STIBO SYSTEMS
MASTER DATA MANAGEMENT

SOLUTION ENABLEMENT

Customer & Supplier Master Data Management

Rel. 10.3-MP4 (March 29, 2022)

Table of Contents

Customer & Supplier MDM Solution Enablement

Customer and Supplier Master Data Management is a technology-enabled discipline which is required to create and maintain an accurate, timely, complete, and unified view of a customer or supplier entity. Customer & Supplier MDM will consolidate, cleanse, enhance, and govern data from disparate siloed systems enabling organizations to identify, consolidate, and link customer and supplier data across heterogeneous data sources in a single, accessible view.

To use the solution implementation, see:

- MDM Solution Overview
- Customer & Supplier MDM Solution Implementation Guidelines

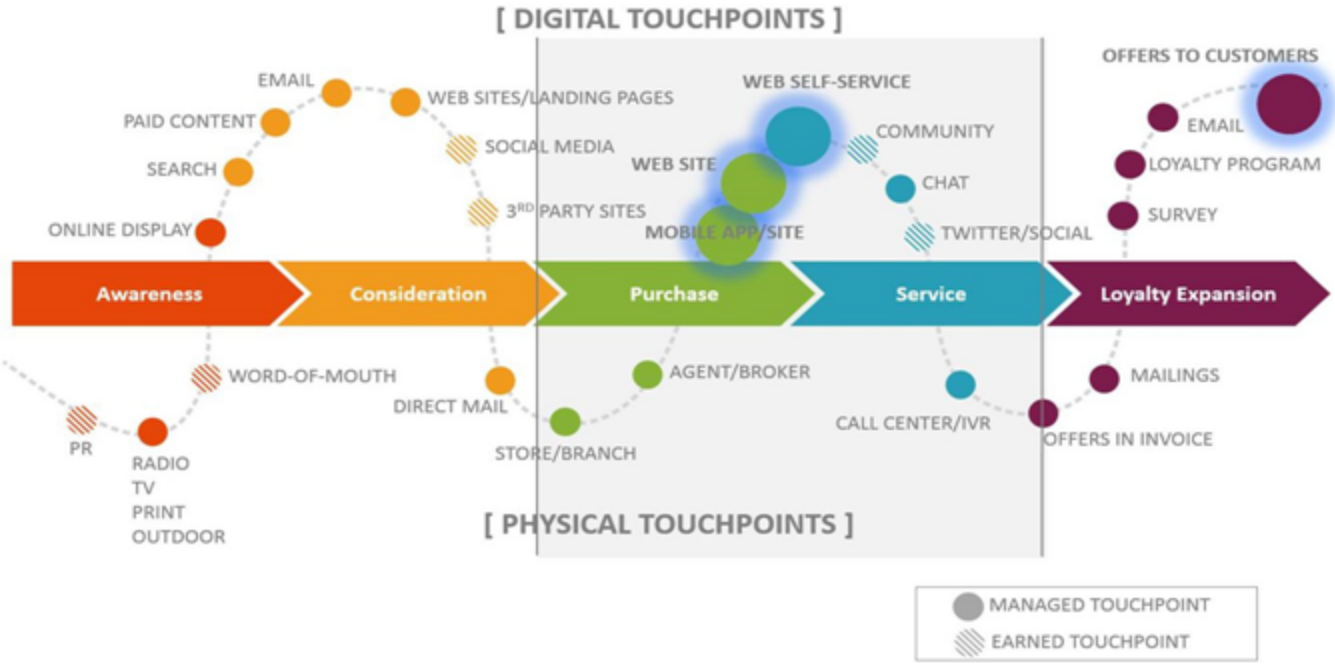
MDM Overview

A Master Data Management solution provides organizations with the means to address complex customer and supplier data challenges and improve their overall experience.

Improving Customer Experience

Customer experience is the holistic journey of a customer that begins the moment they first become aware of a product and continues until the customer is building their own loyalty platform with a product.

This documentation is focused on providing guidance to improve an organization's customer data so informed decisions can improve customer experience.



Data Management Challenges

Organizations are faced with the following primary challenges when managing customer data:

- **Data Quality** - Data quality issues typically arise when attempting to house and manage customer data across multiple platforms. Doing so often leads to inaccurate, duplicated, and incomplete data.
- **Operational Inefficiencies** - Operational inefficiencies can be characterized by an organization's inability to share accurate and up-to-date data across the business, severely impacting downstream systems, processes, and initiatives. For example, some of these inefficiencies can lead to: returned mail, increased costs, payment delays, and unacceptable response time.
- **Reactive Strategic Decisions** - Many organizations struggle to achieve their Customer Relationship Management (CRM) vision. Considering its impact on an organization's success, a concrete CRM strategy

is a high concern for business and IT leaders. Bad data is the most common reason many CRM initiatives fail.

- **Governance** - The most important part of data governance is to consider the availability, usability, integrity, and security of data within an organization. Poor data governance can lead to financial, reputation, and legal risks.

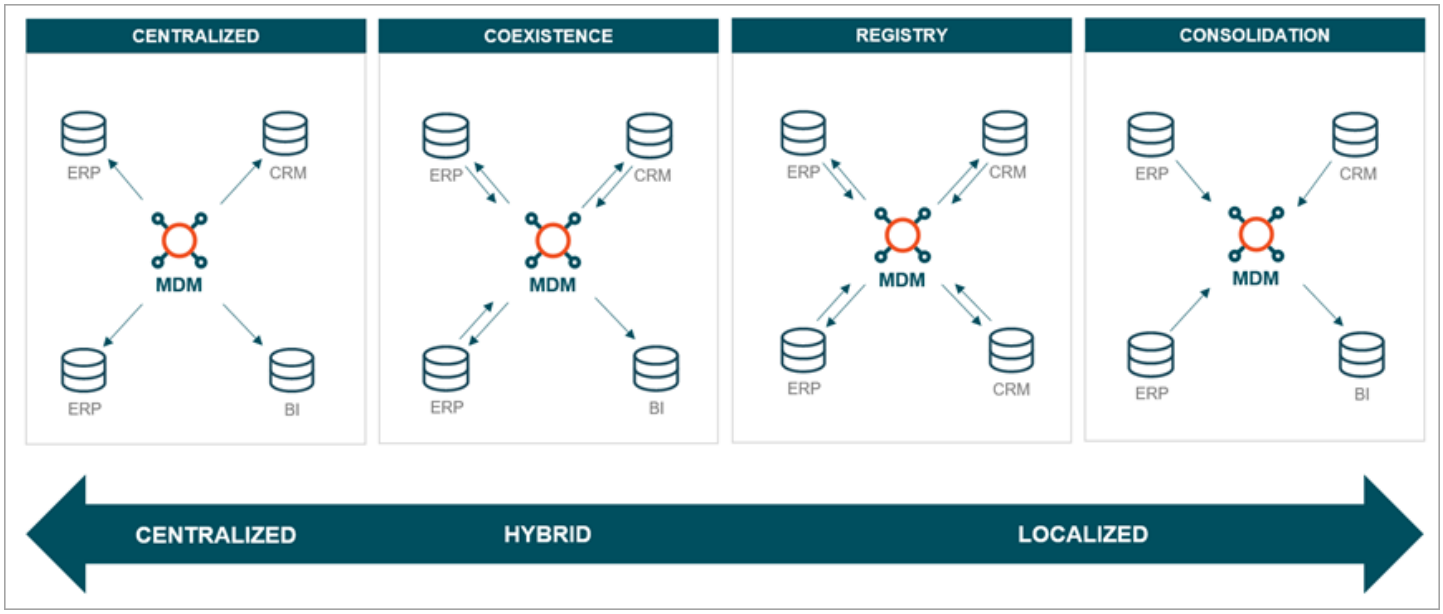
Supplier Data Management Challenges

A Supplier Master Data Management solution provides an organization with the means to address and solve for complex industry challenges when it comes to managing their supplier information to drive the business forward.

- **Supplier Onboarding** - The inability to capture vital and relevant supplier data at the point of data entry causes inefficient management of overall supplier onboarding and management processes, which leads to slow-time-to- market.
- **Single Supplier View** - Without a consolidated, complete, and trusted view of an organization's suppliers, organizations lack the ability to efficiently manage their suppliers across various aspects of the business. Additionally, organizations struggle to validate key supplier information and lack the in-depth insights which makes data more actionable.
- **Hierarchies** - Organizations often lack the ability to visualize and understand the various relationships their suppliers hold today. Hence, it is difficult to identify strategic supplier and organizations are unable to negotiate better pricing and payment terms and miss out on key business opportunities.
- **Data Quality** - As the number of suppliers an organization transacts with grows, data quality suffers because duplicate data increases, supplier performance assessments cannot be performed, overall supplier data cannot be trusted, and all contributes to a lack of transparency to the supply chain.
- **Governance** - The lack of data governance capabilities prevents an organization to standardize and introduce a proactive approach to supplier data governance. This limits the control of supplier risk, reduces the trusted visibility of compliance, and negatively impacts the capability to quickly reach to change.

Implementation Methodologies

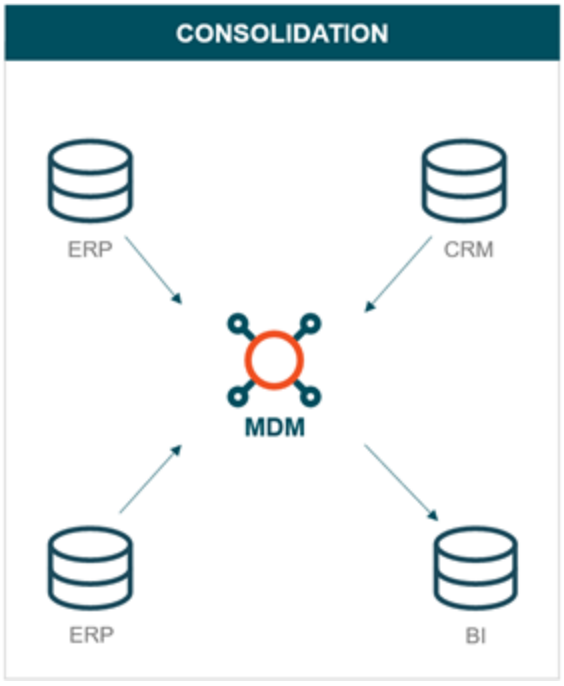
MDM allows the following paths to implementation: Centralized, Coexistence, Registry, and Consolidation.



This implementation guide focuses on a case that implements data management with the **consolidation**, **coexistence**, and **centralized** methodologies.

Consolidation Implementation

This methodology refers to an implementation where MDM is a hub for externally provided data that is funneled into single golden records, which are then made available to external systems. These records are created via matching and merging operations with validation checks throughout the process. The MDM data created from this process is read-only, which is often a necessity as the data originates outside of MDM in an upstream system. A consolidation style implementation does not synchronize data back to the contributing systems.



For example, a company wants to combine all of their sales channel data into one record bank so that data analytics may provide meaningful information. Once all these records are consolidated, marketing initiatives can use this information to target different channels as well as different regions.

Coexistence Implementation

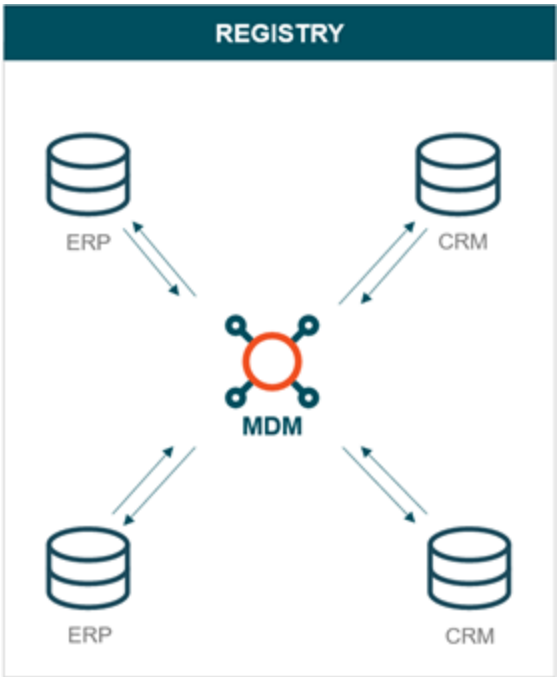
This methodology refers to a scenario where multiple databases containing the master data must operate at the same time. The party data is authored and stored in various external sources while being synchronized to MDM. This process involves deduplication, conflict resolution, and validation operations, resulting in the creation of golden records. Finally, updates to the data can be done in the source system and any external systems that master customer data. MDM synchronizes the content of golden records back to the source systems, keeping every source up to date with data contributed from any source. This allows for a hybrid where data authoring may happen in the source system as well as in the MDM system.

A common situation for coexistence is in companies where each department has their own customer database.

For example, a service department that has data on customers is stored in a separate database than where customer data is stored for the sales and marketing departments. However, using an MDM solution with the coexistence model, if information is changed in the marketing department, it is changed in all other departments.

Registry Implementation

The registry method uses a simple database to reconcile identifiers for all party data across the enterprise. This method allows MDM to serve as a referenced, read-only source of mastered customer data for external systems with minimal data redundancy. MDM will only match on the bare minimum of needed data to confirm a unique record. All externally maintained data quality is the responsibility of the source systems.



For example, a company maintains different types of customer information across various external systems. A registry is used to identify and associate these different types of customer data with external systems. MDM serves as this registry by storing unique records with identifiers that are different from one external system to the next. MDM will ONLY store these records and unique record identifiers, and deduplicate their contact information.

Centralized Implementation

The centralized implementation is based around a well-managed and governed central repository for all master data. This repository holds a set of golden records that are read-only by the operational and analytical systems throughout the enterprise. MDM is the centralized owner of party data, serving as the system of record. This system of record is an organization's single version of truth for customer data. Data is not only centralized in MDM but also created and stored there. The most important aspect of a centralized implementation is that it is a system of record instead of a system of reference.

For example, a company wants to use MDM as the system of record for onboarding, deduplicating, and enriching new suppliers. Over time, MDM could remove duplicate records of these suppliers and then the golden supplier records could be provided to external systems.

Customer & Supplier MDM Implementation Guidelines

The Customer & Supplier Master Data Management Implementation Guidelines provide organizations with baseline configurations and sample data to help get their STEP systems off the ground and guide them in the creation of their Customer and/or Supplier MDM solution.

This user guide offers a number of resources that can help achieve this goal:

- Common MDM use cases that account for both B2C and B2B Customer & Supplier solutions and provides context for system configuration and solution design. For more information, see the **Use Case Overview** section of this user guide.
- A series of component-specific use cases that tie into the common MDM use cases mentioned above. For more information, see the **Use Case Appendix** section of this user guide.
- A comprehensive guide for how data flows in and out of the STEP system, including Matching. For more information, see the **MDM Data Flow** section and the **Matching** section of this user guide.
- A comprehensive guide for modeling the data in STEP. For more information, see the **Data Modeling** section of this user guide.
- A comprehensive guide for data stewardship. For more information, see the **Data Stewardship** section of this user guide.
- For the initial configuration of a base Customer & Supplier MDM solution and how to deploy it, see the **Customer & Supplier MDM Configuration Guide** section of this user guide.

Important: This guide references supporting configuration documentation that exists in other sections of the STEP Online Help.

Use Case Overview

To illustrate how a Master Data Management solution can address common challenges an organization faces, the use cases presented in this section use fictional companies with typical business considerations and problems. Understanding these use case examples will provide additional context to the configuration steps included in this guide, as well as the more granular use cases that concern specific MDM components.

The following use cases are detailed in this section:

- **Centralized Use Cases**
- **Consolidation Business-to-Business Use Cases**
- **Coexistence Business-to-Business Use Cases**
- **Coexistence Business-to-Customer Use Cases**

For more information on component-specific use cases, see the **Use Case Appendix** section of this guide.

Centralized Use Cases

In a Centralized methodology, the ability to onboard new entities directly within MDM allows an organization to more efficiently manage the process of introducing new customers or new suppliers. The efficiency offered is driven by the ability to tightly govern the quality of data at the point of entry, while facilitating any business validations required for new customer / supplier qualification and compliance.

In this documentation, the following use case scenarios are covered:

- Direct Customer Onboarding in MDM
- Onboarding New Customer Retail Store Locations
- Existing Customer in Additional Line of Business
- Direct Supplier Onboarding in MDM
- Supplier Self-Service for Onboarding
- Supplier Self-Service for Supplier-Initiated Maintenance
- Supplier Self-Service for Internally-Initiated Maintenance

Direct Customer Onboarding in MDM

Who is Impacted?

Philip, a Category Manager at BlueBox Retailer, serves as the primary contact for the company's manufacturers and suppliers with a focus on the household cleaning supplies product line.

Lance, a Sales Representative for CleanGoods, has the responsibility of validating the legitimacy of BlueBox Retailer as a prospective customer that CleanGoods Manufacturer wishes to do business with. This may involve validating against third-party syndicated data analytic companies (i.e., Nielsen TDLinx or IRI) which provide insight on the prospect's business operations, retail sales data, corporate structures, and information on individual retail locations.

Kathy, a Controller from the Finance team, runs validations of the prospective customer including a credit check, which may also contain usage of third-party data providers (i.e., Dun & Bradstreet). Such validations enable CleanGoods Manufacturer in identifying any type of fraudulent activity.

George, a Logistics Specialist, is responsible for determining how to fulfill the order. This involves evaluating several factors such as projected order frequency, product assortment, volume, stock, package type, geography, etc. to determine how the customer is to receive the products.

Jennifer, an MDM Specialist, serves as the primary owner of the MDM system and approves all changes submitted by the various departments during the onboarding process.

What is the Business Problem?

BlueBox Retailer has begun negotiations to procure and resell products from two CleanGoods lines of business. BlueBox is especially interested in CleanGoods' line of household cleaning products as well as their charcoal line of business.

Philip has been negotiating with Lance, the sales representative from CleanGoods, in finalizing initial sales terms for their household products line of business. Once finalized, Lance will initiate the new customer request to begin the MDM onboarding / enrichment process.

Part of Lance's responsibility during the onboarding process is to validate the legitimacy of the customer as a legal business entity. Such validations may occur outside of the operational MDM. If the results of the customer validations are favorable, Lance will continue working with Philip to classify / enrich the new customer record with relevant information. Such information that Lance requires includes (not limited to):

- Trade Channel - How do products reach the end-consumer (i.e., In-store retail, e-commerce, mail order, warehouse stores, etc.)?
- Product Lines - What types of products will Bluebox Retailer be purchasing?
- Line of Business - What particular line(s) of business will BlueBox be conducting transactions with?
- Location Setup
    - All Ship To

    - Location information for each physical location CleanGoods Manufacturer must deliver products to (e.g., warehouses, distribution centers, retail stores, etc.).

- All Sell To

    - Location information for each Sell-To location CleanGoods Manufacturer must send invoices to (e.g., Corporate head office, regional offices, accounts receivable department, etc.).

- Financial / accounting information for each Sell-To location, used for invoicing purposes.

Additional processes may be required to be initiated in various operational systems external to the MDM to add additional information in those systems which do not belong in the MDM system, for example, it could be that certain data is local to one process and / or it is most convenient that the data be maintained in the operational systems.

As part of the onboarding process, Kathy from the Finance department must ensure finance-related enrichment activities take place. These activities include conducting credit checks of the customer for any financing approval by leveraging third-party services such as a Dun & Bradstreet company profile. Additionally, Kathy may work with Lance during initial negotiations to provide pricing and trade promotion of products sold to BlueBox. Pricing usually originates and is maintained within an ERP system.

Once the customer sales & financial enrichment process is complete, George from the Logistics department must determine if the customer requires:

- Ship to Store
- Ship to Customer Distribution Center
- Direct Pick-up at the Manufacturer

Considerations must be given to storage requirements as well as any regulatory requirements or hazardous material protocols. For example, some products may only be stored in specific warehouses equipped with refrigeration or hazardous products may require special shipping and handling instructions.

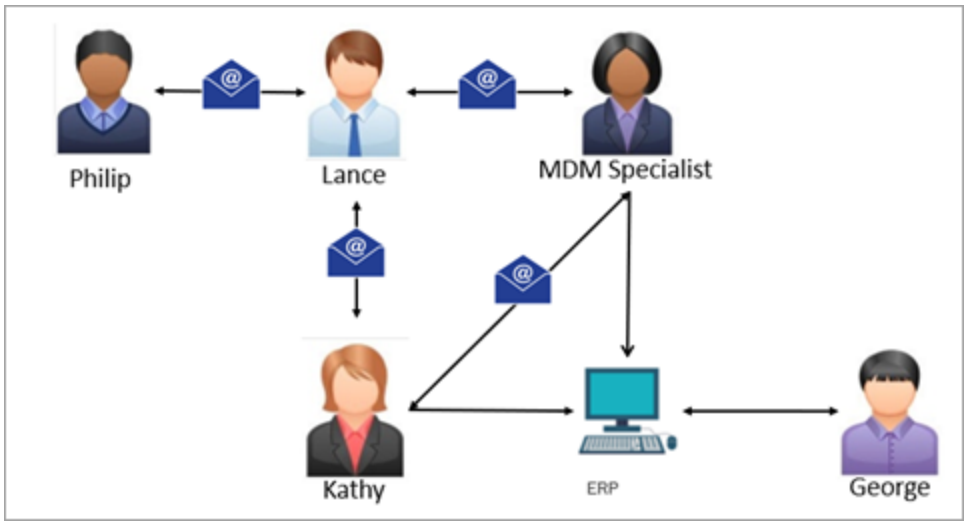
Current State

The current operational landscape of CleanGoods is rather siloed and does not provide a means to onboard and enrich customer data in a centralized way. Due to this, operational data is scattered across disparate systems with facilitation or governance in place. Each role may enact their individual governance to manage data integrity with little to no consciousness of impact to other departments or overall operations.

Currently, Lance must work with the customer to finalize sales terms and other commercial considerations, he must then work internally with a specialist to initiate a new customer request. Kathy from Finance is then notified via email of the new customer request and must then work out of ERP to determine exact pricing and update the financial portfolio. Should Kathy have any questions for Lance, she must likewise correspond with him over email or a phone call. Furthermore, the same process is followed to notify George from Logistics so he may address delivery and warehousing considerations.

Without a centralized ownership of the onboarding process and governance in place to ensure validations and regulatory requirements are adhered to, CleanGoods finds that their data is prone to error, often duplicated, and leads to unreliable business analytics, regulatory fines, and damage to brand reputation.

The current data flow is illustrated below:



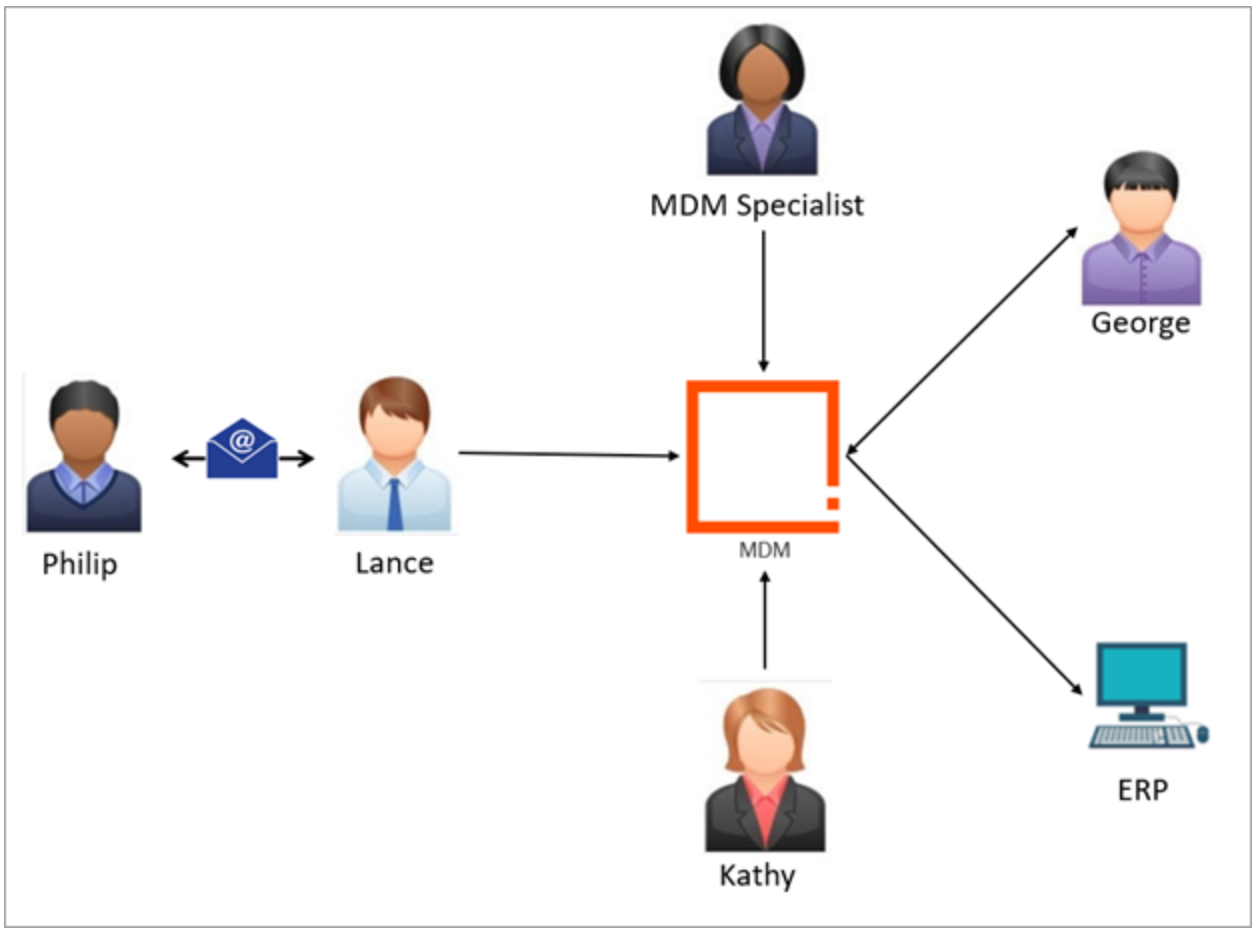
Future State

With an MDM solution in place, the responsibilities and activities of new customer onboarding and enrichment are facilitated in a centralized fashion with all the necessary data integrity and governance in place for a more seamless experience. This ensures operational efficiency within the new customer onboarding process by leveraging the capture and enrichment of the right data by the right people at the right point in time.

Lance will now initiate new customer requests directly within MDM which runs a Find Similar algorithm to proactively ensure uniqueness and eliminates the risk of introducing duplicate or redundant customer records.

Lance will enter the terms of sale and all associated information onto the new customer request, at which point Lance will transfer the onboarding task (within MDM) to Kathy in Finance for further enrichment. Kathy will then see the new task within her task list which informs her that she needs to enrich the same record Lance initiated.

Once the task is assigned to Logistics, George can find all the information in MDM. This allows him to quickly turn around and establish a plan to meet the new order requirements. Once logistical information has been entered a final review is performed by Jennifer and the new customer may be considered as an active customer of CleanGoods.


MDM Specialist
George
Philip
Lance
MDM
ERP
Kathy

Onboarding New Customer Retail Store Locations

Who is Impacted?

Philip, a Category Manager at BlueBox Retailer, serves as the primary contact for the company's manufacturers and suppliers with a focus on the household cleaning supplies product line.

Lance, a Sales Representative for CleanGoods, has the responsibility of validating the legitimacy of BlueBox Retailer as a prospective customer that CleanGoods Manufacturer wishes to do business with. This may involve validating against third-party syndicated data analytic companies (i.e., Nielsen TDLinx or IRI) which provide insight on the prospect's business operations, retail sales data, corporate structures, and information on individual retail locations.

Kathy, a Controller from the Finance team, runs validations of the prospective customer including a credit check, which may also contain usage of third-party data providers (i.e., Dun & Bradstreet). Such validations enable CleanGoods Manufacturer in identifying any type of fraudulent activity.

George, a Logistics Specialist, is responsible for determining how to fulfill the order. This involves evaluating several factors such as projected order frequency, product assortment, volume, stock, package type, geography, etc. to determine how the customer is to receive the products.

Jennifer, an MDM Specialist serves as the primary owner of the MDM system and approves all changes submitted by the various departments during the onboarding process.

What is the Business Problem?

BlueBox Retailer is seeking to strategically expand their footprint in the Southeast region of the United States. In order to beat their competitors to new markets, executive management has devised plans to open 12 new BlueBox Retailer locations in suburban neighborhoods of multiple major metropolitan areas in the Southeast.

The impact of new store locations is that all product line purchase orders must be amended or supplemented to meet the demand of an entirely new geographical region. To meet this demand, Philip has reached out to CleanGoods Manufacturing.

Philip works with Lance, a Sales Support Specialist, to provide the necessary information. Since BlueBox Retailer is a realized customer, Lance does not need to initiate an onboarding process, but rather update the existing customer. To do so Lance must:

- Obtain the physical address of all 12 locations to create the relevant Ship-To addresses and Sell-To addresses.
- Obtain information on what specific product lines the new stores will carry
- Obtain any additional information specific to these new locations (Geographical, new product line availability, order frequency, etc.)

Kathy from Finance will work with Lance reviewing sales terms for the new locations. While new locations may have little to no impact to terms of sale, it is Kathy's responsibility to ensure both parties agree of the negotiated terms. It is up to CleanGoods' discretion whether they would provide special allowances to assist in promoting the BlueBox new store location.

George from Logistics must determine how to fulfill the order. This involves evaluating several factors such as projected order frequency, product assortment, volume, stock, package type, geography, etc. Since BlueBox Retailer requires that all products be shipped directly to their stores, George will leverage the new Ship-To addresses to plan out routes and scheduling.

The challenges faced by CleanGoods are:

- Lack of standardized collaboration mechanism between roles / departments; everything is transacted over email and/or in person notifications
- Lack of traceability of changes to data over time which leads to challenges during investigation and accountability
- Challenges to data governance leads to poor regulatory adherence and decrease in brand reputation

Current State

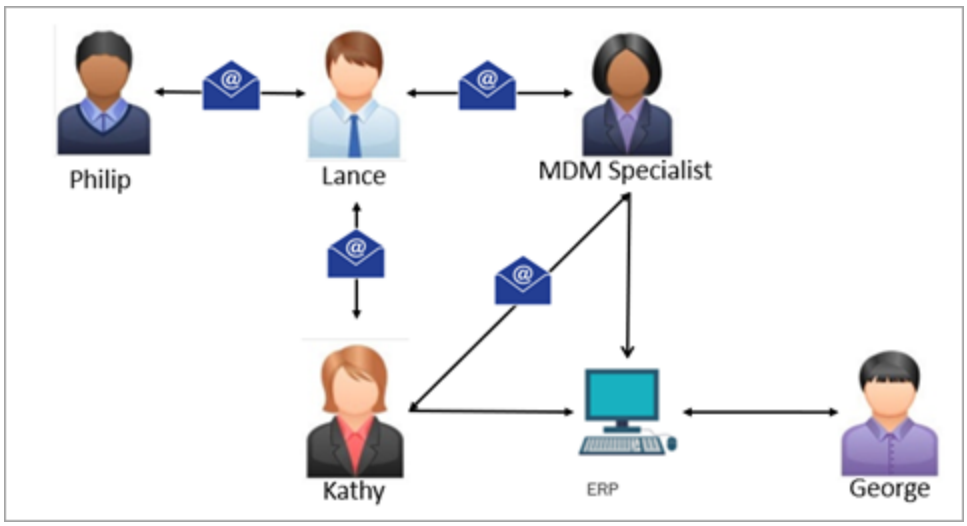
The current data flow of adding new stores and corresponding Bill-to and Ship-to locations does not provide a dynamic for the various teams to easily collaborate. Currently, Lance must email Jennifer (MDM Specialist) to notify of upcoming changes to an existing customer, BlueBox.

Jennifer must then update the customer record in ERP and notify Kathy to evaluate the financial impacts. Should Kathy have any questions for Lance regarding the request, she must contact him separately.

Lastly, George must also be notified so he may consider the logistical impact of multiple new Ship-to locations.

Without a centralized ownership of the maintenance process and governance in place to ensure validations and regulatory requirements are adhered to, CleanGoods often finds that their data is prone to error, duplicated, and leads to unreliable business analytics, regulatory fines, and damage to brand reputation.

The current data flow is illustrated below:



Future State

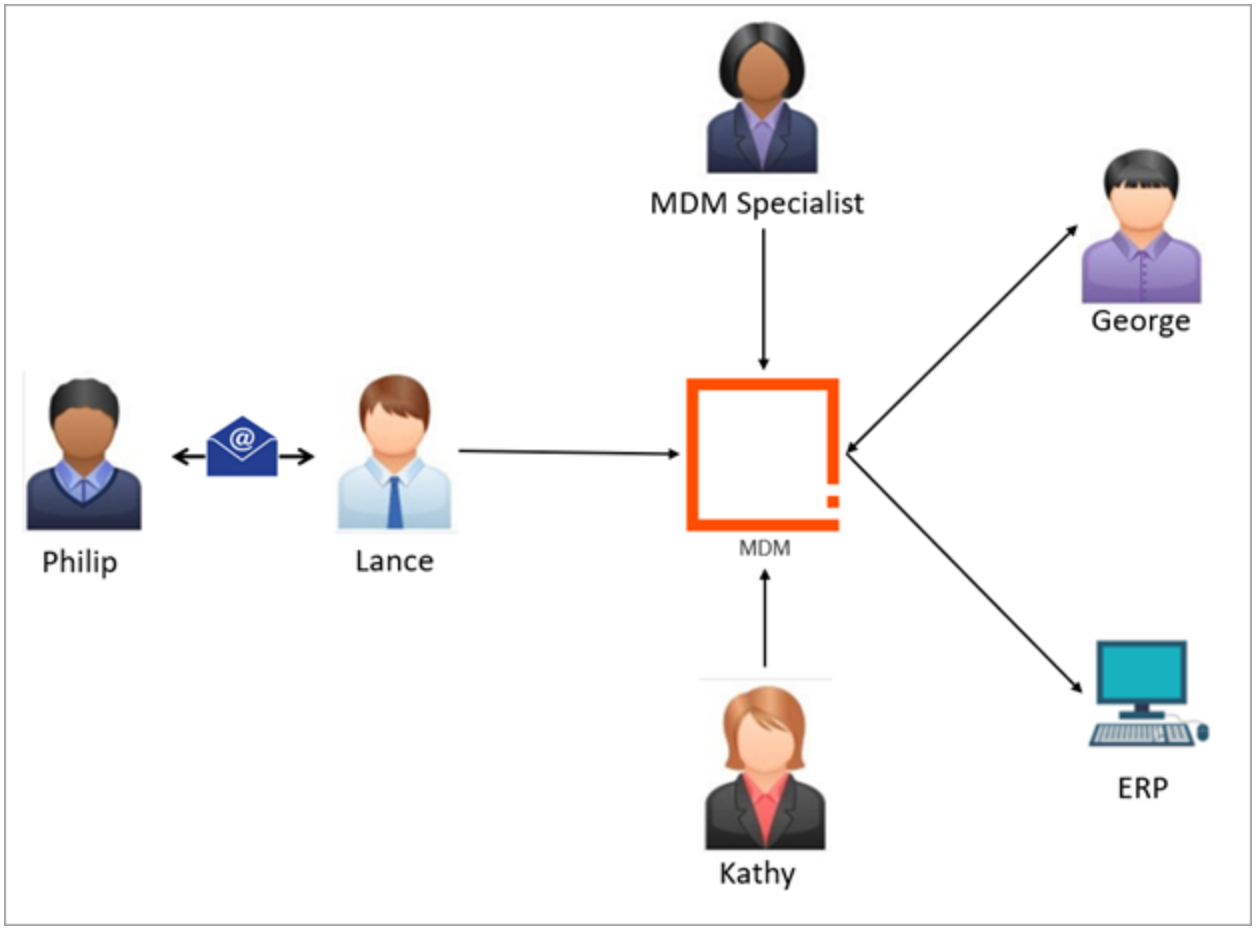
With an MDM system in place, the various departmental responsibilities of updating an existing customer record is now within a centrally governed and collaborative solution.

Lance will now be able to edit the existing BlueBox customer record by adding the necessary location information as well as relevant sales attribution. Lance may then assign a task to Kathy within the application.

Kathy will review and reference the updated information to enable her to finalize any changes necessary to sales terms and related pricing.

Once the task is assigned to Logistics, George is now able to see the updated Ship-To locations which allows him to quickly turn around and establish a plan to meet the new order requirements.

The future state data flow is illustrated below:



Existing Customer in Additional Line of Business

Who is Impacted?

Philip, a Category Manager at BlueBox Retailer, serves as the primary contact for the company's manufacturers and suppliers with a focus on the household cleaning supplies product line.

Lance, a Sales Representative for CleanGoods, has the responsibility of validating the legitimacy of BlueBox Retailer as a prospective customer that CleanGoods Manufacturer wishes to do business with. This may involve validating against third-party syndicated data analytic companies (i.e., Nielsen TDLinx or IRI) which provide insight on the prospect's business operations, retail sales data, corporate structures, and information on individual retail locations.

Kathy, a Controller from the Finance team, runs validations of the prospective customer including a credit check, which may also contain usage of third-party data providers (i.e., Dun & Bradstreet). Such validations enable CleanGoods Manufacturer in identifying any type of fraudulent activity.

George, a Logistics Specialist, is responsible for determining how to fulfill the order. This involves evaluating a number of factors such as projected order frequency, product assortment, volume, stock, package type, geography, etc. to determine how the customer is to receive the products.

Jennifer, an MDM Specialist, serves as the primary owner of the MDM system and approves all changes submitted by the various departments during the onboarding process.

What is the Business Problem?

BlueBox Retailer is seeking to expand their in-store selection of household, consumer-grade charcoal to take advantage of the upcoming BBQ season. To do so, BlueBox has initiated sales negotiations with CleanGoods' charcoal line of business under the sub-brand, CleanCoal. Although BlueBox is an existing customer, onboarding a new line of business requires a similar governing process as onboarding a new customer.

Alexander works with Lance to initiate sales negotiations for CleanCoal's line of business. Since there is no need to validate BlueBox as a legitimate business entity, Lance must determine and update the following information:

- Trade Channel - How products reach the end-consumer (i.e., In-store retail, e-commerce, mail order, warehouse stores, etc.)?
- Product Lines - What types of products will BlueBox Retailer be purchasing?
- Location Set-up - Determine whether existing Ship-To and Sell-To locations are sufficient or if new locations are required to be set up for the new product line.

As part of the process, Kathy from the Finance department must ensure finance-related enrichment activities takes place. This may include working with Lance during sales negotiations to provide pricing and trade promotion of products sold to BlueBox that is specific to the CleanCoal line of business. Pricing usually originates and is maintained within an ERP system.

Once the customer sales & financial enrichment process is complete, George from the Logistics department must determine if the customer requires:

- Ship to Store
- Ship to Customer Distribution Center
- Direct pick-up at the manufacturer

Considerations must be given to storage requirements as well as any regulatory requirements or hazardous material protocols for transporting and storing charcoal.

Current State

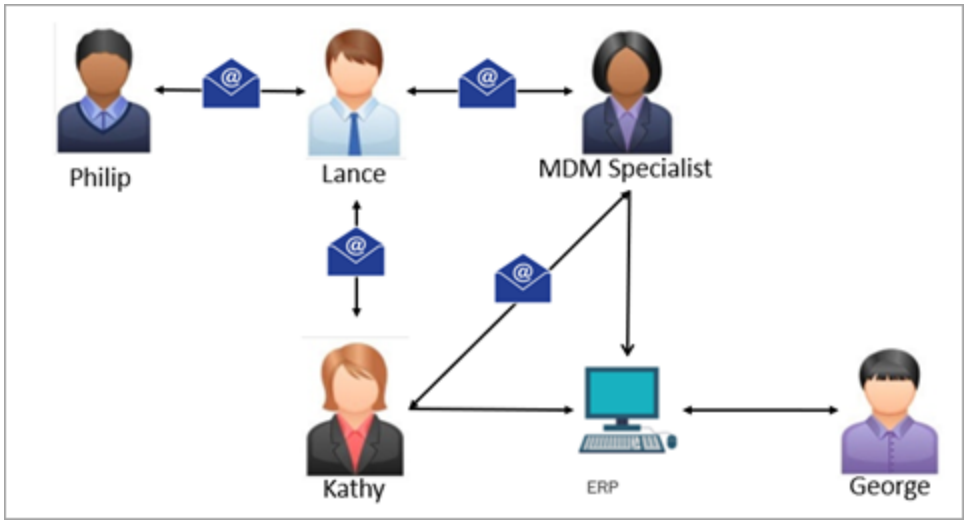
The current data flow of adding a new line(s) of business, relevant sales terms, and corresponding Bill-to and Ship-to locations does not provide a way for the various teams to easily collaborate. Currently Lance must email the Jennifer (MDM Specialist) to notify of upcoming changes to an existing customer, BlueBox.

Jennifer must then update the customer record in ERP and notify Kathy to evaluate the financial impacts. Should Kathy have any questions for Lance regarding the request, she must contact him separately.

Lastly, George must also be notified so he may consider the logistical impact of multiple new Ship-to locations.

Without a centralized ownership of the maintenance process and governance in place to ensure validations and regulatory requirements are adhered to, CleanGoods often finds that their data is prone to error, duplicated, and leads to unreliable business analytics, regulatory fines, and damage to brand reputation.

The current data flow is illustrated below:



Future State

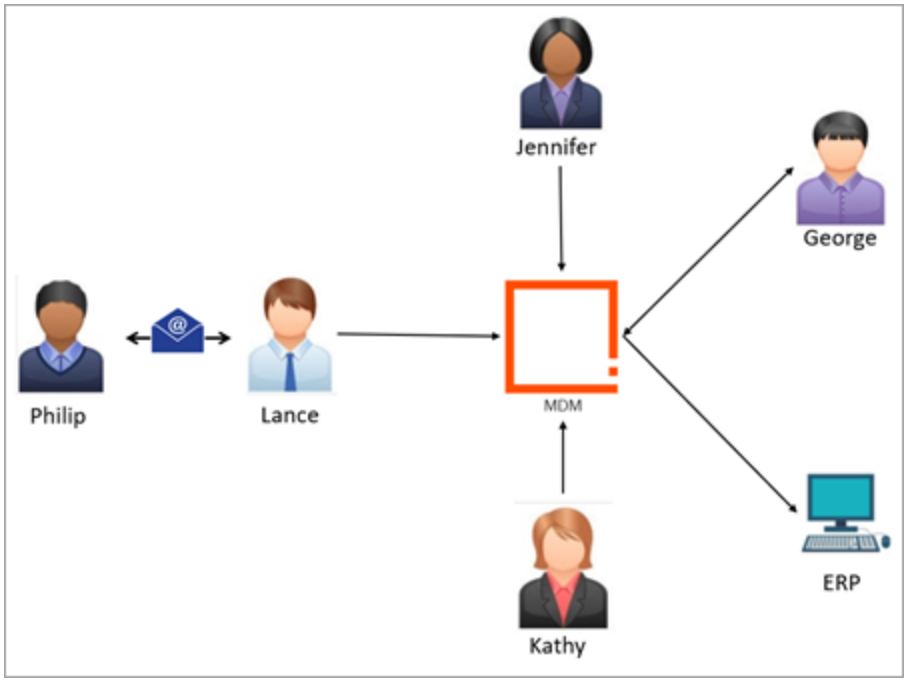
With an MDM system in place, the various departmental responsibilities of updating an existing customer record is now within a centrally governed and collaborative solution.

Lance will now be able to edit the existing BlueBox customer record by adding the new line of business as well as relevant sales attributes; he may then assign a task to Kathy within the application.

Kathy will review and reference the updated information to enable her to finalize any additions or changes necessary to sales terms and related pricing.

Once the task is assigned to Logistics, George is now able to see the updated Ship-To locations which allows him to quickly turn around and establish a plan to meet the new order requirements.

The future state data flow is illustrated below:



Direct Supplier Onboarding in MDM

Who is Impacted?

Arthur, a specialist on the procurement team, is responsible for onboarding new suppliers.

Jennifer is an MDM Specialist who maintains all supplier data for CleanGoods and often supports the procurement team to onboard new suppliers.

Raymond is a contact person for a prospective supplier. Raymond liaises with CleanGoods procurement team to fulfill CleanGoods' new supplier requirements.

Kathy works on the Finance team and corresponds with Procurement in ensuring new Suppliers meet all of CleanGoods' financial requirements.

George, a Logistics Specialist, is responsible for enriching and validating logistical information for suppliers to ensure shipments are moved properly and efficiently through CleanGoods' supply chain.

What is the Business Problem?

As business for CleanGoods continue to expand, demand for more raw materials and finished goods dictates the business' need to grow their supplier base. To facilitate this effort, Suppliers are subjected to a multi-MDM onboarding process which requires the collaboration of many different business teams. Since there is no such collaborative framework within CleanGoods' current ERP platform, the supplier onboarding effort is a challenging and inefficient process that is costly and prone to error.

Current State

Arthur, a specialist on the procurement team, is responsible for onboarding new suppliers. As the business initiator of the onboarding of a new supplier, Arthur provides the basic details required to create a new supplier record.

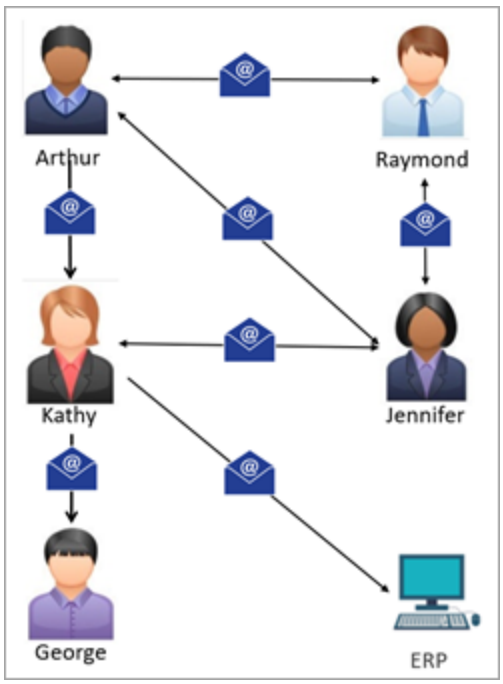
If a supplier is sourced, this information is sent to Jennifer, an MDM Specialist, who works with Arthur and Raymond, the contact person for the new supplier, to carry out further enrichment activities on the new supplier record. Once enrichment is complete, Jennifer informs Kathy on the Finance team.

Kathy conducts her review, additional financial validations, and updates ERP with this new information before assigning the record to George.

George, a logistics specialist, reviews and validates the new supplier information and ensures transportation compliance. George then submits the task back to Jennifer.

At this point, Jennifer runs final validations and provides a final approval. This final approval is the last milestone in onboarding a new supplier and allows for procurement to generate new purchase orders.

Since the ERP application does not provide a collaborative interface, this process is mostly managed over email and attachments, which is extremely inefficient and is prone to error since all business validation checks are manual.



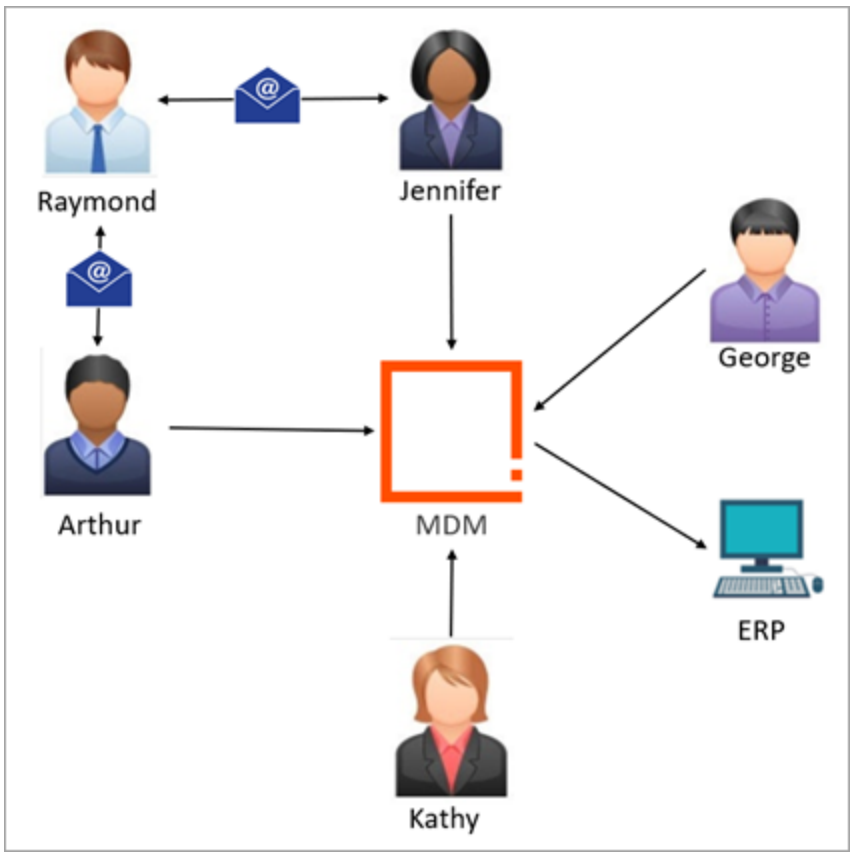
Future State

A Supplier MDM solution will help facilitate this process by providing a centralized hub from which supplier onboarding activities may be conducted in a collaborative fashion.

Arthur logs onto the MDM Web UI to initiate a new supplier record with minimal basic information. MDM will conduct an initial check to make sure the supplier does not already exist.

Arthur then assigns this task to Jennifer who then corresponds with Raymond, the contact person for the new supplier, to provide all further enrichment details, including relevant Ship-To/Bill-To location information. Upon completion, Jennifer then assigns the task to Kathy, a controller on the finance team. During this transition MDM runs a series of business validations to enforce mandatory fields and any compliance rules. If any fields are incomplete or fails validations, the task remains with Jennifer to rectify.

If no errors are presented, Kathy reviews this new supplier and provides additional necessary financial details. Kathy submits the task to George who evaluates all transportation information. Upon completion, George returns the task back to Jennifer who then reviews for a final approval of the new supplier. Upon this approval, MDM re-executes all business validations. Any failed validations prevent a system approval of the new supplier and Jennifer are notified to rectify the error(s). If all validations execute successfully, MDM approves the supplier and automates the publishing of the new supplier to the ERP application.


Raymond
Jennifer
George
Arthur
MDM
ERP
Kathy

Supplier Self-Service for Onboarding

As an extension to Supplier Onboarding within MDM, Suppliers are given the capability to log into the MDM to enrich their supplier profile for the initial onboarding process.

Who is Impacted?

Arthur, a specialist on the procurement team, is responsible for onboarding new suppliers.

Jennifer is an MDM Specialist who maintains all supplier data for CleanGoods and often supports the procurement team to onboard new suppliers.

Raymond is a contact person for a prospective supplier. Raymond liaises with CleanGoods procurement team to fulfill CleanGoods' new supplier requirements.

Kathy works on the Finance team and corresponds with Procurement in ensuring new Suppliers meet all of CleanGoods' financial requirements.

George, a Logistics Specialist, is responsible for enriching and validating logistical information for suppliers to ensure shipments are moved properly and efficiently through CleanGoods' supply chain.

What is the Business Problem?

CleanGoods wishes to expand their current customer onboarding process within MDM by providing prospective suppliers the ability to create, enrich, and submit their supplier profile directly within MDM. This enables a robust supplier self-service onboarding process which greatly efficiently manages the complexities of the onboarding process and improves the MDM user experience.

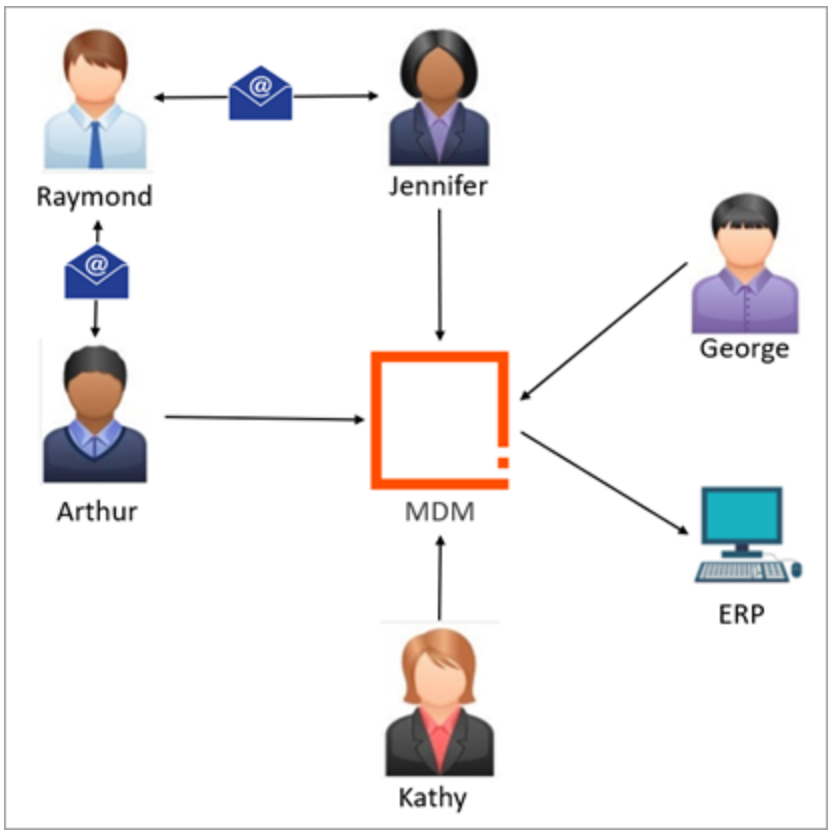
Current State

A Supplier MDM solution will help facilitate this process by providing a centralized hub from which supplier onboarding activities may be conducted in a collaborative fashion.

Arthur logs onto the MDM Web UI to initiate a new supplier record with minimal basic information. MDM will conduct an initial check to make sure the supplier does not already exist.

Arthur then assigns this task to Jennifer who then corresponds with Raymond, the contact person for the new supplier, to provide all further enrichment details, including relevant Ship-To/Bill-To location information. Upon completion, Jennifer then assigns the task to Kathy, a controller on the finance team. During this transition MDM runs a series of business validations to enforce mandatory fields and any compliance rules. If any fields are incomplete or fails validations, the task remains with Jennifer to rectify.

If no errors are presented, Kathy reviews this new supplier and provides additional necessary financial details. Kathy submits the task to George who evaluates all transportation information. Upon completion, George returns the task back to Jennifer who then reviews for a final approval of the new supplier. Upon this approval, MDM re-executes all business validations. Any failed validations prevent a system approval of the new supplier and Jennifer is notified to rectify the error(s). If all validations execute successfully, MDM will approve the supplier and automate the publishing of the new supplier to the ERP application.



Future State

A Supplier MDM solution will help facilitate this process by providing a centralized hub from which supplier onboarding and maintenance activities may be conducted in a collaborative and governed fashion.

Arthur logs onto the MDM Web UI to initiate a new prospective supplier with minimal basic information, including a contact person. MDM will conduct an initial check to make sure the supplier does not already exist. Furthermore, supplier-specific MDM credentials are generated and emailed to the supplier contact person provided by Arthur.

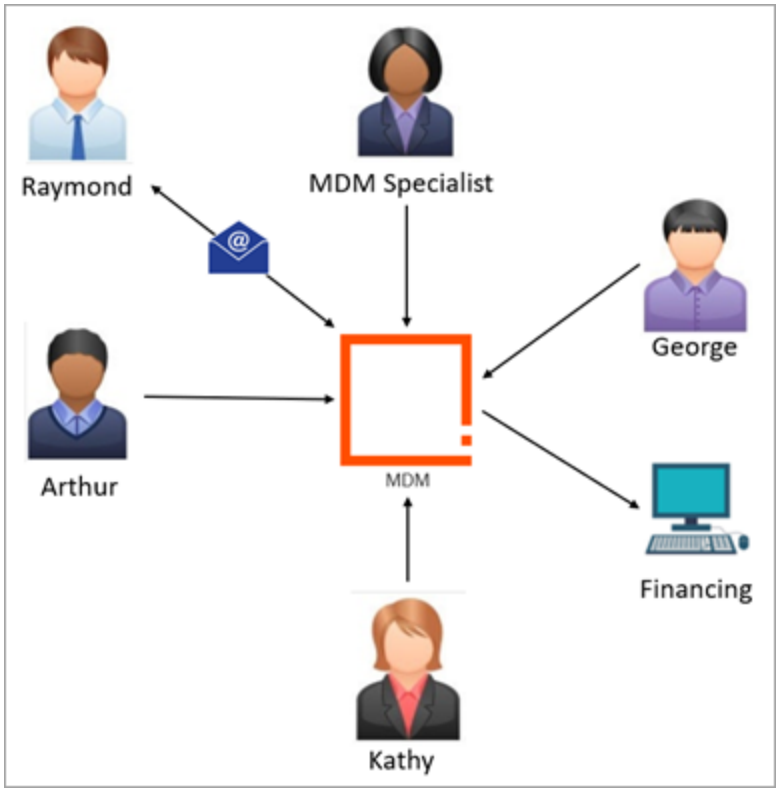
Since the expectation is for supplier users to interface directly with MDM, enforcement of proper user privilege protocols is critical which is why MDM restricts supplier users to only see and edit information for their own supplier group.

Raymond, the contact person on the supplier side, receives the MDM credentials and logs into via the Web UI. Using these credentials, Raymond may only see information specific to his supplier and cannot see or edit data belonging to other suppliers. Additionally, Raymond can create additional users for his colleagues who will also have the same privilege restrictions as Raymond.

Raymond fills out all mandatory information as required by CleanGoods and submits the enriched supplier record for approval by CleanGoods.

The enriched supplier record is assigned to various CleanGoods departments for further enrichment and validations, including Kathy who may review the payment terms and other information for invoicing purposes.

Jennifer, the MDM Specialist, who validates the data sent by Raymond reviews the information and may require Raymond to provide further information upon request. Once all validations have been fulfilled, Jennifer approves the supplier for activation and the new supplier record is sent to the ERP system or any other downstream applications.



Supplier Self-Service for Supplier-Initiated Maintenance

Who is Impacted?

Arthur, a specialist on the procurement team, is responsible for onboarding new suppliers.

Jennifer is an MDM Specialist who maintains all supplier data for CleanGoods and often supports the procurement team to onboard new suppliers.

Raymond is a contact person for a prospective supplier. Raymond liaises with CleanGoods procurement team to fulfill CleanGoods' new supplier requirements.

Kathy works on the Finance team and corresponds with Procurement in ensuring new Suppliers meet all of CleanGoods' financial requirements.

