases. People-related errors and a lack of governance is often the biggest contributor to slip-of-the-finger type errors. On top of that, underlying customer information changes - companies change name post-merger, people move location. To ensure the hard work of data unification is not lost, it's critical to ensure consistent data enrichment. Built-in pipeline enrichment capabilities to ensure critical customer information such as addresses, emails and URLs are kept up to date. While information validation is now standard practice for many, it's vital to establish a streamlined process and consistently integrate enrichment services to reduce pipeline & vendor management complexity.

Say Goodbye to Dirty Dashboards. Guide to Achieving a Single View of Customer Data for Analytics in the Cloud

p11

7 Create a unique ID to consistently identify customer records across systems

Once records are clustered to provide a holistic view of the customers, it's critical to ensure each customer can be uniquely identified and maintained. A golden record view of customers ensures that the data used is the most accurate and complete view of customers based on the data sources available. Often, primary keys don't align across sources. A unique ID that connects customer records within and across systems and enables a golden record view of customers, grouping the underlying records to ensure accurate, consistent information feeds a holistic customer view. When new data sources are added for an existing customer, the cluster is updated and the attributes refreshed, but the unique ID remains as a consistent tracker. These unique IDs (the 'Tamr ID') then feed core operational and analytical systems to ensure the latest information is feeding business decision-making.

Success Stories:

Customer Data Mastering Initiatives That Drove Business Results

Customer data mastering, if correctly done, offers sales and marketing organizations a powerful way to accelerate data-driven selling and enable customer insights for competitive advantage. Below are two examples where enterprise investments in customer data mastering solved longstanding data challenges and made a meaningful difference to the companies' top-line revenue growth.

Littelfuse®

Global high-tech manufacturer unlocks insights from customer data to drive greater upsell and cross-sell opportunities

Littelfuse, a successful $7bn market cap multinational electrical components manufacturer with a vast distributor network, needed to improve insights into its B2B sales process and results—starting with understanding how many customers it actually had. With sales data coming from 200+ transactional data sources, such as SAP and Oracle, understanding the answer to this fundamental question wasn't easy.

Recognizing the problems caused by their poor customer data, Littelfuse chose to master their customer data with machine learning. The company created a repeatable pipeline for tying customer quotes to the millions of resulting transactions through a persistent ID that tracked the customer journey. By successfully mastering its data at scale, the company was able to fuel new on-demand analytics and insight into the

corporate hierarchies of its customers and reinvent its sales process. The results were game-changing.

- **Advanced from one-off reports to on-demand analytics** to analyze sales pyramids and churn reports

- **Enabled prioritization of top accounts & identification of upsell opportunities** by consistently tracking performance

- **Gained insight into channel performance** to better inform distributor exclusivity rights

Not only did Littelfuse streamline and speed up its sales process, critically, the company now trusts its customer data to inform decision making at the highest level of its business.

"Right away the insights that we got into our customer base, our sales orders, and our churn rate were really astounding."

- RYAN STAFFORD, EVP, LITTELFUSE

Read the full story

Santander

One of the world's largest banks cuts lending time in half by solving its customer data problems

Santander UK has 14.4 million customers and £22.3 billion in corporate loans. While the organization had plenty of data, critically it lacked a single, trustable view of its customers. As a result, its lending decisions were slow.

By investing in customer data mastering capabilities, the bank now has accurate, up-to-date 360-degree information on its customers, allowing for much faster lending. The bank combined data from 45 sources, encompassing tens of millions of records, in under four weeks to drive digital transformation.

- **Cut credit decision times in half** with the rollout of the new lending system.

- **Increased the reliability of reporting data** with unified data feeds for financial, credit risk, and regulatory reporting.

- **Empowered Sales with cross-sell and up-sell opportunities** revealed by a holistic view of customers across divisions and products.

" Data mastered by Tamr underpins the entire digital journey - if we didn't have the single customer view, we wouldn't be able to feed information downstream into decision and risk engines "

JONATHAN HOLMAN
HEAD OF DIGITAL TRANSFORMATION, SANTANDER

Read the full story

Summary

Data mastering is a new way to solve the persistent issues of poor customer data. By embracing next-gen technology like machine learning and the cloud and applying a dataops approach, clean and up-to-date customer records become the status quo. Mastered customer data can be fed into operational and analytical systems, putting the data to work and creating a flywheel of positive business results

Tamr Cloud

Tamr Cloud delivers the high quality customer data your team needs. Our data mastering solutions make it easy for you to generate clean, curated, enriched customer data faster than ever before. By leveraging purpose-built machine learning models, Tamr masters data across internal and external data sources for a single view of the customer.

Unify customer data without the pain.

Learn more at **tamr.com/cloud**