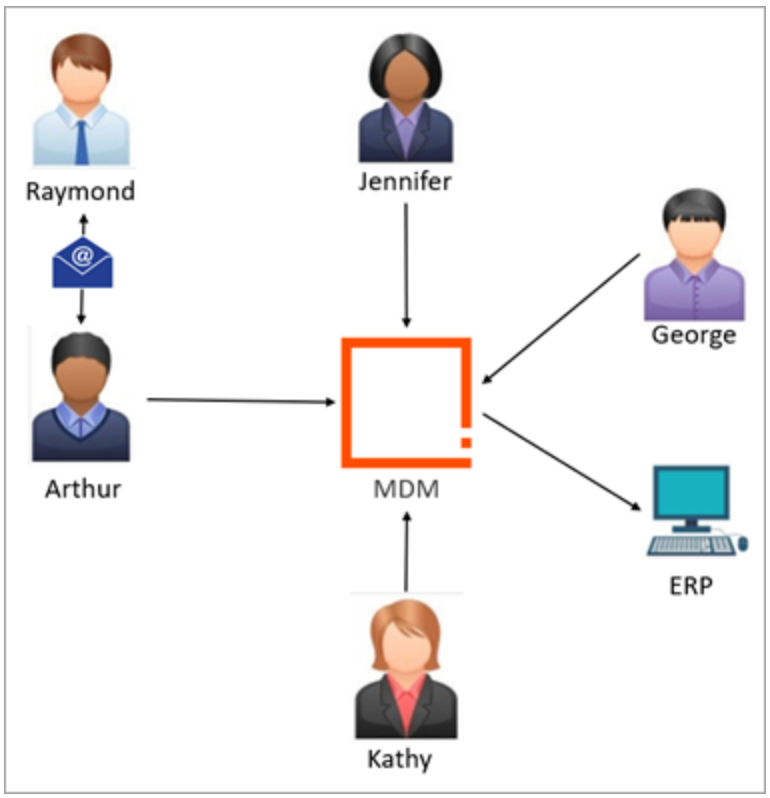
George, a Logistics Specialist, is responsible for enriching and validating logistical information for suppliers to ensure shipments are moved properly and efficiently through CleanGoods' supply chain.

What is the Business Problem?

In addition to providing self-service capabilities to onboard suppliers, CleanGoods wishes to expand the current process by providing self-service capability for maintenance tasks initiated by a supplier user. This extension allows for suppliers to initiate updates to their information, prior to requesting approval by the MDM specialist.

Current State

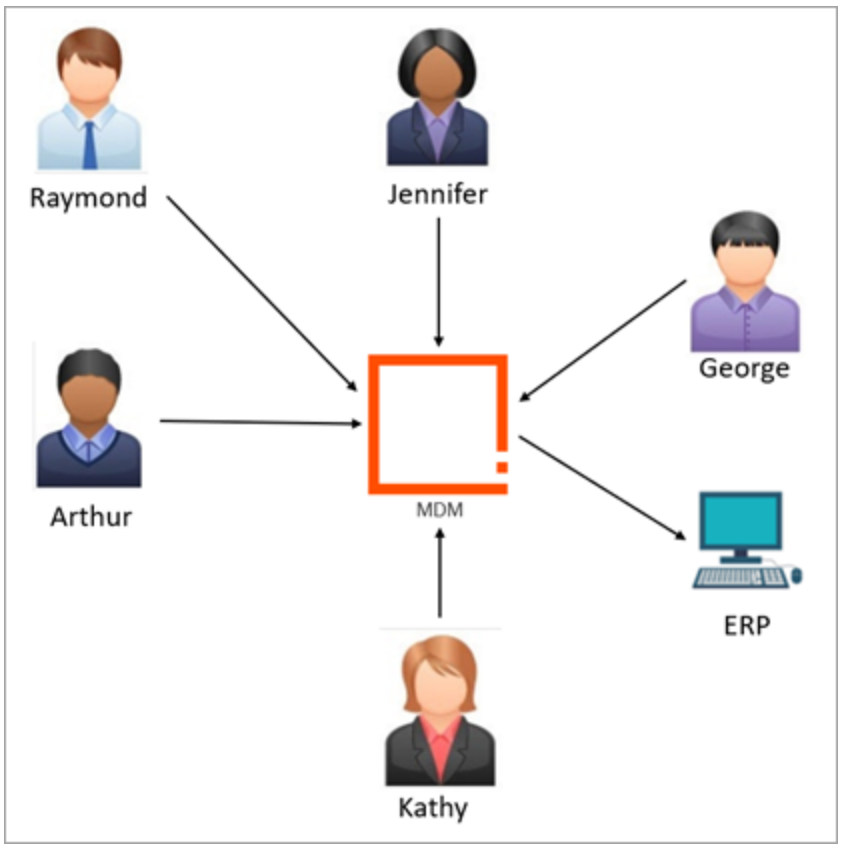
Currently, once a supplier is onboarded with CleanGoods, suppliers must reach out to the procurement specialist and request specific updates to be made on the suppliers' behalf. For example, Raymond must update an email address one of his colleagues who is a contact person at a particular warehouse location. Raymond submits the request to Arthur over email, who will make the update and request Jennifer to approve.



Future State

By extending the MDM workflow, Raymond may log onto the MDM Web UI to update the email address. The supplier record is now a task that is assigned to Arthur for review and then to Jennifer for approval of the change. MDM will automatically send an update to the ERP.

For updates that impacts business operations for Kathy and George, Arthur may assign the task to them for review prior to Jennifer making the approval.


Raymond
Jennifer
George
Arthur
MDM
ERP
Kathy

Supplier Self-Service for Internally-Initiated Maintenance

Who is Impacted?

Arthur, a specialist on the procurement team, is responsible for onboarding new suppliers.

Jennifer is an MDM Specialist who maintains all supplier data for CleanGoods and often supports the procurement team to onboard new suppliers.

Raymond is a contact person for a prospective supplier. Raymond liaises with CleanGoods procurement team to fulfill CleanGoods' new supplier requirements.

Kathy works on the Finance team and corresponds with Procurement in ensuring new Suppliers meet all of CleanGoods' financial requirements.

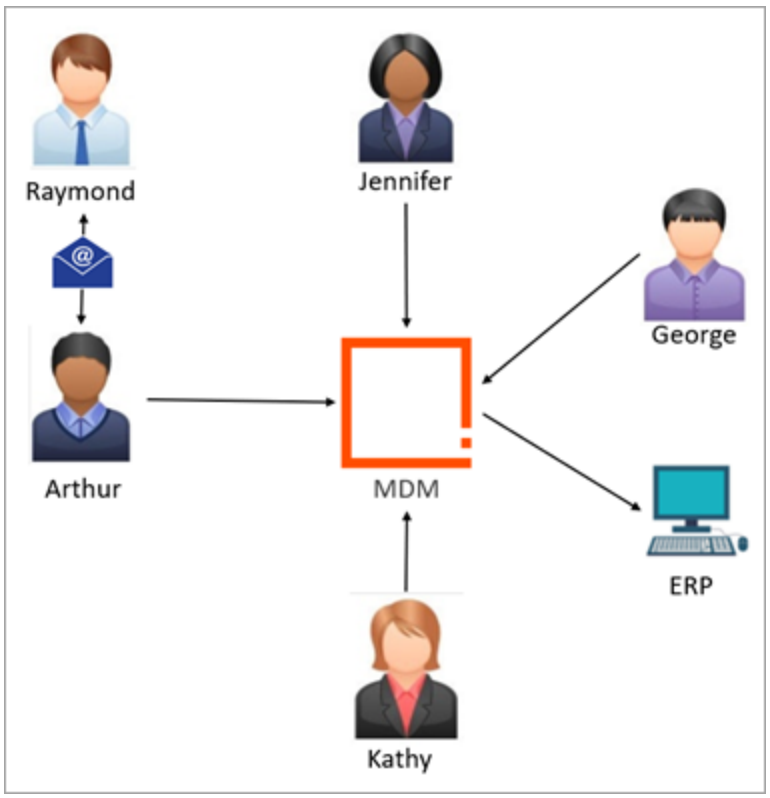
George, a Logistics Specialist, is responsible for enriching and validating logistical information for suppliers to ensure shipments are moved properly and efficiently through CleanGoods' supply chain.

What is the Business Problem?

To further extend the self-service approach, it is quite common that updates to an existing supplier must be made and initiated by an internal user. Such changes must be propagated to the supplier for review before it is approved.

Current State

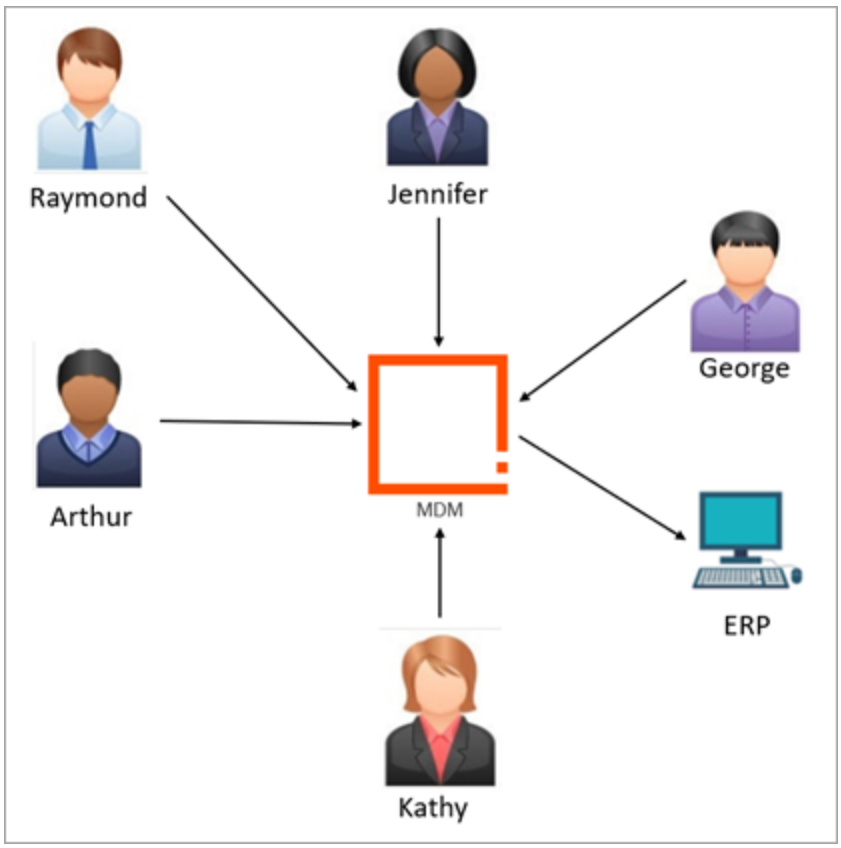
Payment terms that are negotiated with suppliers may require periodic re-negotiates and updates. For example, Arthur must update the payment terms for a supplier due to changes in CleanGoods' fiscal goals and other business factors. To do so, Arthur must first communicate the change to Raymond via email. If Raymond and the supplier approves of the updated terms, Arthur will then make the necessary update within MDM and submit to Jennifer for approval



Future State

By extending the MDM workflow, Arthur can make the change within MDM and submit the request to Raymond for review. If Raymond and the supplier reviews the updated terms and approves, Raymond may submit the task back to CleanGoods and Jennifer, the MDM specialist, will approve the change.

The extended processes provide a means for internal users to request or communicate changes with the supplier, while maintaining ownership of data governance.


Raymond
Jennifer
George
Arthur
MDM
ERP
Kathy

Consolidation Business-to-Business Use Cases

As organizations manage their master data without proper discipline or technologies, it is often the case that the lack of such capabilities introduces duplicate data and other inaccuracies into their source systems. However, with an MDM solution in place, entity data may be consolidated into a single repository which may then be de-duplicated, validated, governed, and operationalized to ensure downstream business processes are empowered with trusted, accurate, and up to date information.

In this documentation, the following use case scenarios are covered:

- Consolidate Customer Data from a Source System
- Consolidate Supplier Data from a Source System
- Customer Data in Disparate Silos across Business Lines
- Supplier Data in Disparate Silos Across Business Lines
- Understanding Customers' Legal Hierarchies
- Understanding Suppliers' Legal Hierarchies

For additional use cases, see the following topics:

- Coexistence Business-to-Business Use Cases
- Coexistence Business-to-Consumer Use Cases

Consolidate Customer Data from a Source System

Organizations often rely on operational reports and business analytics to make strategic decisions. Whether it is to report on supplier performance, determining distributor rebates, or better understanding of consumer purchasing behavior, having reliable master data is foundational in supporting these operational activities.

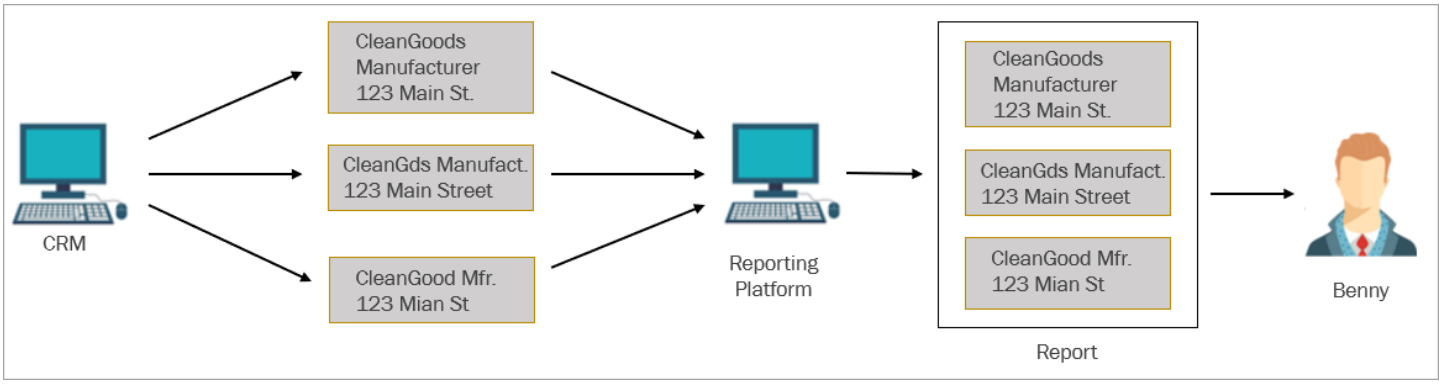
Who is Impacted?

Benny, a representative from the sales team, is responsible for retailers and distributors that have contracts to resell ACME's products. Part of Benny's responsibilities is to ensure the sales team has access to the most accurate and up-to-date sales and financial reports for ACME's retailers and distributors.

What is the Business Problem?

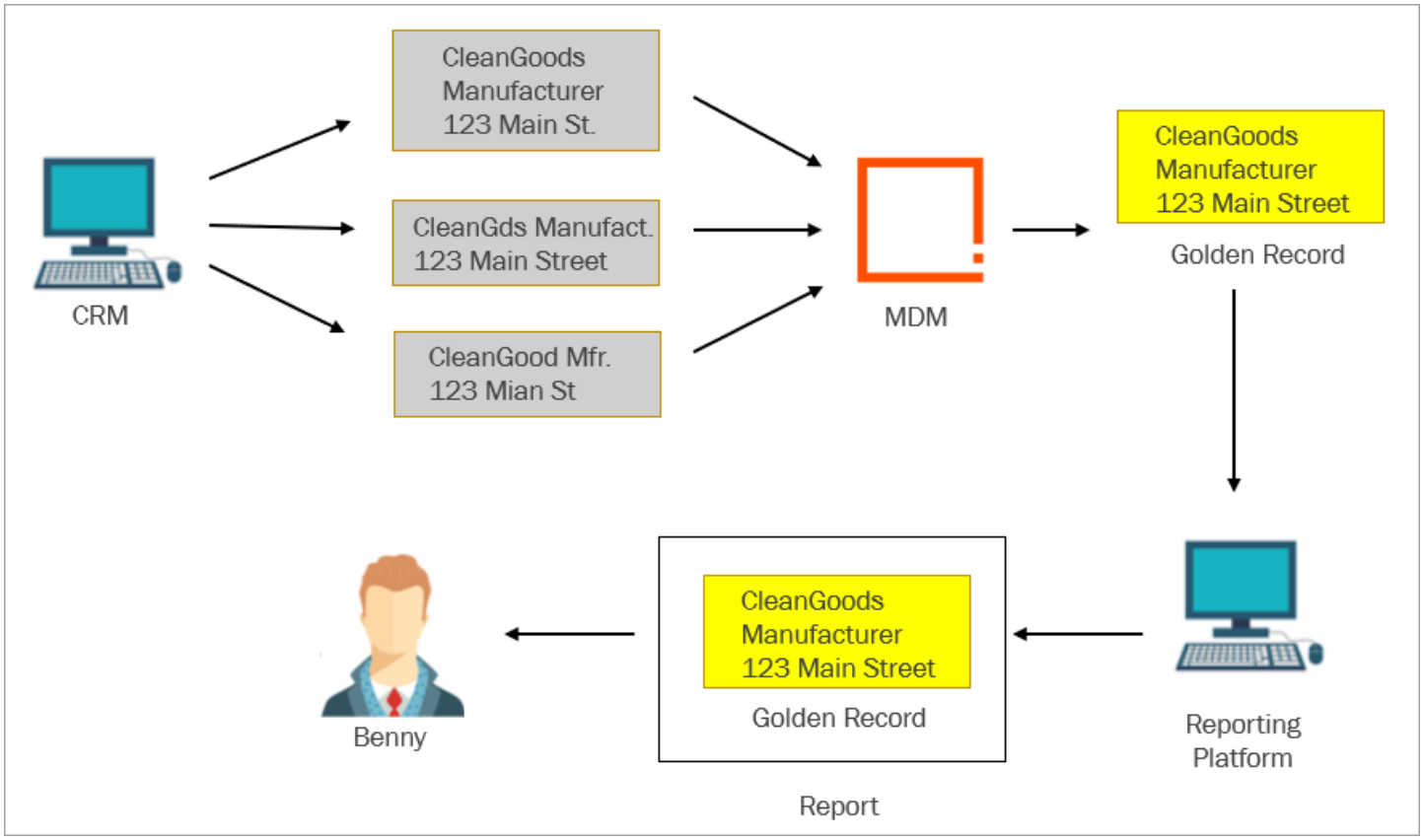
The CRM platform which supports the Western United States sales territory has very limited data standardization and governance capabilities which leads to incomplete and erroneous data being introduced. As data from CRM is synchronized with ACME's reporting platform, this negatively impacts quarterly financial reports from each sales territory and leads to poor decisions and actions being made by the business.

Current Process



As Benny pulls reports on ACME's customers, he notices there are multiple variations of the same customer and even incomplete data are present within the reports. These discrepancies are especially highlighted when attempting to determine who ACME's highest performing distributors are, forecasting demand for particular retailers / regions, etc. This is the result of the lack of governance within the CRM system, which impacts the quality of data within the reporting application.

Future Process



By introducing an MDM solution, ACME can leverage MDM capabilities to identify and connect duplicate customer records, execute standardizations and business logic to validate incoming data, and monitor data completeness. As cleansed data is then synchronized to the reporting platform this enables ACME's B2B customer data to be more accurate and trustworthy for reporting and downstream analytics.

Consolidate Supplier Data from a Source System

Organizations often rely on operational reports and business analytics to make strategic decisions. Whether it is to report on supplier performance, determining distributor rebates, or better understanding of consumer purchasing behavior, having reliable master data is foundational in supporting these operational activities.

Who is Impacted?

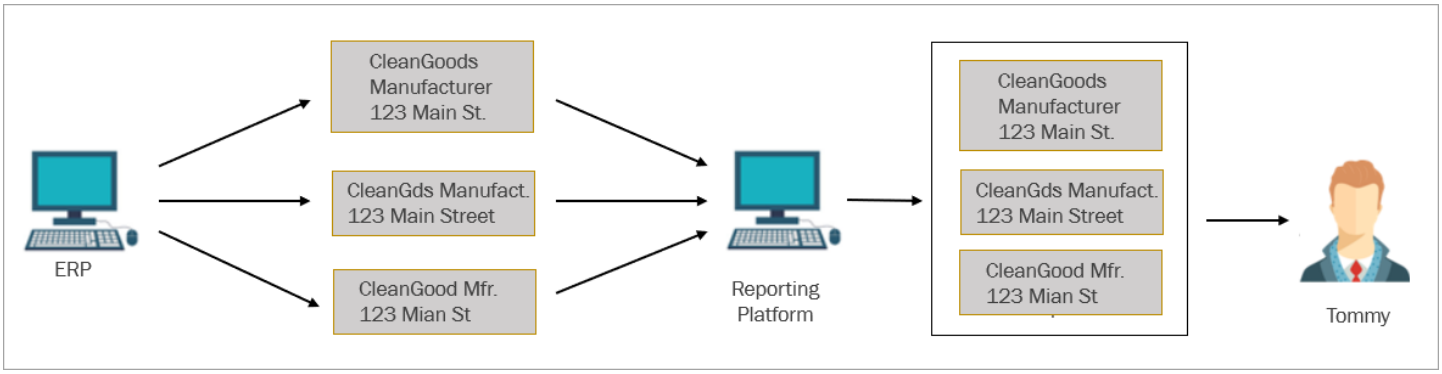
Tommy is an analyst at ACME Company and is responsible for maintaining information on Suppliers that ACME does business with that are in the Western United States sales territory.

What is the Business Problem?

The sales team informs Tommy of increased discrepancies and even conflicting information about ACME suppliers in their sales reports. This is caused by the ERP platform's lack of data standardization and governance capabilities, which leads to duplicate by allowing users to inadvertently create new record for suppliers that already exists.

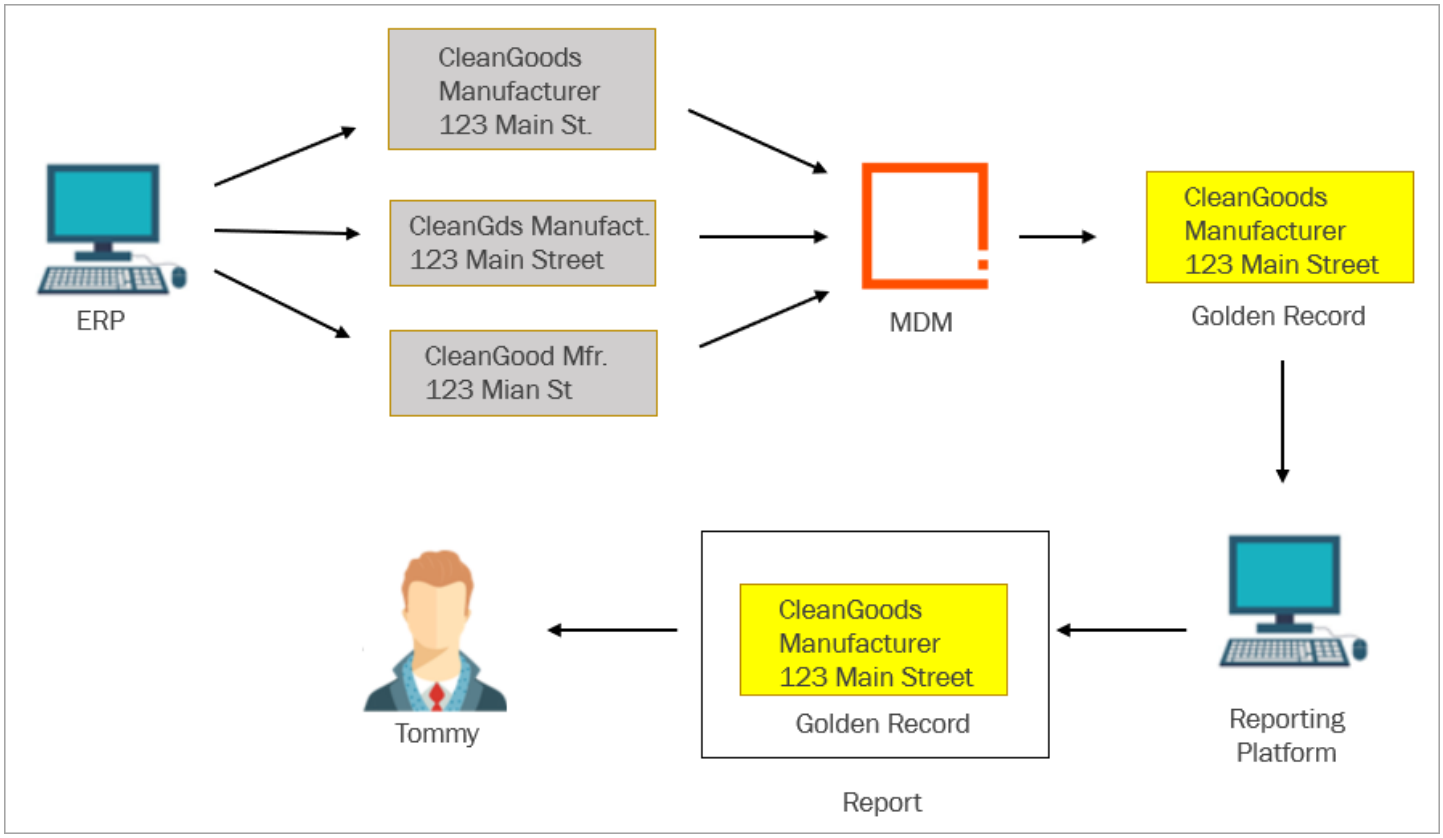
When reports are run for various purposes (ex. chargebacks, compliance reasons, etc.) these data discrepancies are especially accentuated as there is no way to determine which data points on duplicate supplier records are up to date and trustworthy.

Current Process



As Tommy pulls reports for the sales teams on ACME's suppliers, he notices there are multiple variations of the same customer and even incomplete data are present within the reports. This is the result of the lack of governance within the ERP system, which impacts the quality of data within the reporting application.

Future Process



By introducing an MDM solution, ACME can leverage MDM capabilities to identify and merge duplicate customer records, execute standardizations and business logic to validate incoming data, and monitor data completeness. As cleansed data is then synchronized to the reporting platform this enables ACME's supplier data to be more accurate and trustworthy for reporting and downstream analytics.

Customer Data in Disparate Silos Across Business Lines

When organizations operate across multiple lines of business which focus on different offerings and target consumer groups, it is often the case that each line of business is supported by siloed source systems. This presents a challenge as organizations strive to manage their entity data that are maintained in disparate systems yet may overlap from one line of business to another.

Who is Impacted?

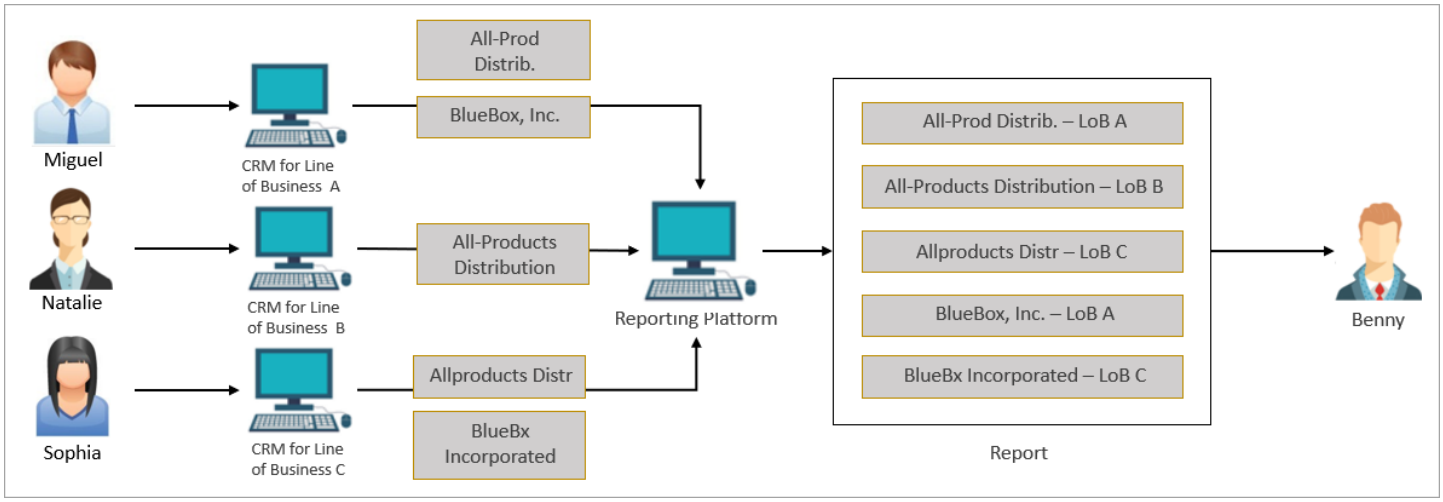
Miguel, Natalie, and Sophia are representatives on the sales team. They are each responsible for managing ACME customer data within the CRM applications that support each line of business.

Benny, also on the sales team, is responsible for compiling reports for ACME's retailers and distributors that have active contracts to resell and distribute ACME products.

What is the Business Problem?

ACME operates across multiple lines of business in which customers may do business with. Miquel, Natalie, and Sophia are updating existing customers and creating new customers within their respective line of business on a daily basis. Since customer data for each line of business is maintained in disparate CRM applications, attempting to aggregate data from multiple CRMs is a major challenge as this aggregation is critical in supporting downstream business operations (ex. Customer Rebates, Financial reporting, behavioral analysis, etc.)

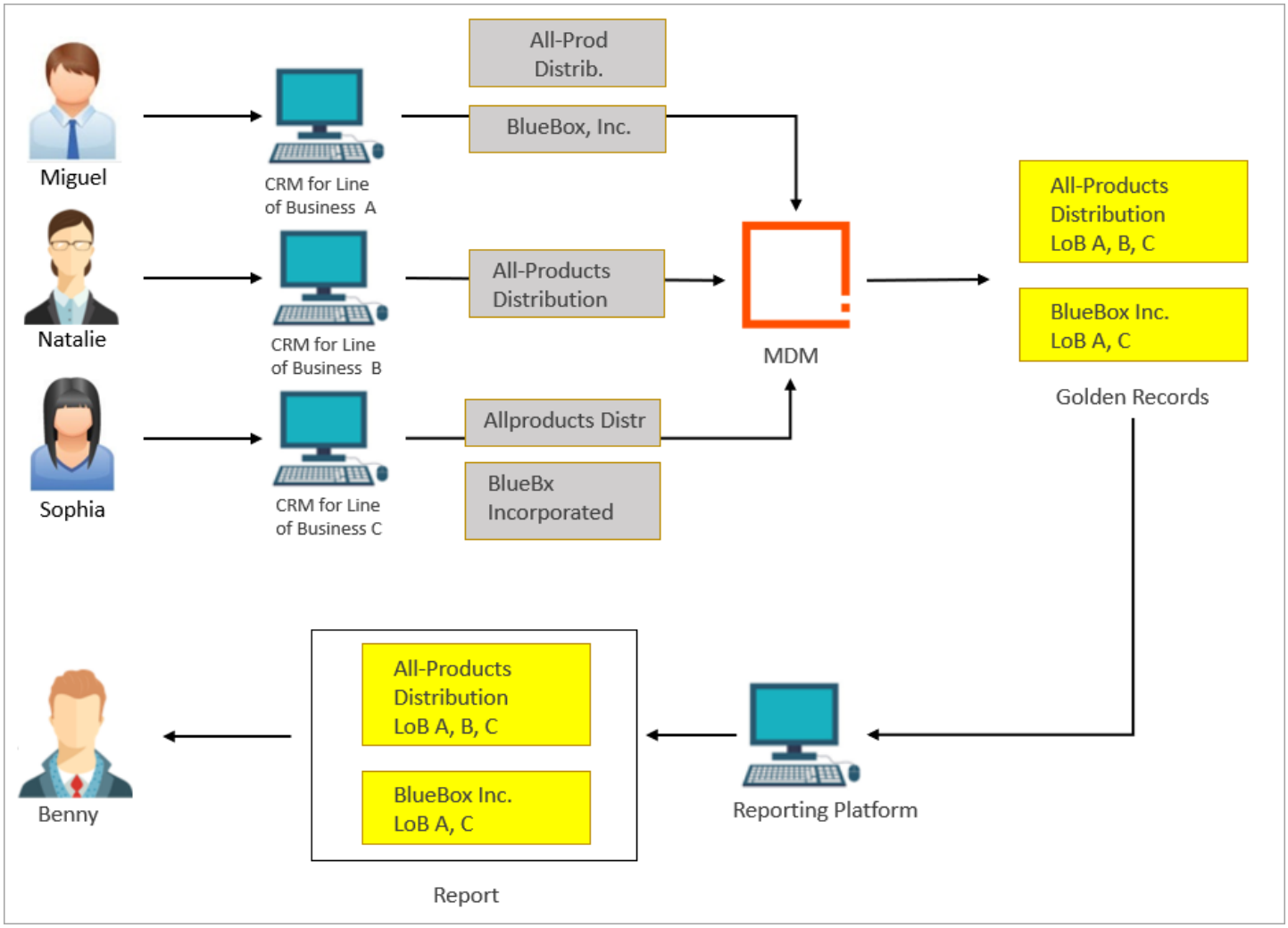
Current Process



With multiple points of data entry in this process, customer data is prone to duplicates across CRMs as well as incomplete and conflicting information being introduced. Furthermore, although data is synchronized from each CRM application to ACME's reporting platform, there is no reliable process for aggregating customer data.

Benny cannot conduct validity checks and other governance measures required by the business; therefore, the reports that he compiles and provides for downstream business operations are prone to inaccuracies.

Future Process



By introducing an MDM solution, ACME can leverage MDM capabilities to identify and connect duplicate customer records, execute standardizations and business logic to validate incoming data, and monitor data completeness. As cleansed data is then synchronized to the reporting platform this enables ACME's B2B customer data to be more accurate and trustworthy for reporting and downstream analytics.

Supplier Data in Disparate Silos Across Business Lines

When organizations operate across multiple lines of business which focus on different offerings and target consumer groups, it is often the case that each line of business is supported by siloed source systems. This presents a challenge as organizations strive to manage their entity data that are maintained in disparate systems yet may overlap from one line of business to another.

Who is Impacted?

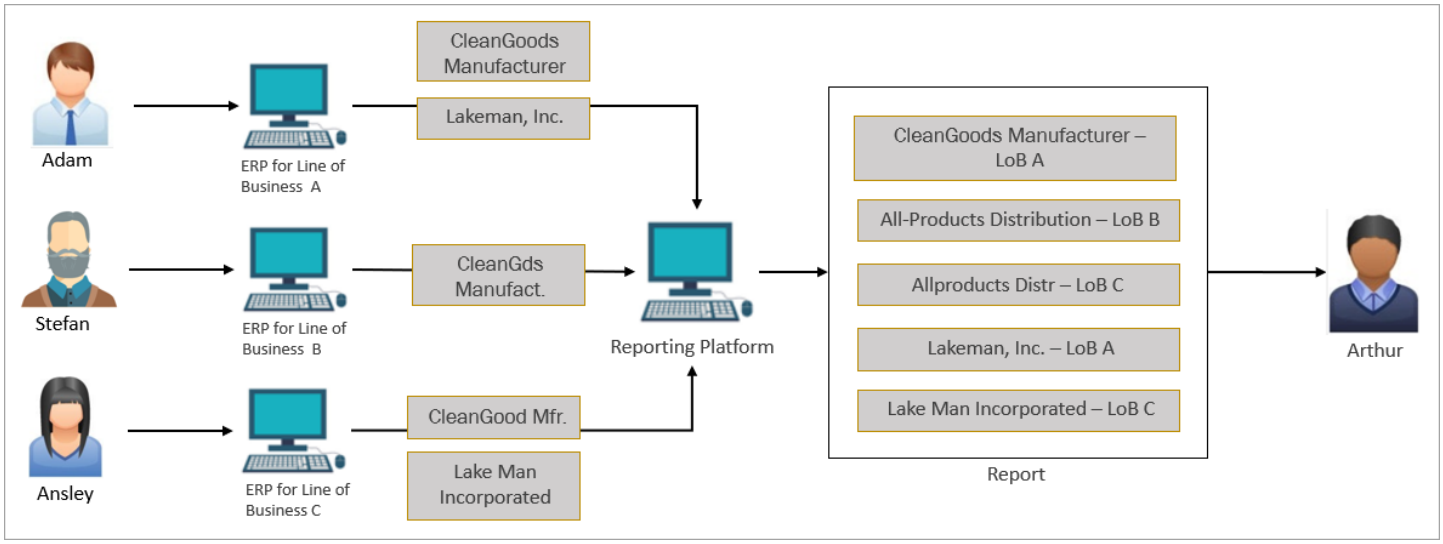
Adam, Stefan, and Ansley are representatives on the procurement team. They are each responsible for managing ACME supplier data within the ERP applications that support each line of business.

Arthur, also on the procurement team, is responsible for compiling reports for ACME's suppliers from whom ACME obtains raw materials and finished goods.

What is the Business Problem?

ACME operates across multiple lines of business in which customers may do business with. Adam, Stefan, and Ansley are updating existing suppliers and creating new suppliers within their respective line of business on a daily basis. Since supplier data for each line of business is maintained in disparate ERP applications, attempting to aggregate data from multiple ERP is a major challenge as ACME Company is not able to produce holistic supplier centric reports.

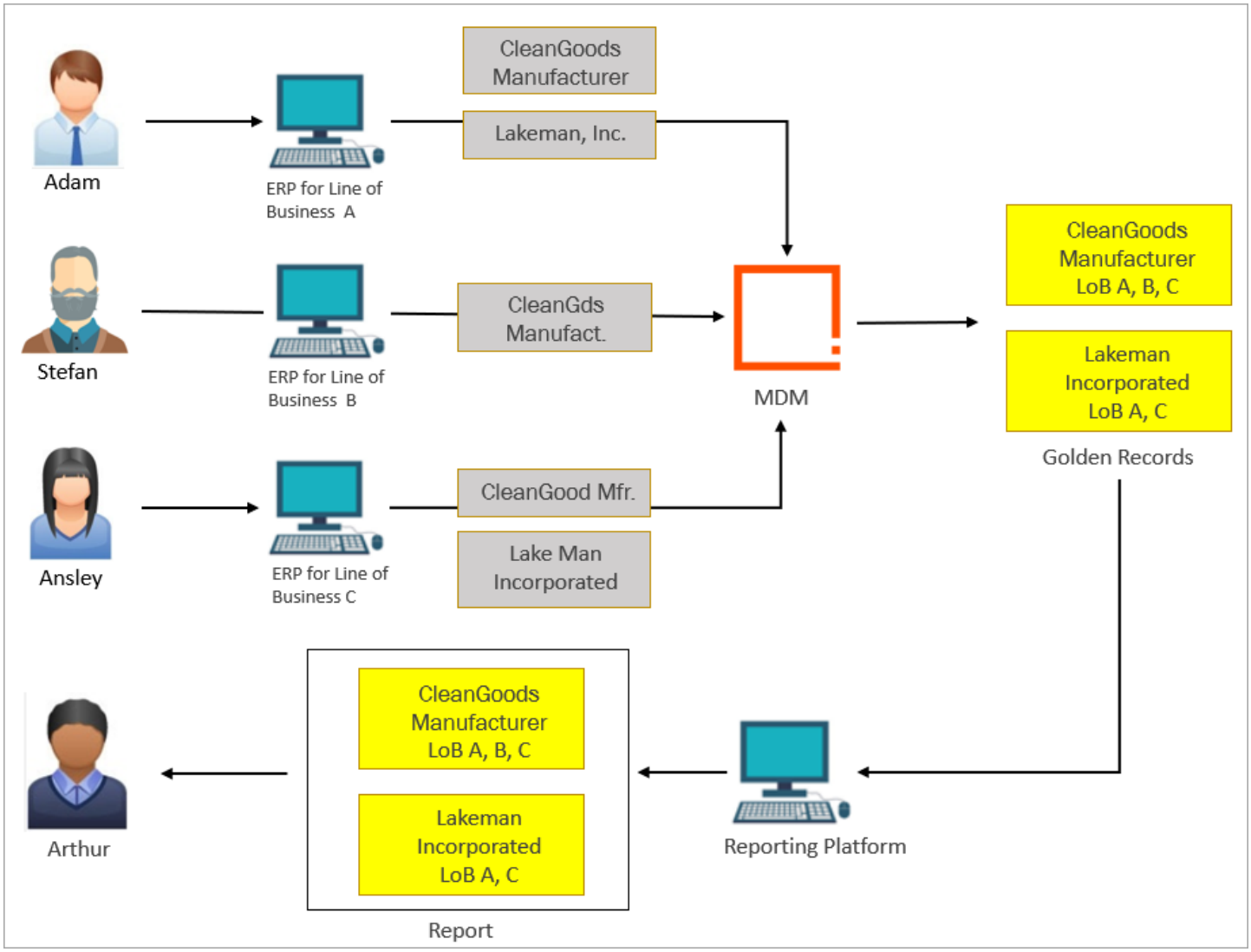
Current Process



With multiple points of data entry in this process, supplier data is prone to duplicates across ERPs as well as incomplete and conflicting information being introduced. Furthermore, although data is synchronized from each ERP application to ACME's reporting platform, there is no reliable process for aggregating supplier data. Arthur

cannot conduct validity checks and other governance measures required by the business; therefore, the reports that he compiles and provides for downstream business operations are prone to inaccuracies.

Future Process



An MDM solution aggregates supplier data from the ERP source systems for all lines of business and automates business validations to ensure completeness and accuracy. MDM can also synchronize with downstream systems to facilitate other business operations. This ensures that duplicate and incomplete supplier records are identified and adjudicated, providing the reporting platform with cleansed, accurate, and up-to-date supplier information.

Understanding Customers' Legal Hierarchies

Who is Impacted?

Benny, an associate on the Sales team, is responsible for managing ACME's retailer/distributor data. It is part of his responsibility to assist in identifying and realizing ACME's sales opportunities. Do to so, Benny relies on ACME's reporting platform to help with his sales analysis.

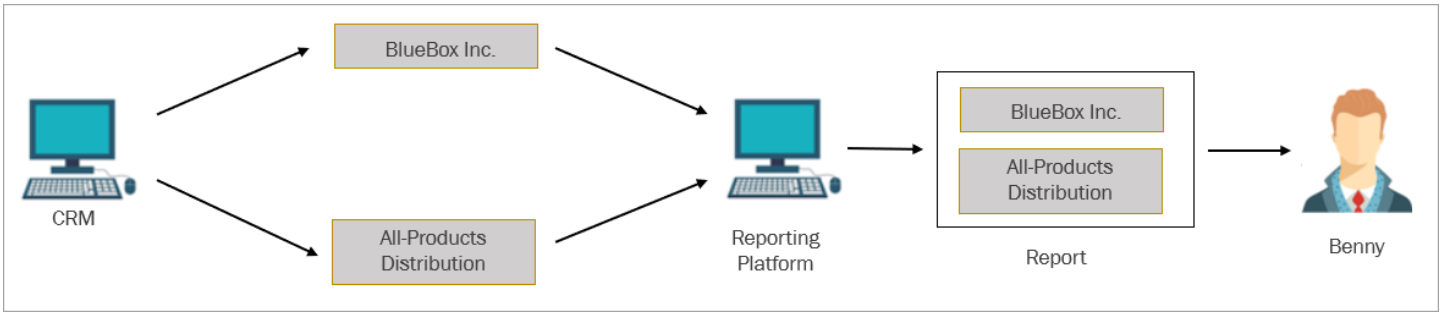
What is the Business Problem?

ACME relies on their reporting platform and sales team to generate new sales leads and maximizing on existing relationships with their current customers. Having a holistic understanding of such relationships and even relationships between customers drives identifying sales opportunities, risk management, and other reporting and analytics exercises.

As a business that relies heavily on distribution channels and retailers to reach the end consumer, it is of particular importance that ACME has a strategic understanding of their relationship with customers and relationships between customers. Realizing how corporations are often related to each other helps sales realize new leads and opportunities. Apart from day-to-day interactions with their customers, ACME lacks third-party industry data to truly maximize their reporting and analytics capabilities.

Furthermore, third-party industry data may also help ACME realize and mitigate risks involving regulatory compliance across various product lines and services that ACME offers.

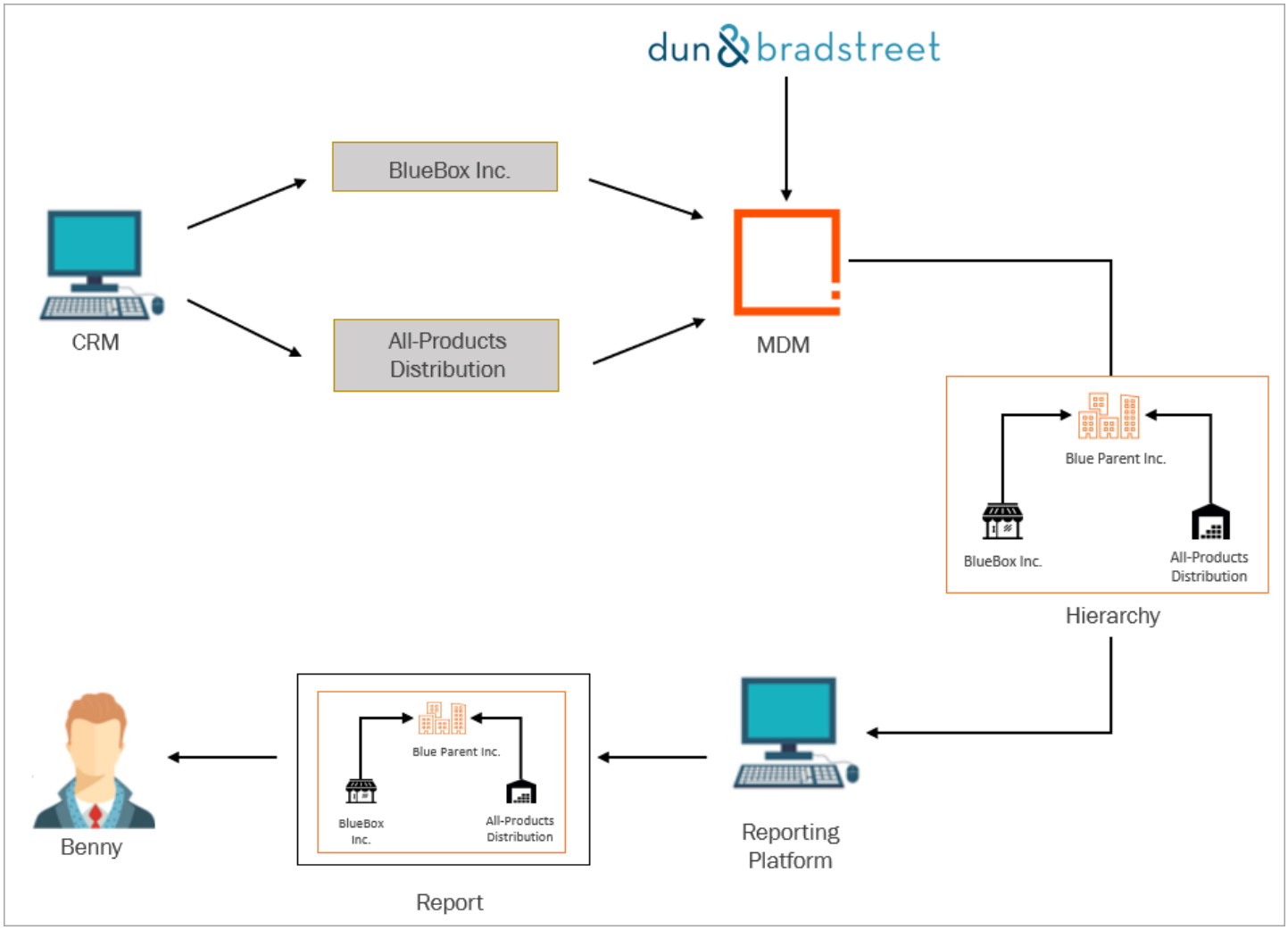
Current Process



In the example of understanding and maximizing customer relationships, two of ACME's customers are subsidiaries of the same parent company. However, this is not realized by the sales team and without the third-party industry profile data, the chances of ACME realizing this relationship is minimal.

As customer data is synchronized from the CRM application to the reporting platform, information provided by the sales team is all that Benny can rely on to extract his sales analysis. Without third-party industry data available to enrich their customer data, potential sales opportunities are not realized.

Future Process



With the introduction of third-party Dun & Bradstreet data, ACME may leverage DUNS numbers to build out customer hierarchies which map newly discovered relationships between existing and prospective customers.

In the example above, BlueBox and All-Products are existing ACME customers. However, the sales team was not aware both customers belong to the same parent company, Blue Parent, Inc. Armed with this information, the sales team can negotiate bulk pricing at a discounted rate and increase their order throughput. Furthermore, ACME can aggregate BlueBox, All-Product, and Blue Parent into a single account grouping and better manage future agreements.

Understanding Suppliers' Legal Hierarchies

Who is Impacted?

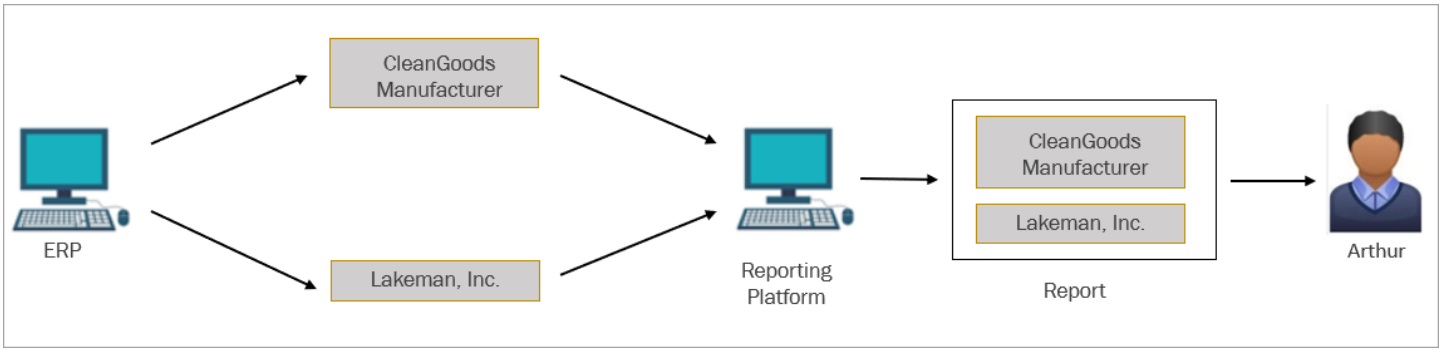
Arthur, a specialist on the procurement team, is responsible for managing suppliers from whom ACME obtains raw materials and finished goods. Among his responsibilities, Arthur strives to streamline ACME's procurement process. To do so, Arthur relies on the reporting and analytics platform to conduct his analysis of supplier sales and performance.

What is the Business Problem?

In working with a multitude of suppliers to procure the necessary products or services for the business, it if often unknown whether there are existing relationships that could be leveraged to optimize ACME procurement processes and costs. Furthermore, being able to understand these supplier relationships also drives downstream reporting/analytical activities. Currently, ACME's ERP application does not maintain such supplier hierarchical relationships as they do not factor in any transactional processes. This greatly limits the value of the reporting and analytics platform without sufficient data.

Without the benefit of third-party industry data, ACME's initiative to better understand their suppliers is greatly limited to their day-to-day interactions with supplier representatives.

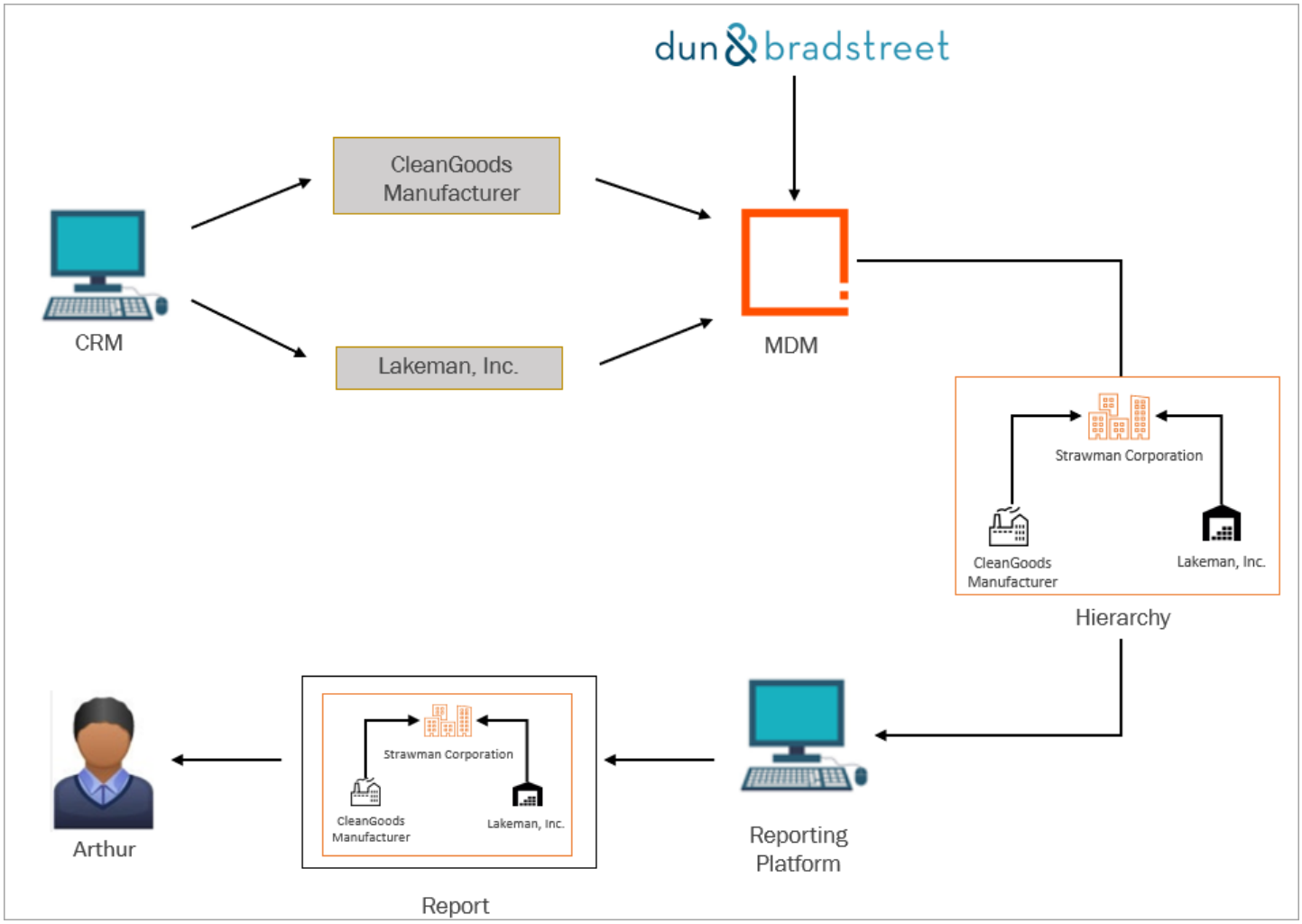
Current Process



For example, two of ACME's current suppliers are subsidiaries of the same parent company. However, this is not realized by the procurement team and without the third-party industry profile data, the chances of ACME realizing this relationship is minimal.

As customer data is synchronized from the ERP application to the reporting platform, information provided by the procurement team is all that Arthur can rely on to extract his sales analysis. Without third-party industry data available to enrich their supplier data, potential cost savings are not realized.

Future Process



By leveraging third-party Dun & Bradstreet data, ACME may utilize DUNS numbers to build out supplier hierarchies which map newly discovered relationships between existing and prospective suppliers.

In the example above, CleanGoods Manufacturer and Lakeman, Inc. are existing ACME suppliers. However, the sales team was not aware both suppliers belong to the same parent company, Strawman Corporation. Armed with this information, the procurement team can negotiate reduced bulk costs for raw materials and goods that ACME require. Furthermore, ACME can aggregate CleanGoods, Lakeman, and Strawman into a single account grouping and better manage future agreements.

Coexistence Business-to-Business Use Cases

As organizations manage their master data without proper discipline or technologies, it is often the case that the lack of such capabilities introduces duplicate data and other inaccuracies into their source systems. However, with an MDM solution in place, entity data may be consolidated into a single repository which may then be de-duplicated, validated, governed, and operationalized to ensure downstream business processes are empowered with trusted, accurate, and up to date information.

In this documentation, the following use case scenarios are covered:

- Synchronization of Customer Accounts
- Synchronization of Suppliers
- Enriching Customer Data
- Enriching Supplier Data

For additional use cases, see the following topics:

- Consolidation Business-to-Business Use Cases
- Coexistence Business-to-Consumer Use Cases

Synchronization of Customer Accounts

When organizations operate across multiple lines of business which focuses on different offerings and target consumer groups, it is often the case that each line of business is supported by siloed source systems. This presents a challenge as organizations strive to manage their customer data that are maintained in disparate systems yet may overlap from one line of business to another.

Having multiple points of data entry comes with data quality and data governance challenges such as data aggregation, cleansing, enrichment, governance, and synchronization. Coexistence MDM builds upon Consolidation MDM by synchronizing mastered customer data back to the source CRM applications. This ensures end users of these applications are presented with trusted, accurate, and up to date customer information.

Who is Impacted?

Miguel, Natalie, and Sophia are representatives on the sales team. They are each responsible for managing ACME customer data within the CRM applications that support each line of business.

Benny, also on the sales team, is responsible for compiling and analyzing reports for ACME's retailers and distributors that have active contracts to resell and distribute ACME products.

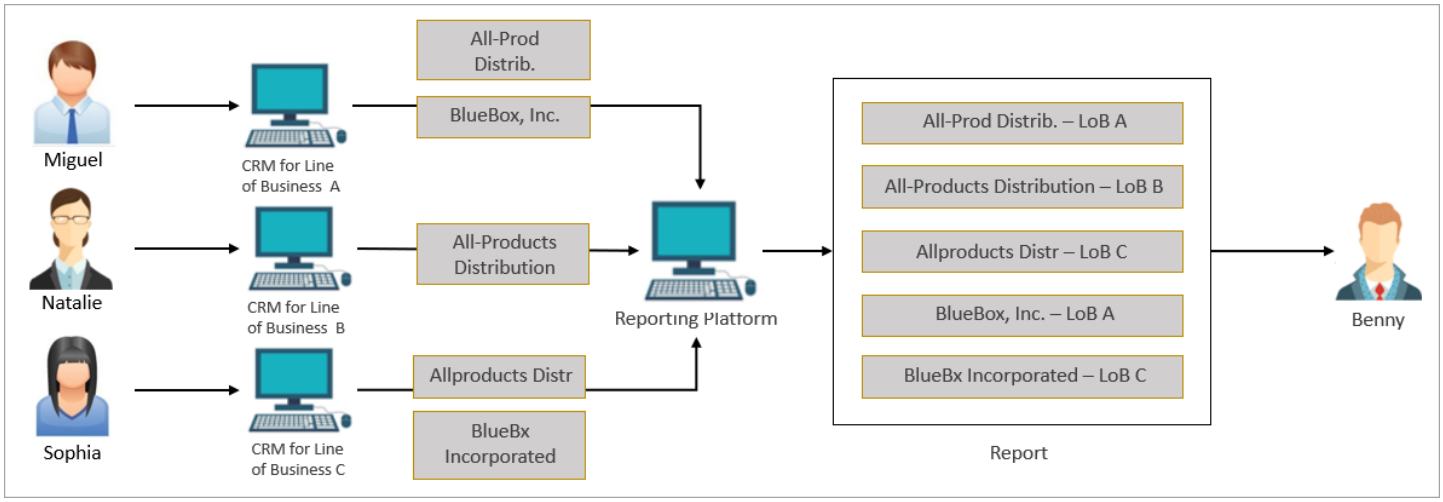
What is the Business Problem?

ACME operates across multiple lines of business in which customers may do business with. Miquel, Natalie, and Sophia are updating existing customers and creating new customers within their respective lines of business on a daily basis. Since customer data for each line of business is maintained in disparate CRM applications which are not in sync with each other, the sales team may manage customer records in their respective CRMs that overlap with each other.

This situation presents a challenge for Benny as ACME frequently compiles sales reports for financial roll-ups that are integral to various operational processes such as strategizing for new sales and marketing campaigns, managing customer charge backs, customer segmentation analysis, etc.

Furthermore, due to this isolated landscape using silos, the lack of synchronization between the various CRMs present operational challenges. For example, a customer's address may have been updated by Natalie within CRM for one line of business but remain outdated for Sophia as she works within CRM for another line of business. This may impact activities such as mailing invoices, business catalogs, or invitations to conferences and promotional events.

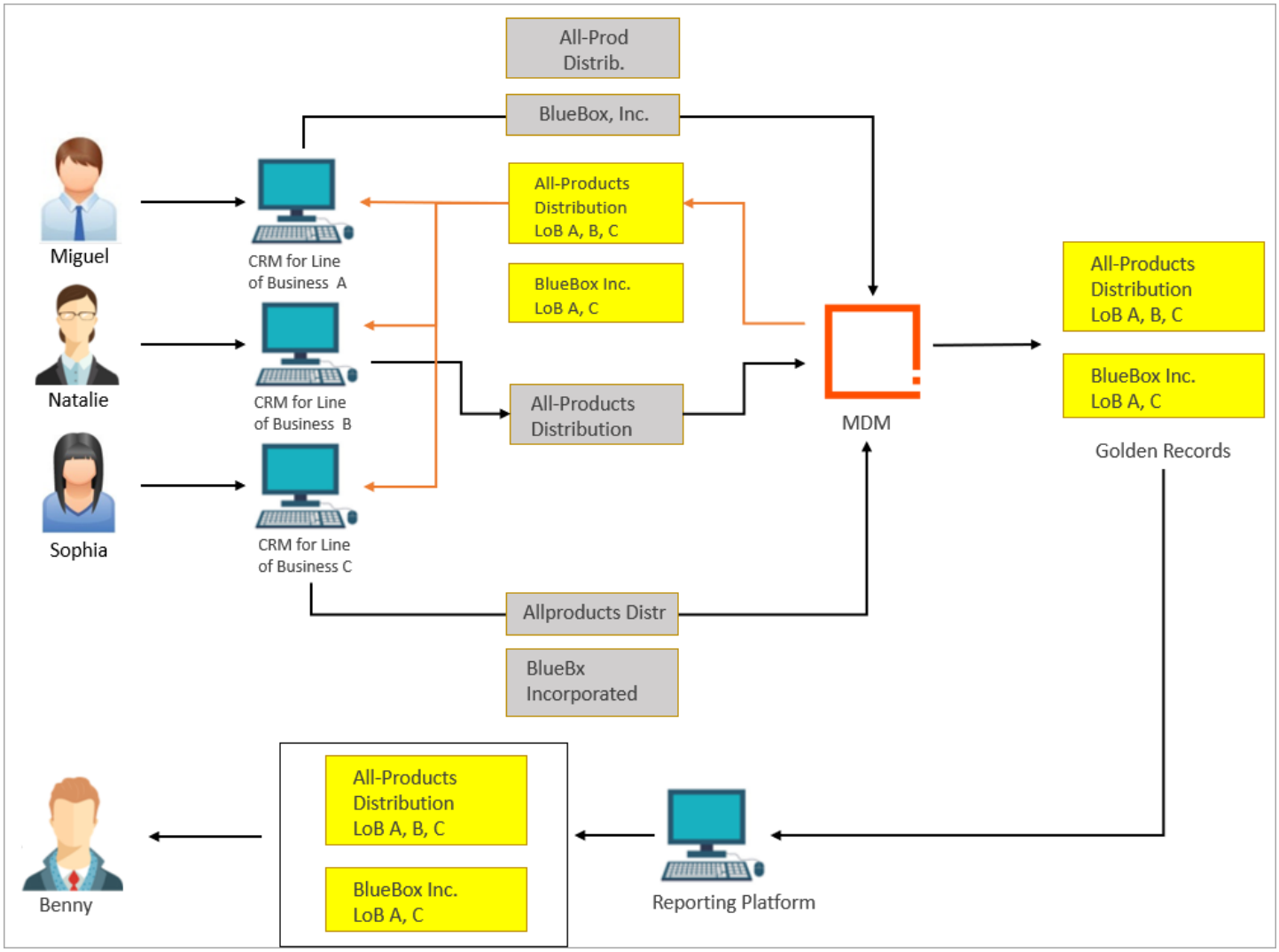
Current Process



With multiple points of data entry in this process, customer data is prone to duplicates across CRMs as well as incomplete and conflicting information being introduced. Furthermore, although data is synchronized from each CRM application to ACME's reporting platform, there is no reliable process for aggregating customer data. Benny cannot conduct validity checks and other governance measures required by the business; therefore, the reports that he compiles and provides for downstream business operations are prone to inaccuracies.

Since there is no synchronization between the three CRM applications, ACME's sales team could be updating address information in one application will not be propagated to other applications, leading to mis-mailed invoices, catalogues, or invitations to industry conferences.

Future Process



An MDM solution will provide the CRM applications the means to preemptively search within MDM for customer records before new records are created in CRM. If the customer already exists based on minimal search parameters, MDM will provide the identifier of the record. If the customer does not exist in MDM, then Miguel, Natalie, and Sophia may proceed with creating the new record in CRM.

MDM will then aggregate customer data from the CRM source systems for all lines of business and automates business validations to ensure completeness and accuracy. In addition to sending the mastered customer data to downstream systems to facilitate other business operations, MDM will also send the mastered customer data back to upstream CRM source systems. This process ensures the synchronization of mastered customer data across all CRM applications which provides Miquel, Natalie, and Sophia with accurate and up-to-date customer data.

Synchronization of Suppliers

Who is Impacted?

Adam, Stefan, and Ansley are representatives on the procurement team. They are each responsible for managing ACME supplier data within the ERP applications that support each line of business.

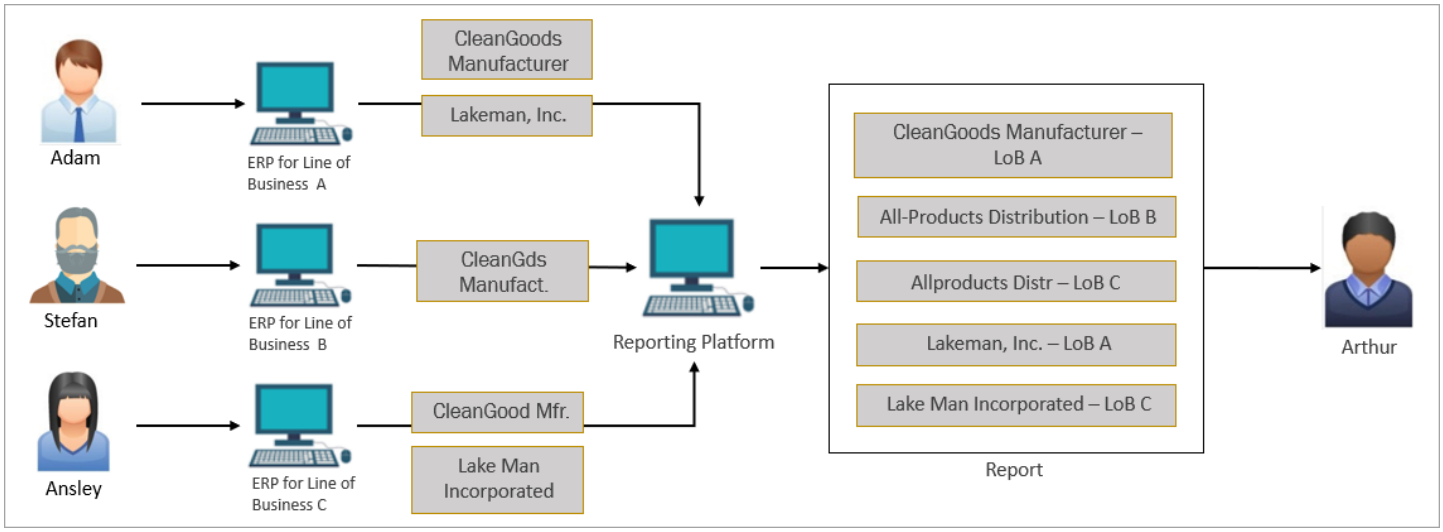
Arthur, a specialist on the procurement team, is responsible for compiling / analyzing reports and performance KPIs for suppliers from whom ACME purchases raw materials, finished goods and services.

What is the Business Problem?

ACME's supplier data is currently managed across multiple different ERP source systems due to differences in territories and lines of business. Since suppliers may conduct transactions across multiple territories, managing the same suppliers across multiple source systems is a challenge as the source systems lack the capability to synchronize with each other.

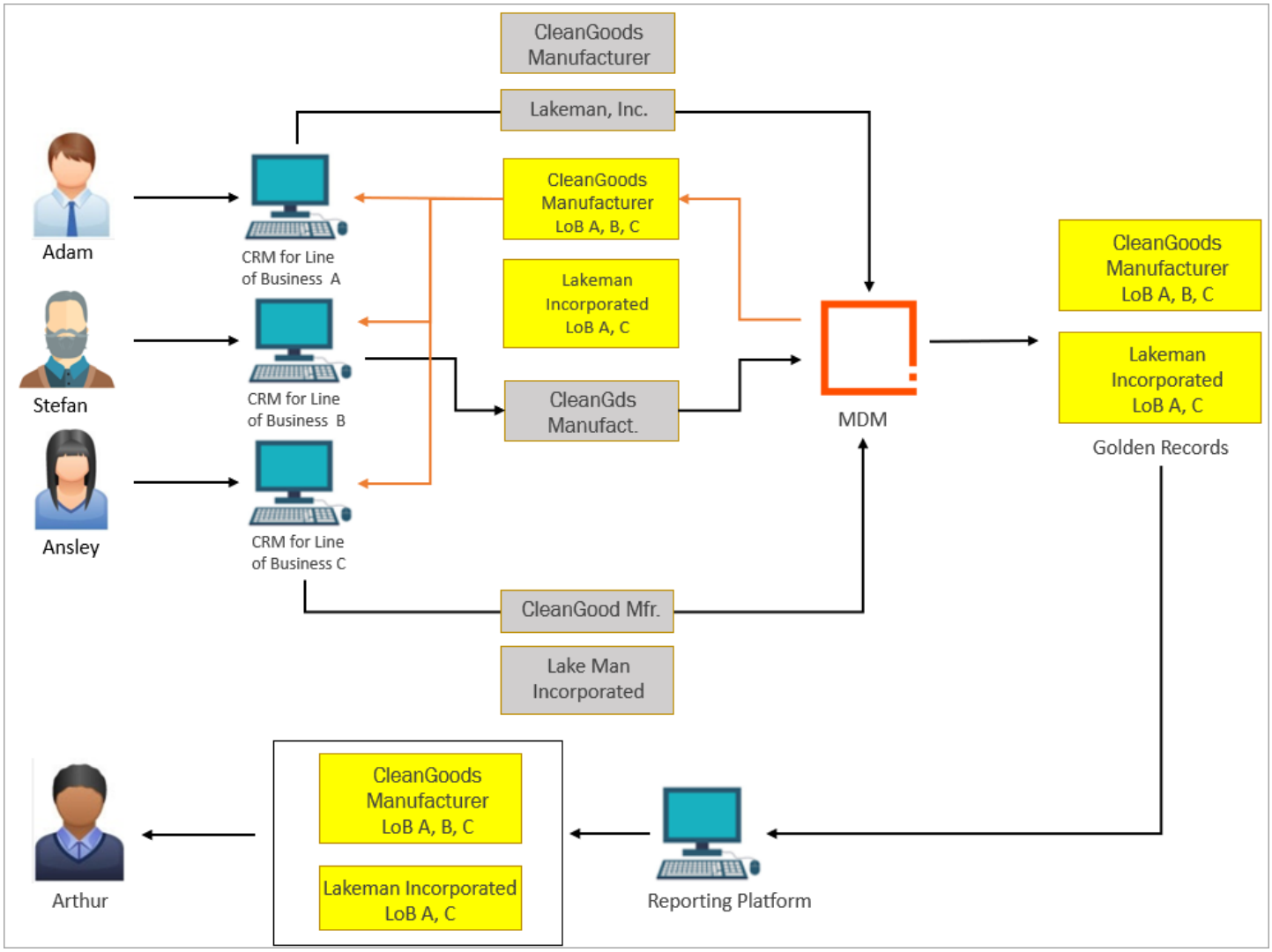
Due to this siloed landscape, the lack of synchronization between the various ERP presents operational challenges. This introduces discrepancy in the customer data which currently cannot be managed or mitigated. For example, Adam may update an address or other contact information for a particular supplier that is also managed by Ansley within a different ERP.

Current Process



With multiple points of data entry in this process, supplier data is prone to duplicates across ERPs as well as incomplete and conflicting information being introduced. Furthermore, although data is synchronized from each ERP application to ACME's reporting platform, there is no reliable process for aggregating customer data. Arthur cannot conduct validity checks and other governance measures required by the business; therefore, the reports that he compiles and provides for downstream business operations are prone to inaccuracies.

Future Process



An MDM solution will provide the ERP applications the means to preemptively search within MDM for supplier records before new records are created in CRM. If the supplier already exists based on minimal search parameters, MDM will provide the identifier of the record. If the supplier does not exist in MDM, then Adam, Stefan, and Ansley may proceed with creating the new record in ERP.

MDM will then aggregate supplier data from the ERP source systems for all lines of business and automates business validations to ensure completeness and accuracy. The resulting golden supplier record(s) are sent to downstream systems to facilitate other business operations. MDM will also synchronize the mastered supplier data back with upstream ERP source systems.

This process ensures the synchronization of mastered supplier data across all ERP applications which provides Adam, Stefan, and Ansley with accurate and up-to-date customer data. As a result, all ERP source systems now have the same, up to date address information available.

Enriching Customer Data

Within a coexistence MDM landscape, operational master data must, at times, be further enriched with additional industry data. Such data is usually procured through third-party channels and provides valuable insight to persons and parties that an organization does business with. Such data is especially useful when operationalized as it provides input to drive more efficient operations which influences strategic business decisions.

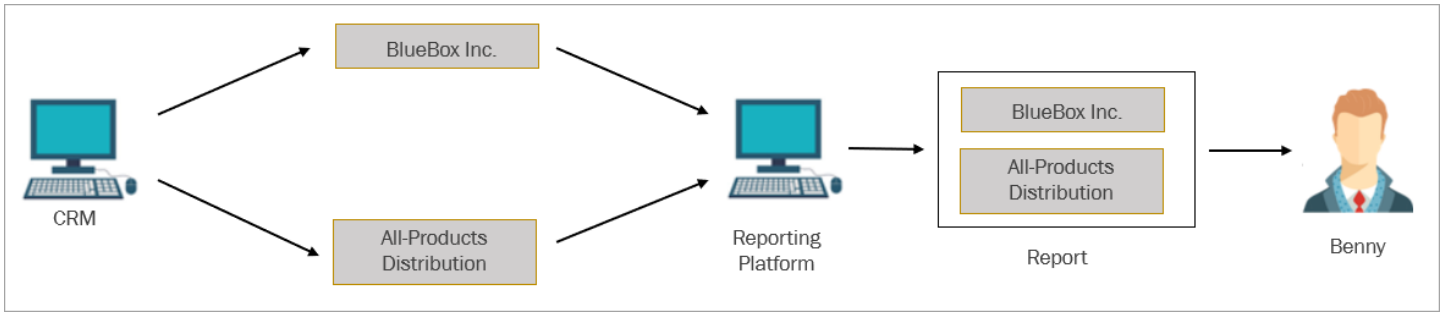
Who is Impacted?

Benny, an associate on the Sales team, is responsible for managing ACME's retailer / distributor data. It is part of his responsibility to assist in identifying and realizing ACME's sales opportunities.

What is the Business Problem?

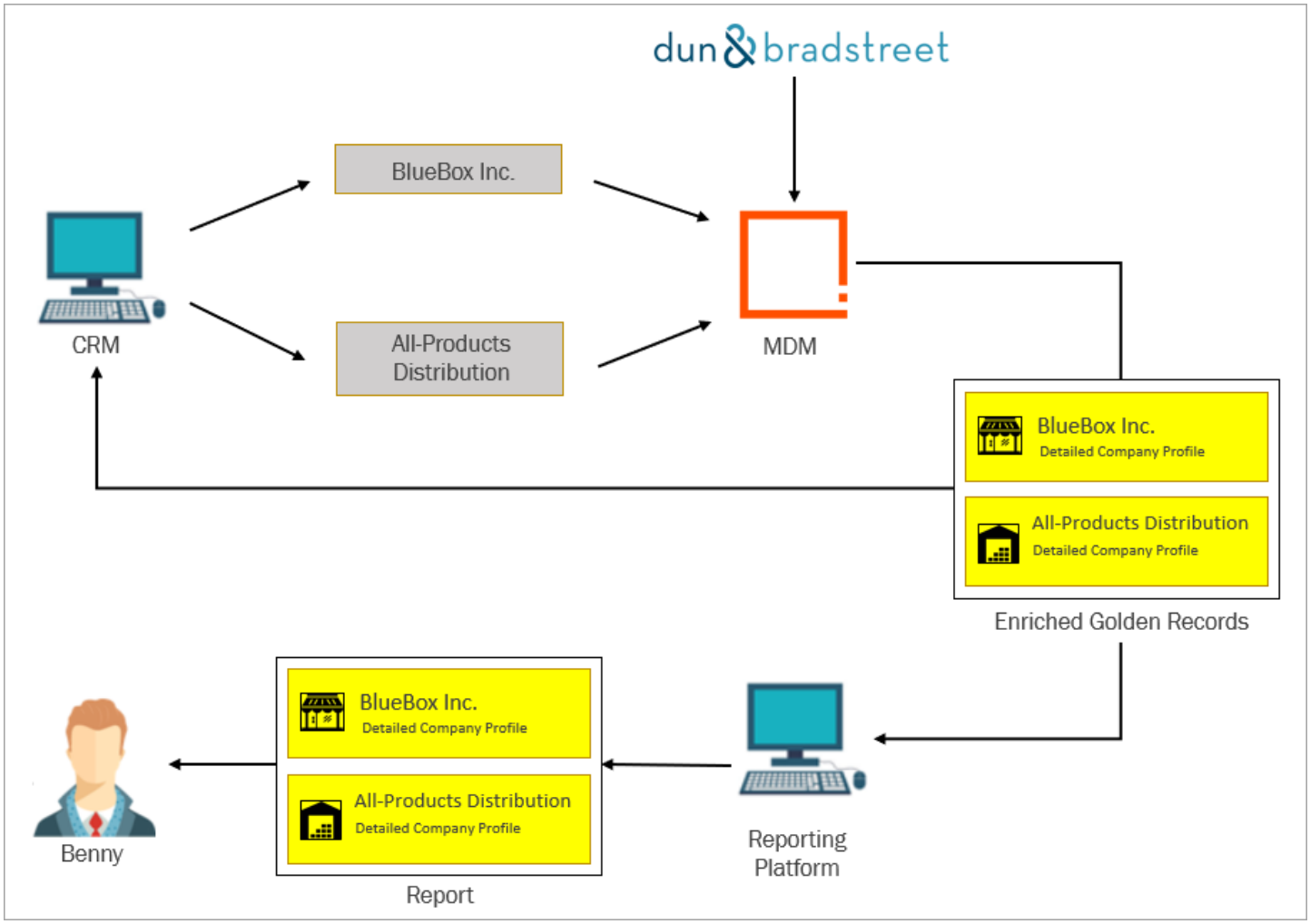
ACME relies on their sales marketing team to generate new sales leads and realize their financial targets. Without the assistance of industry data from third-party sources, it is extremely challenging to conduct customer segmentation and analysis to identify new opportunities

Current Process



Currently, all new sales leads are generated by person-to-person contact by Benny and the sales team. As they gather more information about their existing customers and prospects, they are entered into their CRM system. However, since the sales team does not have access to third-party industry data, the amount of data they are able to gather is limited, and thus hinders their ability to expand their sales opportunities.

Future Process



With the introduction of third-party Dun & Bradstreet data, the sales team now has new insights into prospective customers and can carry out activities such as customer segmentation analysis and other activities driven by sales and marketing. This allows sales to better target prospects by fully understanding their customer profiles and maximizing their cross-sell and up-sell opportunities.

Enriching Supplier Data

Who is Impacted?

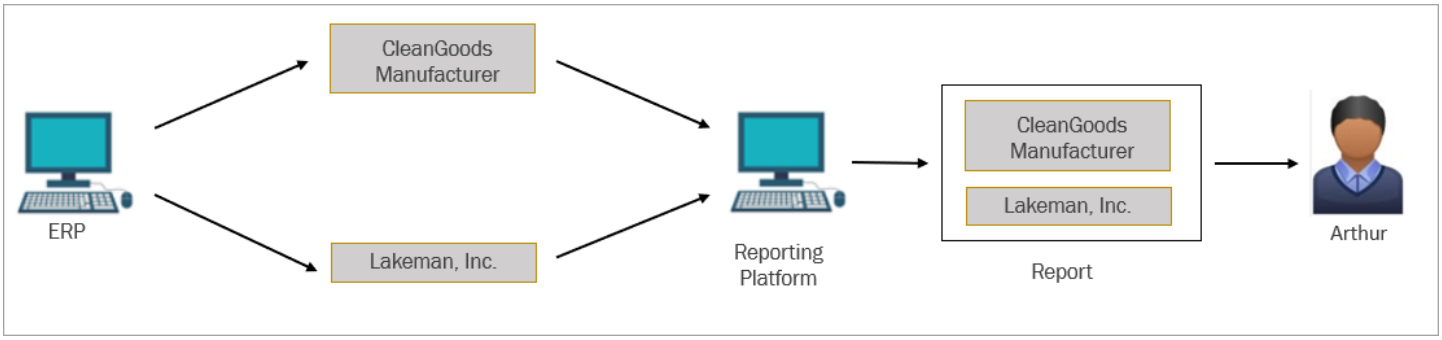
Arthur, a specialist on the Procurement team, is responsible for managing suppliers from whom ACME obtains raw materials and finished goods.

What is the Business Problem?

Beyond information that is provided by our suppliers, our organization really do not know much about our suppliers. From a marketing and procurement perspective key information to determine potential prospects and identify cross-sell and up-sell opportunities is lacking. Such information helps determine target market, leverage market segmentation, and enhances procurement processes to realize financial savings for various products and services.

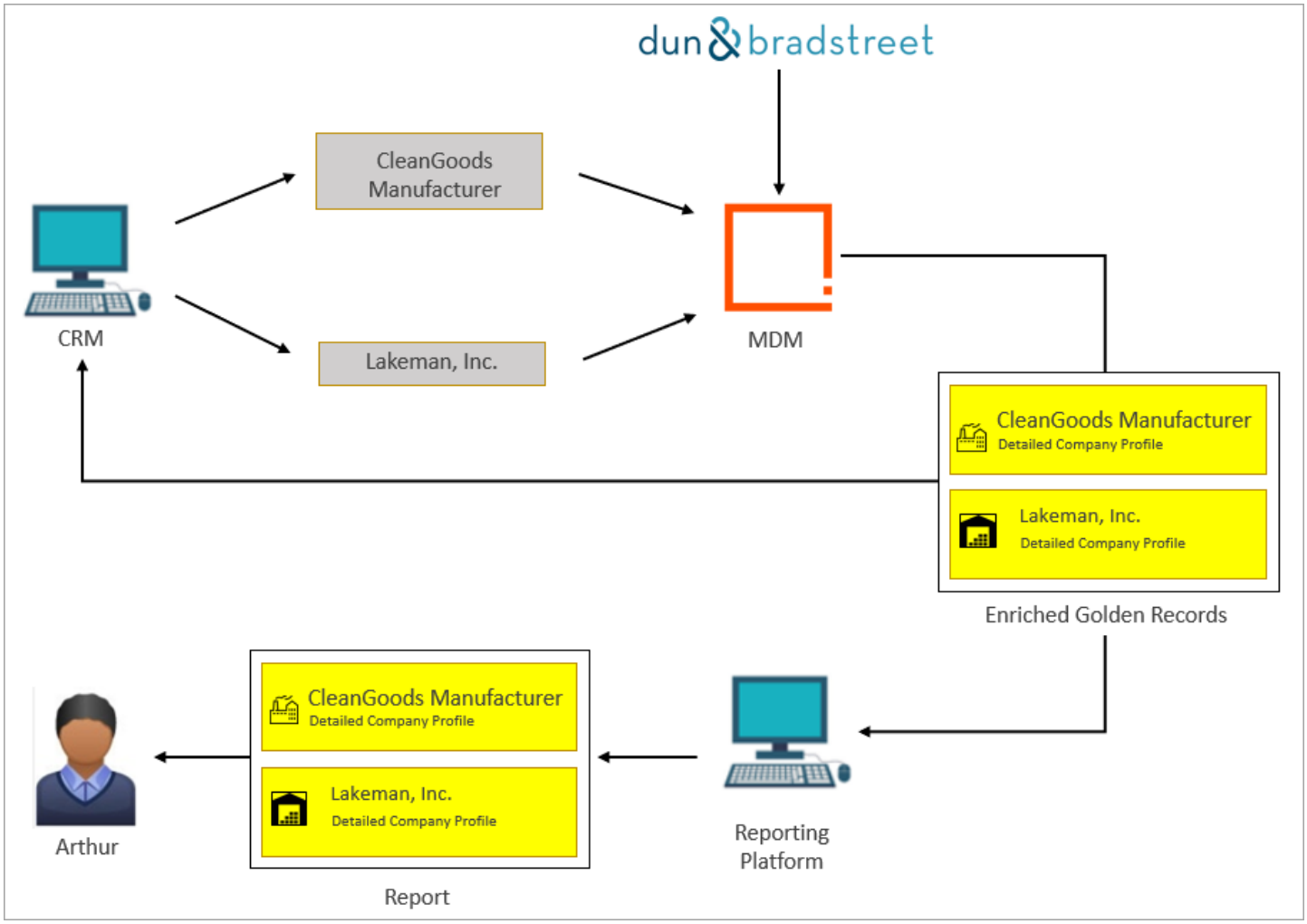
This is because our supplier contacts may not be privy to such industry-related information, or the onboarding and enrichment processes do not require this level of information. In order to maximize our supplier relationships, we need to know more details of who are suppliers are, who they have relationships with, what their financial status is, etc. Such information provides valuable supplier and industry insight.

Current Process



Arthur and the Procurement team evaluates new prospective suppliers and approves them for ACME Company. During their evaluation process, it is often not enough to simply rely on the information the prospective supplier provides, and Arthur needs to conduct some additional research to determine eligibility. Since he does not have access to third-party industry data, it can be challenging for Arthur to make an informed decision on a supplier.

Future Process



With the introduction of third-party Dun & Bradstreet data, Arthur now has new insights into prospective suppliers such as credit and financial information, ownership structure, etc. Such data allows Arthur and his team to make informed decisions on prospective suppliers, as well as grow sales opportunities.

Coexistence Business-to-Consumer Use Cases

As organizations manage their master data without proper discipline or technologies, it is often the case that the lack of such capabilities introduces duplicate data and other inaccuracies into their source systems. However, with an MDM solution in place, entity data may be consolidated into a single repository which may then be de-duplicated, validated, governed, and operationalized to ensure downstream business processes are empowered with trusted, accurate, and up to date information.

In this documentation, the following use case scenarios are covered:

- Synchronization of Consumer Data
- Empowering Consumer E-Commerce

For additional use cases, see the following topics:

- Consolidation Business-to-Business Use Cases
- Coexistence Business-to-Business Use Cases

Synchronization of Consumer Data

As an organization's business grows with their consumer base, customer data must be managed with proper discipline with the proper technologies. This is expounded when businesses require multiple source systems to manage their ever-expanding volume of customer and supplier data. With multiple points of data entry comes with data quality and data governance challenges. This is where coexistence MDMs aggregate master data from the various sources by cleansing, enriching, governing, and synchronizing golden record data back to the individual source systems. In this coexistence approach, MDM becomes a driver for both upstream systems and downstream operational activities.

Who is Impacted?

Catherine Yu lives in Texas and enjoys shopping at ACME's clothing brand store. She signs up for the ACME Clothing's customer loyalty program to take advantage of weekly discounts and exclusive sales. Catherine is also an avid do-it-yourself home improvement hobbyist and frequents ACME's Home Improvement stores. Catherine decides to sign up for ACME Home Improvement's customer loyalty program since it offers similar shopping perks that's exclusive to this brand.

Sandra manages Catherine's order from a B2C e-commerce site which contains no address validation, allows multiple accounts to be created for a single customer, and lacks any form of de-duplication of customer records. Lacking address validation leads to missed order deliveries which damages brand reputation. Allowing multiple customer accounts of a single record due to no deduplication leads to loss of cross sell opportunities as an email or sales catalog may be sent based off the information of an old account. Deduplicating and consolidating these records allow all contact information to be available on a single entity, eliminating errors caused by multiple conflicting entities

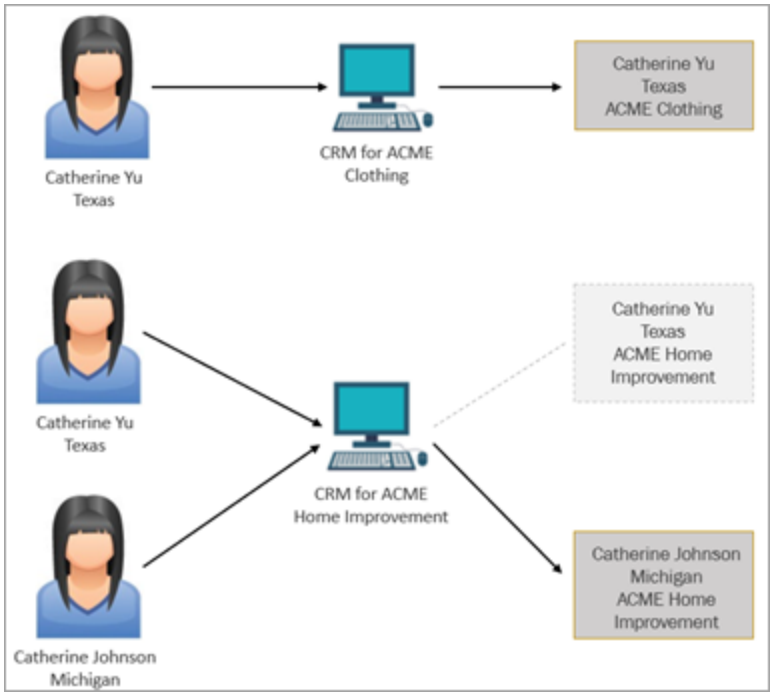
Charles Johnson, Catherine's fiancé, is from Michigan and is an occasional shopper of ACME brand clothing. As a loyalty program member, Charles has decided to take advantage of recent promotions in preparation for their wedding. Upon marriage, Catherine has chosen to take Charles' last name and moved to Michigan with Charles.

What is the Business Problem?

When Catherine becomes married and moves to Michigan, she not only took on a different last name but also a new address. Being a newlywed and living in a new house, Catherine decides to take on some DIY projects and frequents ACME Home Improvement. In order to gather her shoppers' rewards points, she updates her loyalty program profile to reflect her new address and last name.

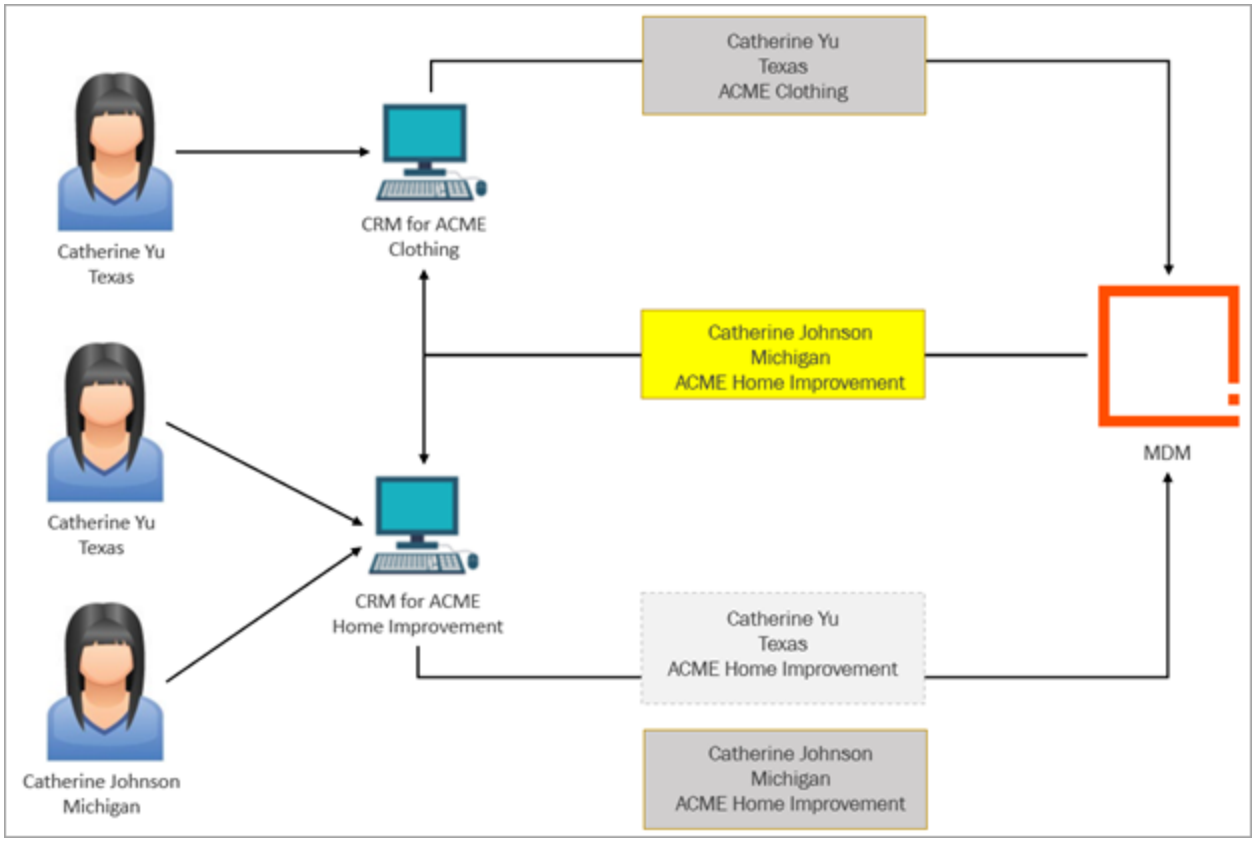
ACME Clothing and ACME Home Improvement customer loyalty programs are managed by different CRM applications that are siloed with no synchronization between the two systems. Although Catherine initially enrolled in both programs under her maiden name and Texas address, ACME Corporation had no mechanism in place to identify and recognize that both records for Catherine were actually for the same customer. The situation is further complicated now that Catherine has updated her profile within one CRM with a different surname and address but not the other. This lack of data governance and management prevents ACME from realizing marketing and sales opportunities across brands, and even introduces unnecessary cost since ACME Clothing marketing ads and other collateral are still being erroneously mailed to the old Texas address.

Current Process



With multiple points of data entry in this process, supplier data is prone to duplicates across ERPs as well as incomplete and conflicting information being introduced. Furthermore, although data is synchronized from each ERP application to ACME's reporting platform, there is no reliable process for aggregating customer data. Arthur cannot conduct validity checks and other governance measures required by the business; therefore, the reports that he compiles and provides for downstream business operations are prone to inaccuracies.

Future Process



An MDM system can identify that Catherine Yu across the two CRMs is actually the same person and merge the two into a single golden record that is validated, standardized, and trusted. The synchronization of the golden record back to the CRMs ensures data is accurate and up to date at the source and any updates to the consumer record such as Catherine's surname and address will also be validated, standardized, and synchronized across both CRMs with MDM as the hub.

This allows ACME's marketing team to effectively identify with their consumer base with tailored marketing campaigns to maximize promotions and cross-sell opportunities across brands.

Empowering Consumer E-Commerce

Master data supports all data touch points an organization has with their customers, including how consumers interact with the organization via e-commerce.

Who is Impacted?

Sandra is responsible for managing orders coming from the new B2C e-commerce site, as well as returns. She succeeds when she can deliver the right product to the right customer at the right time.

What is the Business Problem?

ACME Company's online shopping experience allows consumers to create loyalty accounts by providing their contact information (name, email, shipping, and billing addresses). ACME's infrastructure is such that their e-commerce platform which drives the web shopping experience is directly integrated with their ERP platform. When new consumer accounts are created or new orders are placed, the information is sent to ERP for processing.

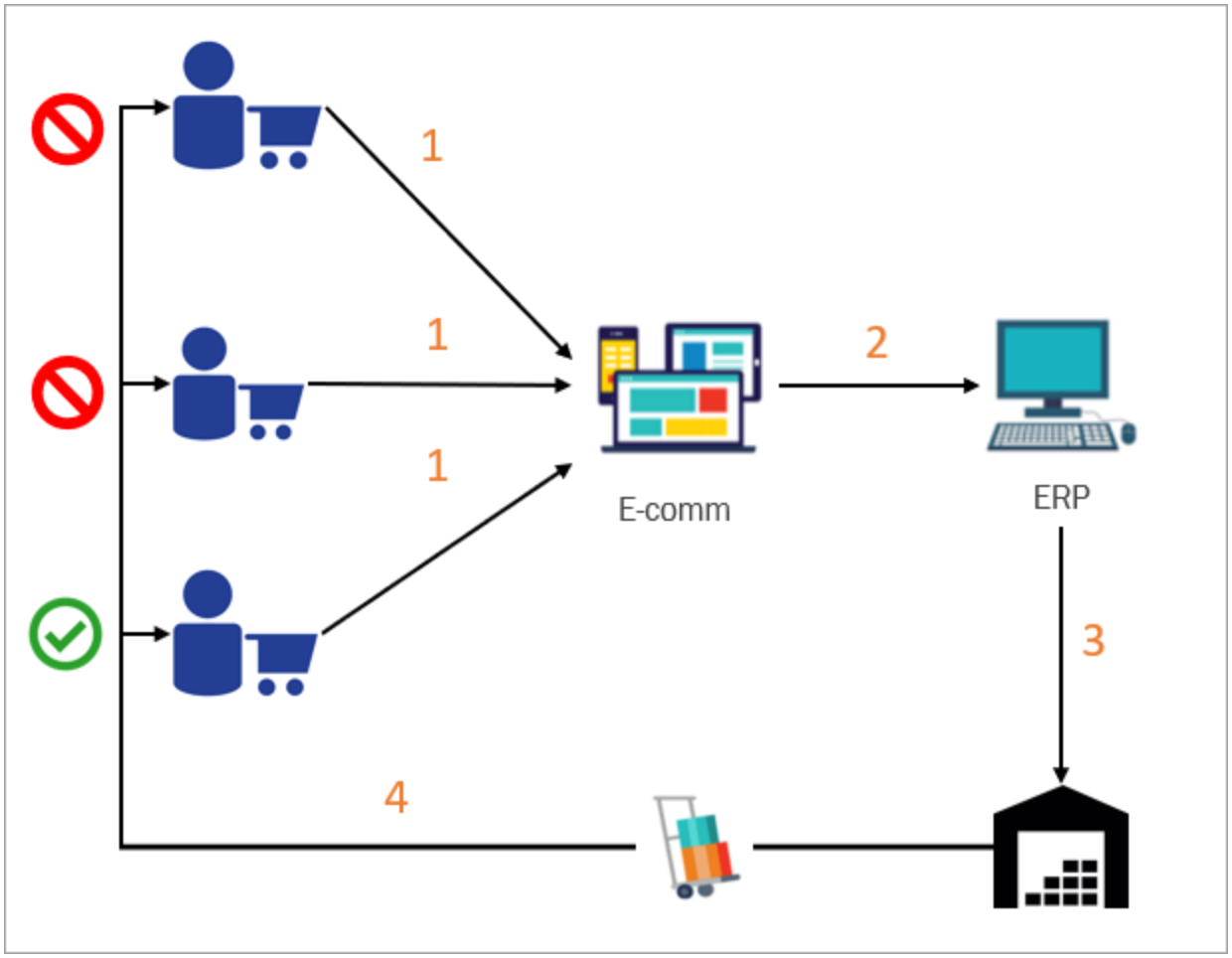
However, both ACME's current e-commerce and ERP platforms lack sufficient data validation and governance capabilities which causes low quality data to be introduced at the point of entry. Poor data quality has created several challenges for ACME Company, preventing them from achieving their goals:

- Addresses are not validated before the order is sent to logistics, leading to a lot of products delivered to wrong addresses or not arriving to the right customer at the right time.
- Sometimes customers accidentally misspell their address or leave out vital information.
- Duplicates are and will continue to be created in the e-commerce and ERP systems.
- Neither the e-commerce system nor ERP can check for duplicates, obfuscating how much a customer actually buys from ACME.
- Customers do not receive their orders in a timely manner (or at all).
- Inconsistent customer experience, as duplicate accounts may lead to orders not being associated between duplicates.

These issues result in:

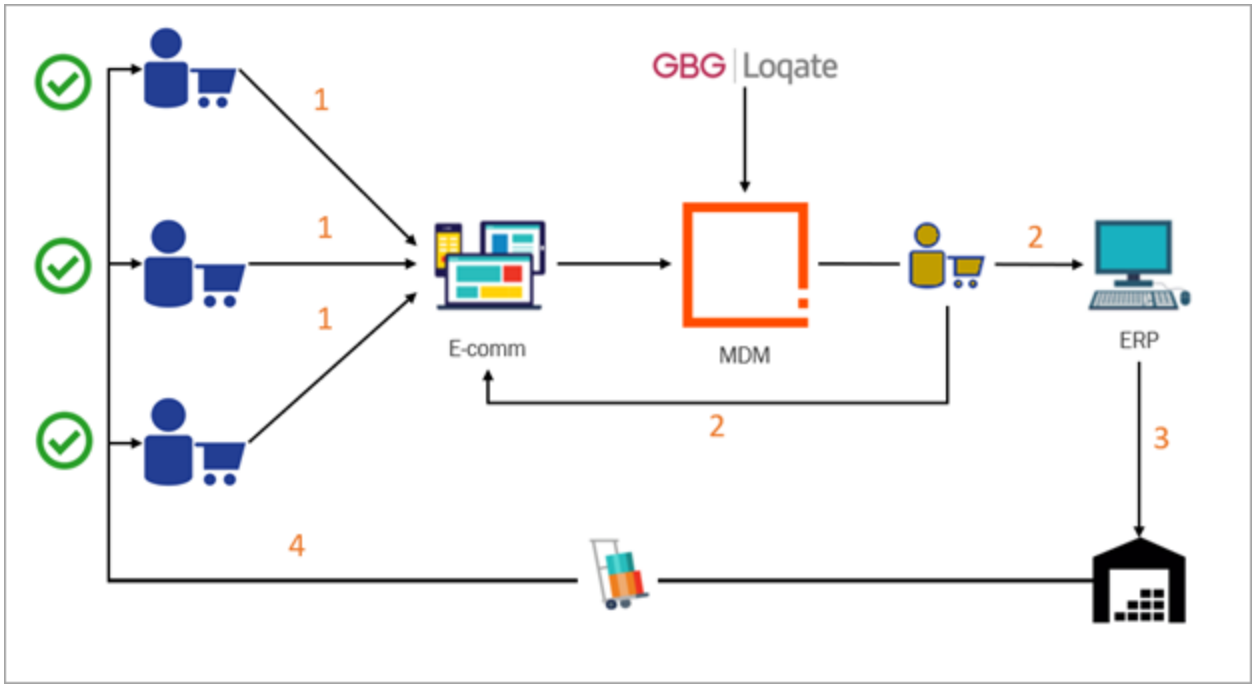
- Low data quality of delivery locations leads to instances of 'return to sender' and double shipping, causing increased costs.
- Poor customer experience due to orders not arriving in a timely manner.
- Lost revenue and loss of customers. If the products are not delivered in a timely manner, customers will buy products from elsewhere.
- Increased product returns.
- Duplicate records carry operational and analytical consequences. Operationally, the correct delivery address may be available on a duplicate record. This leads to customer frustration if ACME is not considering that record when shipping. Analytically, reports may be skewed as duplicate records cause a single customer to be weighed multiple times during analytics.

Current Process



The customer places orders that are processed out of an e-commerce system. The customer and order details are sent to the ERP system for processing and to get shipments out. Currently, there are no validations being done on the provided address. This causes customers to not receive their purchases in a timely manner or not at all. Additionally, shipments may be returned to sender causing increased costs.

Future Process



The MDM solution is integrated with Loqate to verify and standardize addresses. This increases address quality, reducing instances of 'return to sender' and deliveries shipped to the wrong address. Using Loqate as a web service, the addresses can be standardized at the point of entry in the Ecommerce system. Additionally, an e-commerce administrator can manage customer accounts and have addresses standardized without directly interacting with MDM.

Use Case Appendix

The following use cases address several common challenges organizations face when handling customer data, and how they are solved by a MDM solution.

Business-to-Customer Use Cases

- Merging Entity Records via Automatic, Manual, or Advanced merging methods
- Rejecting Potential Duplicates (during Clerical Review)
- Relocation of Individual Customers
- Trigger Clerical Review after Manually Editing Customer Record
- Address Validation
- Unmerging Golden Records
- Find Similar Preventing Duplicates

Business-to-Business Use Cases

- Identifying and Deduplicating Contacts Within the Same Organization and Handling Contacts across Different Organizations
- Find Similar Acquisitions

Merging Entity Records

In these use cases the data steward is responsible for overseeing the deduplication of customer records via different merge processes.

Auto-Merge Records

ACME has several external source systems which contain multiple entity records with separate identifiers (source record IDs). Despite these different identifiers, many of these records are in fact the same individual but with minor variations in demographic data.

These individual records are sent to MDM and automatically deduplicated via the configured matching logic, which identifies and merges identical records. These matched records are then merged, and a single golden record is created.

Original source systems and respective source record IDs are maintained as part of the golden record for source traceability purposes.

Merge Entity Records

ACME has several entity records which are considered potential duplicates of one another, but the system cannot ultimately determine whether or not they are true duplicates without the input of a data steward. In these cases, a data steward must review potential duplicate records and corresponding demographics to intelligently determine whether they are a match or not.

Once STEP determines which records are potential duplicates of one another, a Clerical Review task list is presented to the data steward in the Web UI. Data stewards may review the demographics of each task to determine whether the records should be merged. If it is determined that the tasks are a match, then the **Merge** action will merge the records into a single golden record. If it is determined that the records are not matches of each other, then the **Reject** action will leave the records unmerged.

For matched records, a single resulting golden record will survive that contains the source system and source record ID of all records that were merged. For non-matched records, a Confirmed Non-Duplicate reference is established to prevent future clerical review tasks from being created for the same records. Lastly, the existing task is removed from the Clerical Review task list / workflow.

Advanced Merge Entity Records

As an alternative to the **Merge** function described in the above use case, data stewards can use the **Advanced Merge** option to manually dictate which demographics survive the merge between two or more potential duplicates. In addition, they may choose to exclude certain records from the merge.

Once STEP determines which records are potential duplicates of one another, a Clerical Review task list is presented to the data steward in the Web UI. Data stewards may select a task, and by clicking **Advanced Merge**, they are presented with a table view of the associating records as well as corresponding demographics. As the data steward selects which demographics will survive the merge and be promoted to the golden record, a preview column displays the resulting merged golden record. Should the data steward determine that one or more records should not be considered as a potential duplicate, the option to **Exclude Record** is provided.

The resulting golden record reflects the demographics manually selected by the data steward. The task is removed from the Clerical Review task list once the merge is complete.

Unmerging Golden Records

In these use cases, the data steward Stew is responsible for overseeing the unmerge of entity records within different scenarios.

Simple Unmerge

ACME continues to operate with source systems feeding new entity records and updates into their MDM system. During this time, Stew continually reviews and merges or rejects potential duplicates as needed. ACME's call center receives a complaint from an existing customer stating that their order never arrived at their shipping address. Stew has been asked to investigate to deduce why the order never arrived. In Web UI, Stew searches and finds the associated record. Based on revision date, he can tell that the customer has been recently updated and notices that it has two source record IDs associated with their records. Stew navigates to the source traceability tab and notices that all the address attributes were taken from one source record and the last edited date matches this source system's merging date. This tracing leads Stew to believe that another customer may have been incorrectly merged causing the shipping address to be mistakenly updated.

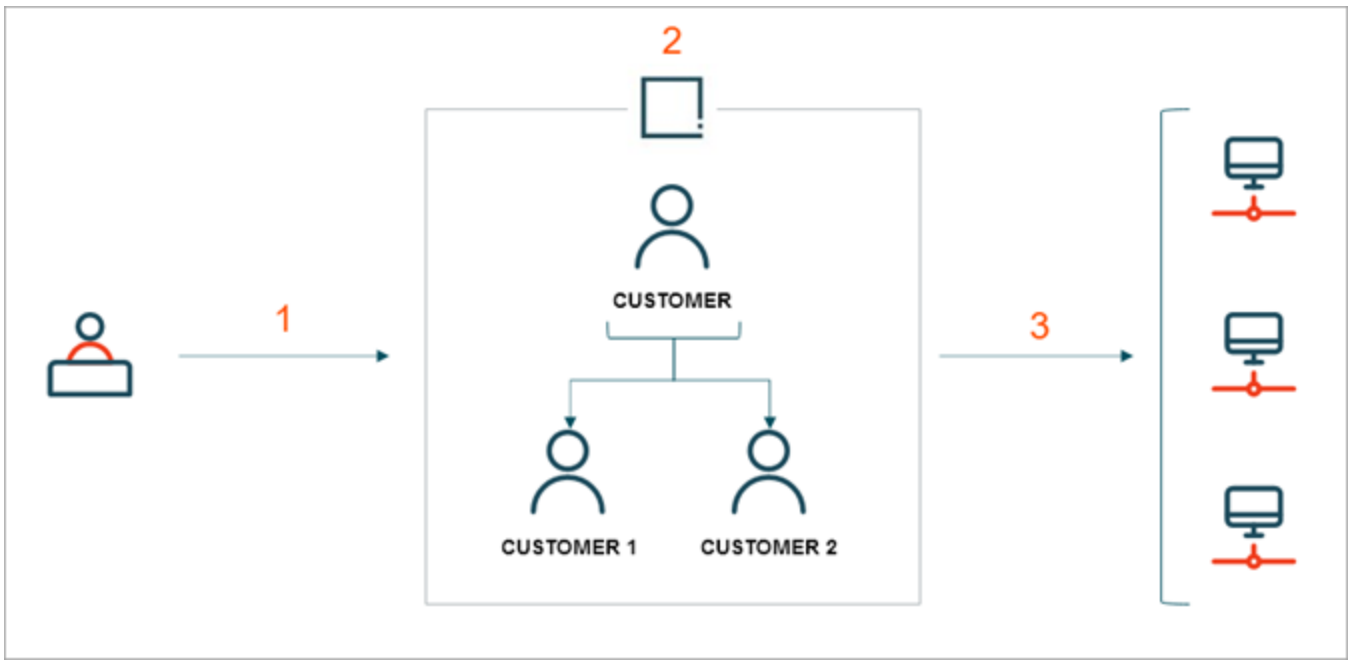
Stew must now unmerge the two records which includes determining which attribute values belong to each newly separated golden record.

Enhanced Unmerge

Due to continued complaints, ACME has decided to implement a new workflow-based procedure to allow for employees in the call center to directly initiate unmerge requests. Stew is still responsible for handling the unmerges but will now have an easier time of monitoring and investigating the requests.

While investigating an unmerge request, Stew notices that the entity involved still has open orders with ACME. After further investigation, Stew realizes this is a valid request, and the entity record needs to be unmerged. Before doing so, Stew must verify with the ERP team to determine which orders and invoices belong to the respective records. With this information obtained, Stew can now make informed decisions to unmerge the record.

With a single golden record now becoming two separate records, the downstream systems must be notified. Within ACME's system landscape, the MDM system will now send out two golden records to the ERP. The ERP must now handle this update by creating two new records, surviving the appropriate invoices to them, and deactivating the previously incorrect entity record. Once the records have migrated to downstream systems, Stew must now verify that they were received correctly. Stew navigates to the new records in each surrounding system and verifies the unmerged records have been accepted. Stew then submits the unmerge task as finalized.



1. A call center employee submits an unmerge request.
2. The request results in a single entity record splitting into two unique records.
3. The two newly created accurate golden records are synced to downstream systems.

Relocation of Individual Customers

In these use cases, the data steward is responsible for overseeing customer record updates in the event that two customers get married or relocate.

Marriage

ACME must update their database in the event that two customers get married. A pre-existing customer named Catherine Yu got married to another pre-existing customer, Charles Johnson, and had her last name and address changed.

In this scenario, the change in status is initially updated in a source system by a customer service representative within a call center. The source system feeds this update to STEP, where it is then matched to an existing record via the source system record ID for Catherine Yu.

Once matched and updated, Catherine Yu is now Catherine Johnson and has a new address. Additionally, the household is updated to reflect that Catherine and Charles belong to the same household.

Relocation

ACME must update their database in the event that a customer moves to another address. In this scenario, Maxie Hadley was a pre-existing customer who moved from KY to UT. At some point a second record was created with the new address. STEP automatically identifies that the two records with different address are the same person.

Once flagged as potential duplicates, a data steward can handle the deduplication task via a clerical review. Concluding that they are in fact the same person, the data steward merges the two records into a single golden record.

Trigger Clerical Review after Manually Editing Customer Record

Sometimes ACME requires entity records to be manually edited in STEP to address errors in data quality. In this scenario, STEP identifies a new potential duplicate after the entity name for a record is found to have two misspellings. Once identified as a potential duplicate, a task is created in a clerical review workflow.

The data steward concludes that they are duplicate records and chooses to merge them into a single golden record.

Address Validation

To improve data quality and matching results, ACME requires that addresses are validated against a third-party service (e.g., Loqate). This applies to any entity record that are either propagated to MDM or onboarded within MDM. This improves the quality of master data through standardization of addresses and identifying duplicates entity records. Other benefits include improved operational processes such as reporting and analytics downstream from MDM.

Having standardized addresses may reduce overhead operational penalties and increase efficiency of existing processes, as well as build organizational trust in MDM.

Rejecting Potential Duplicates

ACME may want to keep two matched records as separate records under certain circumstances. In such a scenario, a data steward can determine if two or more records in a clerical review workflow should merge or stay as separate entities. As an example, Jonathan Dorthy & Jenkin Doorthy are different people but were flagged as potential duplicates due to their similar names.

These potential duplicates are displayed via a clerical review task list in Web UI, where the data steward can reject the task to keep Jonathan Dorthy & Jenkin Doorthy from merging. Once rejected, the customer records remain as separate records.

Managing B2B Contacts

In these use cases, the business user is responsible for managing organizational contacts (i.e., for customers, suppliers, etc.).

Identifying and Deduplicating Contacts Within the Same Organization

ACME must ensure that there are no duplicate contact records within the same organization. In this scenario, Sergio Bennett and Serge Bennet are separate contact person records belonging to the same organization and are identified as potential duplicates.

A supplier's source system initially sends a contact person record with the name of 'Sergio Bennett.' Over time, due to operational inefficiencies, the same source system sends a second record with a different source record ID by the name of 'Serge Bennet.' MDM evaluates the incoming record against the existing record and determines there is a high probability these two records represent the same person. MDM places this record pairing into Clerical Review for a business user to review and confirm these two records are the same person and merges them together.

Handling Contacts Across Different Organizations

In the case where mastering contact persons across different organizations is important to the business, it must be noted that since contact information, such as address, phone, and email, are likely to change from one organization to another, the number data points to compare are greatly reduced. In this case, it is helpful to maintain additional data points such as date of birth to increase the robustness of a match algorithm.

For example, Emma Hoffman was previously a contact person for an ACME supplier but has changed employment, and now works for a different ACME supplier. Since Emma's name has not change but everything else including her supplier, phone number, email, and business address have changed, when the source system updates MDM with the new organization contact, Emma Hoffman's contact person record will not be deduplicated across both organizations.

A new golden record contact for Emma Hoffman is created for that supplier and it is not deduplicated against her previously golden record.

Matching

With a proper Matching strategy, organizations can consolidate their data from a variety of source systems and generate singular, definitive golden records.

Before creating a Matching configuration, it is important to get an understanding of the client organization's data and identify challenges for which the algorithm must account for. With this knowledge, a clear match strategy can be defined.

Once a strategy is defined, the Matching Algorithm is typically configured first, followed by the Match Codes Both should be tailored to the client organization's data. The algorithm should be configured first to know what data points specific to the client organization will provide reliable matching. With the data points identified, you can define what attribute combinations will contribute to the match codes. For more information on configuring these components, see the **Configuring Matching Algorithms and Match Codes** section of this documentation.

When the initial configurations are complete, they can be iterated upon via Match Tuning sessions. The goal of these sessions is to gradually improve the accuracy of the matches being made until you arrive at your final Matching configuration. For more information on Match Tuning, see the **Matching Algorithm Tuning** section of this documentation.

For more information on the Matching process and configuring the relevant components, see the **Matching, Linking, and Merging** documentation.

Configuring Matching Algorithms and Match Codes

The Matching Algorithm is typically configured first, followed by the Match Codes.

Matching Algorithm Configuration

The Matching Algorithm should be tailored to the data and strive for precision. The initial configurations include pre-configured Matching Algorithms that can be used as a foundation to build a client organization's Matching solution:

- Individual Customer
- Organization Customer
- Contact Person
- Household
- Supplier

For more information, see the Algorithm & Match Codes topics.

When configuring the Matching Algorithm, it is important to consider the impact that Thresholds have on match results. If the Clerical Review threshold is set too high, a large number of false negatives may be generated. Similarly, if the Auto Threshold is set too low, false positives could be generated. If the initial Matching configuration produces false negatives and/or false positives, the Thresholds should be reevaluated during the Match Tuning sessions.

For more information on configuring a Matching Algorithm, see the **Configuring Matching Algorithms** section of the **Matching, Linking, and Merging** documentation.

Match Codes Configuration

Match Codes should aim to ensure that anything given a high score by the algorithm is included, and that only those records that may score high get compared. The Match Codes should be considering the same data points that the Match Criteria does. This ensures that the comparison pool that Match Codes generate is relevant to the data points the Match Criteria is matching on.

For example, if the Match Criteria is matching on a combination of Person Name and Address, it is not recommended to generate Match Codes based on unrelated attributes (e.g., Phone and Email).

Example Match Codes

Typical attributes used to generate Match Codes for Individual (B2C) solutions include Name, Address, Email, and Phone. For Organization (B2B), these attributes potentially include Name, Address, Phone, DUNS (D&B Number), and Tax Identifier.

Email

Email is often used to narrow the pool of potential match candidates. The Email Match Code Generator may be selected to work in conjunction with an Email normalizer to auto-generated email match codes.

Example: A customer with the email "InesJung@armyspy.com" becomes the Match Code "EMAIL#INESJUNG@ARMYSPY.COM".

Phone

Phone shares many of the same benefits that email provides.

Example: A customer with the phone number '(931) 839-9039' becomes the Match Code "PHONE#19318399039".

In this example, the phone number normalizer is configured to default the country code to US.

Address

Address is often used in combination with person name or organization name. It is fairly unique if the full address is used but has low uniqueness if only part of the address is used.

The accuracy of addresses varies (i.e., one entry may include suite no. while another may not). When matching, techniques like edit distance is used on city and street. Because of this, and the fact that accuracy varies, the full address does not make a good match code, as it will likely lead to false negatives.

Addresses are often abbreviated ('st' for street or station, etc.) in complex patterns that cannot be uniquely resolved easily. For high accuracy of address matching, it is therefore recommended to use STEP's Address standardization capabilities, which are integrated to Loqate.

The match code generator for addresses provides the following address combinations:

- ZIP code + Street Name
- Metaphone3 City + Street Name

Example: The match codes for "134 Trace Lane, Lawrenceville, GA, 30046" would be "ADDRESS#30046+TRACELANE" or "ADDRESS#LRNSFL+TRACELANE"

Name and Address

Combining elements of a person name and elements of an address is often a good way to create Match Codes that are sufficiently unique, without causing false negatives. However, multiple permutations are often required to avoid false negatives.

Example: Kimberly Kaine resides at 134 Trace Lane, Lawrenceville, GA, 30046. Her corresponding match code would be "INDIVIDUAL#K+K+30046+134 TRACE".

Other examples of Name and Address combination Match Codes:

- First Name initial + Metaphone3 Last Name + ZIP code
- Last Name initial + Metaphone3 First Name + ZIP code
- First Name initial + Metaphone3 Last Name + Metaphone3 City
- Last name initial + Metaphone3 First Name + Metaphone3 City

- First name initial + Last Name initial + ZIP code + Street name
- First Name initial + Last Name initial + Metaphone3 City + Street Name

For more information on configuring Match Codes, see the **Configuring Matching Algorithms** section of the **Matching, Linking, and Merging** documentation.

Algorithm & Match Codes - Individual Customer

Matching Algorithm

The Individual Matching Algorithm delivered as part of the initial configuration is designed as a match algorithm for Individual Customer solutions. The most common data to have available for matching Individual Customers are First Name, Middle Name, Last Name, Address, Emails, and Phone numbers

This algorithm consists of four normalizers and four matchers, with an Auto Threshold of 90.0 and a Clerical Review Threshold of 60.0.

For more information on configuring a Matching Algorithm, see the **Configuring Matching Algorithms** topic of the **Matching, Linking, and Merging** documentation.

Normalizers

Normalizers are used to standardize values that are being compared. This ensures equal formatting is applied, increasing the accuracy of the comparisons being made. For more information, see the **Match Criteria Data Elements** topic of the **Matching, Linking, and Merging** documentation.

Person Name Normalizer

The Person Name Normalizer is configured to normalize the corresponding first, middle, and last name attributes (e.g., FirstName, MiddleName, LastName).

For customers with a large, non-English speaking consumer base, it is recommended to normalize accents and diacritic characters. Such characters are not handled with the phonetic encoding of words (e.g., Soundex or Metaphone 3) during the match process.

Address Normalizer

Because it is recommended to model addresses as data containers, configure the Address Normalizer to normalize data container attributes as defined within the Address component model.

Email Normalizer

Because it is recommended to model emails as data containers, configure the Email Normalizer to normalize data container attributes as defined within the Email component model.

Phone Normalizer

Because it is recommended to model phone numbers as data containers, configure the Phone Normalizer to normalize data container attributes.

Matchers

For general information on configuring Matchers, see the **Match Criteria Data Elements** topic of the **Matching, Linking, and Merging** documentation.

Person Name Matcher

The Person Name Matcher is largely left with the default settings. However, an alias table is used to provide an equivalent names table.

Third-party sources may be leveraged to build and enhance the equivalent names table within STEP. However, it is recommended that the client's expertise with their customer data be consulted to account for industry or business specific patterns. For example, if a company is based in southwestern United States then their equivalent names table may contain an emphasis on Hispanic name equivalents. Furthermore, equivalent name values may also be added as a result of a match tuning exercise with the client's customer data set.

Regarding middle names, an evaluation exercise with the customer is recommended to review the quality of data they have for customer middle names. It is possible that middle names are not collected from the consumers, or only middle initials are required but rarely provided. In such cases, it is recommended to reconsider the weight of MiddleName.

Note: The names in the equivalent names table provided by the initial configuration target the US market.

Address Matcher

The Address Matcher utilizes default configuration values.

Email Matcher

The Email Matcher utilizes default configuration values.

Phone Matcher

The Phone Matcher utilizes default configuration values.

Rules

When considering match rules, the recommended strategy is to dissect the customer's information into the smallest possible portions of data. These rules should not weigh the sum of *all* the customer's input data and should instead be split so that it is possible to optimize each rule. Careful analysis of the customer dataset is required to determine what combinations of attributes present the best chance of uniqueness.

Three rules are provided to calculate the final score of the Individual Matching Algorithm. According to the configured conditions, if all respective Matchers resolve to true (i.e., scores above 70), then the algorithm will take the highest scoring rule as the final score of the algorithm. The rules are comprised of combinations of each Name score, Address score, Email score, and Phone score. The rules are then standardized to resolve to a value between 0 and 100.

These rules specifically include:

- Name & Address
- Name & Email
- Name & Phone

In this scenario, Name is used in all three rules because it is common for family members who use the same email and/or phone number to live at the same address. Thus email, phone number, or address are not reliably unique by themselves. By adding Name to these rules, we are ensuring the uniqueness of the individual's name is taken into consideration, in addition to the other contact information elements.

It is possible to extend a rule by including various other combinations of matchers. This should be considered if there are specific requirements or use cases which requires specific combinations of matchers.

By adding another parameter, you can identify false positives. Unique identifiers such as Social Security, Passport, or Driver's License Numbers may be used as veto rules to further enhance the quality of the match process.

Other extensions include Date of Birth (DOB), which can be used in combination with other rules to be less strict on equality. For example, name and address runs the risk of a father and son having the same name which would be resolved by considering the date of birth. Additionally, you can accept lower scores of names and addresses if DOB is equal.

Survivorship

The following survivorship rules are used by the Individual Customer Algorithm:

- **Value: Most Recent**
  - **Attribute / Attribute Group:** Individual - Most Recent
  - **Last Edit Date Attribute:** Last Edit Date - Record
- **Data Container: Most Recent (Emails)**
  - **Business Condition:** DataContainer Survivorship Email
  - **Data Container Type:** Emails
  - **Last Edit Date Attribute:** Last Edit Date - Email
- **Data Container: Most Recent (Phones)**
  - **Business Condition:** DataContainerSurvivorshipPhone
  - **Data Container Type:** Phones
  - **Last Edit Date Attribute:** Last Edit Date - Phone
- **Data Container: Most Recent (Main Address)**
  - **Business Condition:** DataContainer Survivorship Address
  - **Data Container Type:** Main Address
  - **Last Edit Date Attribute:** Last Edit Date - Main Address

Note: Data Containers require their own survivorship rules. Additionally, each Survivorship rule requires a unique Last Edit Date attribute.

Match Codes

For the Individual Customer entity type, three separate Match Codes are being generated. While based on the demographics of the customer record, these Match Codes are composed of: Email, Phone Number, and a combination of Individual Name and Address.

For information on how to configure Match Codes which are housed in matching algorithms, see the **Configuring Matching Algorithms** section of the **Matching, Linking, and Merging** documentation.

Email Match Code

The Email Match Code is the normalized value of the email address attribute. Both the username (or local part) of the email and the domain are normalized to ensure variations of the same email address (because of differing cases or special characters) are accounted for.

It is recommended to add a discernible prefix to each Match Code so the end-user may easily identify what attribute(s) were used for the Match Code. The email Match Code contains the prefix 'EM~'.

Phone Number Match Code

The Phone Number Match Code is the normalized value of the individual's phone number attribute. The normalization removes any parenthesis and hyphenation in between numbers. Additionally, the area itself has been removed, leaving only the last 7 digits available as the Match Code. Phone number Match Code is prefixed by 'PH~'.

Address and Name Match Code

The Address and Name Match Code is a combination of elements of the individual's name and address. For example, the provided Match Code within the initial configuration is composed of zip code + the first letter of the individual's first name + Metaphone 3 representation of the individual's last name.

The Address and Name Match Code contains the prefix 'ZINM~'.

Configuration Considerations

It is worth considering the use of the Equivalent Values Lookup Table and Anonymous Value Lookup Table. The Equivalent Value Lookup Table is used by both Match Codes and the Match Criteria to ensure that values that mean the same thing are evaluated as such. Equivalent values will score appropriately high, as if the values were actually the exact same.

Example:

- Name: Matt = Matthew

Note: Equivalent Values are only used for person & organization names.

The Anonymous Values Lookup Table is also used by both Match Codes and the Match Criteria to ensure that values that are anonymous, or not meaningful, do not contribute to identifying potential duplicates. Determining what these values should be is highly dependent on the organization's dataset.

Typically, these values are default values that users of a Source System enter when they do not have the correct value, or do not want to enter a value. The actual anonymous values are not included in the baseline build of the Customer MDM configuration.

Examples:

- Phone: 999999999
- Address: DO NOT USE

Algorithm & Match Codes - Organizational Customer & Supplier

Matching Algorithm

The Organization and Supplier Matching Algorithm delivered as part of the initial configuration is designed as a match algorithm for Organization Customer & Supplier solutions. This algorithm is most relevant in a Business to Business (B2B) style implementation and consists of two normalizers and two matchers, with an Auto Threshold of 90.0 and a Clerical Review Threshold of 60.0.

For more information on configuring a Matching Algorithm, see the **Configuring Matching Algorithms** topic of the **Matching, Linking, and Merging** documentation.

Normalizers

Normalizers are used to standardize values that are being compared. This ensures equal formatting is applied, increasing the accuracy of the comparisons being made. For more information, see the **Match Criteria Data Elements** topic of the **Matching, Linking, and Merging** documentation.

Organization Name Normalizer

The Organization Name Normalizer for the Organization and Supplier Algorithm is configured to normalize only the Legal Name (LegalName) attribute for organization and supplier entities.

Address Normalizer

Because it is recommended to model addresses as data containers, configure the Address Normalizer to normalize data container attributes as defined within the Address component model.

Matchers

For general information on configuring Matchers, see the **Match Criteria Data Elements** topic of the **Matching, Linking, and Merging** documentation.

Organization Name Matcher

The Organization Name Matcher is used by the algorithm to match on organization names. This matcher is left with default settings; however, it is recommended that the Organization Name Matcher is tuned to the specifics of the organization.

Address Matcher

The Address Matcher utilizes default configuration values.

When matching entities, keep in mind that multiple business entities may reside at a singular address. For example, Office Buildings where no suite number is provided.

Rules

When considering match rules, the recommended strategy is to dissect the data into the smallest possible portions of data. These rules should not weigh the sum of *all* the customer's input data (which is likely inefficient) and requires a careful analysis of the customer dataset to determine what combinations of attributes present the best chance of uniqueness.

There is only one Rule associated with the Organization and Supplier Algorithm. This rule evaluates the scores of the Name and Address matchers. The resulting score is then weighed against the defined algorithm thresholds to determine whether the organization and supplier record should be created, auto-merged, or if it requires a clerical review.

Note: Common attributes to match for organizations or suppliers include Address, DUNS (D&B Number), and Tax Identifier.

Survivorship

The following survivorship rules are used by the Organization and Supplier Algorithm:

- **Value: Most Recent**
  - **Attribute / Attribute Group:** Organizations - Most Recent
  - **Last Edit Date Attribute:** Last Edit Date - Record
- **Data Container: Most Recent**
  - **Business Condition:** DataContainer Survivorship Address
  - **Data Container Type:** Main Address
  - **Last Edit Date Attribute:** Last Edit Date - Main Address

Note: Data Containers require their own survivorship rules. Additionally, each Survivorship rule requires a unique Last Edit Date attribute.

Match Codes

For the Organization Customer & Supplier entity type, three separate Match Codes are being generated. While based on the demographics of the organization record, these Match Codes are composed of the Email, Phone Number, and a combination of Legal Name and Address.

Since initial demographic data for organizations and suppliers may be limited prior to enrichment (for example, Dun & Bradstreet detailed profile), generating these three Match Codes will help ensure proper identification of organizations or suppliers during the match process.

For information on how to configure Match Codes which are housed in matching algorithms, see the **Configuring Matching Algorithms** section of the **Matching, Linking, and Merging** documentation.

Email Match Code

The Email Match Code is the normalized value of the entity's main contact email address. Both the username (or local part) of the email and the domain are normalized to ensure variations of the same email address (as a result of differing cases or special characters) are accounted for.

It is recommended to add a discernible prefix to each Match Code so the end-user may easily identify what attribute(s) were used for the Match Code. The Email Match Code contains the prefix 'EM~'.

Phone Number Match Code

The Phone Number Match Code is the normalized value of the entity's main contact phone number. The normalization removes any parenthesis and hyphenation in between numbers. Additionally, the area itself has been removed, leaving only the last 7 digits available as the Match Code. Phone number Match Code is prefixed by 'PH~'.

Organization Address and Name Match Code

The Organization Address and Name Match Code is composed of the zip code + Metaphone 3 representation of the entity's legal name.

The Organization Address and Name Match Code contains the prefix 'ZONM~'.

Algorithm & Match Codes - Contact Person

Matching Algorithm

The Contact Person Matching Algorithm delivered as part of the initial configuration is designed as a match algorithm for Contact Person records. This algorithm is relevant in a Business-to-Business (B2B) style implementation and consists of normalizers and matchers with an Auto Threshold of 90.0 and a Clerical Review Threshold of 60.0.

For further details regarding the Contact Persons use case, see the **Data Modeling** and **Use Case Overview** topic s of this documentation.

For more information on configuring a Matching Algorithm, see the **Configuring Matching Algorithms** topic of the **Matching, Linking, and Merging** documentation.

Normalizers

Normalizers are used to standardize values that are being compared. This ensures equal formatting is applied, increasing the accuracy of the comparisons being made. For more information, see the **Match Criteria Data Elements** topic of the **Matching, Linking, and Merging** documentation.

Person Name Normalizer

The Person Name Normalizer is configured to normalize the corresponding first, middle, and last name attributes (e.g., FirstName, MiddleName, LastName).

For customers with a large, non-English speaking consumer base, it is recommended to normalize accents and diacritic characters. Such characters are not handled with the phonetic encoding of words (e.g., Soundex or Metaphone 3) during the match process.

Address Normalizer

Because it is recommended to model addresses as data containers, configure the Address Normalizer normalize data container attributes as defined within the Address component model.

Email Normalizer

Because it is recommended to model emails as data containers, configure the Email Normalizer to normalize data container attributes as defined with the Email component model.

Phone Normalizer

Because it is recommended to model phone numbers as data containers, configure the Phone Normalizer to normalize data container attributes.

Matchers

For general information on configuring Matchers, see the **Match Criteria Data Elements** topic of the **Matching, Linking, and Merging** documentation.

Person Name Matcher

The Person Name Matcher is largely left to the default settings. However, an alias table is used to provide an equivalent names table.

Third-party sources may be leveraged to build and enhance the equivalent names table within STEP. However, it is recommended that the client's expertise with their customer data be consulted to account for industry or business specific patterns. For example, if a company is based in southwestern United States then their equivalent names table may contain an emphasis on Hispanic name equivalents. Furthermore, equivalent name values may also be added as a result of a match tuning exercise with the client's customer data set.

Regarding middle names, an evaluation exercise with the customer is recommended to review the quality of data they have for customer middle names. It is possible that middle names are not collected from the consumers, or only middle initials are required but rarely provided. In such cases, it is recommended to reconsider the weight of MiddleName.

Note: The names in the equivalent names table provided by the initial configuration target the US market.

Address Matcher

The Address Matcher utilizes default configuration values.

Email Matcher

The Email Matcher utilizes default configuration values.

Phone Matcher

The Phone Matcher utilizes default configuration values.

Rules

When considering match rules, the recommended strategy is to dissect the customer's information into the smallest possible portions of data. These rules should not weigh the sum of *all* the customer's input data and should instead be split so that it is possible to optimize each rule. Careful analysis of the customer dataset is required to determine what combinations of attributes present the best chance of uniqueness.

Rules are provided to calculate the final score of the Contact Person Matching Algorithm. According to the configured conditions, if all respective matchers resolve to true (i.e., scores above 70), then the algorithm takes the highest scoring rule as the final score of the algorithm. The rules are comprised of combinations of each Name score, Address Score, Email Score, and Phone score. The rules are then standardized to resolve in a value between 0 and 100.

- Name & Address
- Name & Email
- Name & Phone

In this scenario, Name is used in all rules because it is common for family members who use the same email and/or phone number to live at the same address. Thus email, phone number, or address are not reliably unique by themselves. By adding Name to these rules, we are ensuring the uniqueness of the individual's name is taken into consideration, in addition to the other contact information elements.

It is possible to extend a rule by including various other combinations of matchers. This should be considered if there are specific requirements or use cases which requires specific combinations of matchers.

By adding another parameter, you can identify false positives. Unique identifiers such as Social Security, Passport, or Driver's License Numbers may be used as veto rules to further enhance the quality of the match process.

Other extensions include Date of Birth (DOB), which can be used in combination with other rules to be less strict on equality. For example, name and address runs the risk of a father and son having the same name which would be resolved by considering the date of birth. Additionally, you can accept lower scores of names and addresses, if DOB is equal.

Survivorship

The following survivorship rules are used by the Contact Person Algorithm:

- **Value: Most Recent**
  - **Attribute / Attribute Group:** Surv Contact Most Recent
  - **Last Edit Date Attribute:** Last Edit Date - Record
- **Data Container: Most Recent (Main Address)**
  - **Business Condition:** DataContainer Survivorship Address
  - **Data Container Type:** Main Address
  - **Last Edit Date Attribute:** Last Edit Date - Main Address
- **Data Container: Most Recent (Emails)**
  - **Business Condition:** DataContainer Survivorship Email
  - **Data Container Type:** Emails
  - **Last Edit Date Attribute:** Last Edit Date - Email
- **Data Container: Most Recent (Phones)**
  - **Business Condition:** DataContainerSurvivorshipPhone
  - **Data Container Type:** Phones
  - **Last Edit Date Attribute:** Last Edit Date - Phone
- **Reference: Most Recent (Organization)**
  - **Reference Type:** Contact Person Organization
  - **Last Edit Date Attribute:** Last Edit Date - Org Ref

Note: Data Containers require their own survivorship rules. Additionally, each Survivorship rule requires a unique Last Edit Date attribute.

Match Codes

For the Contact Person entity type, separate Match Codes are being generated. While based on the demographics of the Contact Person record, these Match Codes are composed of the Email, Phone Number, and a combination of Contact Person's Name and Address.

To ensure the algorithm only matches Contact Persons within the same organization, the Match Codes should contain the ID of the reference organization for a given Contact Person. This ensures that only Contact Persons within the same organization are deduplicated.

For information on how to configure Match Codes which are housed in matching algorithms, see the **Configuring Matching Algorithms** topic of the **Matching, Linking, and Merging** documentation.

Email Match Code

The Email Match Code is the referenced organization's ID and the normalized value of the email address attribute. Both the username (or local part) of the email and the domain are normalized to ensure variations of the same email address (as a result of differing cases or special characters) are accounted for.

It is recommended to add a discernible prefix to each Match Code so the end-user may easily identify what attribute(s) the Match Code is based off of.

The email Match Code contains the prefix 'EM~'.

Phone Number Match Code

The Phone Number Match Code is the referenced organization's ID and the normalized value of the Contact Person's phone number attribute. The normalization removes any parenthesis and hyphenation in between numbers. Additionally, the area itself has been removed, leaving only the last 7 digits available as the Match Code. Phone number Match Code is prefixed by 'PH~'.

Address and Name Match Code

The Address and Name Match Code is a combination of the referenced organization's ID and elements of the Contact Person's name and address. For example, the provided Match Code within the initial configuration is composed of the zip code + first letter of the individual's first name + Metaphone 3 representation of the Contact Person's last name.

The Address and Name Match Code contains the prefix 'ZINM~'.

Configuration Considerations

It is worth considering the use of the Equivalent Values Lookup Table and Anonymous Value Lookup Table. The Equivalent Value Lookup Table is used by both Match Codes and the Match Criteria to ensure that values that mean the same thing are evaluated as such. Equivalent values score appropriately high, as if the values were actually the exact same.

Example:

- Name: Matt = Matthew

Note: Equivalent Values are only used for person & organization names.

The Anonymous Values Lookup Table is also used by both Match Codes and the Match Criteria to ensure that values that are anonymous, or not meaningful, do not contribute to identifying potential duplicates. Determining what these values should be is highly dependent on the organization's dataset.

Typically, these values are default values that users of a Source System enter when they don’t have the correct value, or don’t want to enter a value. The actual anonymous values are not included in the baseline build of the Customer MDM configuration.

Examples:

- Phone: 999999999
- Address: DO NOT USE

Algorithm & Match Codes - Household

Matching Algorithm

The Household Matching Algorithm delivered as part of the initial configuration is designed as a match algorithm for Household entities. This algorithm is most relevant in a Business to Consumer (B2C) style implementation and consists of two normalizers and two matchers, with an Auto Threshold of 90.0 and a Clerical Review Threshold of 80.0. The Clerical review threshold is both higher and tighter to ensure only customer records that are extremely likely to belong to the same household are considered.

It is worth noting that the Household Matching Algorithm utilizes the Link Golden Record solution engine rather than the Merge Gold Record approach. This is to ensure that individual customers that are determined to be of the same household are linked to a separate household entity object rather than merged.

For details regarding the Household entity and its purpose, see the **Data Modeling** topic of this documentation.

For more information on configuring a Matching Algorithm, see the **Configuring Matching Algorithms** topic of the **Matching, Linking, and Merging** documentation.

Normalizers

Normalizers are used to standardize values that are being compared. This ensures equal formatting is applied, increasing the accuracy of the comparisons being made. For more information, see the **Match Criteria Data Elements** topic of the **Matching, Linking, and Merging** documentation.

Last Name Normalizer

The Last Name Normalizer uses the Words Normalizer, and only evaluates the Last Name attribute of an individual record. This is because households are referred to only by the Last Name, or household name.

Address Normalizer

Because it is recommended to model addresses as data containers, configure the Address Normalizer to normalize data container attributes as defined within the Address component model.

Matchers

For general information on configuring Matchers, see the **Match Criteria Matchers** topic of the **Matching, Linking, and Merging** documentation.

Name Matcher

The Last Name Matcher is largely left with the default settings. An equivalent names look-up table is not utilized for Household entities since last names generally do not have shortened versions or a nickname alternative.

Address Matcher

The Address Matcher utilizes default configuration values.

Rules

When considering match rules, the recommended strategy is to dissect the customer's information into the smallest possible portions of data. These rules should not weigh the sum of *all* the customer's input data (which is likely inefficient) and requires a careful analysis of the customer dataset to determine what combinations of attributes present the best chance of uniqueness.

There is only one rule associated with the Household algorithm, which evaluates the scores of the Name and Address matchers. The resulting score is then weighed against the defined algorithm thresholds to determine whether the individual record should be created, auto-linked to an existing record, or it requires clerical review.

The initial configuration matches on Last Name and Address for households. Matching on Last Names may be advantageous in identifying family units within a large population such as a retirement home or a school. However, in some cases a married couple within a household may not have the same Last Name. It may be beneficial to make Address weigh more in scenarios like this.

Conversely, sometimes a large number of customer records with the same Address should not be considered the same household. This can include college campuses, retirement homes, and similar institutions. These scenarios should be handled on a case-by-case basis. However, one possible approach is to utilize an anonymous words table to anonymize aforementioned addresses, eliminating addresses from the matching logic for such records.

Survivorship

The following survivorship rules are used by the Household Algorithm:

- **Value: Most Recent**
  - **Attribute / Attribute Group:** Last Name
  - **Last Edit Date Attribute:** Last Edit Date - Record
- **Data Container: Most Recent**
  - **Business Condition:** DataContainer Survivorship Address
  - **Data Container Type:** Main Address
  - **Last Edit Date Attribute:** Last Edit Date - Main Address

Note: Data Containers require their own survivorship rules. Additionally, each Survivorship rule requires a unique Last Edit Date attribute.

Match Codes

The Household Match Code is generated on the Individual Customer entity type along with the Individual Match Codes. Since a household is identified by its last name, the Household Match Code is composed of the zip code + Metaphone 3 representation of the individual's last name.

The Household Match Code contain the prefix 'ZILNM~'.

Matching Algorithm Tuning

Algorithm Tuning begins during the build phase of the implementation. Early on in this process, it is common to find a large number of invalid matches making it past the Auto-Merge threshold while valid matches fall short of the Clerical Review threshold. Thus, the ultimate goal of the algorithm tuning sessions is to perfect the matching logic's accuracy so that good and bad matches score within the appropriate thresholds.

Considerations

When developing a Matching strategy, it is important to consider the client organization's data and the potential challenges the algorithm will have to account for.

The following considerations and challenges are commonly encountered when implementing Matching.

Using Real Data

Algorithm Tuning is highly data-dependent, so real production data from each source system must be made available for analysis. Entity Data is required at key points during the implementation:

- 10-100 records for Data Modeling
- 20% of total data volume for Algorithm Tuning
- 100% of total data volume for go-live

Obtaining Data

Solution Consultants should expect delays in receiving customer / entity data. They should work with the client organization early in the process to define the data to be delivered and push to get it as early as possible. There may be both technical and process reasons for holding up data delivery. The ETL team may have issues in staging data from the source systems, and the legal or security team may introduce their own delays.

Solution Consultants should establish a delivery date that the client organization agrees on and emphasize that delays to that date will delay critical-path tasks.

Note: Having access to real production data is a critical dependency to starting the Algorithm Tuning tasks.

Data from all sources must be included. Data quality and characteristics can vary from source system to source system, so getting samples from all sources is critical. This includes samples of all object types in scope as well as data captured through different means (call center, web, mobile, etc.)

Other considerations for sampling data:

- Data that crosses regions
- The age of records (recently created records vs records created 20 years ago)
- Records updated recently vs records that have not been updated in years

Handling Data

Solution Consultants should work with the client organization to determine the level of security needed around the data they provide. Note that the client may be held to a higher degree of security due to the regulation in their industry.

Data Quality Revelations

Before Algorithm Tuning starts, data should be analyzed to determine its quality and characteristics. This analysis will help in determining the baseline algorithm configuration. The Solution Consultant should analyze the data for attribute completeness (percent populated), bad (or anonymous) data values populated, and any patterns in the data that may be instructive for Algorithm Tuning. The quality and characteristics of data can differ between sources, so data from each source system should be analyzed.

It should be expected that matching will reveal Data Quality issues that were not previously known. This should be viewed as uncovering opportunities to improve data rather than a problem or a setback to the overall project timeline.

Stakeholder Input

It is important to have both the data steward and data owner present during Sample Pair Review sessions. For more information about Sample Pair Review, see the **Process** section below.

In this context, the data steward is the business user who has been tasked with the formation and execution of policies for the management of data and metadata. The data owner is the business user who typically has a direct line of responsibility for a functional area.

There are several personnel considerations to make during the sample pair review process:

- It is recommended that the consultant be present on-site for the Sample Pair Review sessions. Due to the highly interactive nature of these discussions, being on-site helps facilitate the process.
- Data Owners from different functional areas within the organization may have differing opinions on matching requirements. It may be difficult to get consensus among these various Data Owners.
- Some more sophisticated client organizations may introduce a Data Governance Board that can assist with reaching consensus among the Data Owners.

Note: It is important to set expectations with the client organization that while those implementing the initial configuration can provide guidance, it is the organization's responsibility to determine which entity records should match and which should not.

However, Solution Consultants should not expect a client organization to be able to articulate their matching requirements. To arrive at a baseline algorithm configuration, a discussion with the client should occur focusing on what their priorities are for generating Golden Records. The Sample Pair review sessions will help facilitate the discussion around finalizing matching requirements.

False Positives vs False Negatives

Before tuning begins, it is important to discuss with the client organization if false positives or false negatives would be preferred. It is much easier to identify false positives than false negatives in the pair export. Therefore, it is recommended to start with a wide net and narrow the Match Criteria during tuning. However, ultimately client

organizations will generally prefer false negatives over false positives once the algorithm is finalized.

Clerical Review

Keep in mind that any records below the auto-threshold and above the clerical review threshold is placed in clerical review for manual review. It is important to discuss with the client organization what types of potential duplicates are evaluated as part of clerical review and the volume of records that are acceptable. Client organizations with a low volume such as B2B organizations may want a looser algorithm where most or sometimes all records are reviewed manually. Client organizations with a high volume such as B2C organizations may want little to no records reviewed manually.

This discussion becomes increasingly important as volumes become larger. While .5% may sound like an acceptable and small percentage, .5% of 1 million is 5000 tasks that must be reviewed manually. Additionally, if it is not possible to articulate rules that define a match, it probably won't be possible for a human to determine if records are a match using the data provided.

Start Small

Initially tuning with a full data set may not be advisable when volumes are large. Instead, tuning should be done in iterations of increasing volume sizes. As an example, start with 1000 records, eventually go to 100000, and in the end, at least 20% of the full data set should be taken into consideration. Make sure to consider a good sampling of data such as data from all sources and data captured through different means. If the volume is small enough, consider tuning using 20%, 30%, 60% and 100% of the total data set.

Iterations of Review

Each client has a unique data set requiring the match algorithm be tuned specifically to identify matches in that data. Three or more match tuning iterations should be expected. It is not uncommon to have five or many more iterations.

Rule Tips

Rules are a set of criteria that must be true for the result to be assigned. The rules can quickly become complex as varying use cases are identified in the data. The best way to handle this complexity is to make the rules simple and easy to understand, especially since there is no restriction to the number of rules that can be created. Additionally, the rules should be well documented so that changes to the algorithm are easier both during tuning and later after going live.

The result should be calculated as a weighted sum of the matchers in play. Conditions should be used to limit the combinations of matchers in the weighted sum to no more than two matchers, as else it is typically not possible to tune to a threshold.

Re-tuning

It is recommended that the client organization re-engage with the applicable solution consultants post go-live for new Algorithm Tuning sessions when any of the following take place:

- A new source is added that has different demographic or data quality.
- Substantial increase in False Positives or False Negatives.
- The Entity Size or Clerical Review Task Size metrics change significantly over time.

Process

The Matching Algorithm Tuning process is as follows:

1) Configuration

Utilize a Match Tuning configuration to generate a Data Profile. Using this Data Profile, identify key data points to consider when configuring a baseline algorithm (Matching Algorithm and Match Codes).

For more information on Matching Tuning and creating a Match Tuning configuration, see the **Match Tuning** section of the **Matching, Linking, and Merging** documentation.

2) Generate Sample Pair

Once the baseline algorithm is configured, generate the Random Sample Pair spreadsheet via a Match Tuning configuration. This baseline configuration is just a 'best-guess' configuration based on the analysis of the data so far.

Before the sample pair review can kick-off, the raw data from the output file should be formatted to make it human readable. The Sample Pair Formatter Excel sheet can optionally be used on the output file.

Note: To obtain the Sample Pair Formatter Excel sheet, refer to the separately-bundled initial configuration files. Detailed instructions on how to use this Excel sheet are included within the sheet itself.

3) Review Sample Pair

Review the Sample Pairs with the client. Each individual pair gets either a 'Yes', 'No', or 'Not Sure' indicating whether or not they should be considered the same entity by the algorithm and linked together.

The Sample Pair Review process can be a time-consuming task but is critical in getting the algorithm tuned to meet requirements. Typically, review 1,000+ Sample Pairs each cycle with the stakeholders. For some of the iterations, a pair export may be as large as 1,000 records per percentage points of interest.

Once the random sample pair spreadsheet is generated and formatted, it is vital to review the sample pairs to see how the algorithm evaluates them. The primary purpose of the review is to assess the confidence of each merge and modify the thresholds if the scores appear inaccurate. During the review process, it is also important to consider the following:

- The organization should mark each set with a decision as to whether (based on the data available) the records should be considered the same entity.
- It is best to approach this task from a 'human' standpoint as opposed to creating logic to help you achieve a certain score.
- This is not a data cleaning task.

The goal by the end of each Sample Pair Review session is to improve the quality of the matches found. It is much easier to identify false positives than false negatives in the pair export. Therefore, it is recommended to start with a wide net and narrow the Match Criteria during tuning.

4) Tuning the Algorithm

Tune the algorithm based on feedback from the Sample Pair review and generate a new set of Sample Pairs based on the updated algorithm. This goal can be achieved by:

- Adjusting the scoring method and weighting of each scored attribute.
- Adjusting the relative weighting of scoring across all the scored attributes.
- Adjusting the Auto-Merge and Clerical Review Thresholds.

Repeat steps 2 and 3 for two more cycles (or more, as needed).

5) Finalize

Decide on the final Auto Merge Threshold and Clerical Review Threshold.

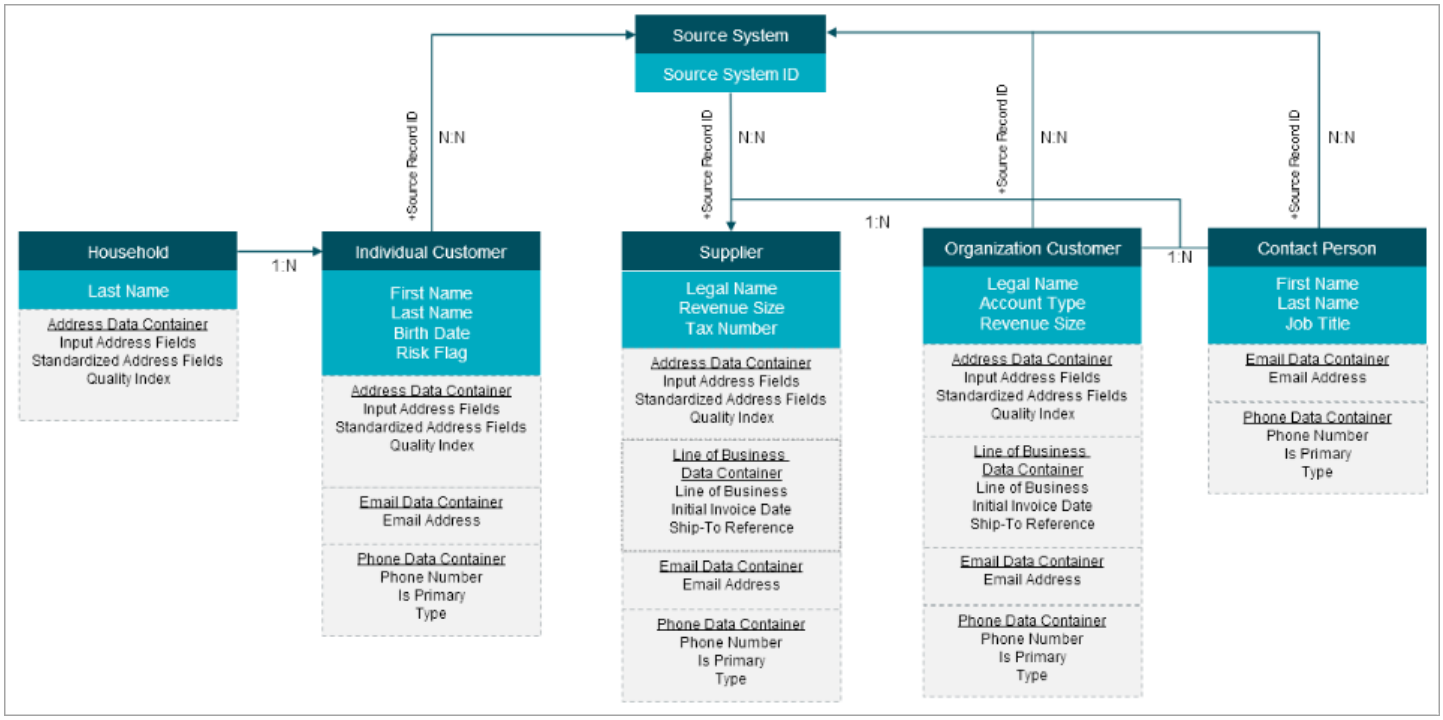
Data Modeling

With the help of this guide, users can model the following common customer data challenges in STEP:

- Individual and Organization Customers
- Individual Customers and Households
- Individuals as Contact Persons
- Suppliers
- Contact Persons in relation to Organization Customers & Suppliers
- Addresses and Contact Information

Note: Additional data may be applicable to include in a Customer and Supplier MDM solution but is not covered by these guidelines.

The collected data model is as follows. Each of the dotted boxes is a data container type.



These Implementation Guidelines assume that customers, contact persons, and organizations are created and maintained by the Merge Golden Record solution. Households are created and maintained by using the Link Golden Record solution.

Customer Object Types

What is a Customer?

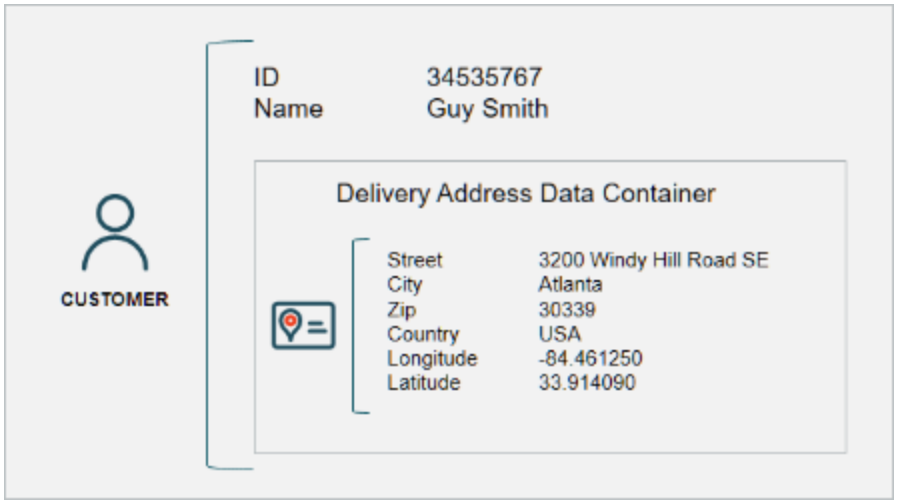
A customer is any individual or organization that purchases from a business. That said, any individual or organization could potentially interact with a business in another capacity (e.g., as a supplier or employee). Because of this, 'Customer' is defined as a role an organization or an individual has in the interaction with a business.

However, businesses do not typically manage information about individuals and organizations in a shared model across vastly different business processes, like procurement (supplier) and sales (customer), for simplicity's sake. To account for this, 'Customer' is used as a catch-all entity type that owns all information about individuals and organizations. For more about customer types, see details about contact persons below.

Individual Customer

For an individual customer, the typical customer attributes are:

- First Name
- Last Name
- Addresses
- Date of Birth
- Gender
- Phone Numbers
- Fax numbers
- Emails
- Social Media IDs
- Loyalty Card Information
- Government IDs (social security, passport, driver's license)



Organization Customer

For an organizational customer, the typical customer attributes are:

- Legal Name of the Organization
- Date of Formation
- Address
- Contact Information
- Main Phone
- Fax
- Main Email
- Website URL
- GLN / ILN
- DUNS Number
- Tax Identification Number
- Risk Category

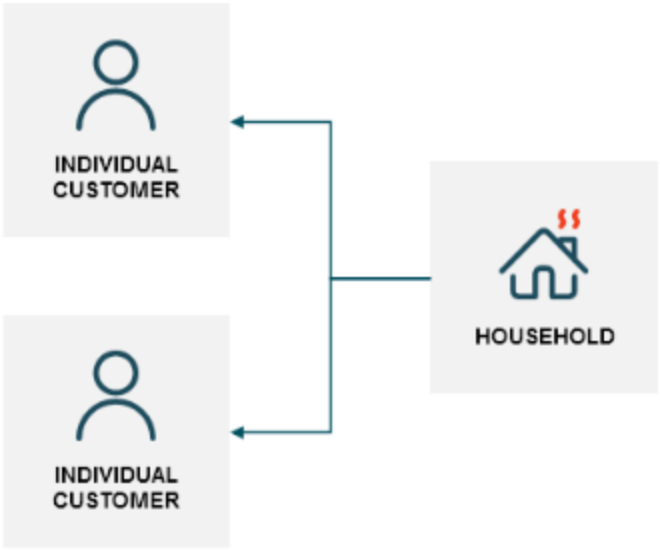
Organization customers are often organized in hierarchies. See section about organization hierarchies below:

What is a Household?

A household is a collection of individuals, which is typically a family, living in the same home. Each member of the household is considered an individual customer.

The purpose of these household constructs is to help businesses target customers that share a household either as a family unit or as individual customers. This allows the company to limit extraneous mailing campaigns and aggregate buying patterns.

In the Customer MDM initial configuration, Households are discovered by matching and created via Link Golden Records.



Accurately identifying the correct members of a household comes with natural limitations, as the definition of household is quite vague and vary from case to case. The data available to determine if specific people are actually accurately identified as part of the same household is often rather sparse.

Address and last name are the two most common matching criteria used to determine households. For more information, see the **Algorithm & Match Codes - Household** section of this documentation.

This approach is intended for situations where accuracy must aim to be good but both false positives and false negatives can be accepted, such as with the case of analytics and mailing campaigns.

Handling Composite Records with Data Containers

A customer record comprises multiple objects. It may have multiple addresses, emails, and phone numbers, which are all considered part of the record by users, as well as surrounding systems.

This is not to be confused with hierarchical structures of separate entities with each containing their own data flow.

Examples of separate entities may include:

- contact persons for organization customers
- store branches
- warehouse locations

Such records are complex structures in themselves that may even reference or be referenced by other entities.

For more information, see the **Data Containers** topic in the **System Setup / Super User Guide** documentation.

Addresses

An address comprises several attributes that define a location. STEP offers specialized functionality around addresses, such as address verification via the Address Component Model.

Modeling addresses using data containers makes it possible to manage more than one address on the same customer / contact person. Data containers simplify the data modeling by providing a reusable address definition across multiple object types. When displayed in Web UI, addresses modeled using data containers can appear as one formatted value, despite being made of several attributes.

Addresses have metadata associated with them via the Address Component Model, the CASS address component model, and potentially solution specific meta data.

Email Addresses and Phone Numbers

Email addresses and phone numbers should be modeled as data containers, as it is typical to have multiple phone numbers and email addresses for the same customer or contact person. It is also common to have meta data associated to each phone number or email address. The initial configurations have PhoneType as a Data container key. This allows a customer to only have one of each type of phone (cell, home, etc.).

Handling Confidential Information

Credit Card Information

Credit Card information should not be stored in or passed through STEP due to compliance requirements. It is simpler to store 'Recurring Charge Subscription IDs' when integrating to third-party services.

Bank Account Information

Handling bank account information does not typically require more security than name and address information.

Privileges

STEP Privileges to control who has access to what attributes and/or attribute groups.

Security

This guide does not cover information security in STEP in a broader sense, like infrastructure recommendations, encryption strategies, etc.

Note: The sample data provided in the initial configuration provides only main addresses. There are no delivery addresses.

Supplier and Classification Object Types

A supplier is a person, business, or entity from which an enterprise purchases raw material, finished goods, or services. Such commodities are procured for the purpose of driving an enterprise's business forward.

The terms 'supplier' and 'vendor' are often used interchangeably.

Suppliers are modeled as a stand-alone entity type within MDM. This allows the flexibility of distinguishing business processes and validations from individual or organization customer entity types.

Some typical attributes associated with suppliers are:

- Legal Name of the Organization
- Date of Formation
- Address
- Contact Information
- Main Phone
- Fax
- Main Email
- Website URL
- DUNS Number
- Tax Identification Number

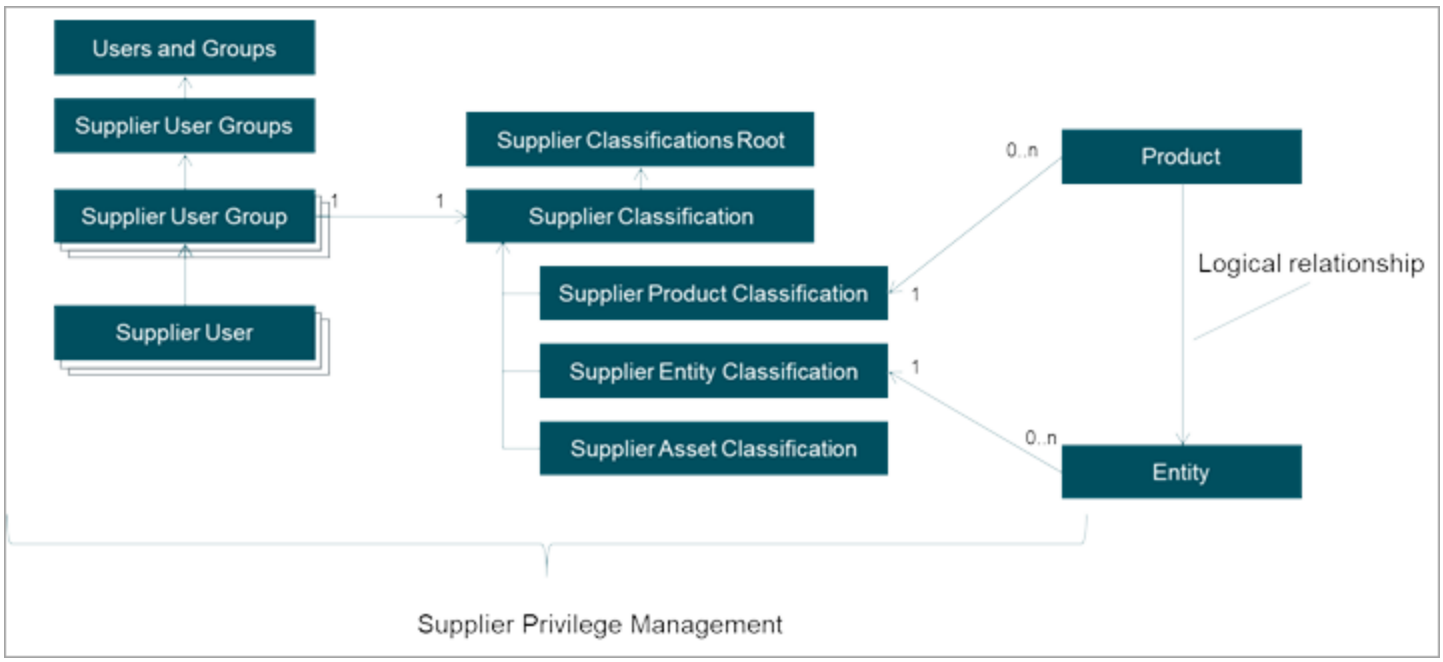
Composite information for suppliers such as addresses, emails, and other contact information should be managed in a similar fashion as for customer object types.

Supplier Self-Service Privileges

In working with supplier entities in MDM, the management of privilege controls is paramount in ensuring supplier-specific data is governed and managed by the intended supplier users. To accommodate this, the Supplier entity type in MDM must contain a Supplier Privileges reference to a Supplier Entities Classification.

Supplier Classifications

The Supplier Entities Classification is a classification (aka 'yellow folder') node type within MDM. Each supplier contains a strict relationship to a Supplier Classification, which ensures suppliers can only interact with objects associated to their unique classification structure. This structure, combined with references to the supplier entity, allow for privilege control between supplier users and supplier entities.



As defined below, privileges can be set once for all suppliers so that a given supplier user cannot see data from other suppliers.

With the following configuration:

- Supplier User Groups are associated to Supplier Classifications
- Users and Group Web UI system setting 'Enable all-view for users that are a member of multiple suppliers' is set to 'Y' (see the **Web UI Settings** topic in the **System Setup / Super User Guide** documentation)
- Access is intentionally granted through privileges on the user group, such as a 'Supplier Users Groups' (see the **Action Sets** topic in the **System Setup / Super User Guide** documentation)

  A supplier user cannot see products, assets, and entities that are:

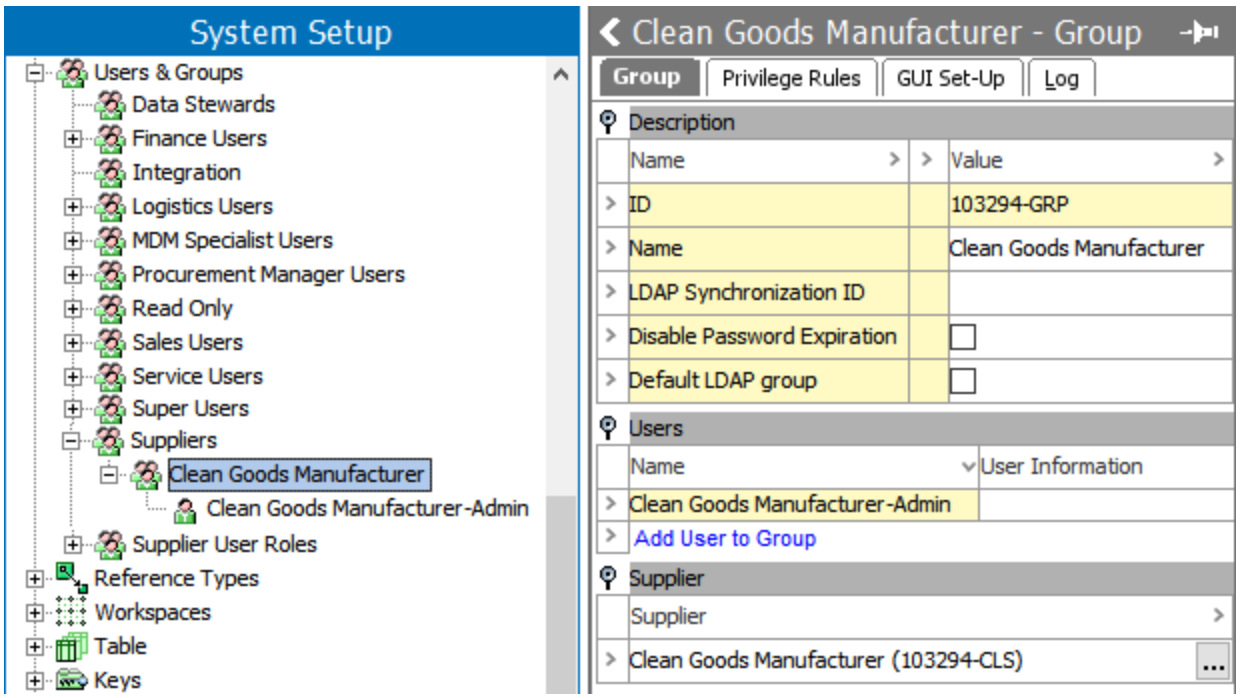
  - linked to a supplier group where the user is not a member.
  - not linked to a supplier group where the user is a member.

  A supplier user can see products, assets, and entities that are:

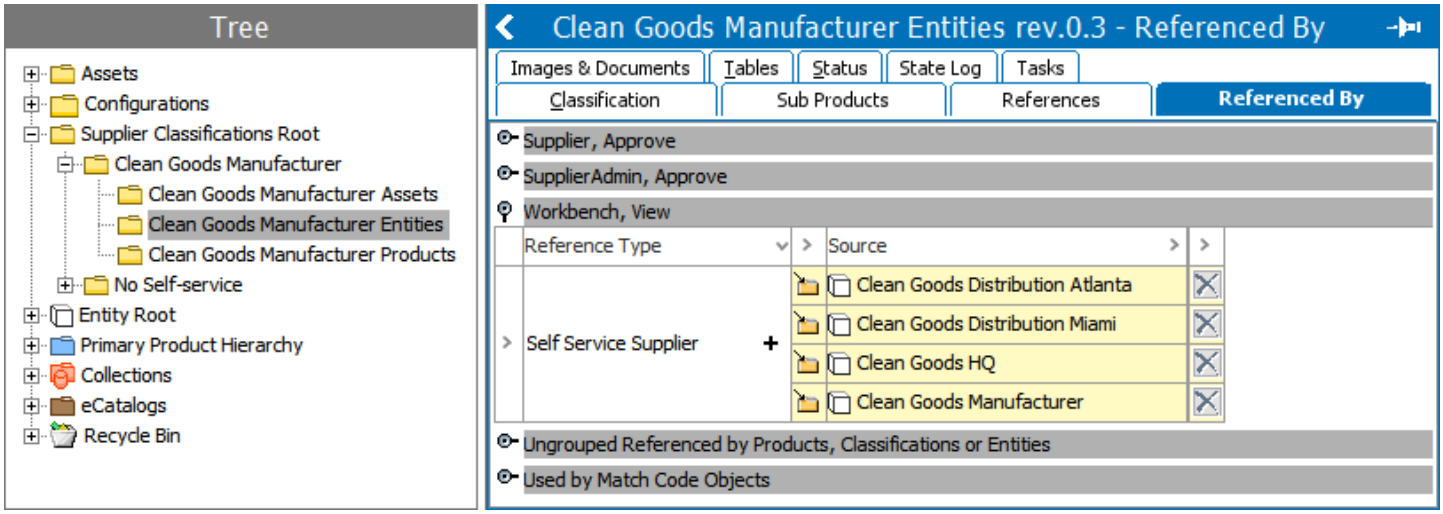
  - not linked to any supplier group.

For example, Clean Goods Manufacturer has been created as a self-service supplier with the configuration defined above.

The Clean Goods Manufacturer user group includes an administrator user and a Supplier link to their classification.



Each classification in the structure includes references from the Clean Goods entities.



This allows the Clean Goods admin user to only have access to objects referenced to any of the super type (entities, product, assets) sub folders within the Clean Goods Manufacturer classification.

Within self-service solutions, it is likely that not every supplier will want to self-service. In this scenario, a 'No Self Service' supplier classification structure and business rules in workflows are used to govern that objects are not unintentionally left visible to all suppliers.

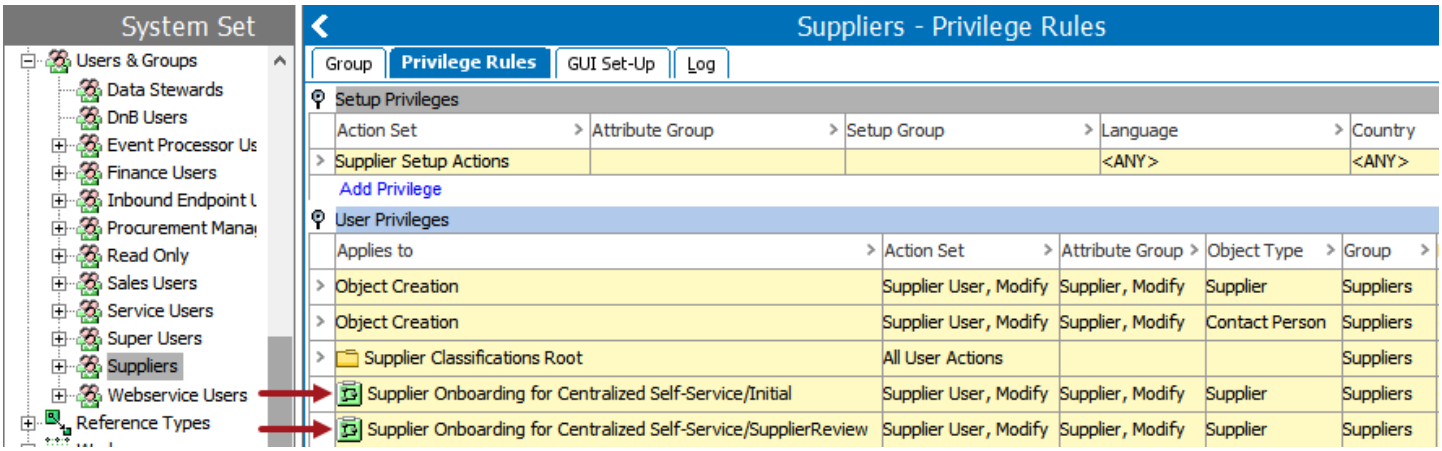
Privilege Setup Considerations

When configuring privilege rules to support self-service, consider the following:

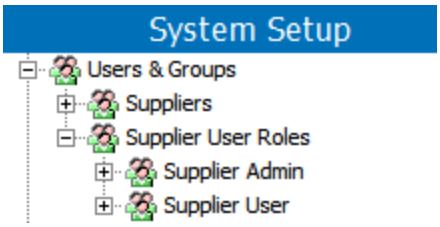
- When should data be editable?
- Who should be allowed to contribute to onboarding or data maintenance?
- Is there an approval process?
- What data should be displayed to the external users?

With the answers those questions, consider the following common configurations:

- **Manage supplier entity onboarding and data maintenance through a workflow.** This allows oversight of data enrichment as well as the ability restrict who can change data and when. To accomplish this, use specific workflow privileges making modify privileges available only when users are interacting with an entity currently in onboarding or maintenance. For maintenance, one approach to provide traceability of user edits is to use revision history by displaying a read-only details screen with a button configured to initialize into a workflow.



- **Allow data enrichment privileges only as needed.** Not all users in an external supplier organization need to nor should be allowed to contribute equally. Apply limited baseline privileges to the parent supplier group so all supplier users have a shared set of privileges. As certain users require additional privileges, configure other user groups to contain the varying added privileges. Link users into the additional groups for the cumulative set of privileges. For example, under the Users & Groups node, an external user added to the Supplier Admin group gains additional privileges over a user only in the Supplier User group. In this case, a supplier admin has the required extra privileges to create new supplier users for his organization.



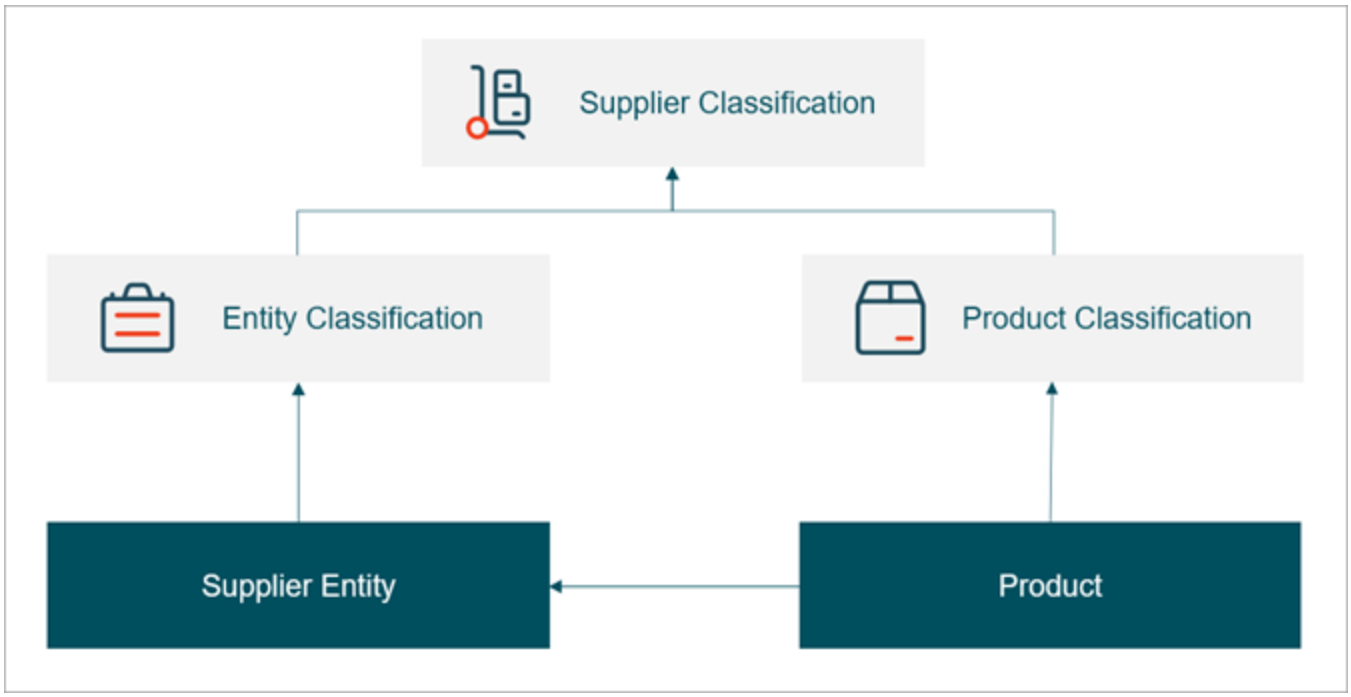
- **Grant approval privileges to only certain user roles.** Allow a specialist user to provide a final review before fully onboarding a new supplier instead of giving all users approval privileges. When required, restrict user roles to only allow view / modify access on certain attributes. For example, a finance user can be allowed to view and edit tax code information but is not allowed to edit address information.

Supplier Products

In a self-service solution, introducing supplier product management requires a similar setup with a classification structure and product privilege links. A product must also have the relevant reference to a specific supplier entity. For example, a supplier's organization may have many distribution centers, however not all products are available at each distribution center.

The following diagram includes:

- A supplier classification with classification sub folders to manage their own supplier entities and products.
- References from supplier entities to the corresponding classification sub folders
- References from products to the corresponding classification sub folder
- Entity references from product to supplier



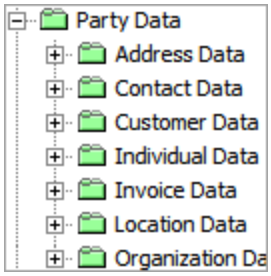
With this setup, each supplier entity within the organization can manage their own unique product offerings by restricting a supplier to view or modify only their own entities and products.

Data Modeling Attributes

The following section details what attributes are used to create the data model.

Attribute Groups

Attributes should be organized into domain specific attribute groups. Typically, these domains include:



- Entity Type: **Address Data Containers**; Domain Folder: **Address Data**

Attribute Name	Attribute ID
Formatted Address	CalcFormattedAddress
Country ISO Code	CountryISOCode
City	InputCity
Country	InputCountry
State	InputState
Street	InputStreet
Zip	InputZip
Latitude	Latitude
Longitude	Longitude
Quality	Quality
Quality Index	QualityIndex
Standardized City	StandardizedCity
Standardized Country	StandardizedCountry
Standardized Country ISO Code	StandardizedCountryISOCode

Attribute Name	Attribute ID
Standardized Formatted Address	StandardizedFormattedAddress
Standardized State	StandardizedState
Standardized Street	StandardizedStreet
Standardized Zip	StandardizedZip
Validation Hash	ValidationHash
Validation Integration Status	ValidationIntegrationStatus
Validation Response	ValidationResponse
Validation Time	ValidationTime

Note: For US addresses, Address1 should always be used since the US reference data does not use a postal box. The PostBox field can be used for countries whose reference data does contain a PO box.

- Entity Type: **Individual Customer**; Domain Folder: **Individual Data**

Attribute Name	Attribute ID
Birth Date	BirthDate
CalcAQI	CalcAQI
CalcHouseholdMembers	CalcHouseholdMembers
Calc Individual Name	CalcName
Name on Collection	CalcNameCollection
Creation Date	CreationDate
Credit Limit	CreditLimit
Deactivated Record	DeactivatedRecord
First Name	FirstName
Gender	Gender
Golden Record ID	GoldenRecordID
Good Person Flag	GoodPersonFlag

Attribute Name	Attribute ID
Income	Income
Income Update Date	IncomeUpdateDate
Last Name	LastName
LastUpdated	LastUpdated
Matching Source	MatchingSource
Middle Name	MiddleName
Nationality	Nationality
Past Due Days	PastDueDays
Past Due Dollars	PastDueDollars
Risk Flag	RiskFlag
Sibling Delinquency Flag	SiblingDelinquencyFlag

Note: For US addresses, 'Address1' should always be used because US reference data does not use postbox. The 'PostBox' field can be used for countries whose reference data does contain a PO box.

- Entity Type: Organization Customer; Domain Folder: Organization Data

Attribute Name	Attribute ID
Account Type	AccountType
CalcAQI	CalcAQI
Name on Collection	CalcNameCollection
CompanyVATNumber	CompanyVATNumber
Creation Date	CreationDate
Deactivated Record	DeactivatedRecord
Employee Size	EmployeeSize
Golden Record ID	GoldenRecordID
LastUpdated	LastUpdated

Attribute Name	Attribute ID
Legal Name	LegalName
NAICS Code	NAICSCode
Revenue Size	RevenueSize
SIC Code	SICCode

- Entity Type: Supplier; Domain Folder: Supplier Data

Attribute Name	Attribute ID
Account Type	AccountType
Bulk Cleanse	BulkCleanse
Certification	Certification
CompanyVATNumber	CompanyVATNumber
Duty Paid	DutyPaid
Duty Unpaid	DutyUnpaid
Employee Size	EmployeeSize
Language	Language
Legal Name	LegalName
NAICS Code	NAICSCode
Revenue Size	RevenueSize
Status	Status
Tax Number	TaxNumber

- Customer Source System Reference Type

Attribute Name	Attribute ID
Source Record ID	SourceRecordID

- Entity Type: **Household**

Attribute Name	Attribute ID
Last Name	LastName

Note: Data container types should be created under the domain with which they are used.

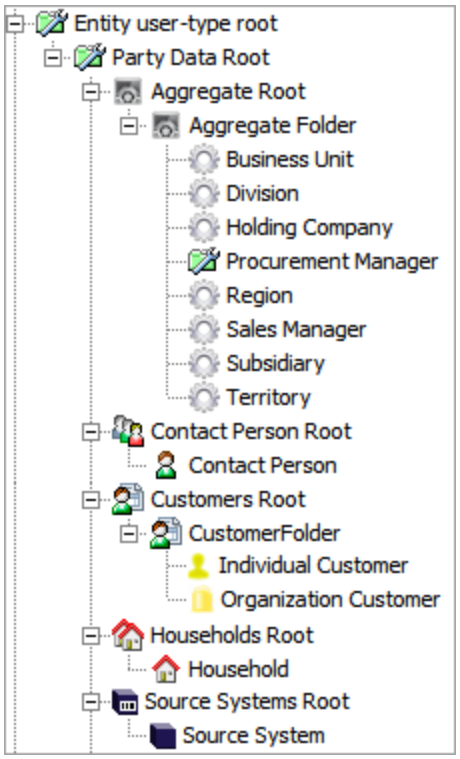
These attributes can be used across different object types, like where Individual and Organizational Customer records, as well as Contact data on Individual Customer, Organization Customer, and Contact Person records.

Per the general guidelines for organizing attributes in attribute groups, separate attribute groups should also be created for:

- Displaying in Web UI, to ease Web UI component setup.
- Each survivorship rule, to provide a clear definition of survivorship rules.
- Each data quality policy, to provide a clear overview of which attributes are part of each data quality metric.
- Each Object Type for each export definition, to provide a clear definition of what is exported and where.
- According to the guidelines of privilege configuration.
- Each Object Type, to describe what is included in data profiling, to ease the profiling configuration.

Other Object Types

Domain Folder constraints are less important than hierarchy constraints, as it typically does not require a user to select a folder. Due to this, it is recommended to limit complexity of Domain Folder Object Types. To preserve a basic integrity of data, it is recommended to create separate Object Types per data domain.



ID Patterns

Use auto-generated STEP IDs for Entities and Data Containers managed by STEP. For information on auto IDs, see the **Autogenerate Using Name Pattern and ID Pattern** topic of the **System Setup / Super User Guide** documentation.

Line of Business Data Model

Introduction

The term 'Line of Business' is defined by Gartner as 'a corporate subdivision focused on a single product or family of products.' It is quite common for Business-to-Business (B2B) organizations to be able to manage their operational master data across multiple lines of businesses (LOB).

For example, CleanGoods Manufacturer, operates across a number of different LOBs:

- Household Cleaning Supplies – Cleaners and chemical cleaning agents manufactured and marketed towards typical household consumers.
- Consumer Charcoal – Consumer-grade charcoal of varying types. Manufactured and marketed for us with barbecue grills.
- Professional Cleaning Supplies – Cleaners and chemical cleaning agents manufactured and marketed towards businesses and institutions such as schools, professional contractors, government entities, etc.

Generally, one LOB may be a subdivision focused on perishable goods that are manufactured, marketed, and sold to average household consumers, whereas another LOB may be focused on selling to professional businesses such as hospitals, corporations, or government entities.

To further illustrate this concept, BlueBox Retailer is a fictitious retail business that is a direct customer to CleanGoods, specifically for CleanGoods Household Cleaning Supplies and CleanGoods Consumer Charcoal.

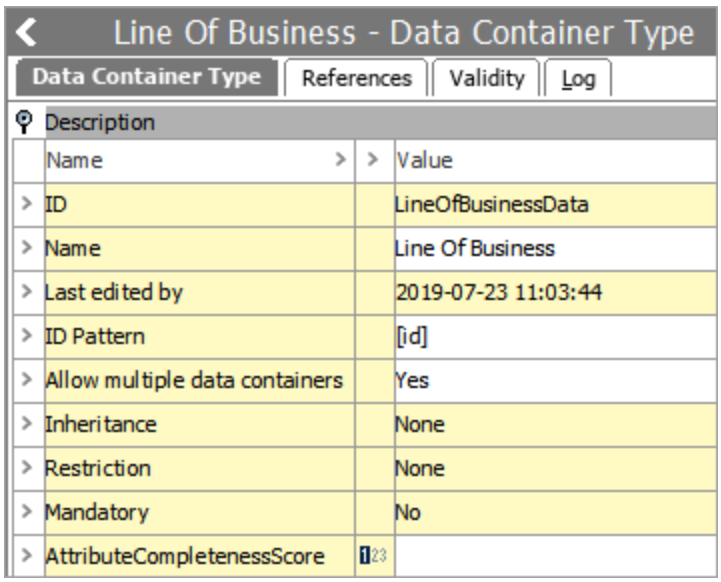
For the sake of ease, CleanGoods should master the customer data of a particular BlueBox retail store, as one Organization Customer Entity, which is shared amongst the different LOBs. This model makes it possible to keep an overview of all interactions with BlueBox, across the LOBs and optimizes the process of having one accurate and complete version of information about BlueBox, across the enterprise.

While this implementation works for these use cases, CleanGoods Household Cleaning Supplies and CleanGoods Consumer Charcoal do not handle all interactions with BlueBox the same way. Instead, they use different invoice schedules and ship-to destinations. This information is treated as customer master data while being LOB specific.

Data Model Explained

LOB-specific customer data should be modeled as data containers on an Organization Customer Entity.

In the example of CleanGoods and BlueBox, exactly one Entity would exist for one particular BlueBox retail store. That Entity would have two data container objects of type 'LineOfBusinessData,' as outlined below.



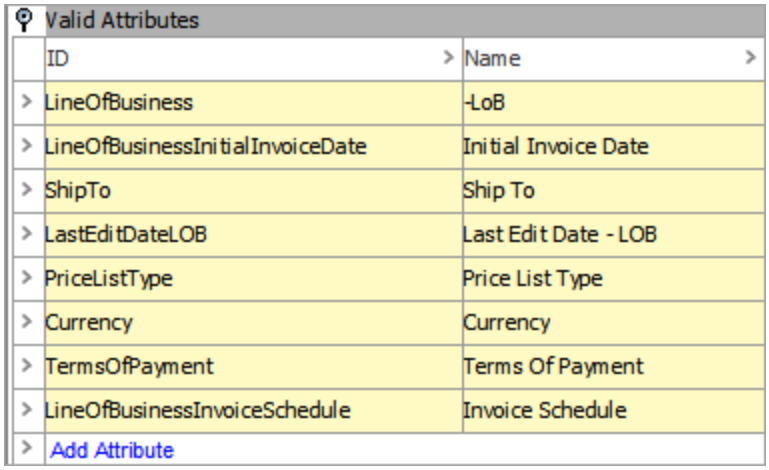
Line Of Business - Data Container Type

Data Container Type | References | Validity | Log

Description

Name	Value
ID	LineOfBusinessData
Name	Line Of Business
Last edited by	2019-07-23 11:03:44
ID Pattern	[id]
Allow multiple data containers	Yes
Inheritance	None
Restriction	None
Mandatory	No
AttributeCompletenessScore	

The initial pre-configurations will include the following attributes associated with LOB data containers:



Valid Attributes

ID	Name
LineOfBusiness	LoB
LineOfBusinessInitialInvoiceDate	Initial Invoice Date
ShipTo	Ship To
LastEditDateLOB	Last Edit Date - LOB
PriceListType	Price List Type
Currency	Currency
TermsOfPayment	Terms Of Payment
LineOfBusinessInvoiceSchedule	Invoice Schedule
Add Attribute	

Data containers are used to model LOB data as associated to a customer record. Each LOB is its own data container object with corresponding attributes and references. The initial configuration for the LineOfBusiness attribute is a data container key. This configuration allows only one entry of each LineOfBusiness per customer.

SAP R/3 Customer, Supplier, and Enterprise Organizational Data Structures

When addressing Business-to-Business (B2B) Customer and Supplier Master Data Management, this documentation considers the following industries:

Customer MDM Domain

- Consumer Packaged Goods (CPG) industry
- Distribution and Manufacturing industry

Supplier MDM Domain

- Consumer Packaged Goods (CPG) industry
- Distribution and Manufacturing industry
- Retail industry

The customer and supplier master data includes the following elements:

- Complexity within the customer's or supplier's organization
- Complexity within the CPG company's internal organization – in SAP called the 'Enterprise Structure'
- MDM must couple the customer's complex organization with the CPG company's complex internal organization

SAP has a structure that is closely aligned with the reality of these organizations and SAP is also likely the ERP of most of these companies. Therefore, SAP is an important and very useful reference point for understanding these organizations and processes, as well as for understanding what surrounding MDM systems must be able to coexist.

The MDM platform must be able to accommodate the complexities of SAP data structures and processes with basic capabilities for data modeling, user privileges, processes, integrations, and usability.

Background

Many companies in the CPG and manufacturing industries, for example, manage a wide array of products, brands, categories, and channels, and serve diverse customer segments across multiple countries worldwide. The end consumers of products are either individual consumers or businesses like hospitals, schools, restaurants, etc. For CPG and manufacturing companies, typically revenue is generated by selling through a network of retail chains, distributors, and master wholesalers. The company considers these parties their customers. Any CPG or manufacturing company has the following functions as part of the sales side of their business:

- Sales (including promotion planning, for example)
- Logistics
- Accounting and finance

While companies require mastering customer data for understanding and maintaining supplier data as well as for sales purposes, it is equally critical to the business to facilitate purchasing activities, logistics, and other business processes.

Companies often have the following functions as part of the procurement side of the business:

- Purchasing (including sourcing, soliciting, etc.)
- Accounting and financing
- Logistics

It is common for each of these functions to be managed locally (in each country) and to some extent, in each line of business. In many cases, this is a reasonable decentralization of the organization due to the need for regional agility in combination with relatively small optimization gain a centralized organization could have.

Example

Consider CleanGoods Manufacturer, a fictitious CPG company that specializes in cleaning products and operates across two major business units:

- Traditional household cleaning products - where the end consumer buys the products in a retail store.
- Clinical products for professionals - where the end consumers are hospitals, doctors, physicians, etc.

Business units sell different and/or the same products and these products are sold and distributed differently via various sales organizations and distribution channels. The same customer may buy products from different product lines. The two business units may also source from the same suppliers.

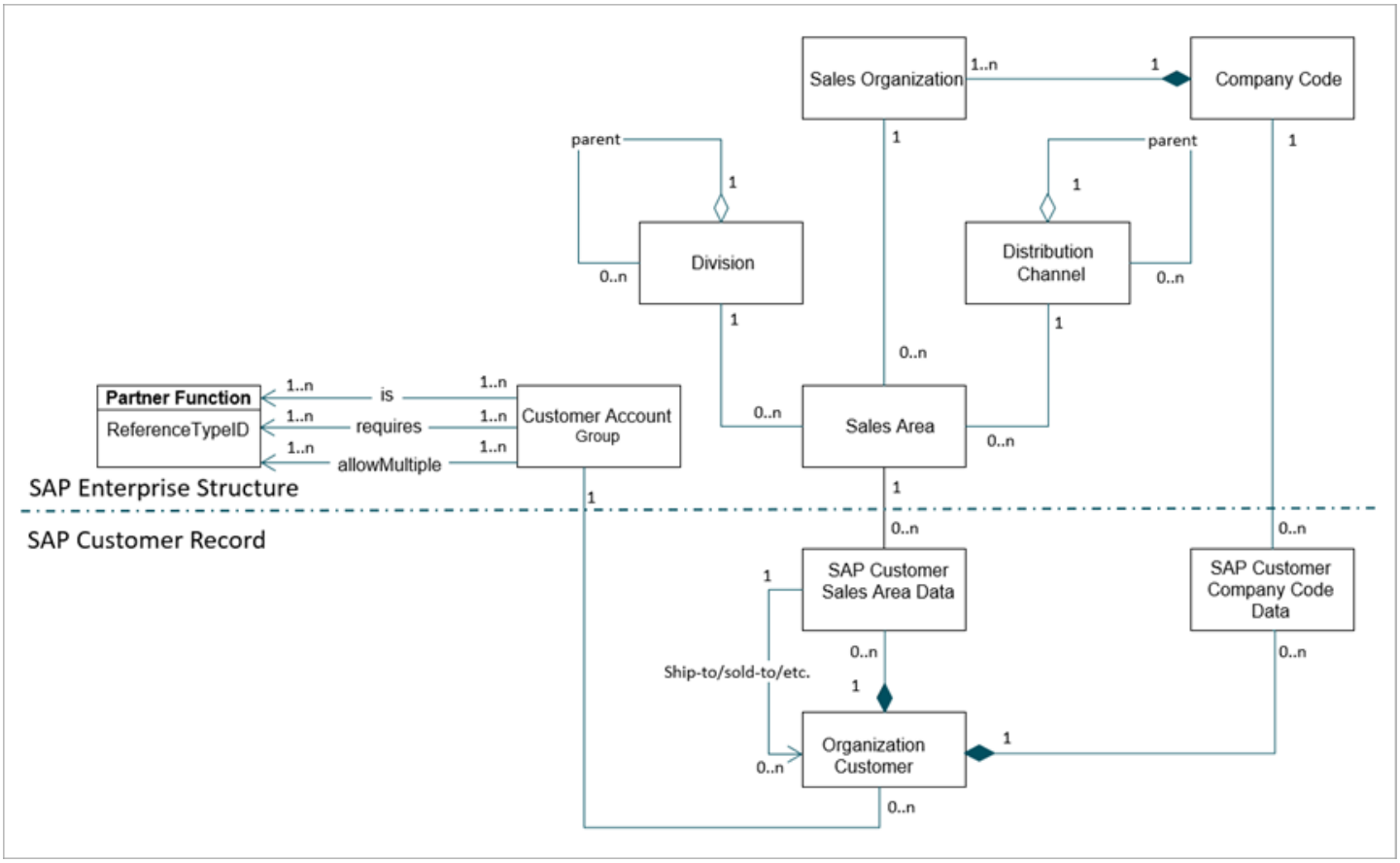
CleanGoods Manufacturer operates in many countries. Even though CleanGoods deals with international and local companies, it is common for customers to have separate business units in each country.

By mastering supplier data, CleanGoods is better equipped to support the onboarding and maintenance of suppliers across lines of business and to facilitate activities such as negotiating terms of payments, managing supplier rebates, etc.

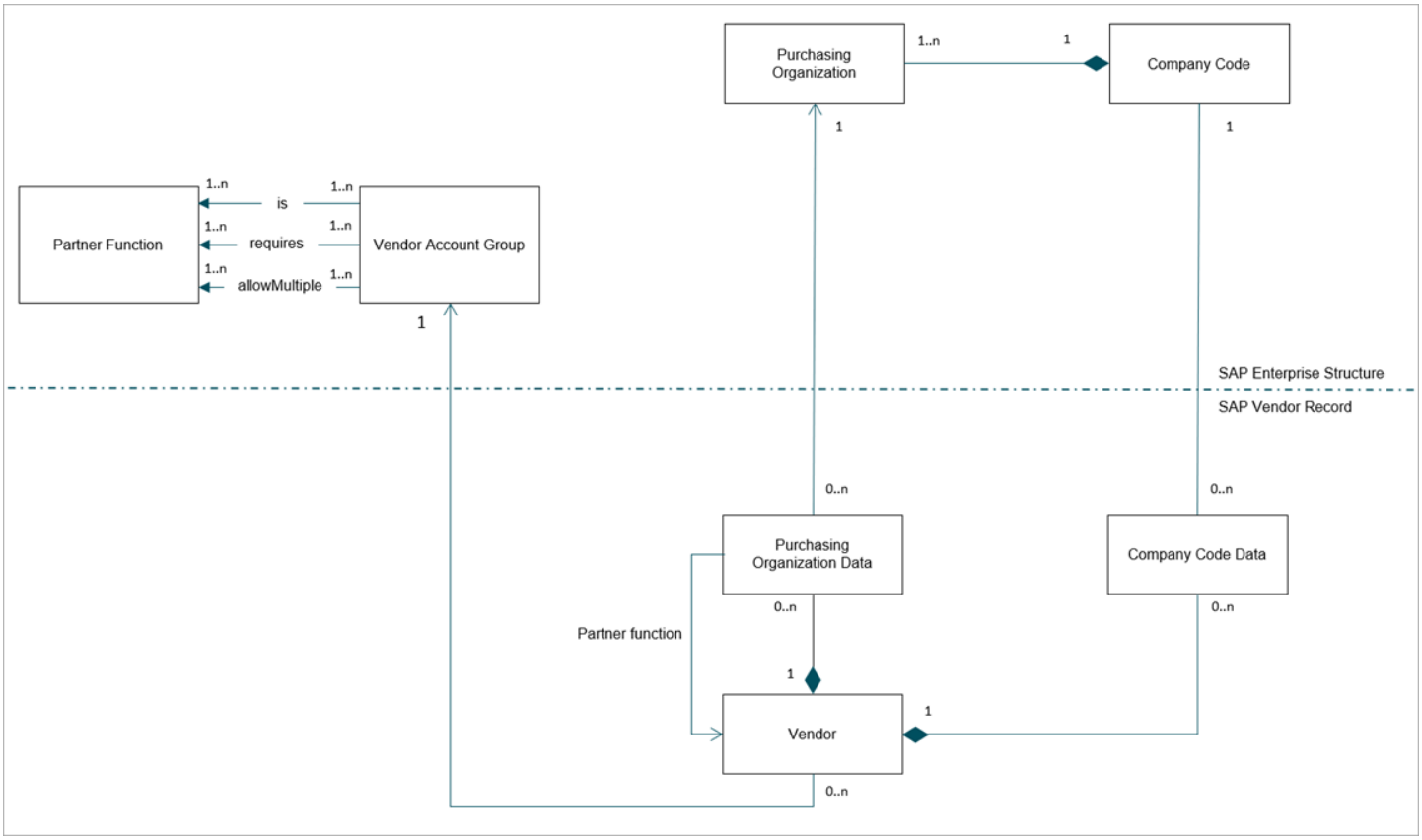
For information on the general relationships, constraints, and terms used, see the **SAP R/3 Customer & Supplier, and Enterprise Structure Definitions** topic. For information on the implementation structures of SAP in STEP, see the **SAP Customer, Supplier, and Enterprise Structures** topic.

SAP R/3 Customer, Supplier, and Enterprise Structure Definitions

This documentation assumes an understanding of generic SAP structures. The following logical model diagram shows an example of the data structure within SAP in a Customer MDM implementation.



The following diagram shows an example of the data structure within SAP in a Supplier MDM implementation.



Because this topic uses SAP terminology, all readers are recommended to be familiar with these terms to better understand SAP data structures.

For the following terms, see the SAP Online Documentation for further details:

https://help.sap.com/viewer/e43f5774d00640d9adbf39734fe5933c/6.00.29/en-US/9a95c7536e8e2a4be10000000a174cb4.html

- Company Code
- Sales Organization
- Distribution Channel
- Division
- Sales Area
- Partner Functions
- Account Group
- Purchase Organization

For Partner Functions and Account Groups, refer to this SAP Online Documentation for further details:

https://help.sap.com/viewer/e79854f090014378a89d74024923dbab/6.00.29/en-US/8b6fbd534f22b44ce10000000a174cb4.html

SAP Customer and Supplier Data Constraints

The following sections expand on the logical model which details the relationships between entities and elements that include SAP customer, supplier, and enterprise structures.

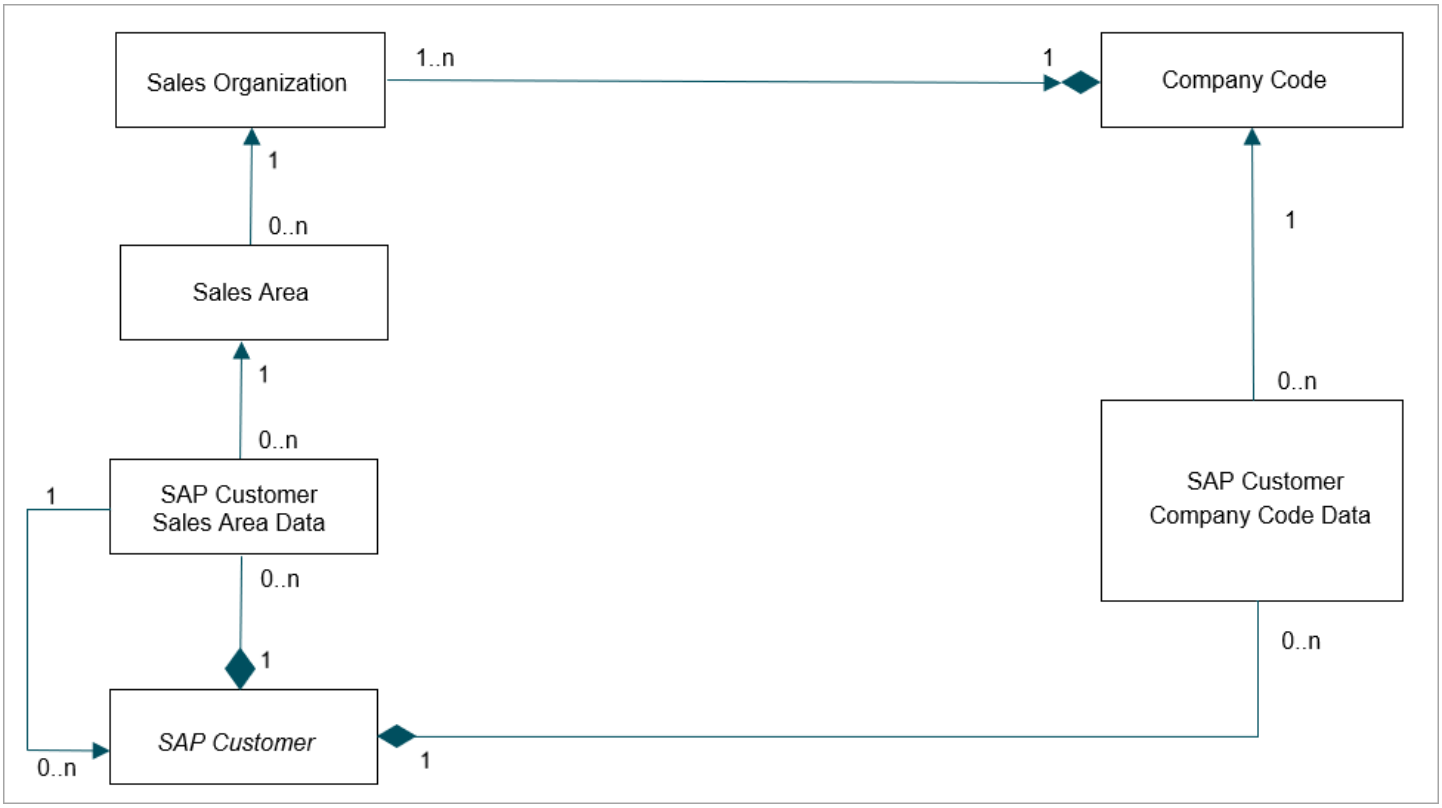
Uniqueness Constraints

The following is a summary of the uniqueness constraints presented in the logical model diagram at the beginning of this topic:

- The following types are unique by ID:
  - Customer Account Group
  - Vendor Account Group
  - Company Code
  - Sales Organization
  - Distribution Channel
  - Division
  - Customer
  - Supplier
  - Purchase Organization
- Sales Area is unique by the combination of sales organization, distribution channel, and division.
- Customer Sales Area Data is unique by customer and sales area.
- Customer Company Code Data is unique by customer and company code.
- Supplier Company Code Data is unique by supplier and company code.
- Purchase Organization Data is unique by supplier and purchase organization.

Sales Area Constraints

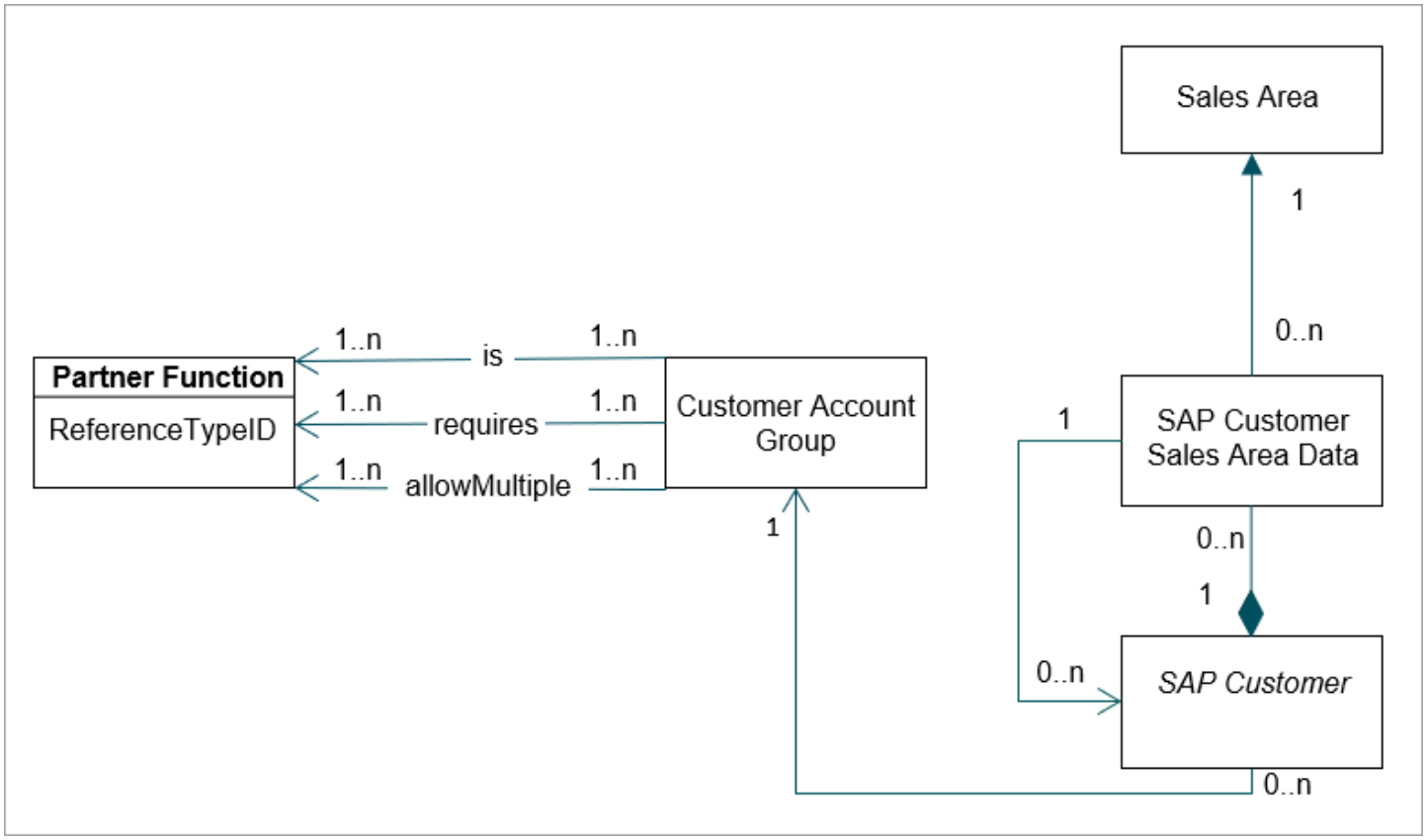
A Sales Area defines the sales organization and distribution channel a division uses to sell products. Therefore, sales areas are comprised of distribution channel, division, and sales organization.



In STEP, sales areas are referenced from the SAP Customer object using a reference (ID= SAPCustomerSalesAreaDataSalesArea) from the Sales Area data container (ID= SAPCustomerSalesAreaData). A Business Partner (BP) may contain zero-to-many Sales Area data containers (ID=SAPCustomerSalesAreaData).

An SAP Customer may only have Sales Area Data if the corresponding Company Code Data is present such that the Sales Area must be associated to a Sales Organization that is associated to a Company Code for which the SAP Customer has Company Code Data.

Partner Functions Constraints



A Customer has exactly one Customer Account Group and a Supplier has exactly one Vendor Account Group. An account group has one-to-many valid partner functions. Not all customers and suppliers can act in all partner functions.

Since a Customer has a given account group, it must also have one or multiple Partner Function relations.

- **Valid Partner Functions** - Not all customers or suppliers can act in all partner functions. The Account Group defines one-to-many valid Partner Function relations for a Customer / Supplier.
- **Mandatory Partner Functions** - For each Sales Area Data or Purchasing Organization Data, the Account Group defines the required Partner Function relations for a Customer / Supplier.
- **Allow Multiple of a Partner Function** - For each Sales Area Data or Purchasing Organization Data, the Account Group defines which Partner Function relations allow multiple Customers / Suppliers.

Common Distribution Channel and Division Constraints

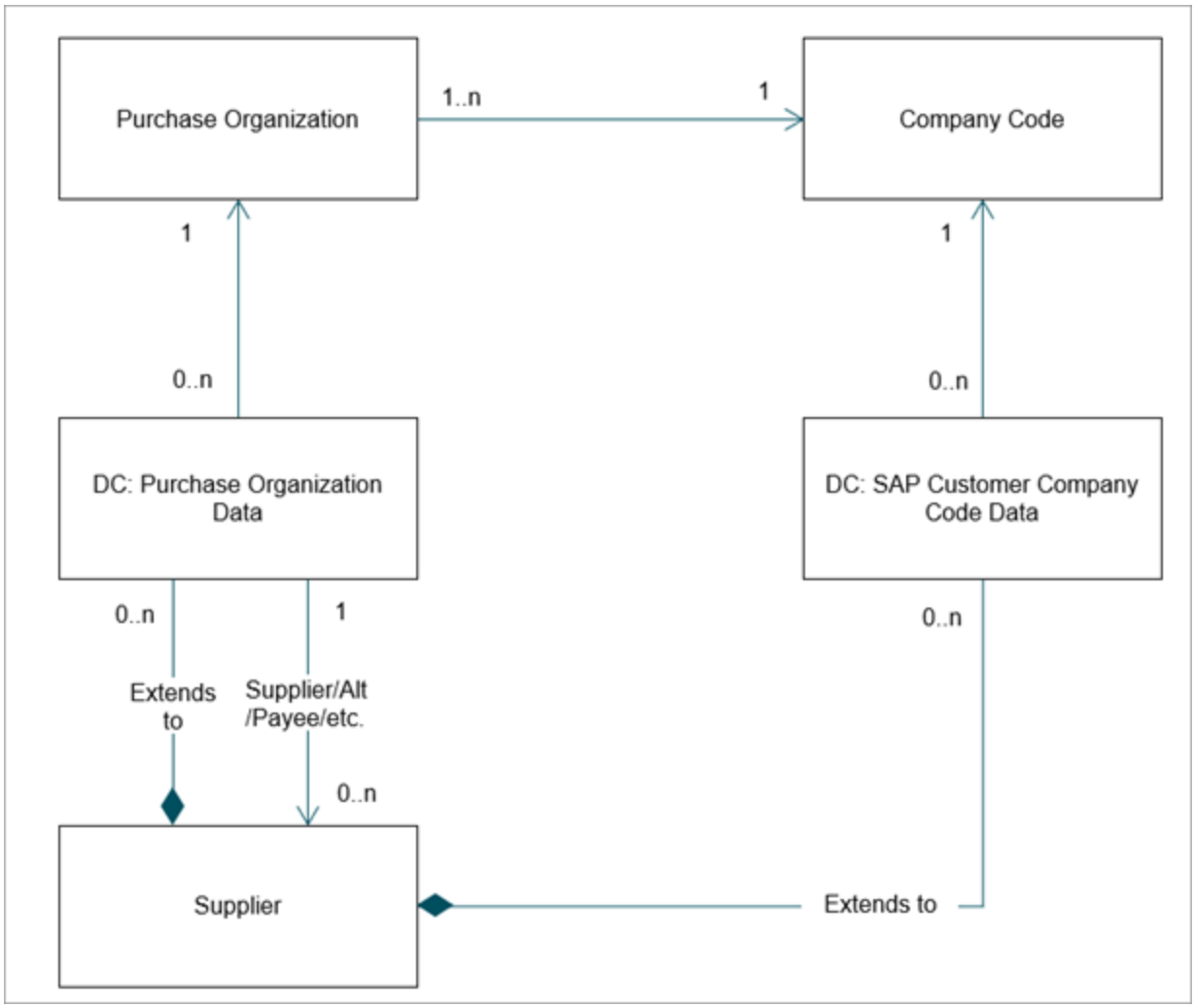
To avoid redundant sales area master data across distribution channels and divisions, distribution channels and division may have a common 'parent,' which is logically resolved into what may be termed as 'common sales areas.' This concept means that sales areas associated with a particular distribution channel and division may have a common distribution channel and common division in another sales area within the same sales organization. This hierarchy is nested.

- If a customer is active in a sales area that is common for other sales areas, this customer cannot have master data in the other sales areas.
- If a customer is active in a sales area, a user is not allowed to activate that customer in the sales area's common sales areas.

Purchase Organization Constraints

A BP may contain zero-to-many Purchasing Organization Data Containers (ID=SAPPurchasingOrganizationData).

A BP may only have Purchase Organization Data if the corresponding Company Code Data is present such that the Purchase Organization (also called a 'purchasing organization') must be associated with a Company Code for which the Supplier has Company Code Data.



SAP R/3 Customer, Supplier, and Enterprise Structures

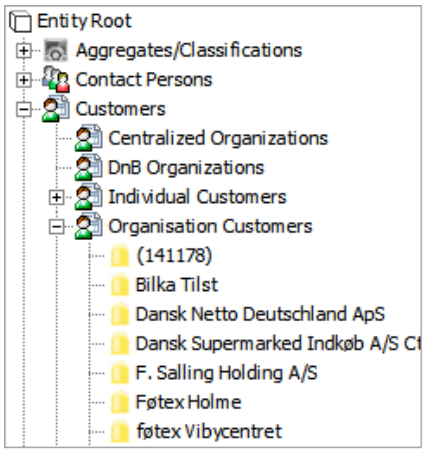
The sections below provide an overview of the Customer MDM implementation of SAP data structures.

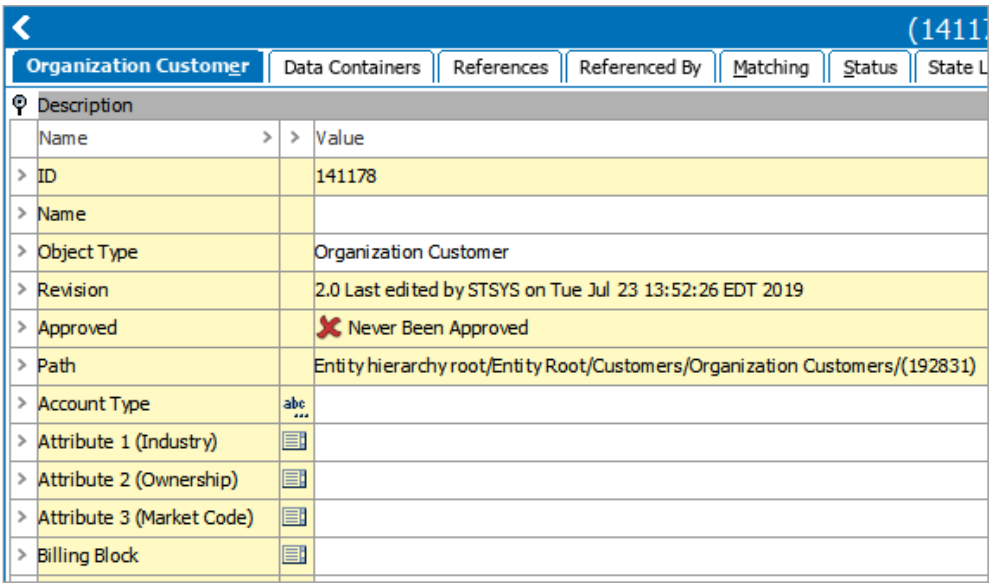
For additional information, see the following topics:

- **SAP R/3 Customer, Supplier, and Enterprise Organizational Data Structures** topic includes SAP implementation.
- **SAP R/3 Customer, Supplier, and Enterprise Structure Definitions** topic includes the general relationships, constraints, and terms used.

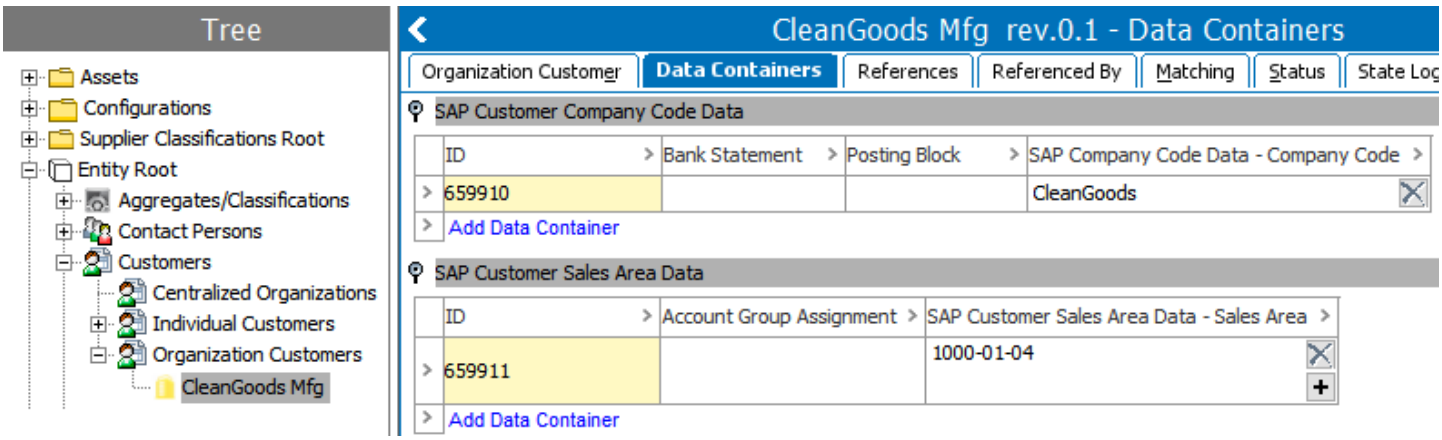
SAP Customer

The SAP Customer is the direct customer of the CPG company. For example, BlueBox Retailer is a customer of CleanGoods Manufacturer. The SAP Customer is represented by the Organization Customer entity type in Customer MDM.





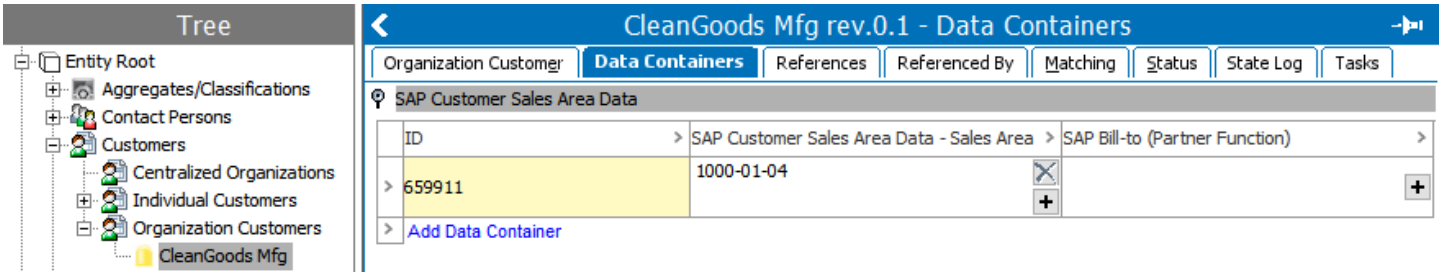
- **Customer Company Code Data** and **Customer Sales Area Data** are data containers on the Organization Customer entity.



- **Data Container-to-Entity References**

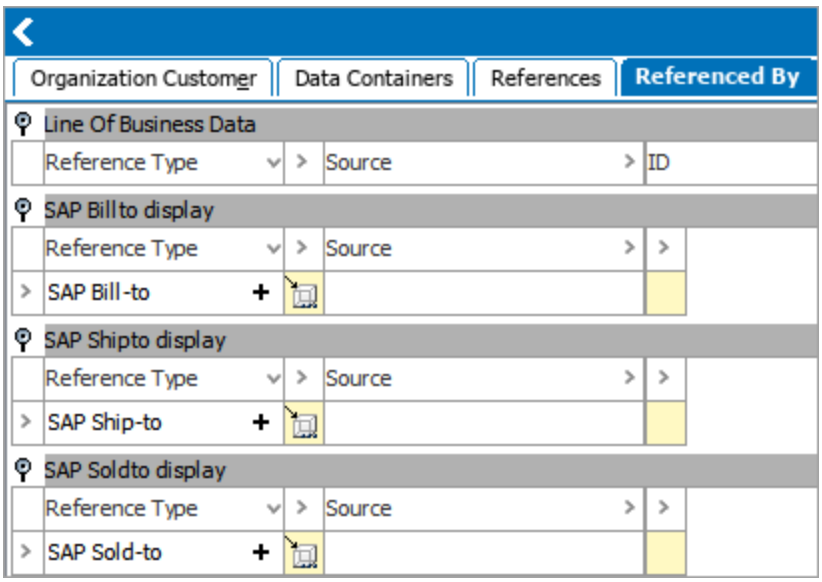
  For each ‘Customer Company Code Data’ data container, a reference points to the target Company Code entity, such as 'SAP Company Code Data' in the image above.

  For each ‘Customer Sales Area Data’ data container, a reference points to the target Sales Area entity, such as '-SAP Sales Customer Sales Area Data' in the image above.



  For each ‘Customer Sales Area Data’ data container, Partner Function references point to an Organization Customer record. Conversely, each Organization Customer record may be referenced by a ‘Customer Sales Area Data’ data container.

  Additionally, on the Organization Customer object, SAP Customer Account Group reference points to an Account Group entity.

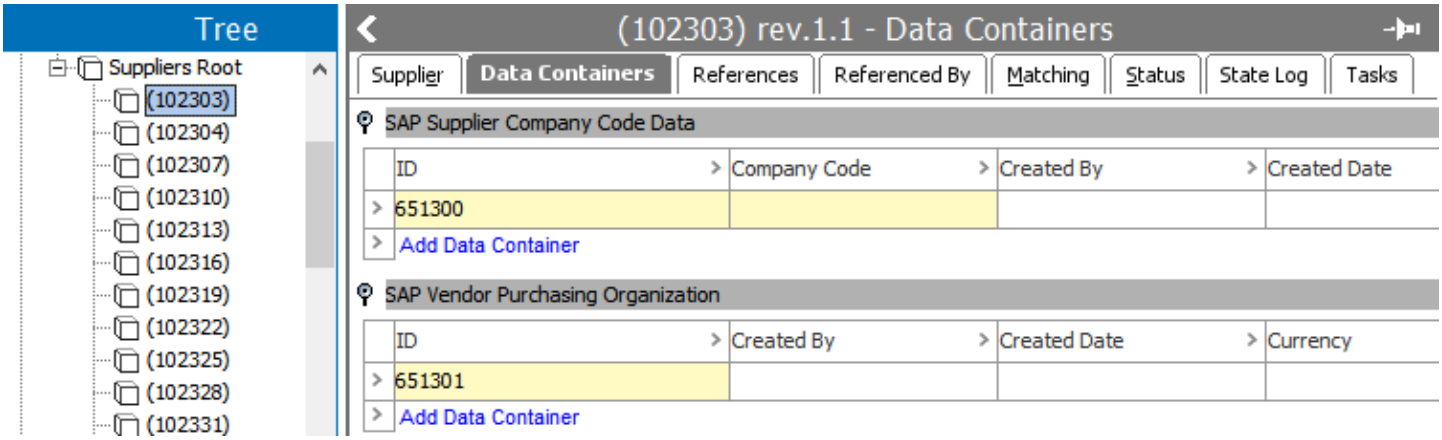


SAP Supplier

The SAP Supplier is the supplier of the CPG company and is represented by the Supplier entity type in Supplier MDM.

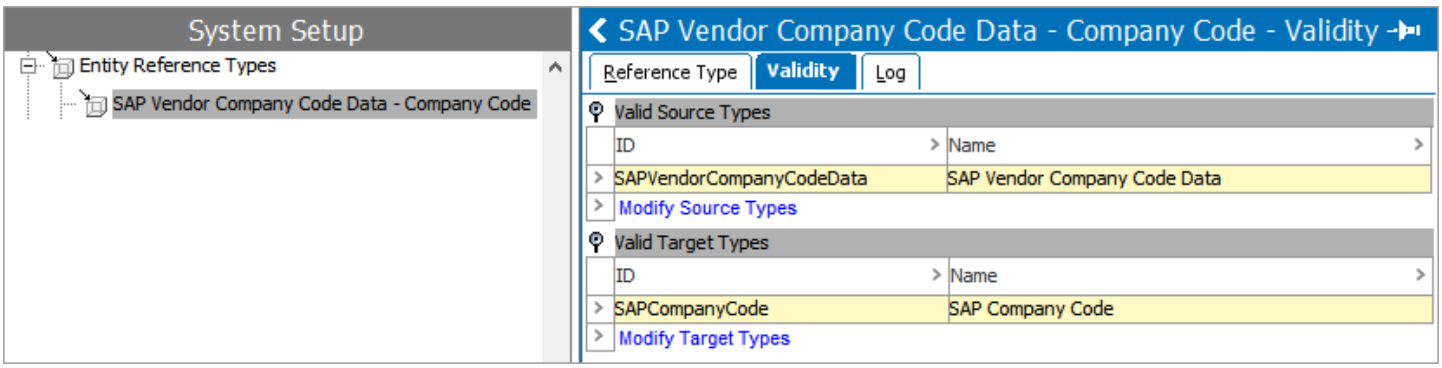
Important: Since this topic uses SAP terms, being familiar with this terminology will help you to better understand SAP data structures.

- **Supplier Company Code Data and Supplier Purchasing Organization Data** - data containers on the Supplier entity.

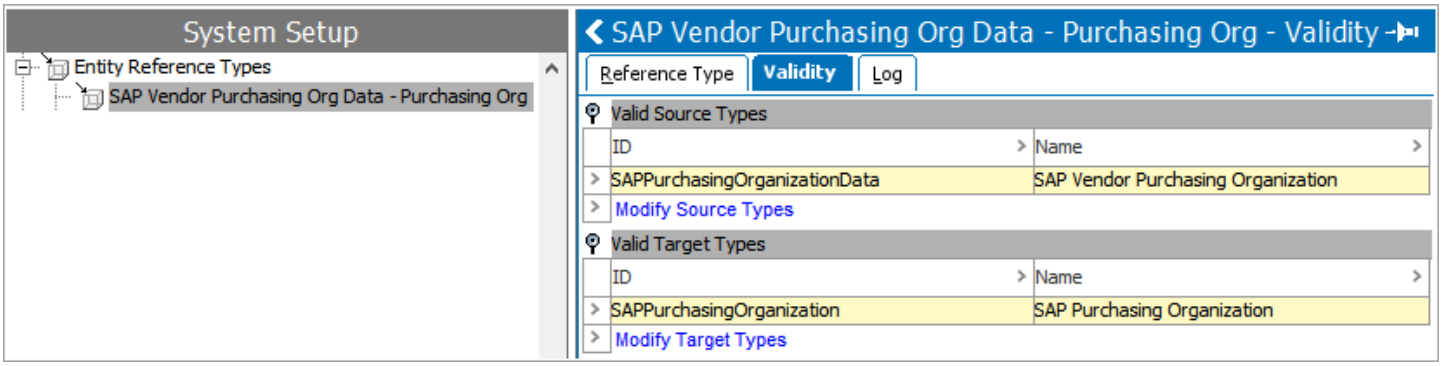


- **Data Container-to-Entity References**

  For each 'SAP Vendor Company Code Data' data container, a reference points to the target Company Code entity.



For each 'Supplier Purchasing Organization Data' data container, a reference points to the target Purchasing Organization entity.



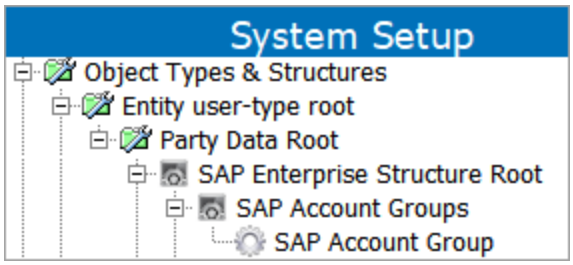
SAP Enterprise Structure

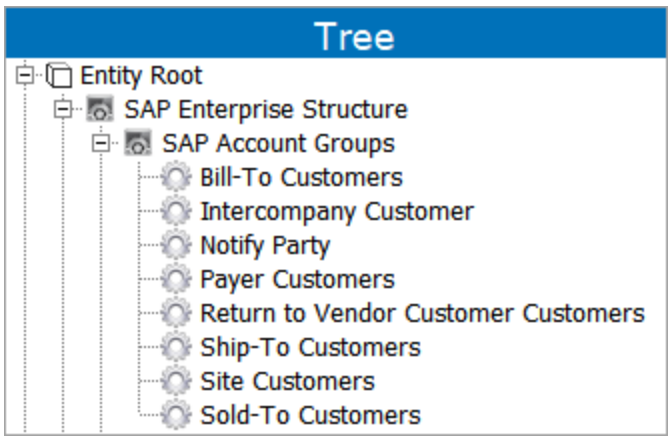
A JavaScript business condition is used to govern the Customer Entity data integrity constraints. For more information, see the **SAP Customer, Supplier, and Enterprise Structure Definitions** topic in this documentation.

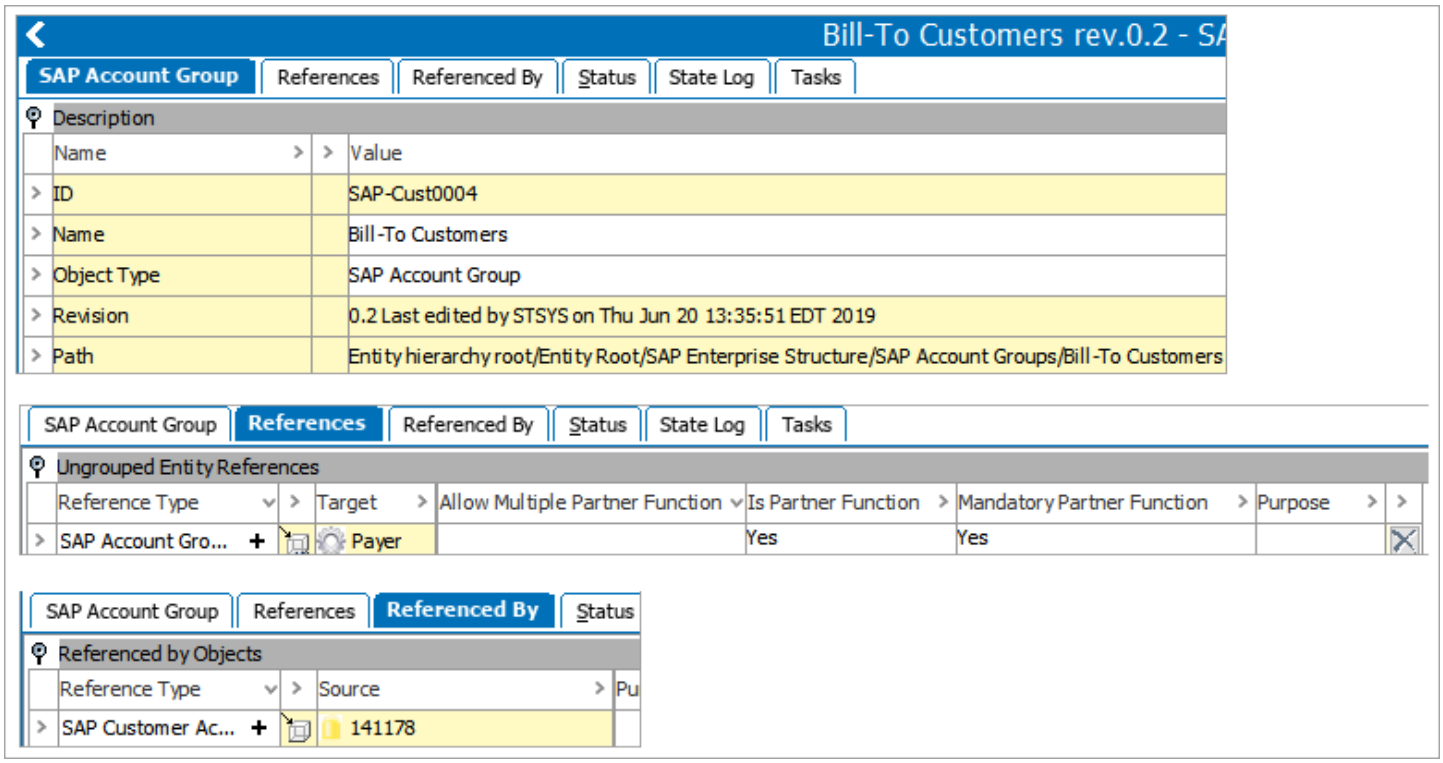
The business condition assesses if the Customer Entity is valid for the configured Enterprise Structure. The execution of this business condition should happen in a workflow as part of a transition state and on approval of customer data.

The Customer and Supplier MDM enablement pre-configurations for SAP data structures include the basic SAP data model for object types, attributes, and references.

Account Group

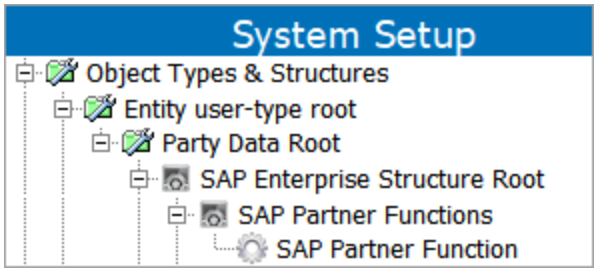


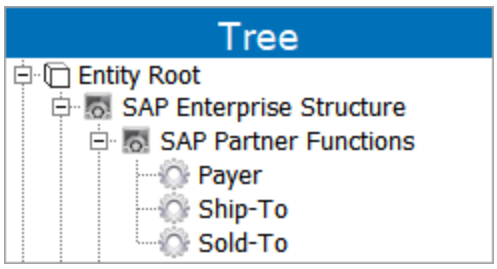


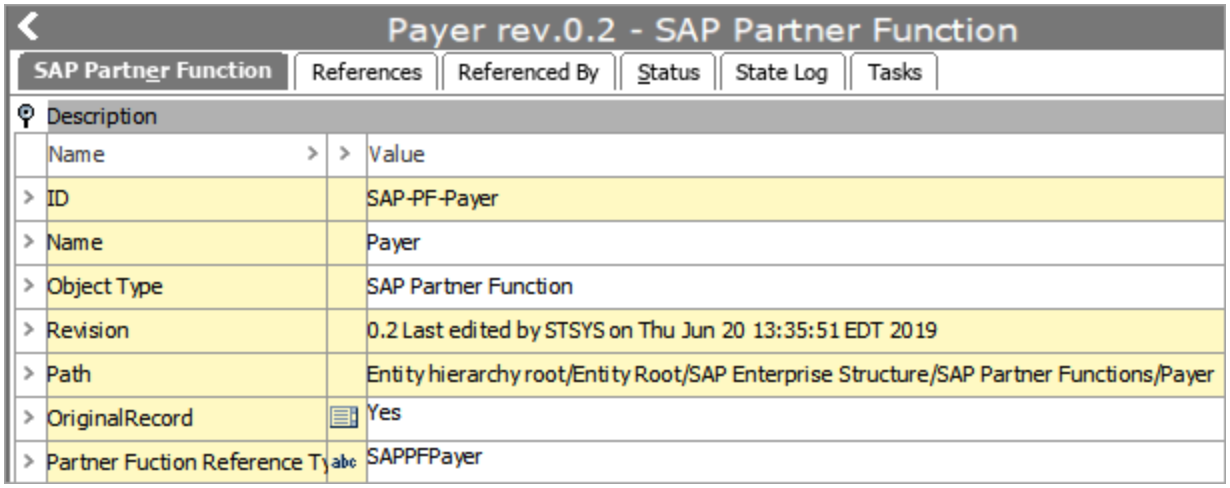


Partner Function

The Customer MDM representations of SAP partner functions are via a dedicated entity type.

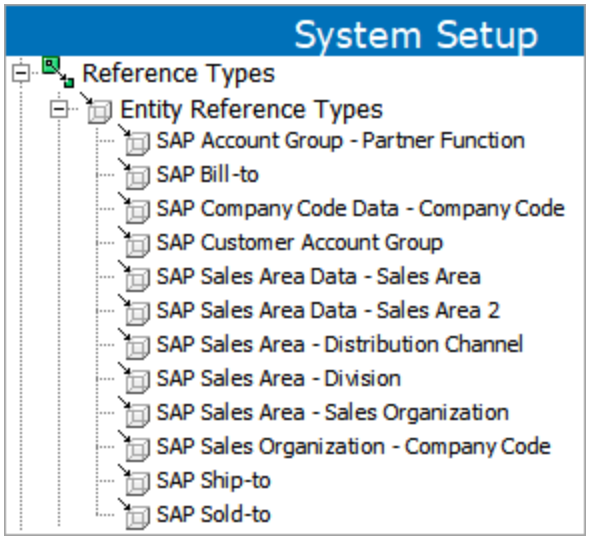






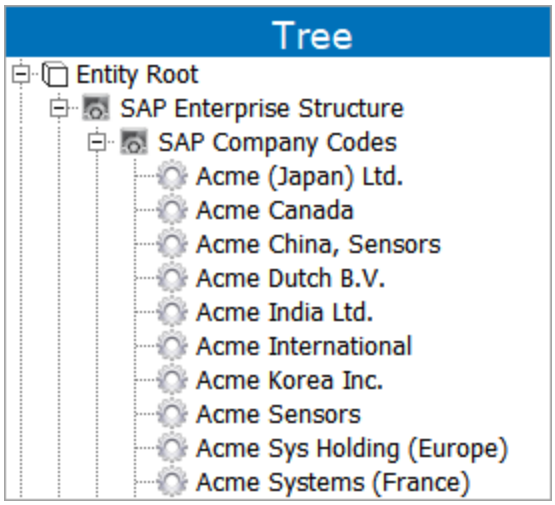
Reference Types

References from 'Customer Sales Area Data' data container on Organization Customer entity to other Organization Customer entity or reference to itself.



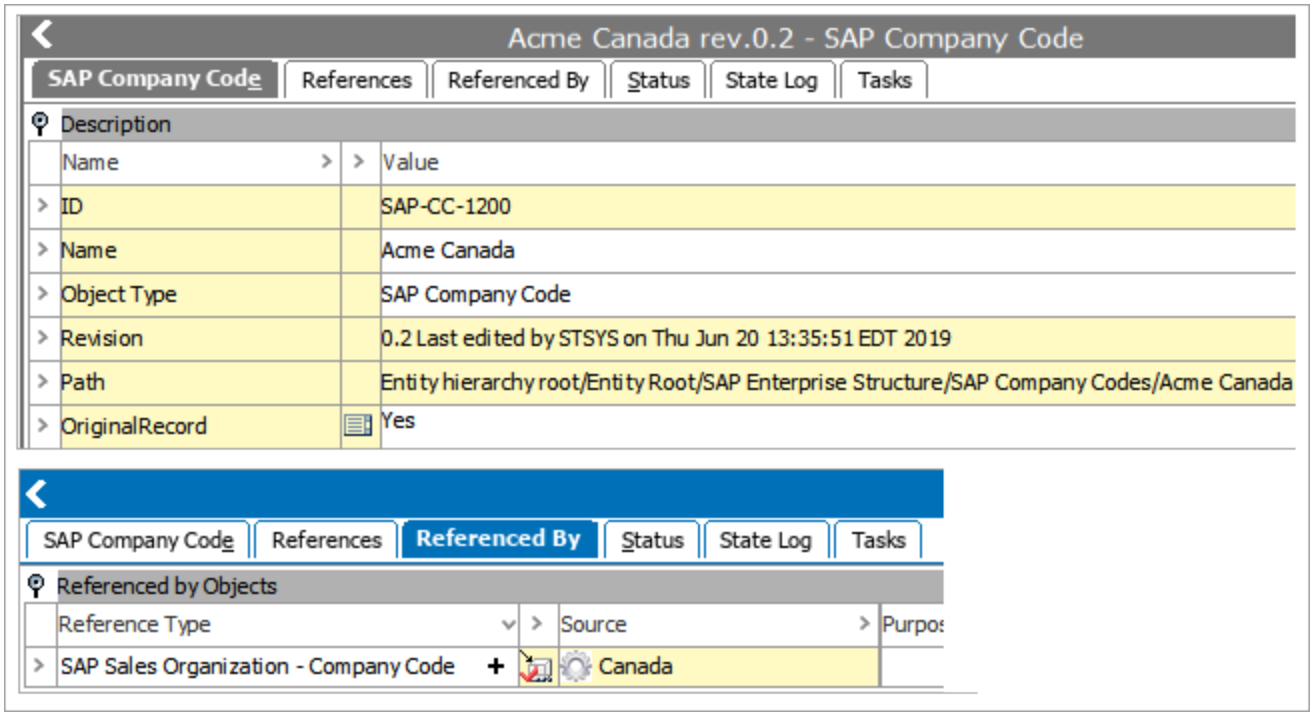
SAP Company Code

As a global company, CleanGoods has separate operating entities determined by geography. Each operating entity is responsible for its own sales and financial reporting, hence the need for separate organization representation as Company Codes.

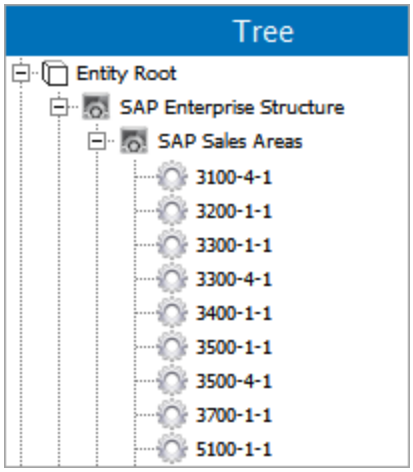


In Customer MDM, Company Code is its own entity type and is a reference from:

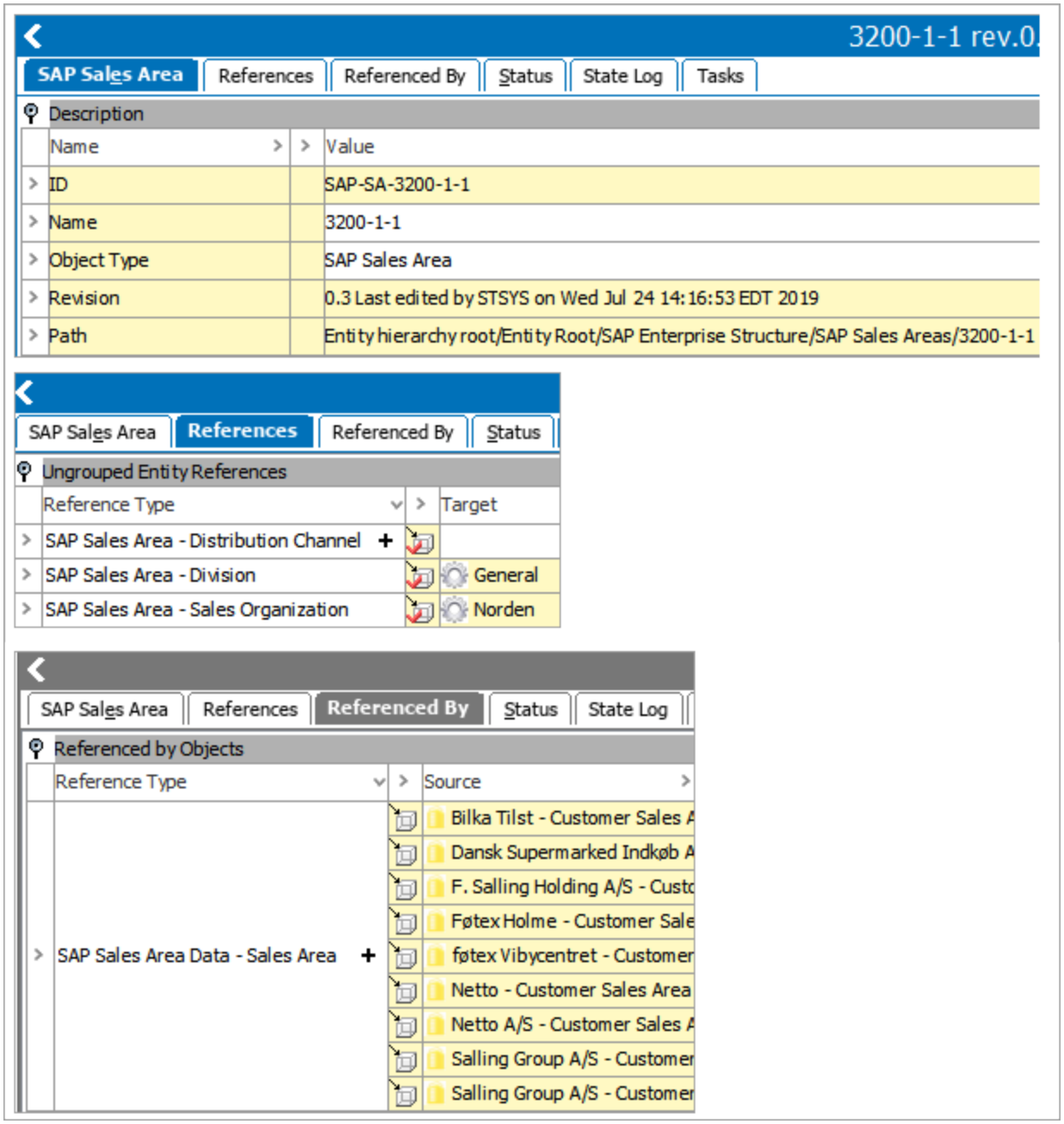
- 'Customer Company Code Data' data container via 'SAP Company Code Data – Company Code' reference type.
- 'Sales Organization' entity via 'SAP Sales Organization – Company Code' reference type.



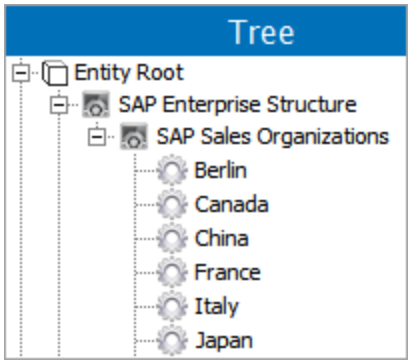
Sales Area

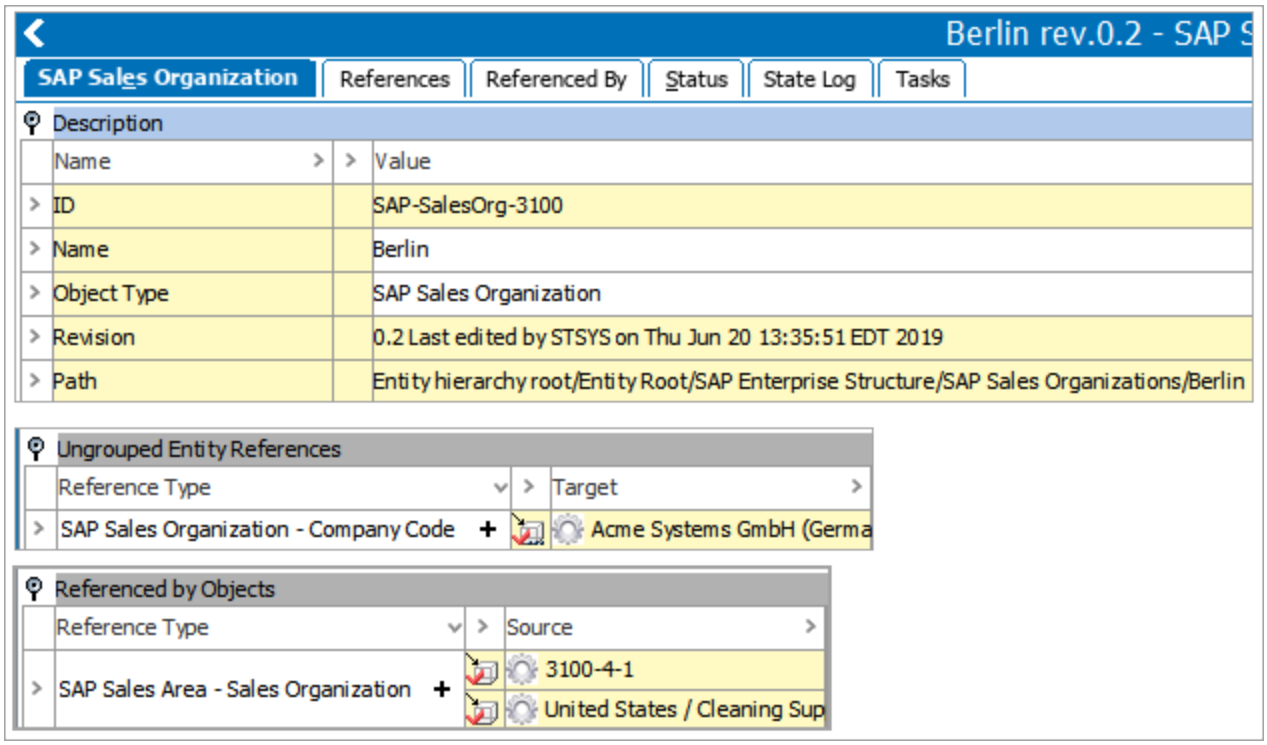


SAP Sales Area Data – Sales Area is a reference from a Sales Area data container.

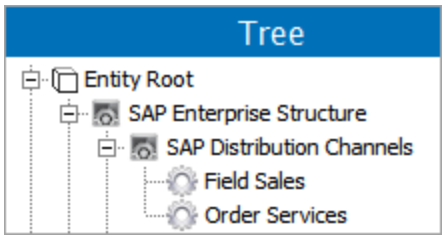

3200-1-1 rev.0
SAP Sales Area
References
Referenced By
Status
State Log
Tasks
Description
Name
Value
ID
SAP-SA-3200-1-1
Name
3200-1-1
Object Type
SAP Sales Area
Revision
0.3 Last edited by STSYS on Wed Jul 24 14:16:53 EDT 2019
Path
Entity hierarchy root/Entity Root/SAP Enterprise Structure/SAP Sales Areas/3200-1-1
SAP Sales Area
References
Referenced By
Status
Ungrouped Entity References
Reference Type
Target
SAP Sales Area - Distribution Channel
SAP Sales Area - Division
General
SAP Sales Area - Sales Organization
Norden
SAP Sales Area
References
Referenced By
Status
State Log
Referenced by Objects
Reference Type
Source
SAP Sales Area Data - Sales Area
Bilka Tilst - Customer Sales A
Dansk Supermarked Indkøb A
F. Salling Holding A/S - Custo
Føtex Holme - Customer Sale
føtex Vibycentret - Customer
Netto - Customer Sales Area
Netto A/S - Customer Sales A
Salling Group A/S - Customer
Salling Group A/S - Customer

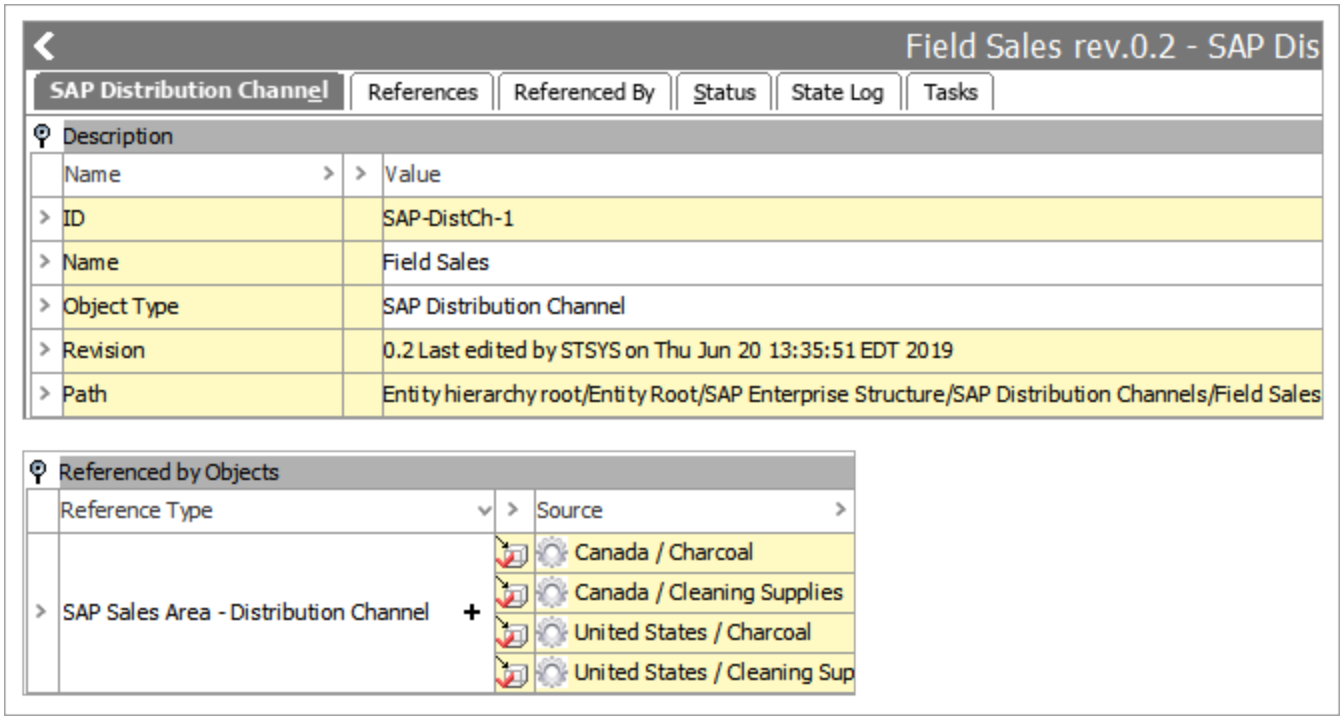
Sales Organization



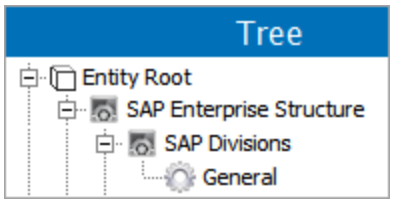


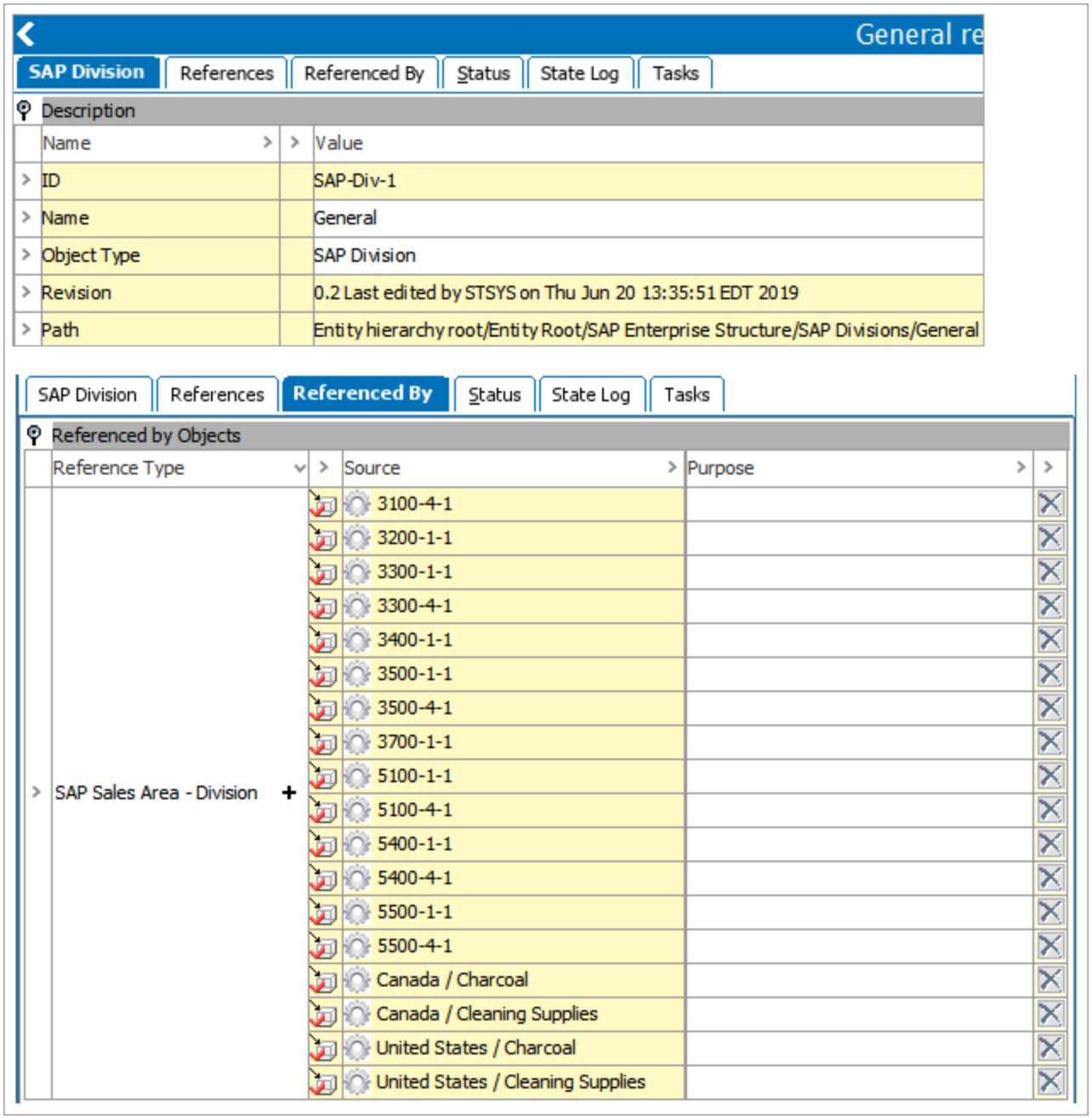
Distribution Channel



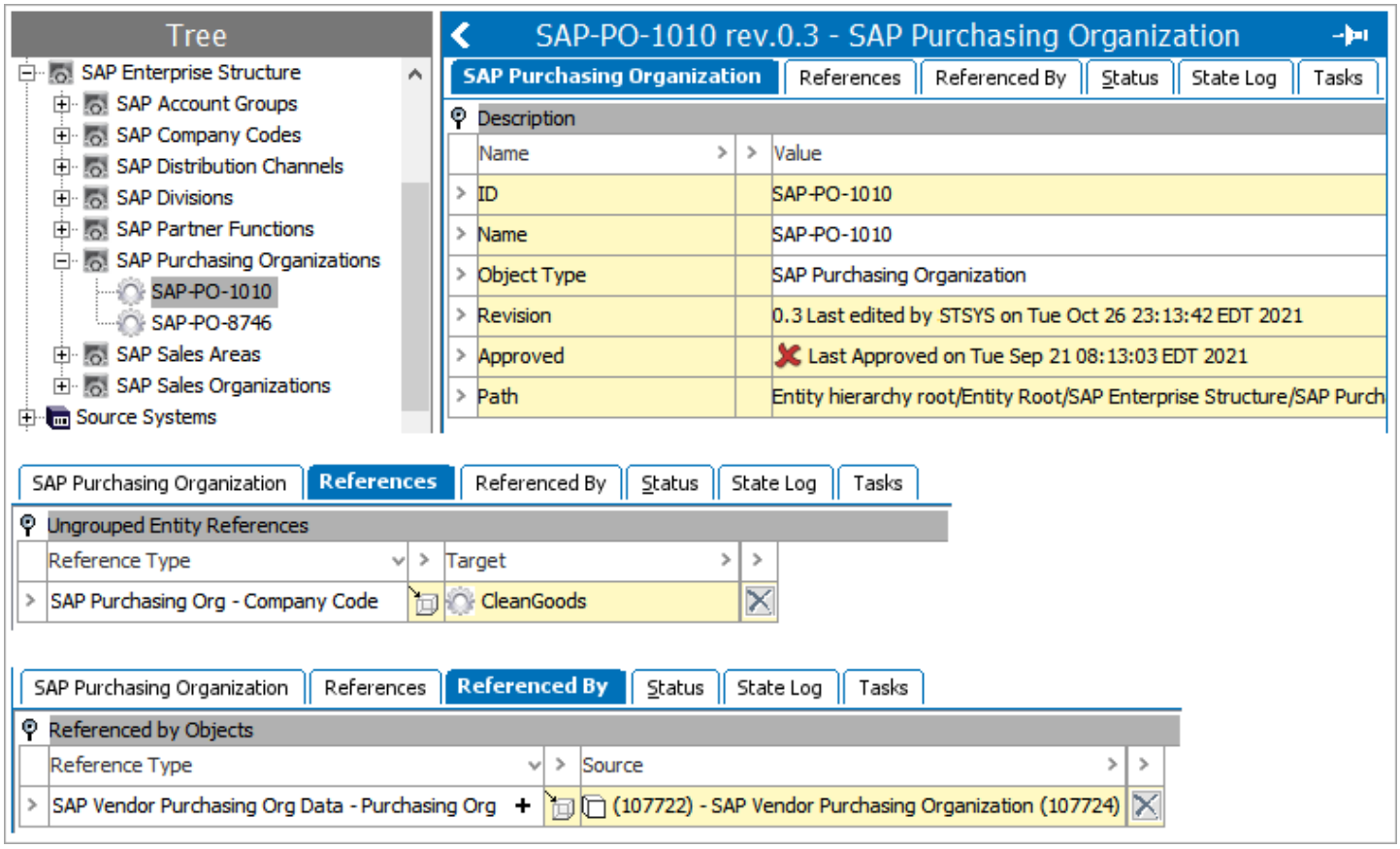

Field Sales rev.0.2 - SAP Dis
SAP Distribution Channel
References
Referenced By
Status
State Log
Tasks
Description
Name
Value
ID
SAP-DistCh-1
Name
Field Sales
Object Type
SAP Distribution Channel
Revision
0.2 Last edited by STSYS on Thu Jun 20 13:35:51 EDT 2019
Path
Entity hierarchy root/Entity Root/SAP Enterprise Structure/SAP Distribution Channels/Field Sales
Referenced by Objects
Reference Type
Source
SAP Sales Area - Distribution Channel
Canada / Charcoal
Canada / Cleaning Supplies
United States / Charcoal
United States / Cleaning Sup

Division


Tree
Entity Root
SAP Enterprise Structure
SAP Divisions
General


General re
SAP Division
References
Referenced By
Status
State Log
Tasks
Description
Name
Value
ID
SAP-Div-1
Name
General
Object Type
SAP Division
Revision
0.2 Last edited by STSYS on Thu Jun 20 13:35:51 EDT 2019
Path
Entity hierarchy root/Entity Root/SAP Enterprise Structure/SAP Divisions/General
SAP Division
References
Referenced By
Status
State Log
Tasks
Referenced by Objects
Reference Type
Source
Purpose
SAP Sales Area - Division
3100-4-1
3200-1-1
3300-1-1
3300-4-1
3400-1-1
3500-1-1
3500-4-1
3700-1-1
5100-1-1
5100-4-1
5400-1-1
5400-4-1
5500-1-1
5500-4-1
Canada / Charcoal
Canada / Cleaning Supplies
United States / Charcoal
United States / Cleaning Supplies

Purchase Organization



SAP S/4HANA Business Partner Data Structures

S/4HANA is the next evolutionary product and successor to SAP's R/3 and ECC offerings. SAP has simplified the customer and supplier modeling approach into a single, unified concept referred to as Business Partner (BP). A BP within SAP is categorized as either a person, organization, or group. This unified approach allows the introduction of roles as well as a shared address data between customer and supplier categories.

For additional information, see these topics:

- **SAP Business Partners and Enterprise Structure Definitions**
- **SAP Business Partner Data Integrity Constraints**

Background

Many companies in the CPG and manufacturing industries, for example, manage a wide array of products, brands, categories, and channels, and serve diverse customer segments across multiple countries worldwide. The end consumers of products are either individual consumers or businesses like hospitals, schools, restaurants, etc. For CPG and manufacturing companies, typically revenue is generated by selling through a network of retail chains, distributors, and master wholesalers. The company considers these parties their customers. Any CPG or manufacturing company has the following functions as part of the sales side of their business:

- Sales (including promotion planning, for example)
- Logistics
- Accounting and finance

While companies require mastering customer data for understanding and maintaining supplier data as well as for sales purposes, it is equally critical to the business to facilitate purchasing activities, logistics, and other business processes.

Companies often have the following functions as part of the procurement side of the business:

- Purchasing (including sourcing, soliciting, etc.)
- Accounting and financing
- Logistics

It is common for each of these functions to be managed locally (in each country) and to some extent, in each line of business. In many cases, this is a reasonable decentralization of the organization due to the need for regional agility in combination with relatively small optimization gain a centralized organization could have.

Example

Consider CleanGoods Manufacturer, a fictitious CPG company that specializes in cleaning products and operates across two major business units:

- Traditional household cleaning products - where the end consumer buys the products in a retail store.
- Clinical products for professionals - where the end consumers are hospitals, doctors, physicians, etc.

Business units sell different and/or the same products and these products are sold and distributed differently via various sales organizations and distribution channels. The same customer may buy products from different product lines. The two business units may also source from the same suppliers.

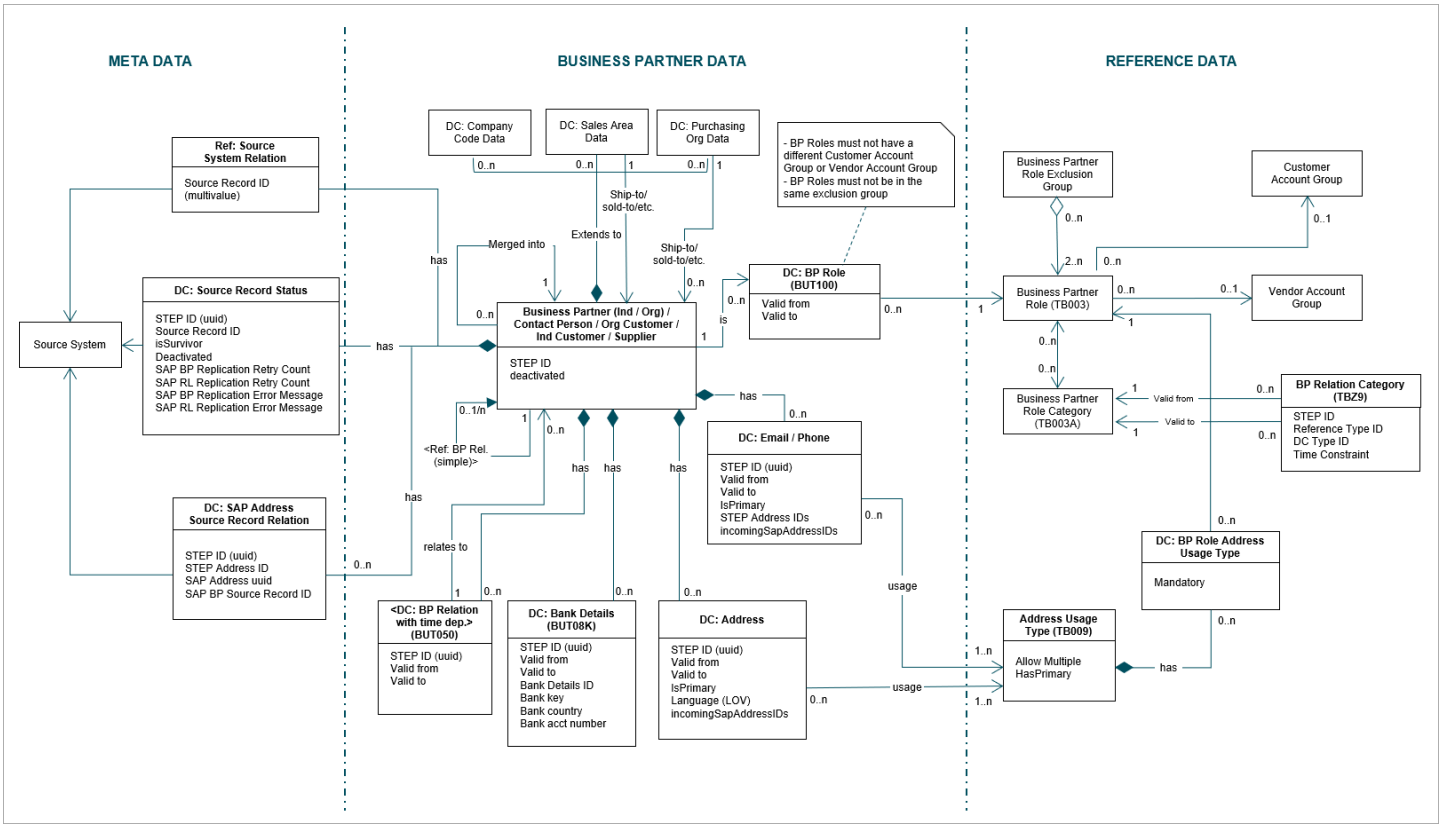
CleanGoods Manufacturer operates in many countries. Even though CleanGoods deals with international and local companies, it is common for customers to have separate business units in each country.

By mastering supplier data, CleanGoods is better equipped to support the onboarding and maintenance of suppliers across lines of business and to facilitate activities such as negotiating terms of payments, managing supplier rebates, etc.

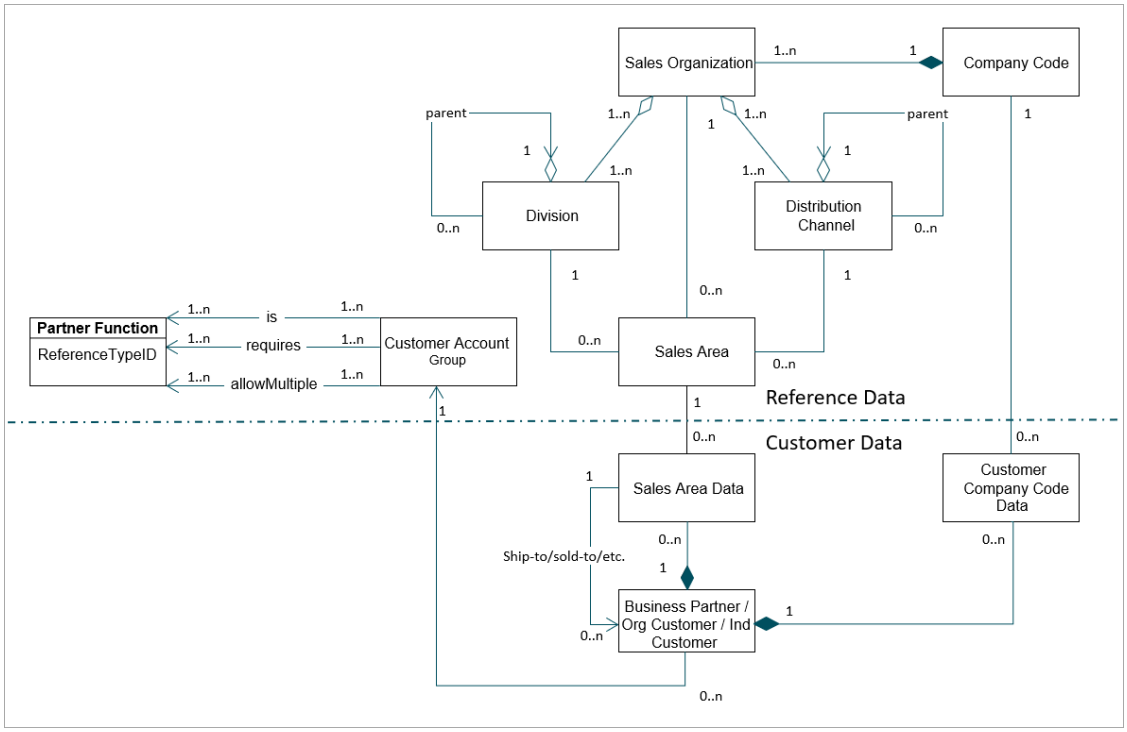
SAP Business Partners and Enterprise Structure Definitions

The following entity relationship diagram outlines a logical model for a STEP-to-SAP integration. The individual elements are defined in the sections below.

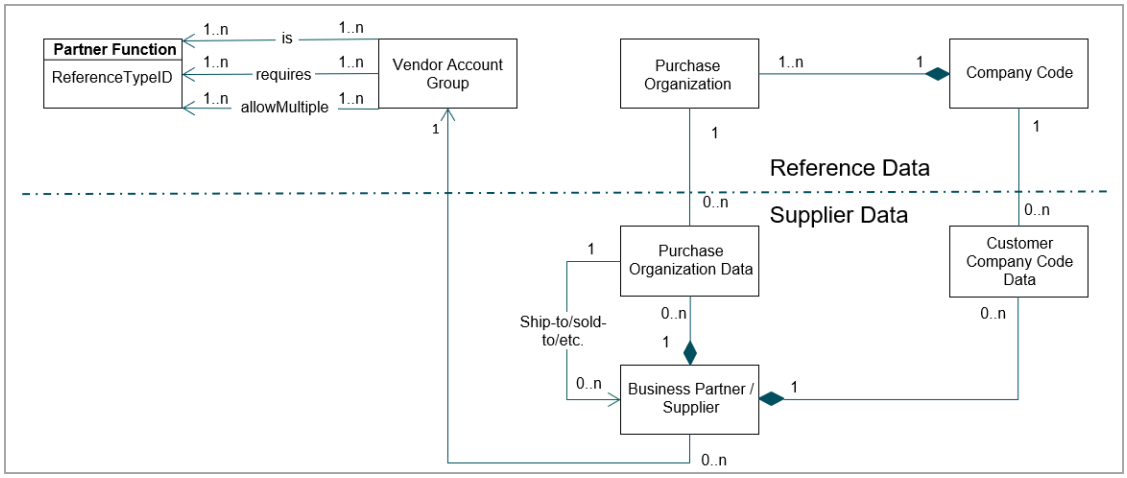
The entity relationship diagram below is best viewed in online help.



When a Business Partner (BP) is a customer, the data model extends as shown below:



When a Business Partner (BP) is a supplier, the data model extends as shown below:



Business Partner vs Customer and Supplier Object Types

Through Business Partner (BP) roles, SAP allows the same BP record to act as a customer as well as a supplier.

Deciding to implement this same construct in STEP or not must be determined from a data governance point of view instead of a system setup reduction point of view.

Answer this question to determine the most beneficial construct: Is it required to share the management of data for the same legal entity or the same person across the buying and the selling part of the organization?

- Yes - manage objects as business partners.
- No - have distinct object types for customer and supplier, which makes it significantly simpler to manage all system setup, like attribute validity, workflows, privileges, and Web UI configuration.

The initial configuration includes distinct object types for customer (individual, organization, contact person) and supplier.

Contact Persons

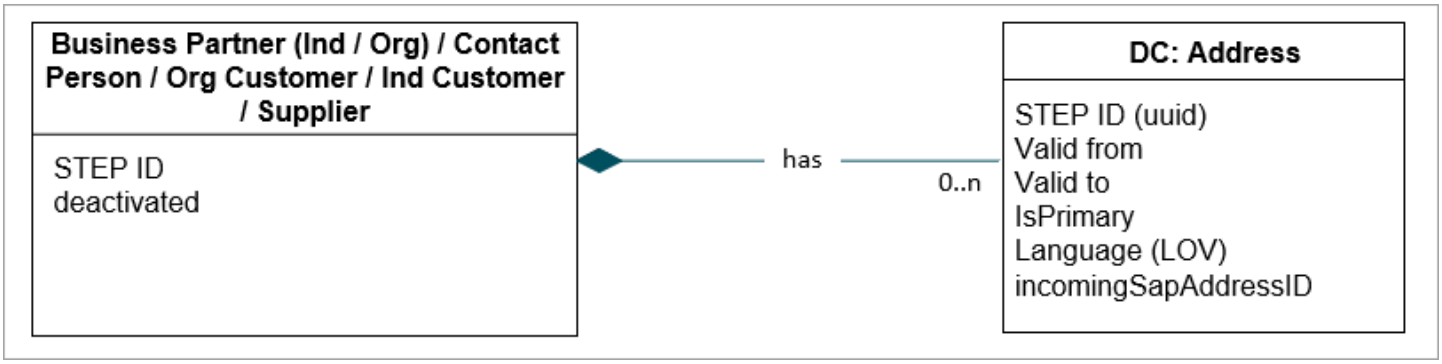
Through Business Partner Roles, SAP supports that the same BP record may act as a Customer as well as a Contact Person. The considerations for reflecting this generalization in STEP are the same as for Customer and Supplier.

Since the most common approach is that Individual Customers and Contact Persons for an organization are managed separately, the initial configuration contains Contact Person as a separate object type.

Addresses, Emails, and Phone Numbers

SAP offers a general service for managing address, email, and phone number data for many purposes, including BPs. This makes it possible for SAP to optimize the functionality around managing this data.

Since STEP takes a different approach to making general capabilities around this data available in relation to different business objects, to achieve this reuse it is not necessary to model this data as a separate object type. Addresses, emails, and phone numbers are modeled as data containers on the BP object.



Ownership of an Address

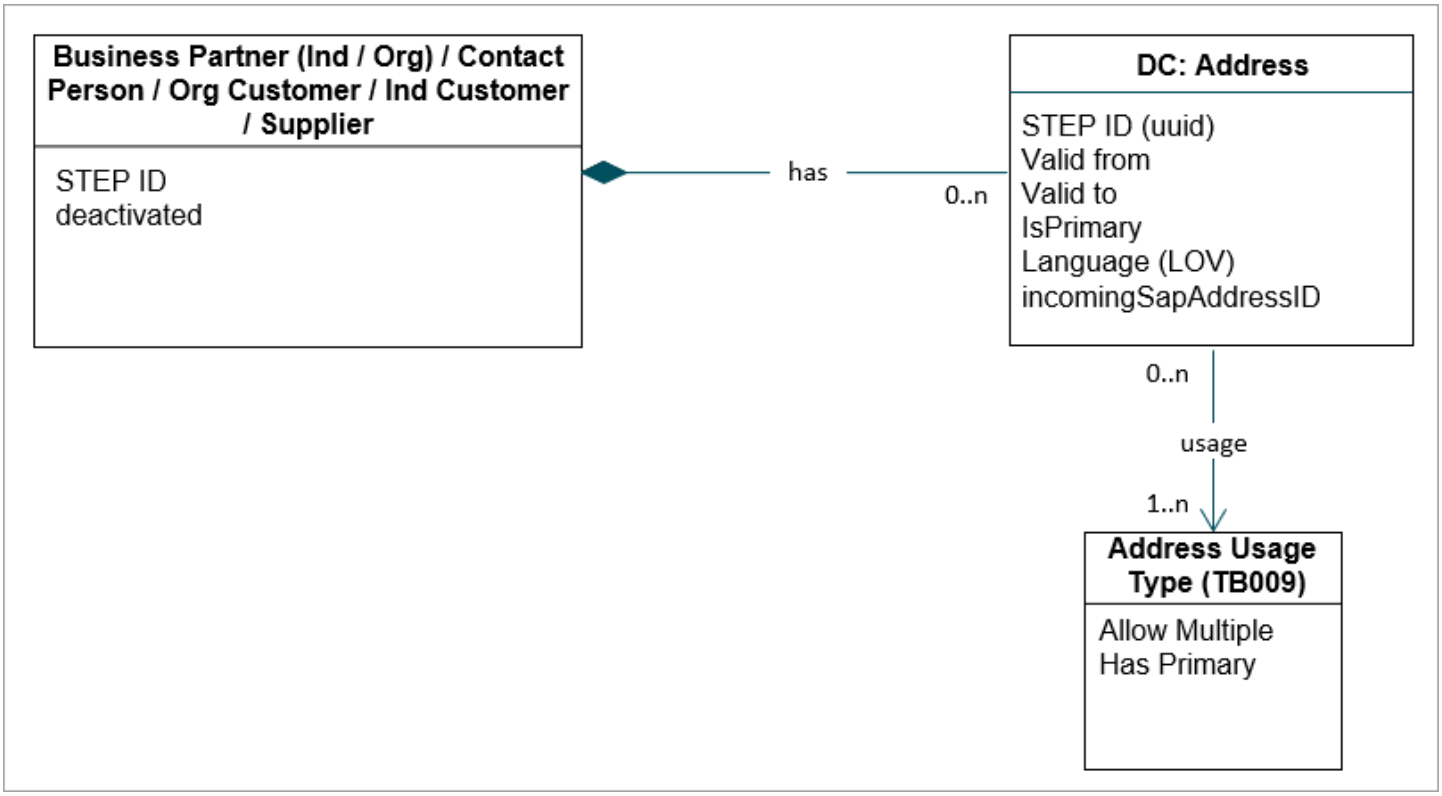
In SAP, one address belongs to exactly one BP and is managed in context of one specific BP.

It makes sense to implement the same pattern in STEP, as it is the BP (or customer, supplier, and contact person) objects that are understood by users as the owners of the lifecycle of the relevant data.

Multiple Addresses and Usages

In SAP, one BP may have multiple addresses and each address may have multiple usages. The address itself is valid from a start date to an end date. In addition, each usage of the address has a validity period that must be within the validity period of the address itself.

In STEP, this is simplified, so that each address may have a validity period, but the usages of the address inherit the validity period of the address. A BP object contains Address data containers, which are referenced (ID=SAPUsage) to separate Address Usage (TB009) entities (ID=SAPBPAddressUsage).



Multiple Versions

In SAP, one address can exist in multiple versions. A version of an address is ultimately a language of the address. For example, the same address may exist in English and Chinese (for international trade in China) or in German and Italian (demographics of northern part of Italy).

In STEP, this is handled with a simpler model where each address is assigned a language, but the relationship between the two versions of the same address is not expressed. However, it is possible to logically identify which address should be used for a specific usage in a specific language at a given point in time if that is unique. When multiple addresses of the same usage are valid at a given point in time, it is also possible to identify the set of those that are in a specific language.

Addresses in Contact Person to Organization Relations

In SAP, as contact persons are related to organizations, as BP Relations, SAP makes it possible to associate the contact person to one of many addresses owned by that organization BP.

In STEP, this relation is expressed logically to the address usage so that if a specific reference type requires that the organization has an address of a specific usage, it is validated as the relationship is established.

The limitation is that if the organization has multiple valid addresses of the same usage, a specific address cannot be identified. It is recommended to avoid such complexities, as they are difficult for data owners to understand,

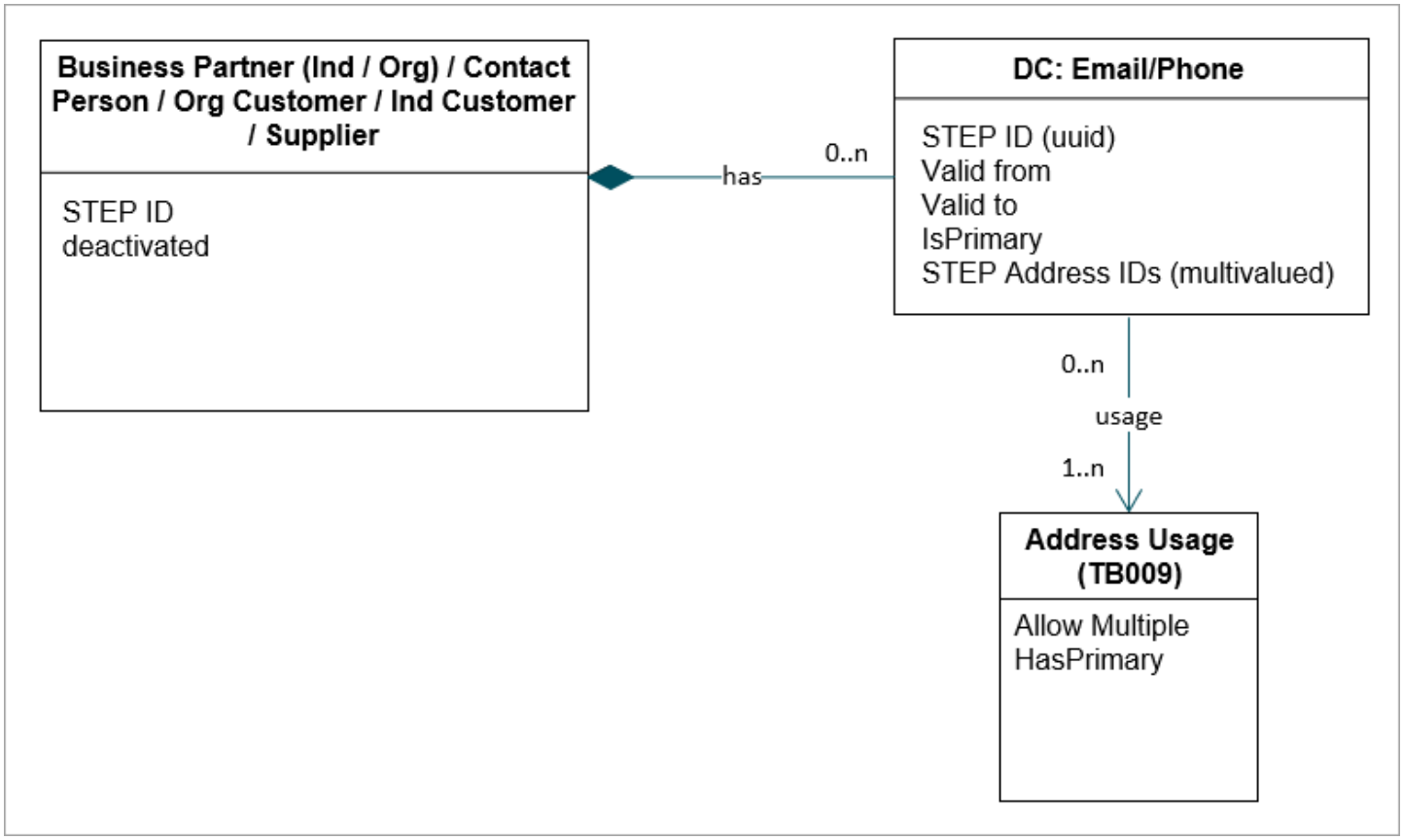
and they also severely complicate integration with other systems that do not support such complexities. Instead, it is recommended to express such complexities as multiple related BP objects.

Emails and Phone Numbers and the Approve Trigger

In SAP, emails and phone numbers belong to addresses, meaning that two addresses of the same BP may have different emails and phone numbers.

In STEP, the relationship between emails and phone numbers and addresses are through usages since that simplifies the understanding of the data. The limitation is that when there are multiple addresses of the same usage, STEP expresses that the emails and phone numbers with the same usage belong to all of those addresses. As emails and phone numbers are commonly identifiers of a person and not an organization, this is not usually a problem. When it is necessary to express such complexities of multiple addresses of the same usage, it is recommended to do this as multiple BPs, as that also defines a clearer lifecycle and governance of such data.

The BP object includes data containers for emails (ID=EmailDataContainer) and phone numbers (ID=PhoneDataContainer). Each data container contains a reference (ID= SAPUsage) to an address usage object (ID=SAPBPAddressUsage) which indicates the usages of each email or phone number.



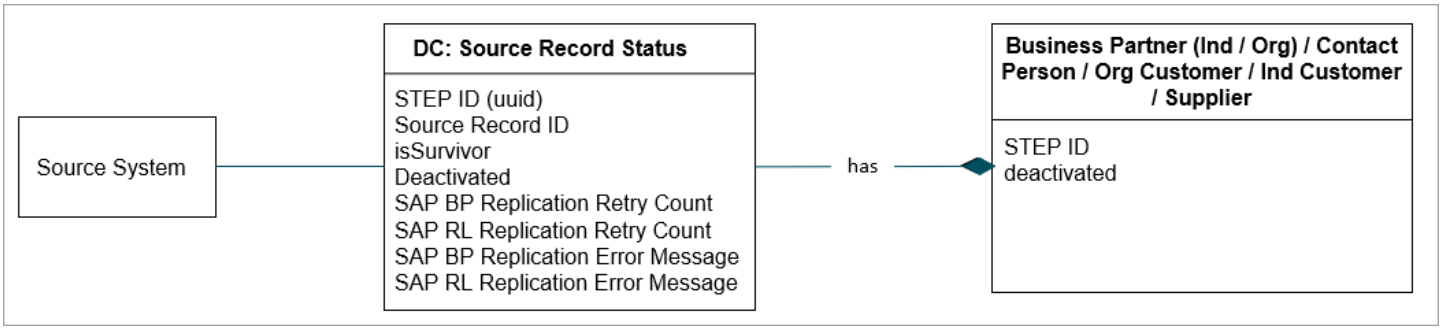
To simplify the data model transformations in the outbound BP and Relationship Link (RL) replication, the resolution of which address UUIDs of DC:Emails and DC:Phone that belong to which DC:Address, an approval trigger must resolve the logical relationships through Usages in to UUIDs and add those to the SAP Address IDs

attribute on the DC:Emails and DC:Phone objects and afterward also maintain the SAP Address Source Record Relations.

For details on this approve trigger, see the **Email and Phone Number Approve Trigger** section in the **SAP Publishing From STEP** topic.

Source Record Status

Source Record Status is a data container (ID= SAPBPSourceRecordStatusData).



- **Source Record ID** - As inbound integrations or merges of existing golden records add source records to the source system relation, survivorship rules must also add those source record IDs as source record status objects, so that 'Is Survivor' and 'Replication Error Messages' can be managed.

SAP Address Source Record Relations

In SAP, address objects are assigned a UUID so they can be uniquely identified in integrations and so that they can be referenced in certain cases.

- In one SAP system, two addresses must not have the same UUID.
- In SAP, one address belongs to exactly one Business Partner.

In STEP, as multiple source records merge into one golden record or as one golden record is exported to multiple SAP systems, one address in STEP may represent multiple addresses in one or multiple SAP systems.

To identify existing addresses for update and reference, it is necessary to keep track of which UUID the address has per Source Record ID.

Inbound integrations and merges of existing golden records can add address, email, and phone objects and can associate incoming objects to existing objects. To ensure that future exports can consistently make updates to the right address objects, survivorship rules must add the UUID of the incoming address to the DC: SAP Address Source Record Relation objects.

For more information on survivorship rules, see the **Survivorship of Addresses, Emails, and Phones** section of the **SAP Survivorship Rules** topic.

Business Partner Role

BP roles are used to classify BP entities in their specific role(s). The roles that are assigned indicate the business functions and transactions allowed for a BP.

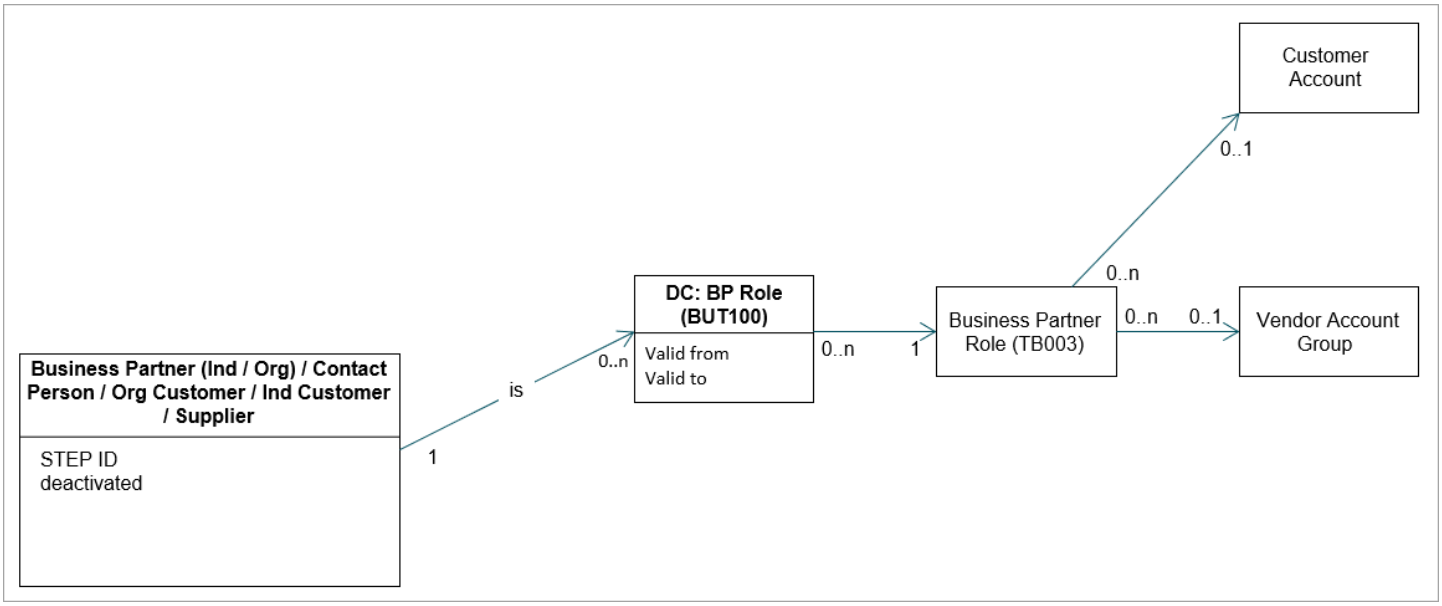
A BP may be assigned more than one BP role. This indicates that the BP may be involved in different business transactions with different roles.

Examples of BP roles include:

- Sold-to Party
- Ship-to Party
- Contact Person
- Employee
- Vendor

In STEP, the BP entity contains a BP Role data container (ID=SAPBPRoleData) that references (ID=SAPBPRoleDataRole) a BP Role object (ID=SAPBPRole).

The Customer Account Groups and the Vendor Account Group of a Business Partner can be derived from its Business Partner Roles. One Business Partner can have 0..1 Customer Account Group and 0..1 Vendor Account Group. For more information, see the **Business Partner Role Constraints** section of the **SAP Business Partner Data Integrity Constraints** topic.

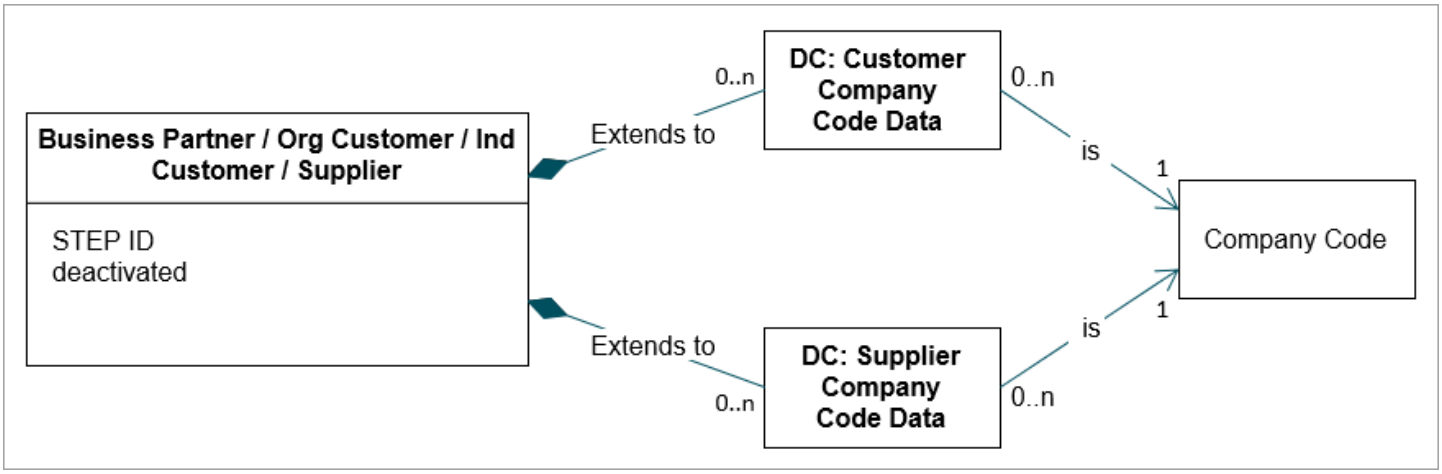


Company Code Data

In SAP, the company code is the central organizational unit of external accounting within the SAP System. A BP may be extended to one or more Company Codes.

In STEP, company codes are separate entity objects that are associated to a business partner using a sales organization referenced by a Customer Company Code Data Container (ID= SAPCustomerCompanyCodeData) for customers or using a purchasing organization referenced by a Supplier Company Code Data Container (ID= SAPSupplierCompanyCodeData).

When a BP is both a customer and a supplier, both data container types are valid and as an indirect / direct reference to a company code. In this case, some attributes may overlap between the two data container types.

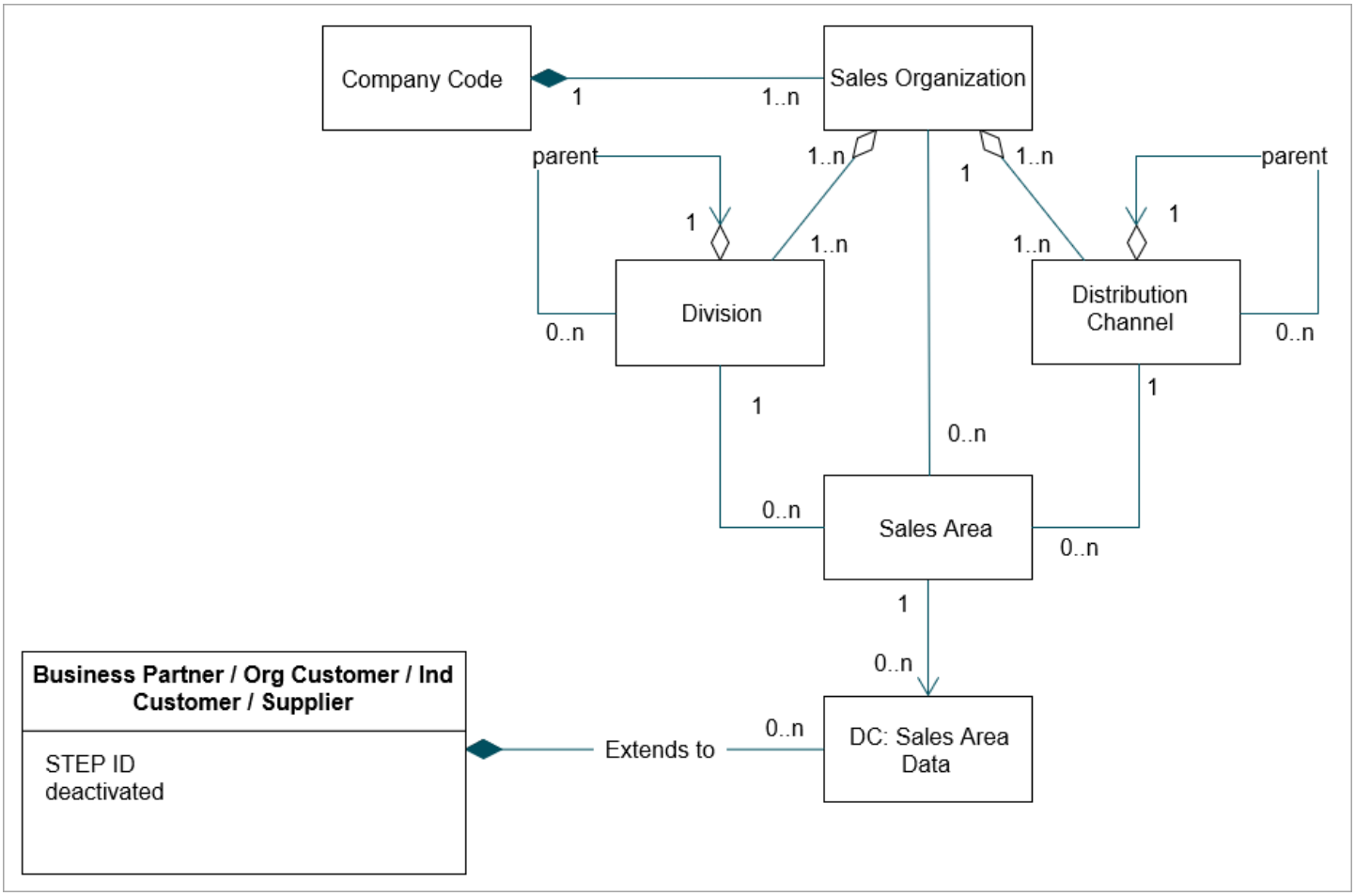


Sales Area Data

In SAP, a Sales Area defines the Sales Organization and Distribution Channel that a Division uses to sell products. Therefore, Sales Areas are comprised of:

- Distribution channel: Channel through which materials or services reach customers, for example, direct sales, retail, or wholesale. Distribution channels may be assigned to one or more sales organizations.
- Division: Product groups can be defined for a wide-ranging spectrum of products. Customer-specific agreements (for example, partial deliveries, pricing, and terms of payment) are allowed for each division. Within a division you can carry out statistical analysis or set up separate marketing.
- Sales organization: The organizational unit of a business that negotiates sales terms and distributes products. Sales organizations may be referenced to a company code.

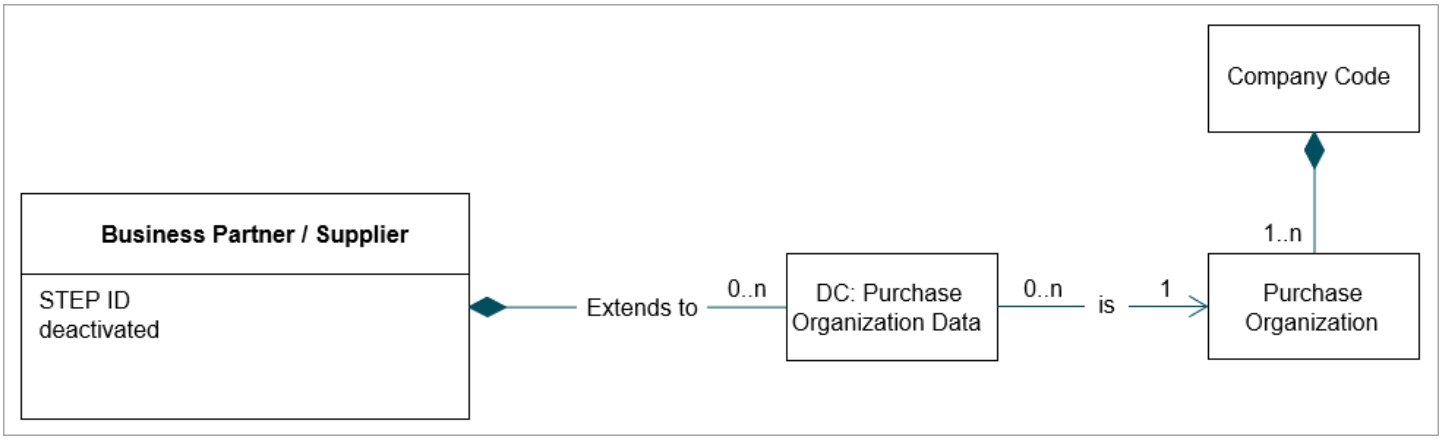
In STEP, Sales Areas are referenced from the business partner object using a reference (ID= SAPCustomerSalesAreaDataSalesArea) from the Sales Area data container (ID= SAPCustomerSalesAreaData).



Purchase Organization Data

In SAP, a Purchase Organization (also called a 'purchasing organization') is responsible for the procurement of materials or services with negotiation terms and conditions from vendors. Purchase organizations are also responsible for pricing agreements with external vendors and can only be assigned to one Company Code.

In STEP, purchasing organizations (ID=SAPPurchasingOrganization) are referenced from a BP using a reference (ID= SAPSupplierPurchasingOrgDataPurchasingOrg) from the Purchase Organization Data Container (ID= SAPPurchaseOrganizationData).



Partner Functions

In SAP, Partner Functions are designations that describe the rights and responsibilities of each individual and/or organization BP with whom you conduct business transactions.

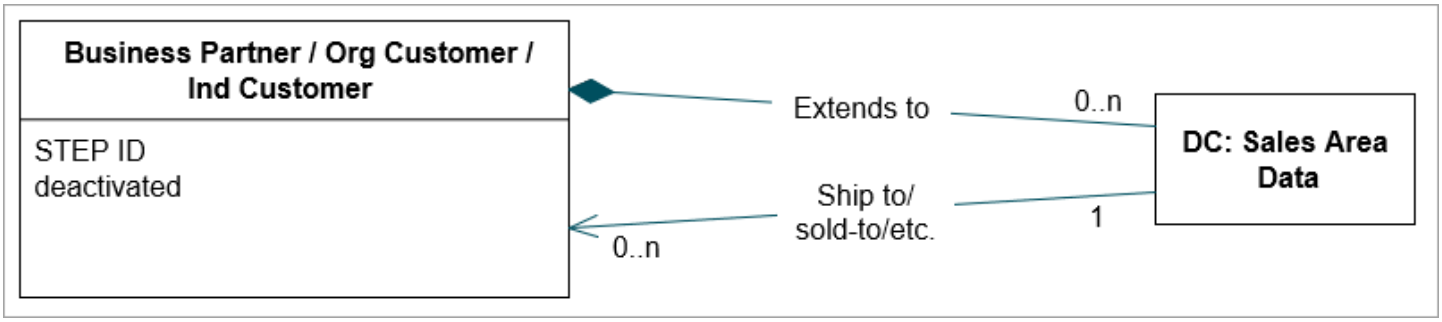
In STEP, partner functions are modeled as references between Sales Area- and Purchase Organization Data Containers to BPs.

Each Partner Function type is a dedicated reference type.

Examples of Partner Functions for Customer BPs include:

- Sold-to Party
- Ship-to Party
- Bill-to Party
- Payer

Partner functions for customer BPs are referenced from the Sales Area Data Container to a BP object.

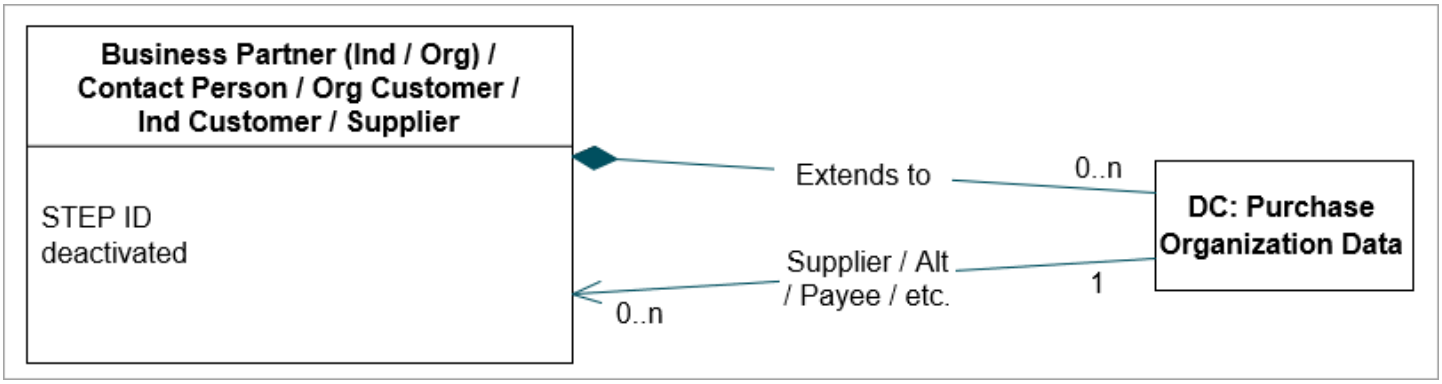


Examples of Partner functions for supplier BPs include:

- Ordering address
- Invoice represented by

- Goods supplier
- Alternative Payee

Partner Functions for supplier BPs are referenced from the Purchase Organization Data Container to a BP object.

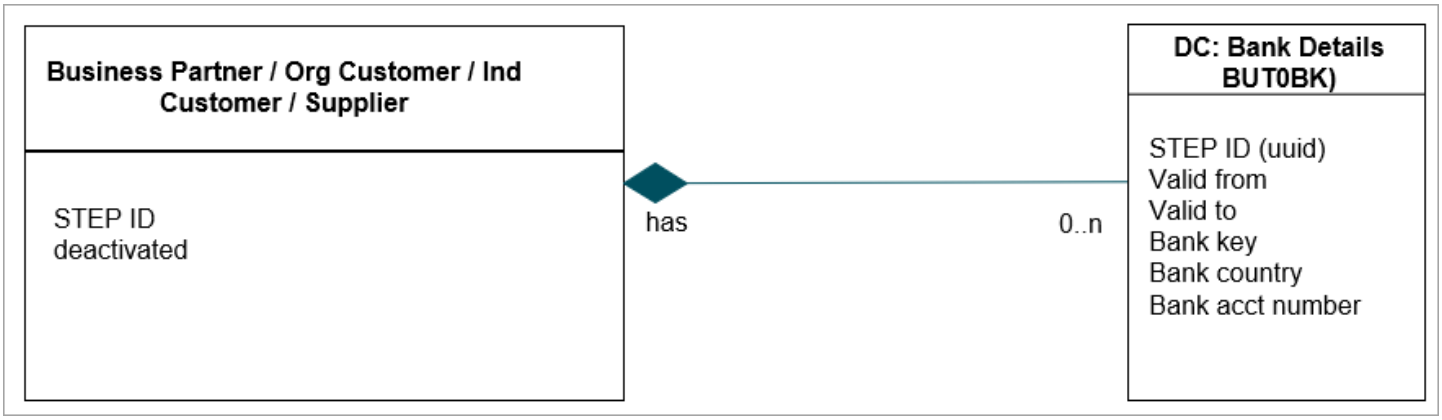


A BP may fulfill a certain Partner Function to itself. For example, a BP placing an order can be the same BP receiving the goods. Each partner function reference type may be referenced from a data container back to the same BP that owns the data container.

Bank Data

In SAP, for financial transactional purposes, BPs must maintain banking information. The bank account of a BP can be identified using country key of the bank, bank key, and the account number.

In STEP, bank details are modeled as a data container (ID= SAPBankDetailsData) on the BP object.



Business Partner Relationships

In SAP, BP Relationships are all managed through the same data model (SAP table BUT050, BUT051, etc.), regardless of the complexity of their cardinalities, constraints, etc. The complexity of cardinalities, constraints, etc., of a particular kind of BP relations are controlled through BP Relationship Categories (SAP Table TBZ9).

In STEP, each BP Relationship Category should be managed as its own reference type so that STEP's built-in mechanism for managing cardinalities, constraints, privileges, UI components, etc., can be used.

For BP Relationships where time dependency must be managed, each BP relationship should be modeled as a separate data container type, so that valid-from and valid-to sales can be managed.

Examples of common BP Relationship Types are:

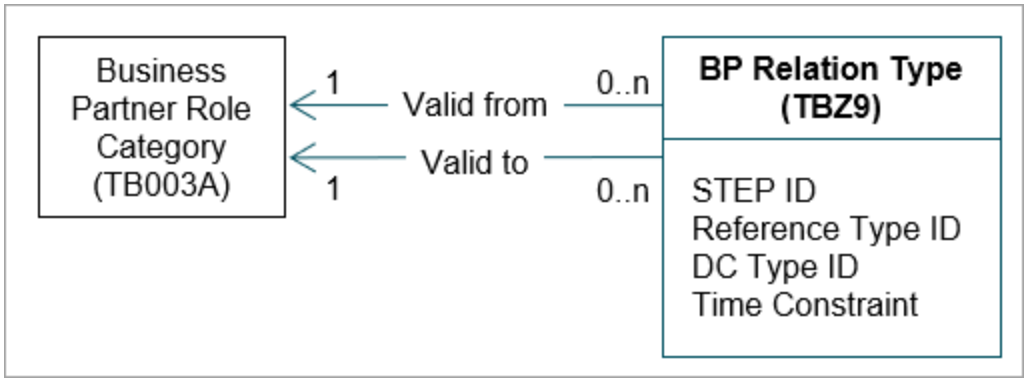
- FB8P002 - Has the Invoicing Patty
- FBIJROOI - Has Contact Person
- FBIJR002 - Has Activity Partner
- FBIJROIO - Has the Employee
- FBIJROII - Has the Employee Responsible
- FBIJR013 - Is Replaced By
- FBUR020 - Has Department
- FBUR025 - Has Service Provider
- FBIJRCOI - Is Shareholder Of
- FCRMH04 - Has the Bill-To Party
- FEWMOOI - Has the Dock Appointment Scheduling Planner
- FFS0030 - Borrower Entity Member
- FFSBOOI - Is A Shareholder Of

Business Partner Relation Category

In SAP, in some cases, the validity of Reference Source and Reference Target is more specific than what constitutes the Valid Object Types of a Reference Type in STEP. This may depend on the strategy chosen for Object Types of Business Partner Objects in STEP. See the **Business Partner vs Customer and Supplier Object Types** section above.

To enable the validation of such constraints, BP Relationship Category Entities are created with the relevant configuration.

See the **Business Partner Relationship Constraints** section in the **SAP Business Partner Data Integrity Constraints** topic.


Business Partner Role Category (TB003A)
1
Valid from
0..n
Valid to
1
0..n
BP Relation Type (TBZ9)
STEP ID
Reference Type ID
DC Type ID
Time Constraint

SAP Business Partner Data Integrity Constraints

A variety of constraints exist and are illustrated in the logical model diagram in the **SAP Business Partners and Enterprise Structure Definitions** topic. The constraint types are outlined in the following sections.

Uniqueness Constraints

The following are uniqueness constraints as defined:

- Unique by ID
  - Business Partner Role
  - Customer Account Group
  - Vendor Account Group
  - Company Code
  - Sales Organization
  - Distribution Channel
  - Purchase Organization
  - Business Partner
  - Business Partner Role Exclusion Group
  - Business Partner Role Category
  - Business Partner Relation Type
  - Business Partner Grouping
  - Address Usage Type
- Sales Area is unique by the combination of sales organization, distribution channel and division.
- Data Container Sales Area Data is unique within each business partner by sales area.
- Data Container Customer- and Vendor Company Code Data are unique within each business partner by company code.
- Data Container Purchase Organization Data is unique within each business partner by purchase organization.
- Data Container BP Role is unique within each business partner by business partner role.
- Data Container Source Record Status is unique within each business partner by source system and source record ID.
- Data Container SAP Address Source Record relation is unique within each business partner by STEP address ID and SAP address UUID.

Business Partner Role Constraints

One BP can reference multiple BP Roles. The Roles of a BP determine:

- constraints for BP Relations. See the **Business Partner Relationships** section of the **SAP Business Partners and Enterprise Structure Definitions** topic.
- the Customer- and/or Vendor Account Group of the BP.

The Customer- and Vendor Account Group of the BP determines constraints for Partner Function relations in Sales Area- and Purchasing Organization Data of the Business Partner. See the **Partner Functions** section of the **SAP Business Partners and Enterprise Structure Definitions** topic.

A BP references a BP Role within a validity period.

At a given point in time, the following constraints apply to the Roles of a Business Partner:

- The Roles of a BP must not have different Customer Account Group or Vendor Account Group.
- The Roles of a BP must not be in the same BP Role Exclusion Group.

Business Partner Relationship Constraints

To the extent possible, it is recommended to use the basic configuration of reference types and data container types to control the constraints of Business Partner (BP) Relationships.

However, the following constraints may require additional logic, which can be implemented in business rules.

Note: When additional constraints are required beyond those mentioned below, manage them by extending the data model for BP Relationship Categories and using STEP business rules to govern the constraints.

BP Role Validity on BP Relationships

When the setup of object types for BP entities (see the **Business Partner vs Customer and Supplier Object Types** section of the **SAP Business Partners and Enterprise Structure Definitions** topic) results in one BP entity having multiple BP Roles of multiple BP Role Categories, the object type of a BP entity is not sufficient to determine if it is a valid source and/or target of a particular BP Relationship reference type.

The BP Role Category Entity and its valid from / valid to references to BP Relationship Categories should be used to express this additional constraint and a STEP business rule should be used to govern this constraint.

Validity Periods

In SAP, a BP Relationship may have a validity period.

These are the validity period constraints:

- 0 / no value: No validity period constraint
- 1: Record must have no gaps, no overlap

- 2: Record may have gaps but no overlap
- 4: Record exists at least once from minimum to maximum

This should be managed in STEP as a Time Constraint Attribute (attribute ID=SAP-TIMDP) on the BP Relationship Category Entity and STEP business rules should be used to govern the constraint.

Address, Email, and Phone Usage Constraints

Address Usage Types define which usages an address, email, and/or phone number that a BP, at a given point in time, in a specified language:

- can have multiples of
- has a primary of

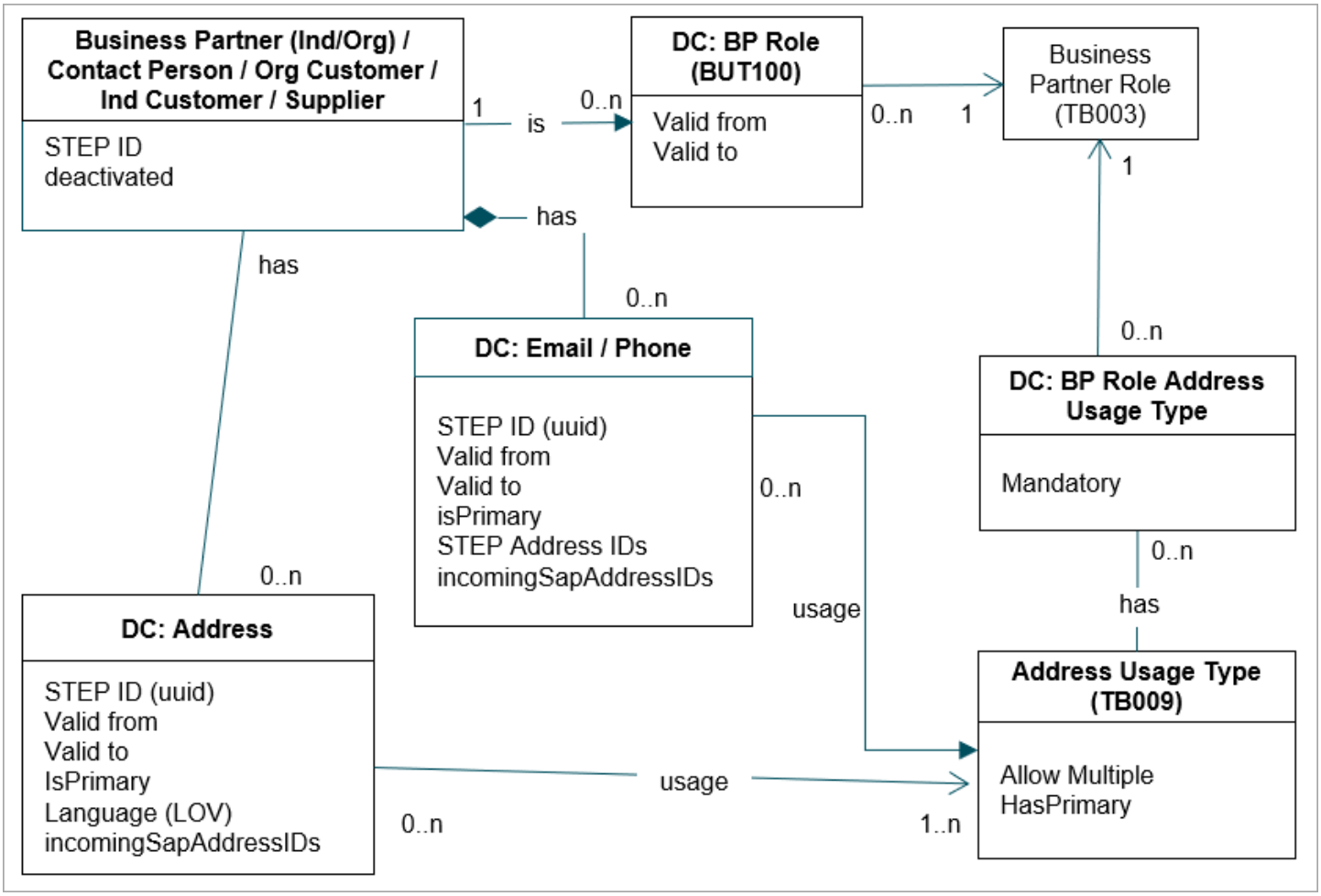
This is expressed by the Address, Email and Phone Data Containers having references to Address Usage Types. One address, email, or phone may have multiple usages.

The Roles of a BP determines which Address Usages it:

- can have now and in the future
- must have now and in the future

This is expressed by Address Usage Type Entities being related to BP Roles.

Use STEP business rules to govern such constraints.



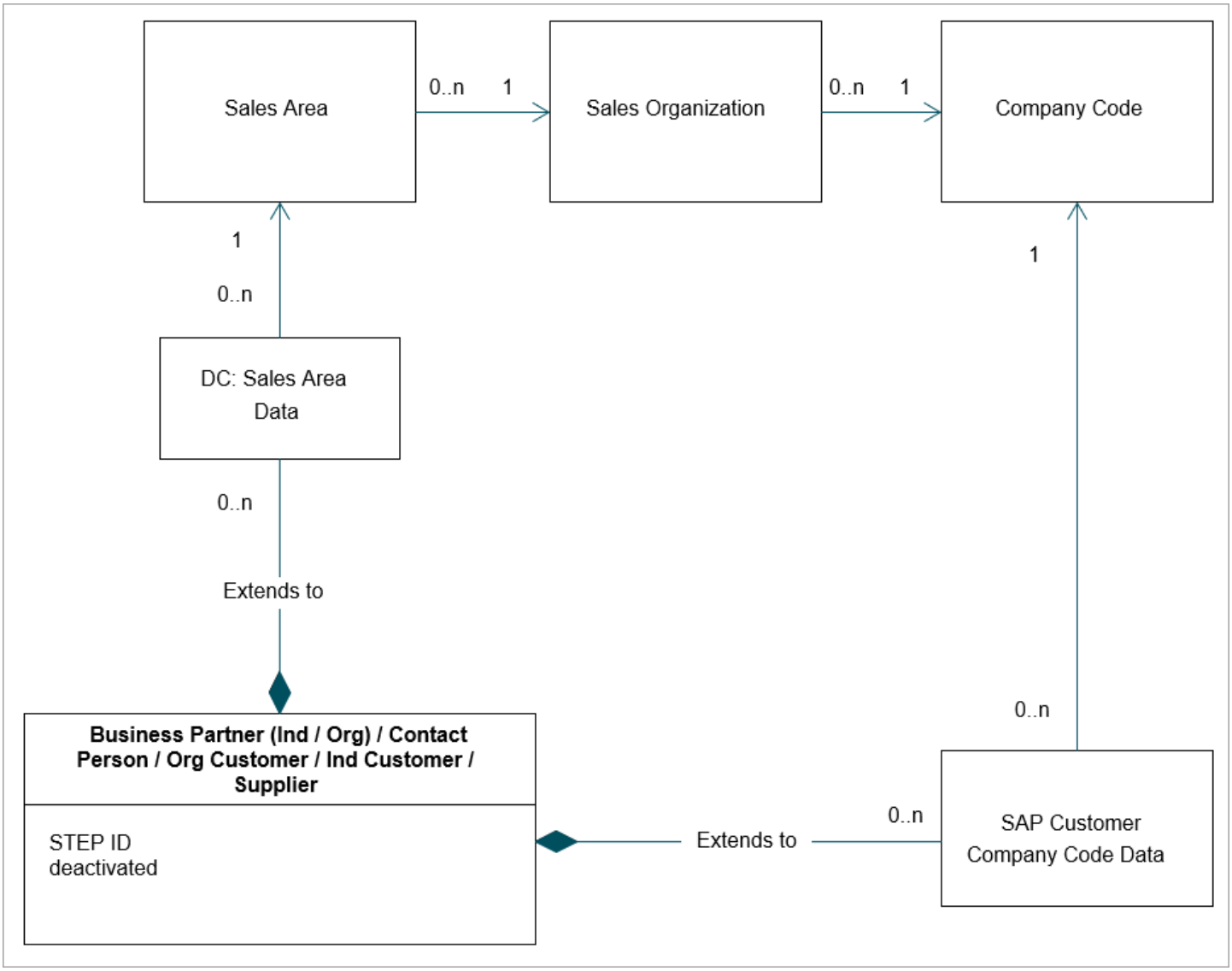
Sales Area Data Constraints

In SAP, a Sales Area defines the Sales Organization and Distribution Channel that a Division uses to sell products. Therefore, Sales Areas are comprised of distribution channel, division, and sales organization.

In STEP, Sales Areas are referenced from the BP object by way of a reference (ID= SAPCustomerSalesAreaDataSalesArea) from the Sales Area data container (ID= SAPCustomerSalesAreaData).

A BP may contain zero-to-many Sales Area data containers (ID=SAPCustomerSalesAreaData).

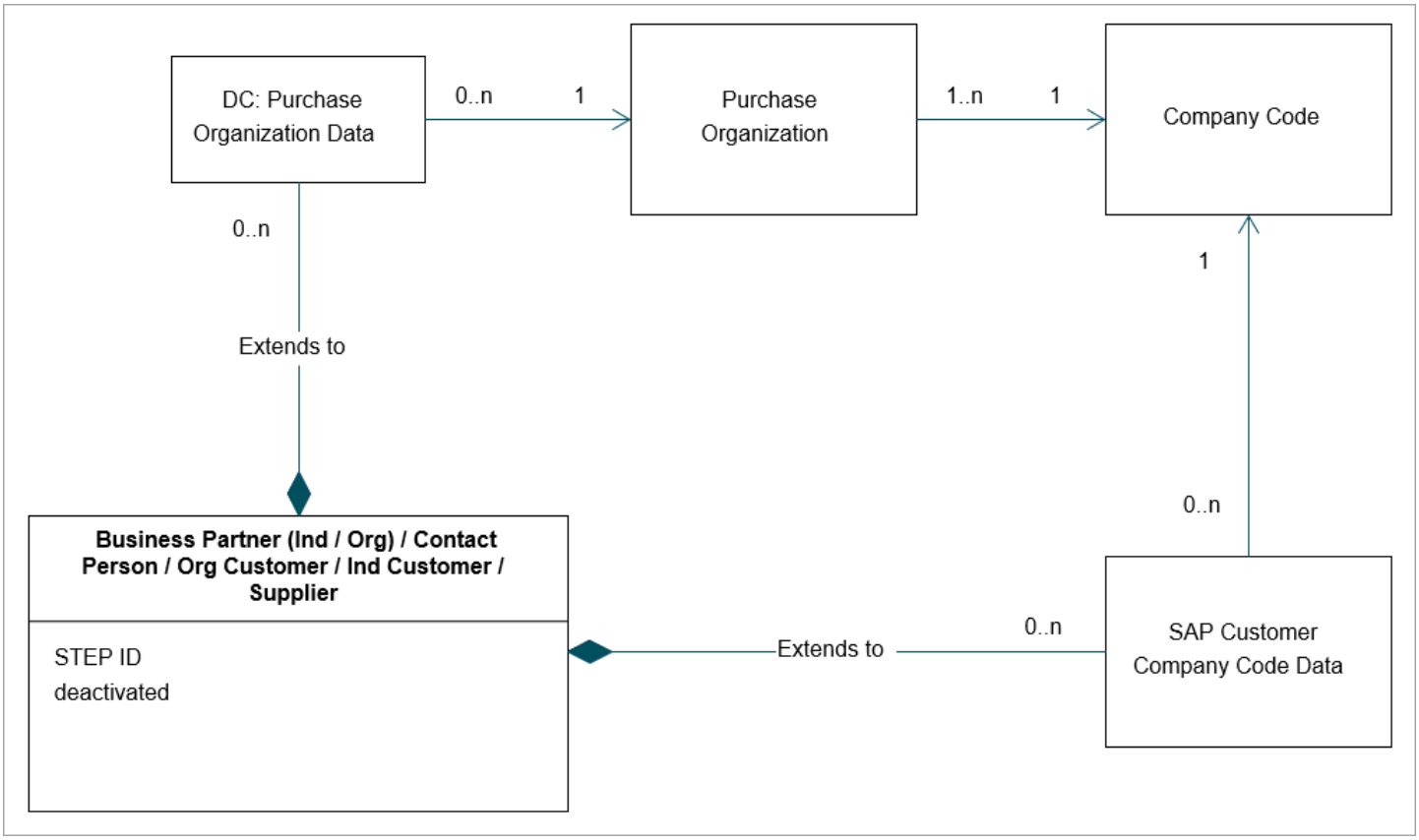
A BP may only have Sales Area Data if the corresponding Company Code Data is present such that the Sales Area must be associated to a Sales Organization that is associated to a Company Code for which the BP has Company Code Data.



Purchase Organization Data Constraints

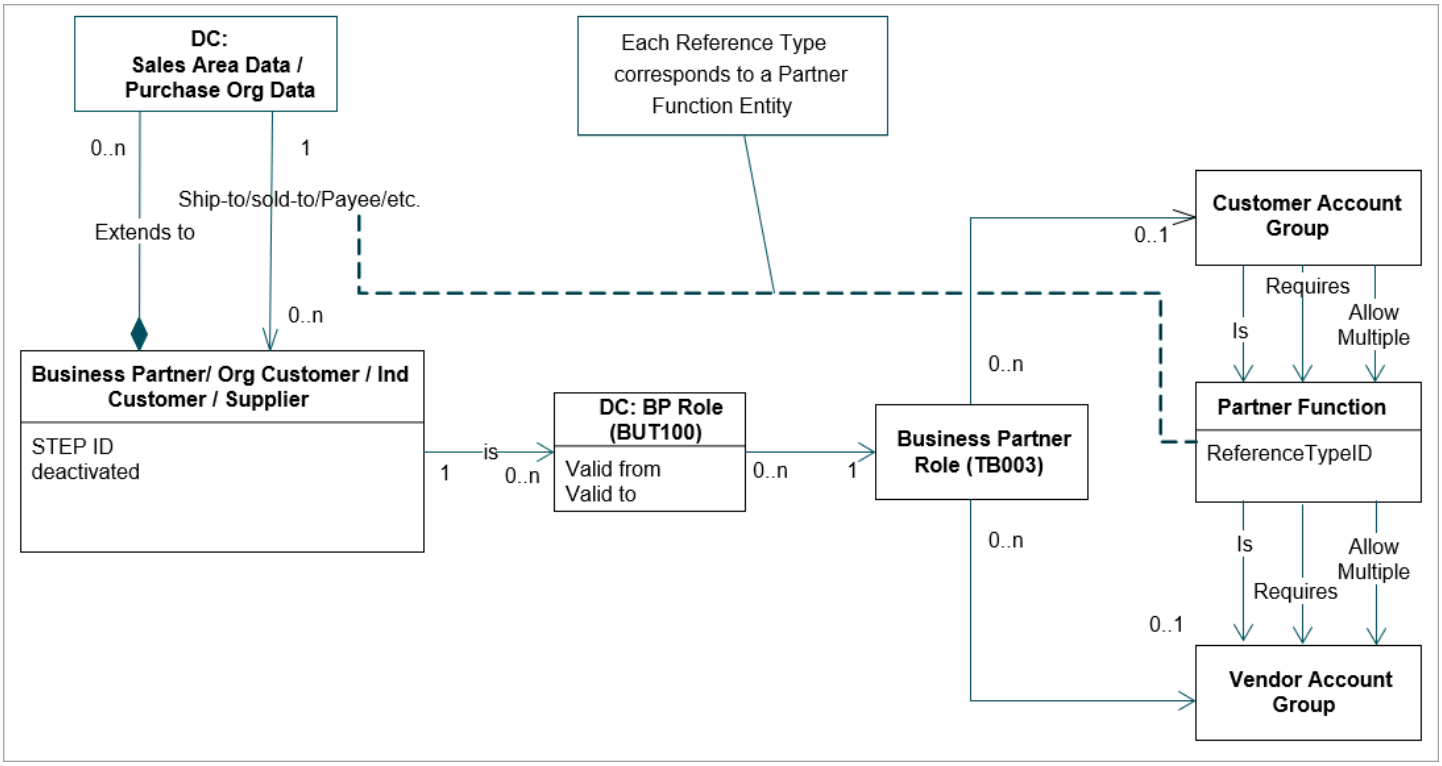
A BP may contain zero-to-many Purchasing Organization Data Containers (ID=SAPPurchasingOrganizationData).

A BP may only have Purchase Organization Data if the corresponding Company Code Data is present such that the Purchase Organization (also called a 'purchasing organization') must be associated to a Company Code for which the Business Partner has Company Code Data.



Partner Functions Constraints

Through its BP Role references, a Customer has exactly one Customer Account Group and a Supplier has exactly one Vendor Account Group. An account group has one-to-many valid partner functions. Not all customers and vendors can act in all partner functions.



Since a BP has a given account group, it must also have one or multiple Partner Function relations.

- **Valid Partner Functions** - Not all customers or suppliers can act in all partner functions. The Account Group defines one-to-many valid Partner Function relations for a Customer / Supplier.
- **Mandatory Partner Functions** - For each Sales Area Data or Purchasing Organization Data, the Account Group defines the required Partner Function relations for a Customer / Supplier.
- **Allow Multiple of a Partner Function** - For each Sales Area Data or Purchasing Organization Data, the Account Group defines which Partner Function relations allow multiple Customers / Suppliers.

Common Distribution Channel and Division Constraints

To avoid redundant sales area master data across distribution channels and divisions, distribution channels and division may have a common 'parent,' which is logically resolved into what may be termed as 'common sales areas.' This concept means that sales areas associated with a particular distribution channel and division may have a common distribution channel and common division in another sales area within the same sales organization. This hierarchy is nested.

- If a customer is active in a sales area that is common for other sales areas, this customer cannot have master data in the other sales areas.
- If a customer is active in a sales area, a user is not allowed to activate that customer in the sales area's common sales areas.

Entity Super Types

STEP uses four types of object super types to model data: products, classifications, assets, and entities. Entities are the suggested object for modeling customer, suppliers, contacts, and household data. Some benefits of entity object types include:

- May be configured as globally revisable, which requires no approvals or workspace revisions, and instantaneous updates.
- References can be configured as parent-child relationships.
- Support for Matching and Merging.
- Support for Dun & Bradstreet integration.
- Support for Experian Email Validation integration.
- Support for Policy Data Monitoring

Some limitations on using entities are that they do not support inheritance, translations, specification attributes, and specification data container types.

For more information on entities in STEP, see the **Entity Maintenance** topic in the **Getting Started / User Guide** documentation.

Domain Folders

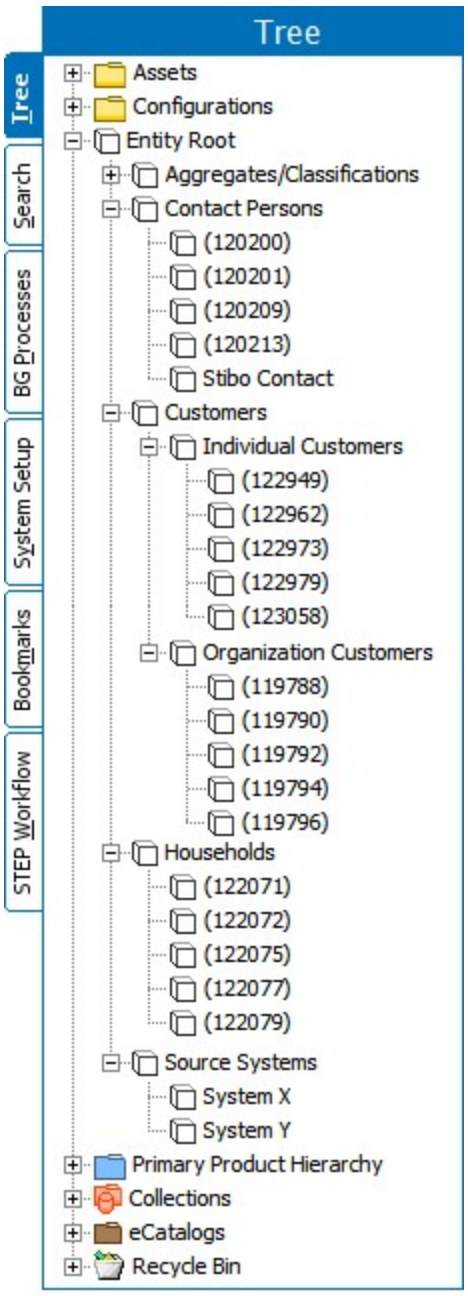
In STEP, any Entity must have a parent Entity. In an MDM solution, this relationship does not provide any semantic meaning. Therefore, very little effort should be put in to organizing parent/child relationships.

The recommended approach is to create 'Domain Folders.'

For traceability in the Workbench tree and when searching for records, it is recommended to split records governed by different matching algorithms into separate domain folders.

Source System Entities

Place all source systems into one folder as there are often very few to consider.



Dimensions and Contexts

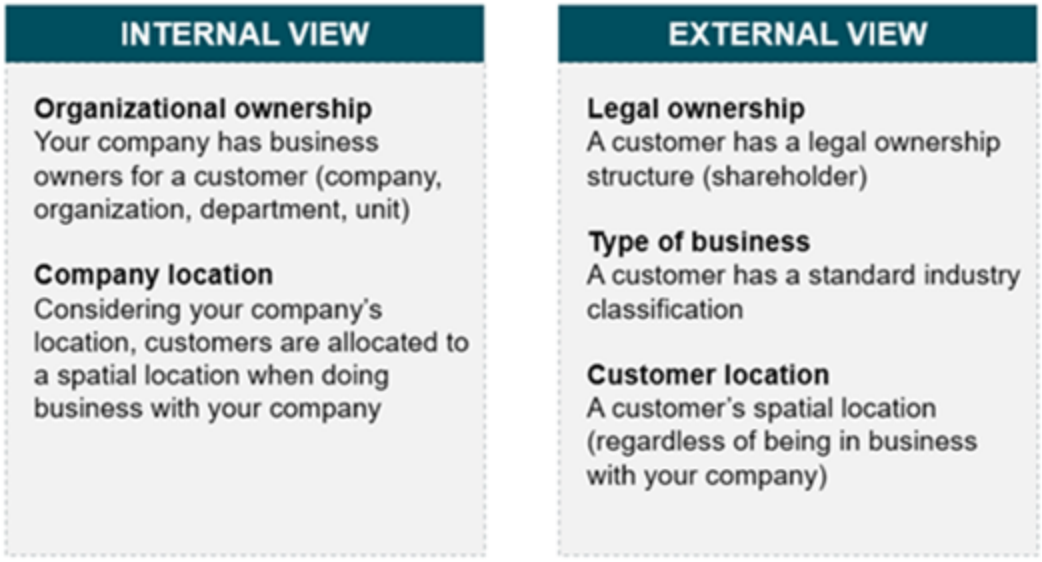
Dimension specific attribute values and references should not be used in an MDM solution. However, a language dimension can be used for the purpose of translating data definitions as Attribute names, Reference Type names, Object Type names, and Data Container Type names in Web UI and workbench. This also means that the use of multiple Contexts only allows for translation of data definitions in the UI.

Hierarchies

The purpose for constructing and managing entity hierarchies within MDM is to establish and model relationships from an entity to other entities. This provides a view of entity relationships at a particular point in time. There are many reasons for this, some of which are to drive operational activities for procurement, sales, marketing, and financial reporting in applications external to MDM.

Customer & Supplier Hierarchies

The purpose for a hierarchy in MDM is for proper grouping of customers or suppliers. The avenues for application include an internal view and an external view.



With internal view, organizational ownership shows business owners as the customers and allocate customers to actual locations when doing businesses with your company. This affords a view from your own company's perspective as well as the ability to group customers according to your internal structure.

With external view, the customer has a legal ownership structure, works as a standard industry classification, and it shows the customer location, without regard to the company's location. This affords a view on customers from their company's perspective.

Constructing Hierarchies

There are three primary methods of constructing a customer hierarchy in STEP:

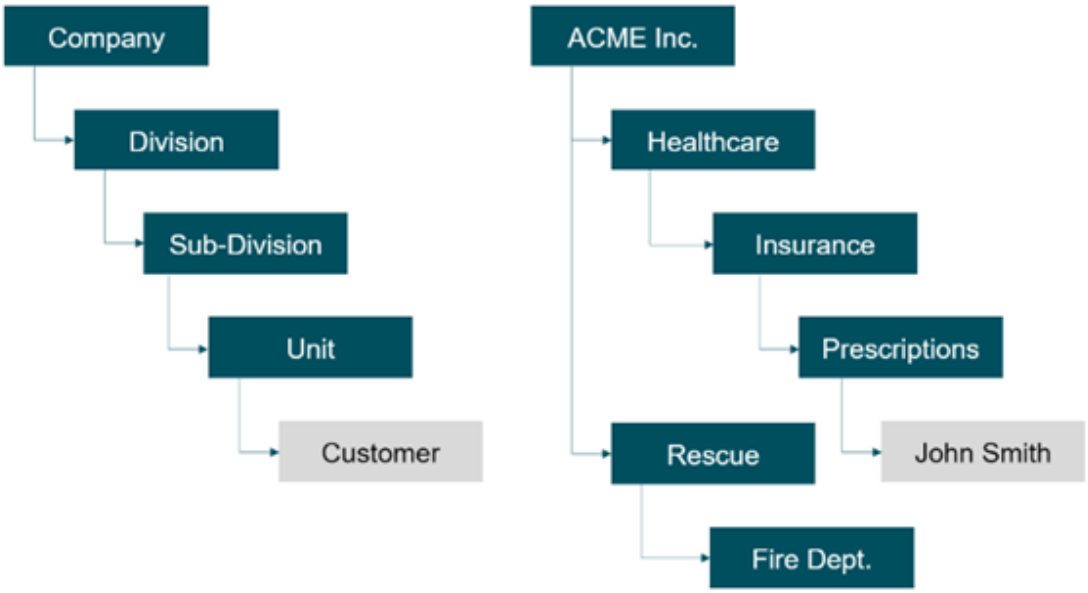
- **Manually** - This is the approach delivered with the implementations package configurations.
- **Import** - Relevant reference types and parent ID's can be loaded via an import to construct hierarchies based on external data.
- **Business logic** - D&B could potentially be leveraged to construct hierarchies based on linking of DUNS number and parent DUNS.

Hierarchy Benchmark

When considering recommended practices in constructing customer hierarchies, customers must pass the 'is grouped by' test. This is a general analysis of grouping based on common characteristics. Some examples of characteristic groupings are as described in the following sections.

Organizational Characteristics

Organizational hierarchies may be characterized by customers that are owned by the same organizational unit, such as with a Financial or Sales organization. Organizational hierarchies are typically used by the business to organize customers according to the organizational setup of their company (division, sales organization etc.) and may also be used in sales transactions, revenue reporting, and/or commissioning. The same customer may also be referenced in different locations within this hierarchical structure.

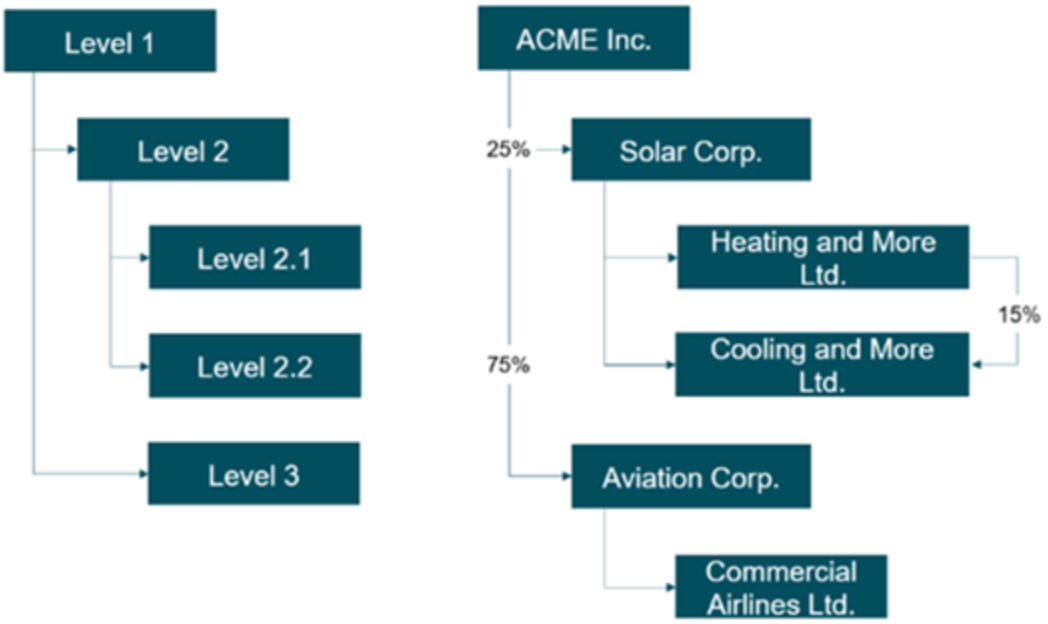


A division is regarded as an organizational unit for sales and distribution with a specific product portfolio to a customer. For every division, customer-specific agreements can be made like for partial deliveries, pricing, and terms of payment. A division can create an exclusion list of products for other divisions which can only be sold to a customer from that division.

In large organizations, one or more divisions may be responsible for a customer, such as if the company is selling cleaning equipment such as disinfectant dispensers (healthcare division) and laundry machines (textile care division) to the same hospital.

Legal Characteristics

Legal hierarchies may be characterized by customers that are owned by the same legal entity. Legal hierarchies are typically used by the business to categorize customers according to legal ownership, i.e., between holdings and subsidiaries and/or shareholders. A common usage is for legal reporting and analytics such as whitespace, risk analysis and credit checks (ex. roll up from subsidiary to holding). The same customer is not typically placed under different locations.

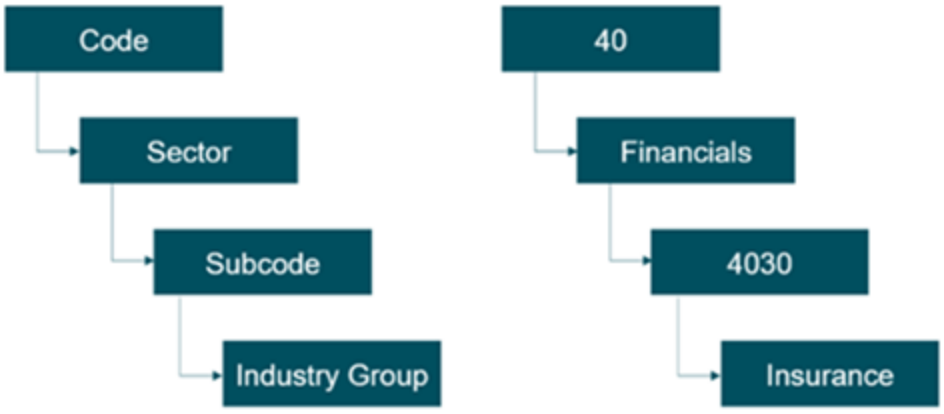


The above example illustrates a legal hierarchy for ACME, Inc.

Industry Characteristics

Industry-specific hierarchies may be characterized by customers that are conducting the same type of business. Industry hierarchies are typically used by the business to categorize customers by type of business using the Standard Industry Classification (SIC). This hierarchy may be used for analytics, i.e., customer/ market segmentation (customer allocated to industry, organizational unit allocated to industry).

The same customer is not typically placed under different locations.



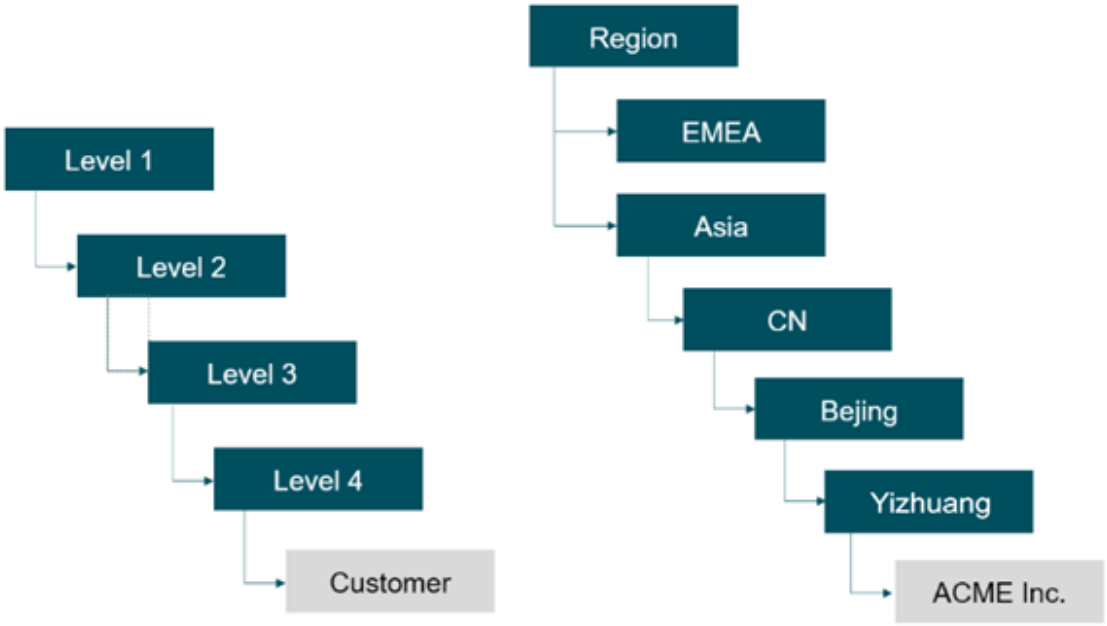
Companies of the same industry are assigned the same SIC code. For example, General Motors, Ford Motor Corporation, Honda, and Toyota were all assigned the SIC code of 3711.

Note: Entities have limited functionality when using in combination with Classifications:

- Entities are not shown as 'sub-products' (or children) of the Classification node
- Attribute values cannot be inherited from the Classification node to the Entity object.
- Classification references to Entities are not inherited from parent Classification to child Classification.

Spatial Location

Location-based hierarchies may be characterized by customers that are allocated at the same spatial location. Location hierarchies are typically used by the business in sales transactions (ordering address), delivery planning (routing), regional revenue reporting, whitespace analysis, segmentation (demographics). This hierarchy may mix different customers under the same location.



In this example, the hierarchy is purely spatial:

- The Region contains lower categories called 'EMEA' and 'ASIA'
- Below 'EMEA' are the countries of Europe and Middle East
- Below 'ASIA,' are the countries of 'CN,' 'JPN,' etc.
- Below 'CN' could be provinces like 'Beijing' and cities like 'Yizhuang,' etc.
- Below the 'Yizhuang' node are all customers which are allocated on this location

When grouping customers, consider what the usage of the hierarchy is (organizational, legal, spatial, or other). Furthermore, hierarchy names should be designed to organize customers used in that specific context.

Customer hierarchies are a way of structuring customer data so that the data is related to and connected to other objects. This is not the same structuring as with classifications for products, nor will it be the same structure as a web hierarchy where each sales channel has their own web sites. Finally, customer hierarchies are not structured for searching, sorting, and browsing data, though these can be built to apply to customer hierarchies.

Customer Hierarchy Use Cases

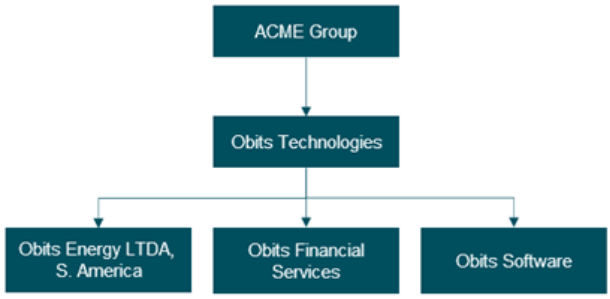
Use Case 1 - ACME Group Company Financial Hierarchy

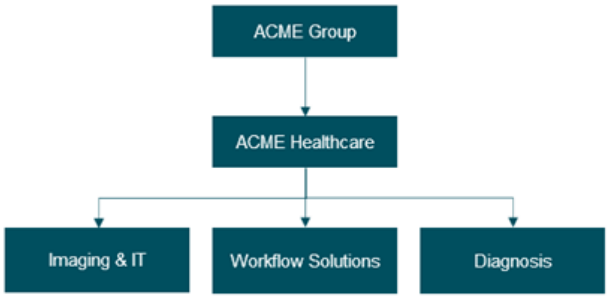
Acme Group consists of fourteen daughter companies across the Americas region. Among these, the relationships across sister companies may or may not exist. In each of these companies there may also exist subsidiaries which operate independently of each other and of the parent company. These subsidiaries are a combination of acquisitions and spin-offs of individual divisions / business units:

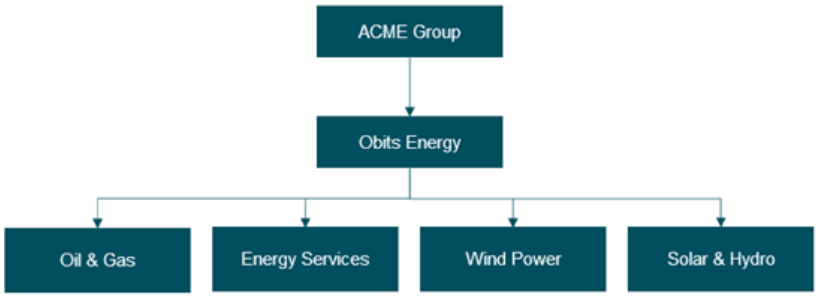
- **ACME Group** - Parent holding company with multiple subsidiary
- **ACME Inc** - Apparel Manufacturer and Distributor
- **ACME Healthcare** - Manufacturer and Distributor of healthcare equipment and supplies
- **Imaging & IT** – Focus on healthcare imaging systems
- **Workflow Solutions** – Focus on business workflow solutions either as a standalone solution or to compliment an ERP system.
- **Diagnostics** – Focus on healthcare chemical diagnostic supplies
- **ACME Healthcare Supplies LTDA, Brazil** - Healthcare Supply Division headquartered in Brazil and caters to South America
- **Obits Technologies** – Subsidiary focusing on enterprise software
- **Obits Financial Services** – Focus on financial sector: banking, insurance, wealth management
- **Obits Energy Innovations** – Think tank for innovations with a focus on energy generation, usage, and consumption
- **Wind Power** – Leader in wind turbine technology
- **Solar & Hydroelectric** – Leader in alternative energy solutions for businesses and municipalities
- **Oil & Gas** – Focus on technology to safely transport crude oil and natural gas
- **Energy Service** – Energy Solutions Professional Service offerings
- **Obits Energy LTDA, Brazil** – Energy division headquartered in Brazil and caters to South America

Walter is a Level I analyst for Acme Group, the primary holding company of Acme Inc and all daughter companies and subsidiaries. In alignment with a new initiative, one of Walter's tasks is to assist with the quarterly reporting for Acme Group. To do so, Walter requires running consolidated (roll-up) reports to provide an aggregate view of all of Acme Group's holdings and investments.

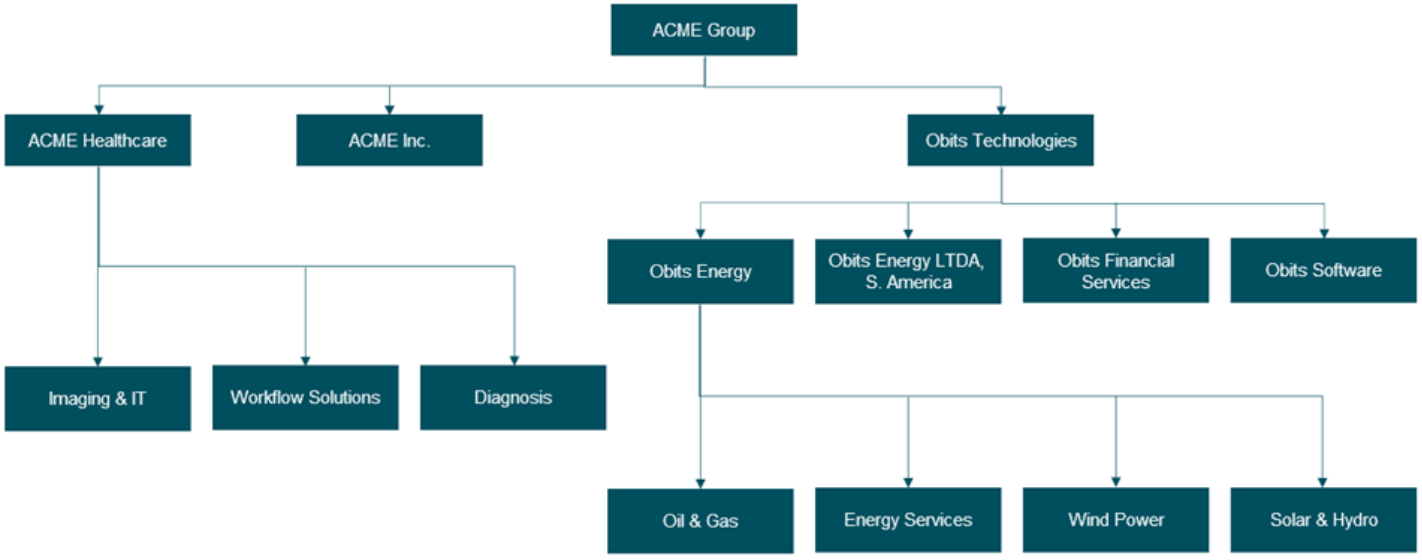
The challenge that Walter faces today is due to the growing number of holdings and subsidiaries over time, each individual company may have implemented their own accounting solutions to track various performance metrics and KPIs. As such, each company has different views of the truth this makes it very difficult to track and report on the relationships between daughter-to-subsidiary, daughter-to-daughter, subsidiary-to-subsidiary companies.





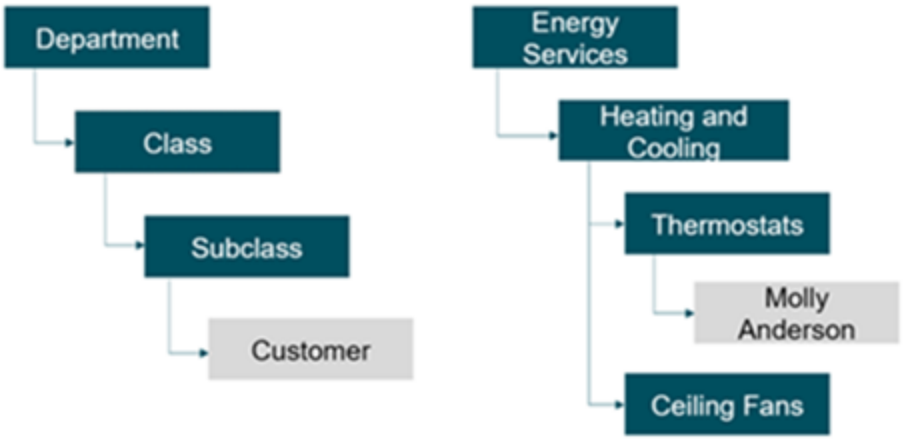


Every quarter, Walter is tasked with generating a quarterly financial report broken down by industry and subsidiaries. However, with no centralized view of the corporate hierarchy and since finances are maintained in separate, isolated ERPs based on industry and divisions. Walter is operationally challenged due to lack of governance.



However, in using Customer MDM, Walter is now able to utilize the hierarchy maintenance tools to visualize and maintain a golden hierarchy which models the complete company hierarchy for Acme Group. By taking in consolidating the various information provided by each company's source system, Customer MDM may now establish the proper relationships between companies within Acme Group.

This golden hierarchy information may be synchronized with Acme Group's accounting application, so Walter now has an accurate and complete view of all holdings and daughter companies. Walter can confidently select which subsidiaries to be included for various roll-up reports to drive Acme Group's quarterly analysis.

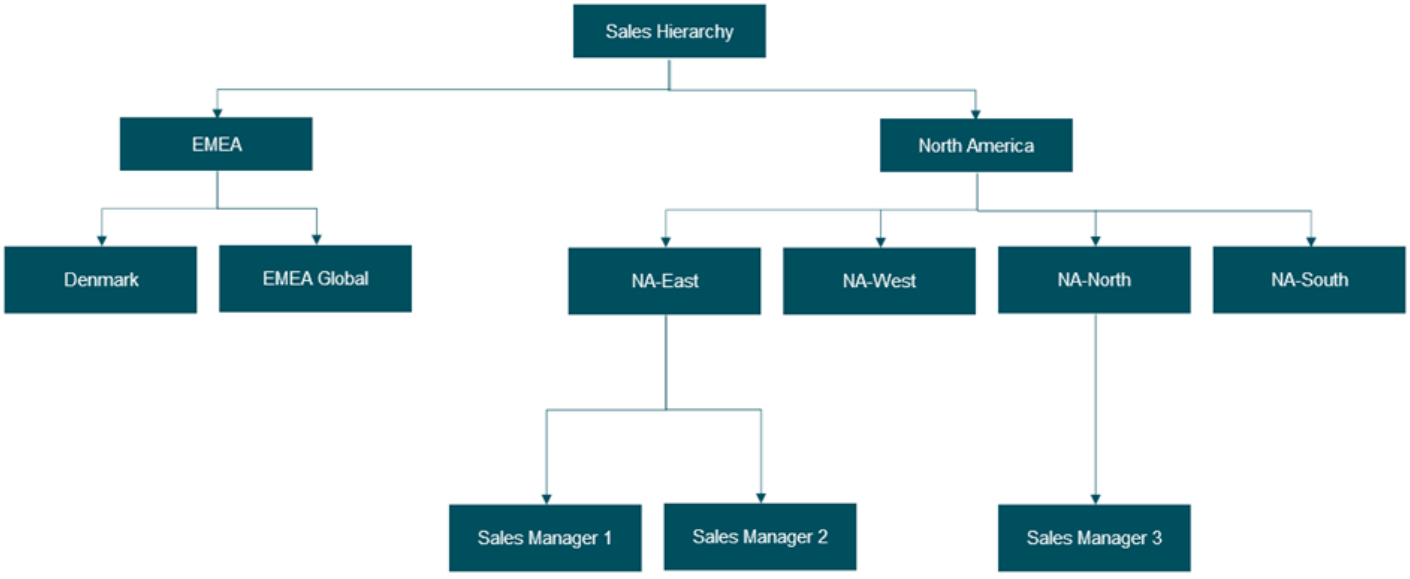


Use Case 2 - ACME Group Company Sales Hierarchy

Whitney, a data steward, is responsible for maintaining internal customer hierarchies at ACME. A new sales manager has been hired to replace a retiree, and Whitney has been asked to update the sales hierarchy. She will need to locate the appropriate territory for the retiree and remove him before adding the new sales manager.

Without a structured internal view, however, ACME struggles with the ability to accurately view and maintain their internal sales structure. To overcome this obstacle, Whitney can use STEP's customer hierarchy capabilities to easily identify the correct region for the new sales manager and create a new entity to represent the sales manager directly below.

With STEP, ACME can now easily view their internal hierarchies, and data stewards are empowered to maintain them.



Recommended Practices for Hierarchies

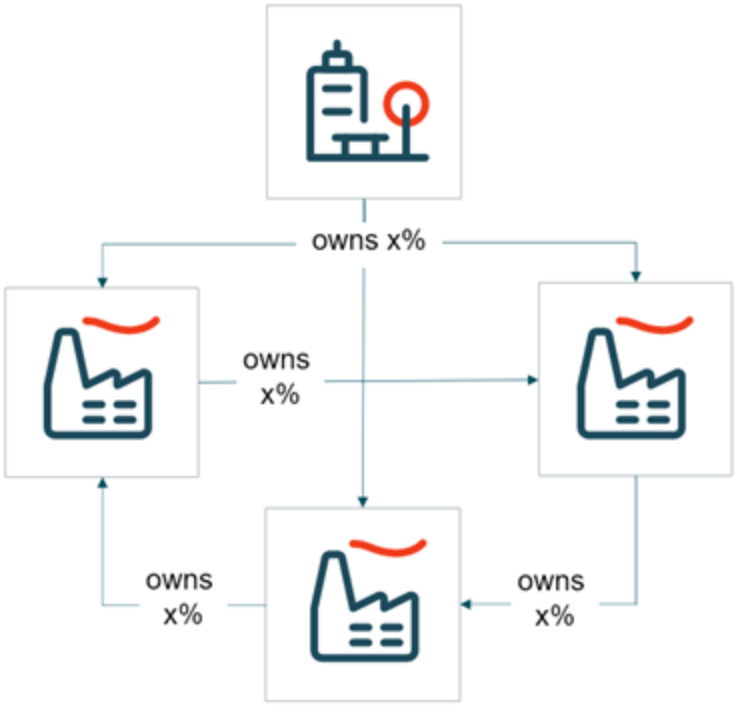
Navigation

The Web UI should be optimized for navigating hierarchies:

1. The Global Navigation Panel should include 'Hierarchies' as a Tree Navigator menu item.
2. All hierarchy roots should be included in the within the Hierarchies menu item. Customer entities should be visualized within the company hierarchy screen but should not be included in the Hierarchies Tree in the navigation panel.
3. A hierarchy tab page should be available on all Details screens for valid hierarchy object types.

Referenced Based Hierarchy

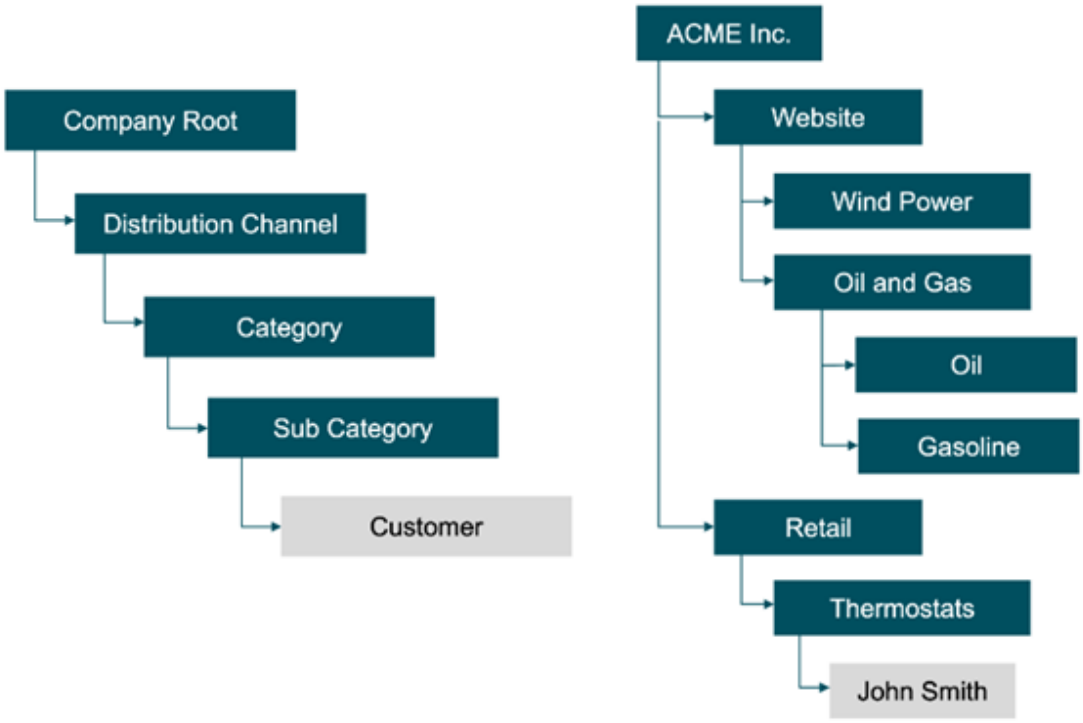
Referenced based hierarchies are structured to where customers reference other customers. Customers that are assigned to higher customer levels specified by the reference type used to appropriate attribution. One benefit of this hierarchy over the parent-child hierarchy is that reassigning a customer doesn't not require related customers to have to be moved, though the relationship change does apply to them.



Web Hierarchy

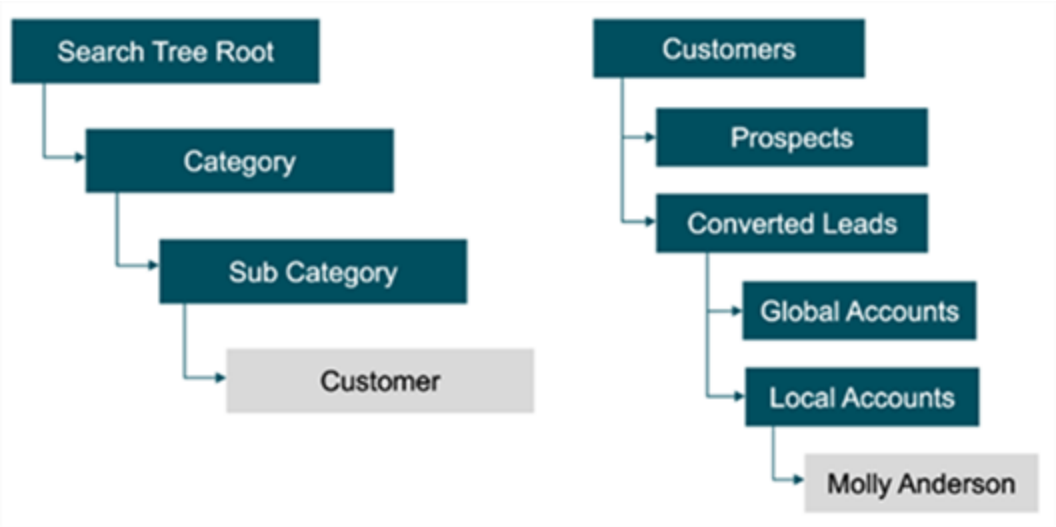
It is recommended to keep the web or sales hierarchy for products clean from customer data. A web or sales hierarchy for products is not intended to be customer specific but rather to represent the product portfolio intended for a specific sales channel.

Customers are typically assigned to a distribution channel in ERP in sales transactions. The distribution channel is an entity through which sales items (products or services) reach customers. They may include wholesale, retail and direct sales. Sales organization is the legal entity that structures the company according to its sales requirements (selling unit) and defines the basic sales and distribution structure. A sales organization is responsible for selling products and services.



Search Hierarchy

Customer hierarchies is not structured a searching, sorting, or browsing hierarchy. If necessary, these structures may be built in addition to customer hierarchies.



Supplier Hierarchies

The purpose of Supplier-centric hierarchies is for the grouping of supplier entities and maintaining those relationships. Examples of how such hierarchies may be used to support business operations include:

- Creating, maintaining, and understanding legal relationships amongst supplier entities which may be integral for negotiating pricing and terms, thereby reducing cost.
- Managing supplier territories for streamlining procurement processes.
- Supporting cost center roll-ups to track supplier performance.

In the case of creating, maintaining, and understanding supplier legal hierarchies, MDM leverages reputable third-party industry data sources, such as Dun & Bradstreet. By utilizing information provided by D&B's Detailed Company Profile offering, MDM can identify and relate entities to their Global and Domestic Ultimate Parent companies as well as immediate parent information. This allows for the building out of hierarchical relationships of suppliers within their family structures, boosts purchasing power, and cut costs.

MDM Data Flow

This section of the enablement materials details the flow of customer data into and within MDM for Consolidation, Coexistence and Centralized methodologies. Depending on the implementation style, the origination and flow of data differs with respect to the MDM and any corresponding source systems.

For more information on the various implementation methodologies of MDM, see the **Master Data Management Overview** section of this documentation.

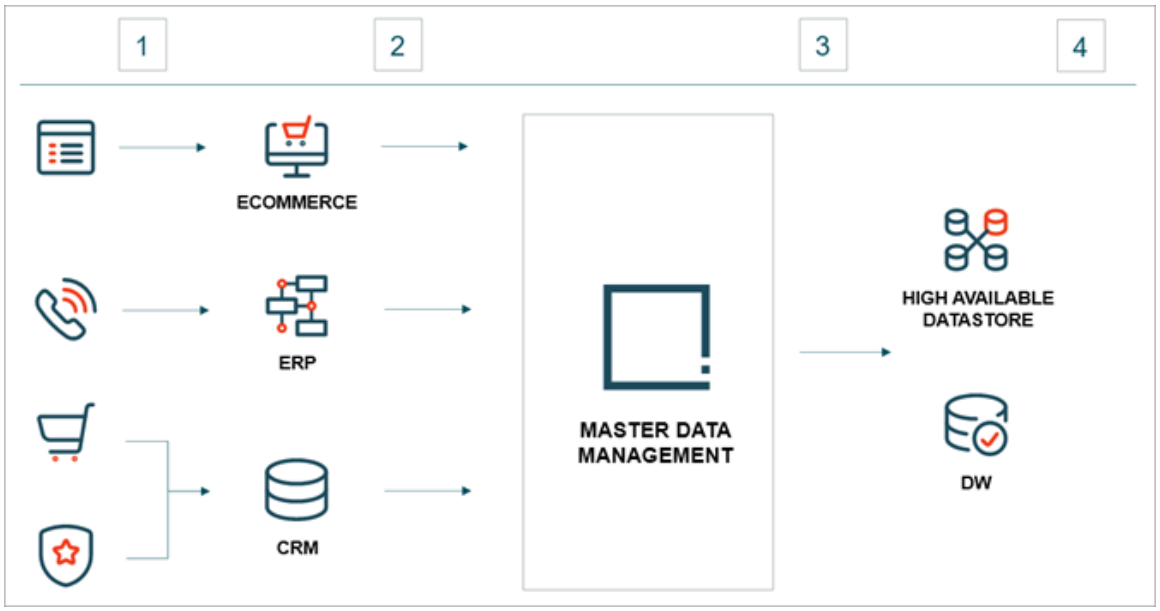
For examples of how the data flows into STEP within the context of a Coexistence solution, see the **Data Flow for Consolidation and Coexistence MDM** topic in this documentation.

For examples of how the data flows into STEP within the context of a Centralized solution, see the **Data Flow for Centralized MDM in B2B Solutions** topic in this documentation.

Data Flow for Consolidation and Coexistence MDM

Having entity data authored and stored across separate source systems creates a barrier for achieving a unified and trusted view of customer or supplier data. When data has multiple sources, it can become inconsistent, inaccurate, outdated, and cause conflicts.

ACME Company has multiple source systems from which entity data originates.



The flow of data for ACME Company begins at these points of entry, each of which is associated with a variety of channels. These channels include: the company website, phone direct, in-store, etc. (shown in column 1 of the image above). ACME contains the source systems (column 2):

- E-commerce system for consumers who place their orders and self-registers via the company's website.
- ERP system for suppliers from whom ACME currently procures their raw materials and finished goods.
- CRM system for driving seasonal marketing campaigns and financial data, which drives in-house financing as well as managing consumer loyalty programs. CRM may also manage existing distributor profiles as well as direct and in-direct B2B customers.

To understand how data flows into STEP and between applications, see the consolidation and coexistence implementation methodologies as defined in the **Master Data Management Overview** topic of this documentation.

Data enters MDM through external channels where customers / suppliers provide their information. In the B2C example image above, channels can include a telephone order, a web order, a consumer signing up for a loyalty platform, or buying in a store.

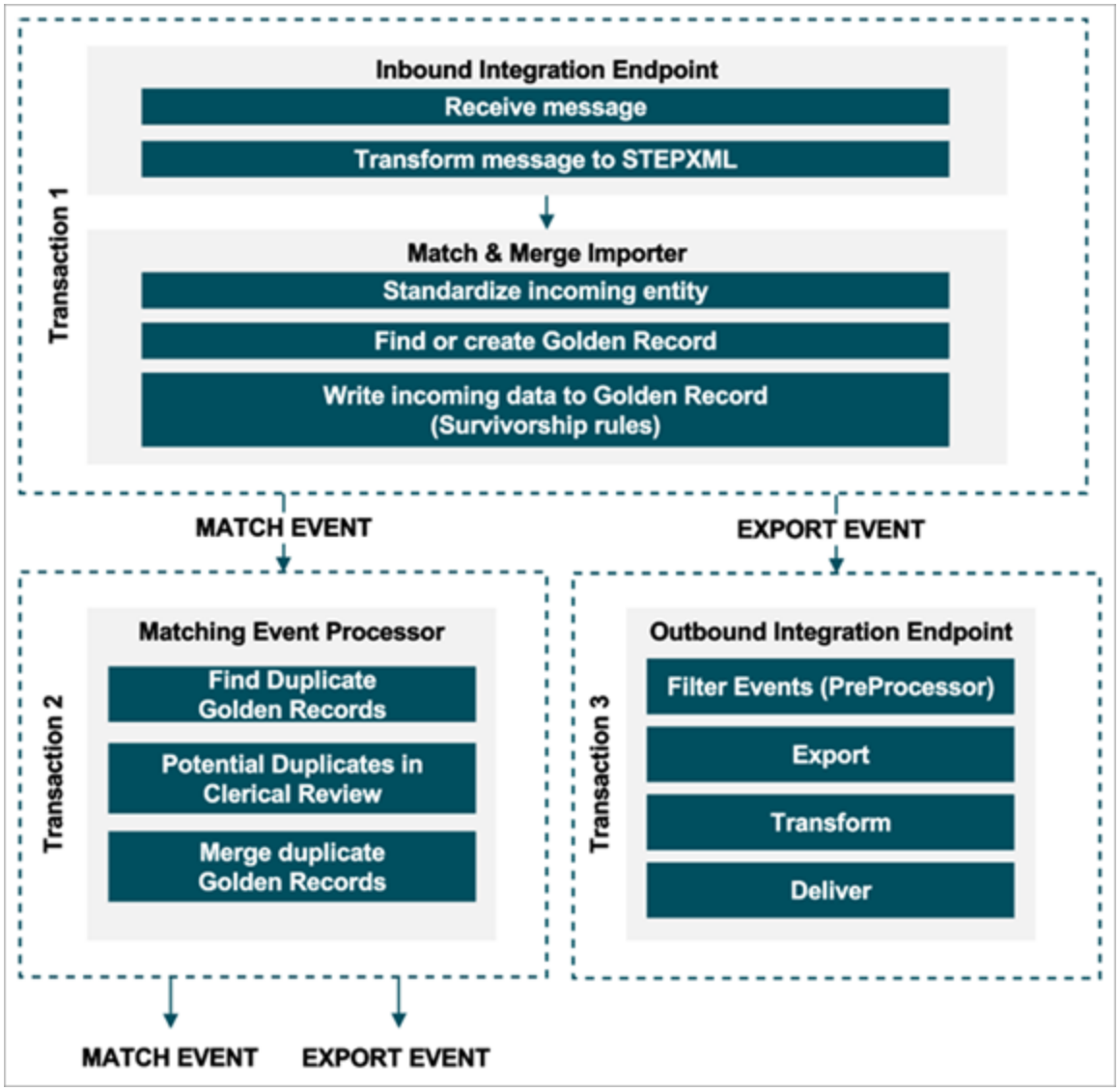
After the data is captured, it is saved to various source systems, which are the systems that store data to be delivered to the MDM solution. In this use case, the source systems are the CRM, ERP, and E-commerce platforms (column 2 in the image). Once data is within the MDM application, it is standardized, matched with existing records, and then golden records are created / maintained. Data is then stored in a data store / data warehouse.

For more details on what happens in MDM solution (column 3), see the **Transaction Examples for Consolidation and Coexistence MDM** topic. Also, see the **Matching** topic and the **Loqate** topic of this documentation for more information.

In a coexistence implementation, the data matched and merged in the MDM solution (column 3) is synchronized to the source systems (column 2) as well as downstream solutions (column 4).

Transaction Examples for Consolidation and Coexistence MDM

This section provides a process overview of the primary data flow throughout the STEP system using a Consolidation and Coexistence approach.



Transaction 1: Inbound Integration Endpoint

An inbound integration endpoint will receive a message and transform it into STEPXML.

Once transformed, the Match & Merge Importer will:

- Execute standardizations and other associated business rules to cleanse the data. Cleansed data allows for the Match algorithm to more accurately identify duplicate records.

- Determine the target Golden Record. For more information, see the **IIEP - Configure Match and Merge Importer** section of the **Data Exchange** documentation.
- Survivorship rules will write any updates from the incoming data to the golden record. For more information, see the **Golden Records Survivorship Rules** section of the **Matching, Linking, and Merging** documentation.
- To the extent that a Golden Record was created, or an existing Golden Record was updated, events are written to event queues triggering on those changes. In this case, a Matching event processor as well as an Outbound Integration Endpoint.

Transaction 2: Matching Event Processor

When new golden records are created or an existing golden record is updated, this will trigger a matching event, with the purpose of comparing golden records to each other.

Prior to matching, the golden record under treatment will have its match codes updated.

Matching golden records may result in a clerical review workflow initiation or a merge, depending on the score.

For more information on Clerical Reviews, see the **Clerical Review** section of the **Matching, Linking, and Merging** documentation.

Transaction 3: Outbound Integration Endpoint

Whenever Golden Records are created or changed, an export event will trigger. This allows the mastered data to be synced back to the various source systems.

For more information about Outbound Integration Endpoints, see the **Outbound Integration Endpoint** section of the **Data Exchange** documentation.

Note: Clerical Review is not a blocking state.

Use Case 1: New Golden Record

1. **Transaction 1: Import**
   - IIEP: Receive message
   - IIEP: Transform message to STEPXML
   - M&M: Standardize Address
   - M&M: Check if Source System and Source Record ID is associated to existing, active Golden Record (it is not)
   - M&M: Generate Match codes for incoming data and look for matching record using Match Code and Matching Algorithm (no match is found)
   - M&M: Create new Golden Record
   - M&M: Survivorship Rules Write data to new Golden Record
2. **Transaction 2: Handle duplicate Golden Records**

The Matching Event Processor polls for events on a schedule.

- Update match codes for changed Golden Record and look for matching Golden Records using Match Code and Matching Algorithm
    - No duplicate golden records found
    - No potential duplicate golden records found

3. **Transaction 3: Export**

Outbound Integration Endpoints poll for events on a schedule. Notice that transaction 2 and 3 are independent.

- OIEP: No event PreProcessor in use
- OIEP: Export data to STEPXML
- OIEP: Transform to target message format
- OIEP: Deliver message to target

Use Case 2: Update Existing Golden Record with Known Source Record ID

1. **Transaction 1: Import**

- IIEP: Receive message
- IIEP: Transform message to STEPXML
- M&M: Standardize Address
- M&M: Check if Source System and Source Record ID is associated to existing, active Golden Record (it is)
- M&M: Survivorship Rules Write data to existing Golden Record

2. **Transaction 2: Handle Duplicate Golden Records**

The Matching Event Processor polls for events on a schedule.

- Update match codes for changed Golden Record and look for matching Golden Records using Match Code and Matching Algorithm
    - If an absolute match is found, auto-merge
        - MA: Survivorship Rules Write data to existing Golden Record
    - If potential duplicate golden records are found, initiate into clerical review

3. **Transaction 3: Export**

Outbound Integration Endpoints poll for events on a schedule. Notice that transaction 2 and 3 are independent.

- OIEP: No event PreProcessor in use
- OIEP: Export data to STEPXML
- OIEP: Transform to target message format
- OIEP: Deliver message to target

Use Case 2.1: Update Existing Golden Record with Known Golden Record ID

1. **Transaction 1: Import**

- IIEP: Receive message
- IIEP: Transform message to STEPXML

- M&M: Standardize Address
- M&M: Find existing Golden Record based on STEP ID (golden record id)
- M&M: Survivorship Rules Write data to existing Golden Record

2. **Transaction 2: Handle Duplicate Golden Records**

The Matching Event Processor polls for events on a schedule.

- Update match codes for changed Golden Record and look for matching Golden Records using Match Code and Matching Algorithm
  - If an absolute match is found, auto-merge
    - MA: Survivorship Rules Write data to surviving Golden Record
  - If potential duplicate golden records are found, initiate into clerical review

3. **Transaction 3: Export**

Outbound Integration Endpoints poll for events on a schedule. Notice that transaction 2 and 3 are independent

- OIEP: No event PreProcessor in use
- OIEP: Export data to STEPXML
- OIEP: Transform to target message format
- OIEP: Deliver message to target

Use Case 3: Add new source record to existing golden record by matching

1. **Transaction 1: Import**

- IIEP: Receive message
- IIEP: Transform message to STEPXML
- M&M: Standardize Address
- M&M: Check if Source System and Source Record ID is associated to existing, active Golden Record (it is not)
- M&M: Generate Match codes for incoming data and look for matching record using Match Code and Matching Algorithm (match is found)
- M&M: Update Existing Golden Record with additional source record ID
- M&M: Survivorship Rules Write data to existing Golden Record

2. **Transaction 2: Handle Duplicate Golden Records**

The Matching Event Processor polls for events on a schedule.

- Update match codes for changed Golden Record and look for matching Golden Records using Match Code and Matching Algorithm
  - If an absolute match is found, auto-merge
    - MA: Survivorship Rules Write data to surviving Golden Record
  - If potential duplicate golden records are found, initiate into clerical review

3. **Transaction 3: Export**

Outbound Integration Endpoints poll for events on a schedule. Notice that transaction 2 and 3 are independent.

- OIEP: No event PreProcessor in use
- OIEP: Export data to STEPXML
- OIEP: Transform to target message format
- OIEP: Deliver message to target

Use Case 3.1: Add New Source Record to Existing Golden Record by Golden Record ID

This use case is relevant when the Find Similar web service has been used to identify that a golden record for a specific entity already exists and a new source record needs to be associated with it.

1. **Transaction 1: Import**

   - IIEP: Receive message
   - IIEP: Transform message to STEPXML
   - M&M: Standardize Address
   - M&M: Find existing Golden Record based on STEP ID (golden record id)
   - M&M: Update existing Golden Record with new source record ID
   - M&M: Survivorship Rules Write data to new Golden Record

2. **Transaction 2: Handle Duplicate Golden Records**

The Matching Event Processor polls for events on a schedule.

- Update match codes for changed Golden Record and look for matching Golden Records using Match Code and Matching Algorithm
  - If an absolute match is found, auto-merge
    - MA: Survivorship Rules Write data to surviving Golden Record
  - If potential duplicate golden records are found, initiate into clerical review

3. **Transaction 3: Export**

Outbound Integration Endpoints poll for events on a schedule. Notice that transaction 2 and 3 are independent.

- OIEP: No event PreProcessor in use
- OIEP: Export data to STEPXML
- OIEP: Transform to target message format
- OIEP: Deliver message to target

Use Case 4: Add New Source Record to Existing Golden Record by Matching, Clerical Review

1. Transaction 1: Import

   - IIEP: Receive message
   - IIEP: Transform message to STEPXML

- M&M: Standardize Address
- M&M: Check if Source System and Source Record ID is associated to existing, active Golden Record (it is not)
- M&M: Generate Match codes for incoming data and look for matching record using Match Code and Matching Algorithm
- M&M: Match Rules will not find certain matches
- M&M: Create new Golden Record
- M&M: Survivorship Rules Write data to new Golden Record

2. **Transaction 2: Handle Duplicate Golden Records**

The Matching Event Processor polls for events on a schedule.

- Update match codes for changed Golden Record and look for matching Golden Records using Match Code and Matching Algorithm
    - Identify that newly created golden record is within clerical review threshold and put oldest record in clerical review workflow

3. **Transaction 3: Merge**

- User review duplicate records in clerical review
- User merge potential duplicates
- Source record references from non-surviving record gets written to surviving record
- References TO the non-surviving record gets moved to the surviving record
- Survivorship rules promote updates from non-surviving- to surviving golden record
- Non-survivor gets deactivated and Merged-Into reference is established from non-survivor to survivor

4. **Transaction 4: Export**

The deactivated record is exported, as the 'Merged-Into' reference was added and the 'Is Deactivated' was changed to 'true'. The surviving record is exported, as source records were added to it. Any other object having a reference target moved from the non-surviving record to the surviving record will get exported, to the extent triggering definitions are defined for it.

- OIEP: No event PreProcessor in use
- OIEP: Export data to STEPXML
- OIEP: Transform to target message format
- OIEP: Deliver message to target

Use Case 5: Manually Modify Golden Record

1. **Transaction 1: Manual Edit**

- Data Steward: Manually makes edits on an existing Golden Record via the Web UI.

2. **Transaction 2: Handle Duplicate Golden Records**

The Matching Event Processor polls for events on a schedule.

- Update match codes for changed Golden Record and look for matching Golden Records using Match Code and Matching Algorithm
  - If an absolute match is found, auto-merge
    - MA: Survivorship Rules Write data to existing Golden Record
  - If potential duplicate golden records are found, initiate into clerical review
    - MA: If match is confirmed, survivorship Rules Write data to existing Golden Record
    - MA: If match is rejected, create a 'Confirmed Non-Duplicate' reference between the two (or more) Golden Records.

3. **Transaction 3: Export**

Outbound Integration Endpoints poll for events on a schedule. Notice that transaction 2 and 3 are independent.

- OIEP: No event PreProcessor in use
- OIEP: Export data to STEPXML
- OIEP: Transform to target message format
- OIEP: Deliver message to target

Use Case 6: Import Cross References by Source Record ID

1. **Transaction 1: Import**

- IIEP: Receive message
- IIEP: Transform message to STEPXML
- M&M: Standardize Address
- M&M: Check if Source System and Source Record ID is associated to existing, active Golden Record (it is for the target entity)
- M&M: Generate Match codes for incoming data and look for matching record using Match Code and Matching Algorithm
- M&M: Match Rules will not find certain matches
- M&M: Create new Golden Record (Source entities)
- M&M: Survivorship Rules Write data to new Golden Record including creating cross references

2. **Transaction 2: Handle Duplicate Golden Records**

The Matching Event Processor polls for events on a schedule.

- Update match codes for changed Golden Record and look for matching Golden Records using Match Code and Matching Algorithm
  - If an absolute match is found, auto-merge
    - MA: Survivorship Rules Write data to surviving Golden Record
  - If potential duplicate golden records are found, initiate into clerical review

3. **Transaction 3: Export**

Outbound Integration Endpoints poll for events on a schedule. Notice that transaction 2 and 3 are independent.

- OIEP: No event PreProcessor in use
- OIEP: Export data to STEPXML
- OIEP: Transform to target message format
- OIEP: Deliver message to target

When the Match and Merge Importer reads cross references in the STEPXML, the target Entity may be specified with a source system and a source record ID, instead of a STEP ID. This allows import of references to objects, where the STEP ID is not known.

Match and Merge Web Service

In large heterogeneous enterprise landscapes, customer data is authored and stored in various systems such as the e-commerce system, the customer relationship management (CRM) system, the Enterprise Resource Planning (ERP) system, etc. These systems provide valuable functions making it necessary that users continue to work in them. However, business processes, such as placing an order for a product, can often span multiple applications. For instance, an order may come in via the e-commerce system but be fulfilled by the ERP. This can often cause problems due to significant dependencies on the consistency and timeliness of data synchronization between applications. If the delivery address for an entity is changed in the e-commerce system, it must also be changed in the ERP, so the order can successfully be delivered. Therefore, it is essential to create a tighter coupling between operational systems and the MDM system, so that the entity data is synchronized as needed. This tight coupling is possible with a Coexistence-Style MDM implementation, where centrally governed records are created and updated in a large heterogeneous enterprise landscape.

In the Coexistence methodology, multiple systems containing the master data must operate at the same time. The entity data is authored and stored in various external sources while being synchronized to the MDM system. This is accomplished in STEP through the synchronous, request-response based web service for golden record creation and update. This process involves deduplication, conflict resolution, and validation operations, resulting in the creation / update of golden records or rejection of the create / update. Updates to the data can be done in the source system and any external systems that masters entity data. STEP synchronizes the content of golden records back to the source systems, keeping every source up to date with data contributed from any source. This can be done either via direct integration with source systems or publishing to a data warehouse both for synchronization and for further analytics.

For the ability to only search STEP for existing records before new ones are created in external systems, see the **Find Similar Web Service** topic in this documentation.

Match and Merge Use Cases

In both Business to Business (B2B) and Business to Customer (B2C) interactions, users will access customer relationship management (CRM) systems and Enterprise Resource Planning (ERP) systems, which often do not communicate with one another. As a result, users accessing client records will frequently face instances of insufficient, inaccurate, and/or duplicate data. Due to insufficient records or user error, users can unintentionally produce data that is unusable or causes inefficiencies. They could create or update data in the CRM system which conflicts with or duplicates data in the ERP system. This in turn results in inefficiencies which cause substantial cost overhead and creates a poor customer experience.

Through Stibo Systems' request-response based web service for golden record creation and update, users circumvent this problem. The golden record contains all the data your customers have provided by drawing information from all connected silos. When users create or update data in one silo, that data also gets added to the golden record. The user can now standardize, validate, and/or reject data based on their quality criteria, without fear of creating inconsistencies or duplicates.

For examples of golden record creation and update via web service, see the following topics:

- Use Case 1: B2B
- Use Case 2: B2C

Use Case 1: B2B

Pedro from the marketing team is responsible for converting marketing leads into retailer accounts. He needs to create and update retailers in the CRM and ERP based off results from a marketing campaign.

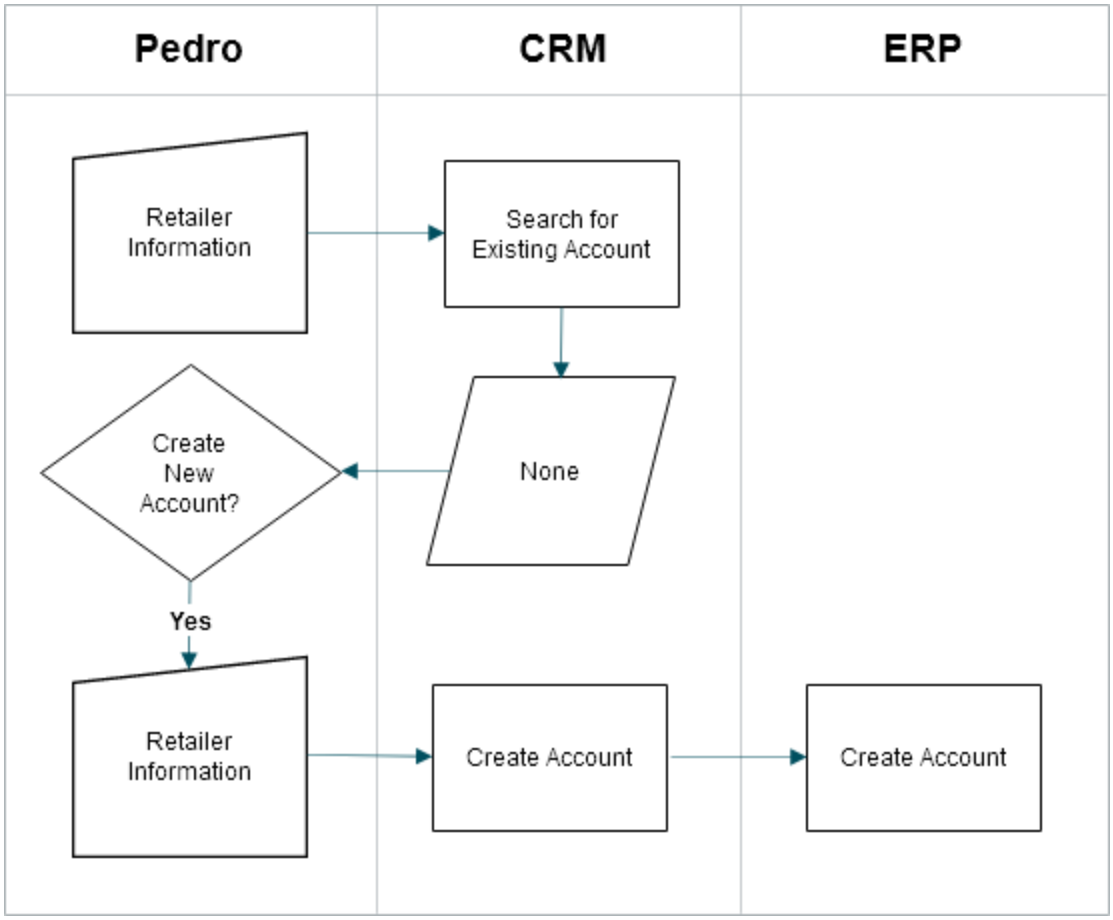
Pedro thinks the address information he gathered is sufficient for future business needs. He doesn't realize how poor some of the address data is and is unaware that some of the retailers gathered during the marketing campaign are duplicates of existing retailers that have active orders. Pedro believes they are all new retailers to ACME and thinks the leads collected are useable. He doesn't realize some of the retailers he collected during his campaign are no longer in operation under the name collected. The retailers have been acquired or have gone out of business. The CRM and ERP are disjointed systems at ACME, with Pedro only having access to the CRM. Changes to data are not always propagated efficiently between the two systems.

As a result, Pedro causes operational inefficiencies with substantial cost overhead in returned mail due to the poor address data. He creates a poor customer experience when creating duplicate retailers in the landscape. Duplicate retailers may be contacted by multiple different salespeople, unaware of the existing relationship with the retailer. Additionally, updates to existing retailers may change the records in a negative way. Pedro produces operational inefficiencies when retailers that do not meet the quality criteria are created. This costs the sales team both time and money when they pursue non-useable retailers. ACME provides a poor customer experience as it struggles to keep retailers in sync, accurate, and up to date in both the CRM and ERP.

The following are examples of some solutions:

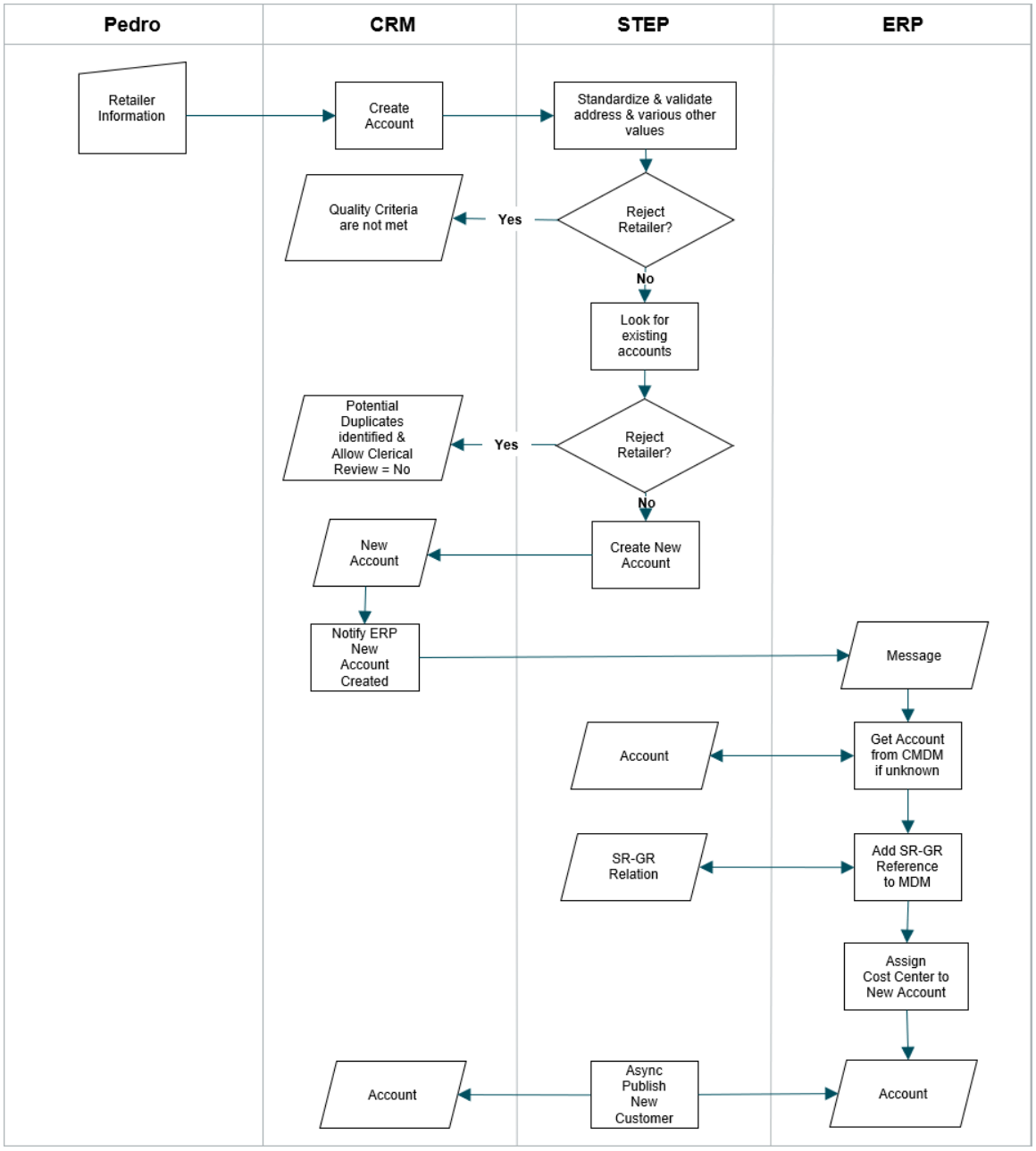
As-is

With the current System landscape, Pedro cannot reject the creation of records in the CRM when records that don't meet the quality criteria are being created. He also cannot reject duplicates during creation. Additionally, when changes are made to records in the CRM, the ERP does not always reflect the most recent version.



To-be

With the request-response web service for coexistence style Customer MDM in place, Pedro can reject the creation or updating of records during the process. This allows the business to reject potential duplicates and records that do not meet quality criteria. The problem is addressed at the root, ensuring smoother operations throughout the entire landscape. Additionally, when the CRM notifies the ERP of changes to records, the ERP can pull the missing master data from STEP with the confidence that it is accurate and up to date.


Pedro
CRM
STEP
ERP
Retailer Information
Create Account
Standardize & validate address & various other values
Quality Criteria are not met
Yes
Reject Retailer?
No
Look for existing accounts
Potential Duplicates identified & Allow Clerical Review = No
Yes
Reject Retailer?
No
New Account
Create New Account
Notify ERP New Account Created
Message
Account
Get Account from CMDM if unknown
SR-GR Relation
Add SR-GR Reference to MDM
Assign Cost Center to New Account
Account
Async Publish New Customer
Account

Use Case 2: B2C

Maria, from the call center, is responsible for taking ACME orders over the phone. She needs to associate a customer to orders in the CRM by either updating an existing customer or creating a new customer based off the information provided during her phone calls.

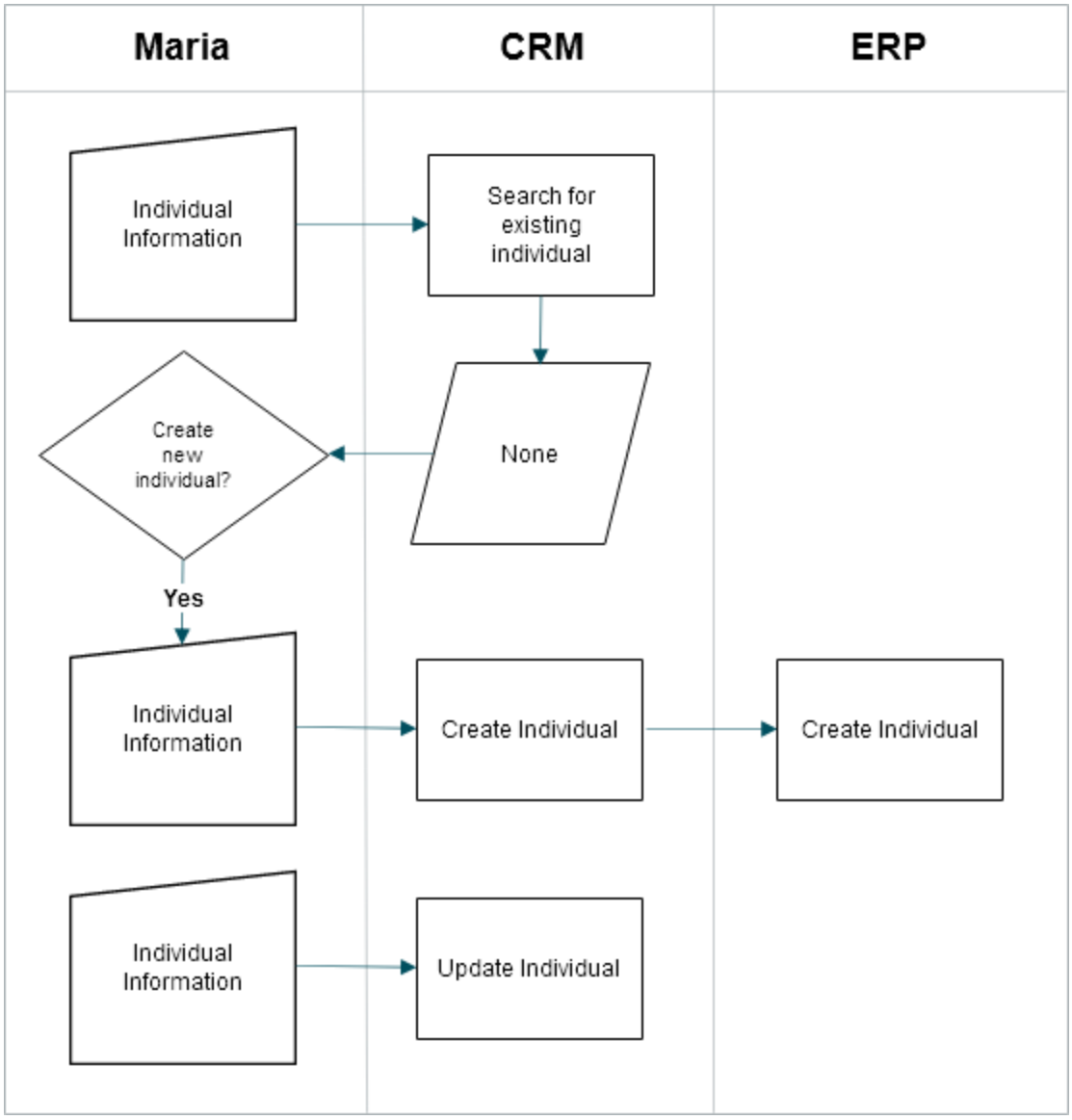
Maria believes that the delivery address and the email address information she enters for the customers are correct. However, she and the other customer service representatives frequently make typos or don't enter the information properly such as Peachtree Street vs Peach Tree Street. She is not able to determine if the customer already exists in the enterprise landscape with existing orders. Maria believes that any changes she makes to existing customers in the CRM are automatically updated for their orders.

Maria causes operational inefficiencies with substantial cost overhead in undelivered orders and in billing issues due to the poor address data. Operational inefficiencies also exist with missed cross-sell and up-sell opportunities when email promotions are not delivered. She provides a poor customer experience due to her lack of knowledge of other engagements with the same customer. ACME customers have a poor customer experience because despite calling the call center to update their information they are not seeing the changes reflected in their orders.

The following are examples of some solutions:

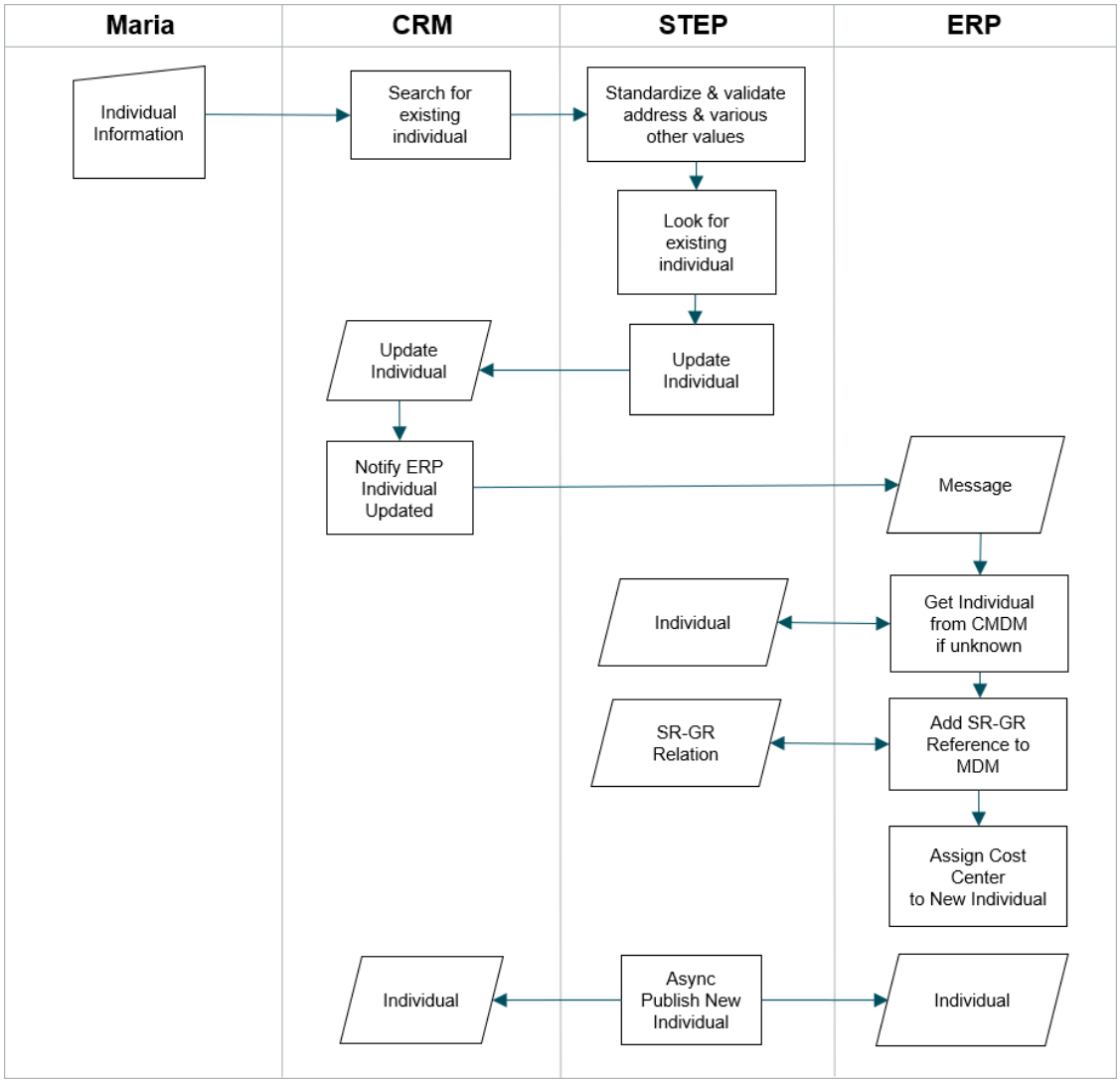
As-is

With the current system landscape, Maria cannot standardize and validate incoming information. She has a difficult time determining if the customer already exists and syncing changes to existing customers between the CRM and ERP.



To-be

With the request-response web service for coexistence style Customer MDM in place, Maria can standardize and validate incoming information during the creation / update process. From there, STEP can determine if duplicates exist for the customer, then update the existing customer based on the survivorship rules in place and send the changes to the customer back to the CRM. The CRM can then notify the ERP which can poll STEP for the current customer information. This keeps the customer accurate and up to date throughout the landscape.


Maria
CRM
STEP
ERP
Individual Information
Search for existing individual
Standardize & validate address & various other values
Look for existing individual
Update Individual
Update Individual
Notify ERP Individual Updated
Message
Individual
Get Individual from CMDM if unknown
SR-GR Relation
Add SR-GR Reference to MDM
Assign Cost Center to New Individual
Individual
Async Publish New Individual
Individual

Coexistence Examples

The synchronous, request-response based web service for golden record creation and update employs the matching functionality, business actions, and business conditions to produce the information returned in the response. For this reason, certain aspects of the solution must be appropriately paired for the call to function.

For more information, see the **Match and Merge Web Service Endpoint** topic of the **Matching, Linking, and Merging** documentation.

Configuration

It is recommended that the same algorithms be used for the match and merge web service requests.

The provided web service endpoints utilize Loqate to standardize address, while also include an 'address check' business condition. This business condition allows for the rejection of records if a certain data quality standard is not met.

See the following examples below:

- Create New Contact Person
- Update Existing Individual Customer
- Reject Organization Customer Based on Failed Validation
- Reject Organization Customer as Potential Duplicate
- Update Existing Supplier

Create New Contact Person

The following section includes sample REST and SOAP XML request call as well as the XML response.

Required values to identify match:

- Name
- Address
- Email
- Phone

New Contact Web Service Request

See the online version of this topic to view the web service request.

New Contact Web Service Response

See the online version of this topic to view the web service response.

Update Existing Individual Customer

The following section includes sample REST and SOAP XML request call as well as the XML response.

Required values to identify match:

- Name
- Address
- Email
- Phone

Update Existing Individual Customer Web Service Request

See the online version of this topic to view the web service request.

Update Existing Individual Customer Web Service Response

See the online version of this topic to view the web service response.

Reject Organization Customer Based on Failed Validation

The following section includes sample REST and SOAP XML request call as well as the XML response.

Required value for successful validation:

- InputStreet

Reject Organization Customer Based on Failed Validation Web Service Request

See the online version of this topic to view the web service request.

Reject Organization Customer Based on Failed Validation Web Service Response

See the online version of this topic to view the web service response.

Reject Organization Customer as Potential Duplicate

The following section includes sample REST and SOAP XML request call as well as the XML response.

Required values to identify match:

- Legal Name
- Address
- Email
- Phone

Reject Organization Customer as Potential Duplicate Web Service Request

See the online version of this topic to view the web service request.

Reject Organization Customer as Potential Duplicate Web Service Response

See the online version of this topic to view the web service response.

Update Existing Supplier

The following section includes sample REST and SOAP XML request call as well as the XML response.

Required values to identify match

- Legal Name
- Address

Update Existing Supplier Web Service Request

See the online version of this topic to view the web service request.

Update Existing Supplier Web Service Response

See the online version of this topic to view the web service response.

Find Similar Web Service

Business users in external systems throughout the enterprise may need to check if customer records already exist. By using the Find Similar web services (for both REST and SOAP), the business can search for existing similar records and prevent duplicates from being created at the point of entry.

For information on configuration, see the **Find Similar** topic of the **Matching, Linking, and Merging** documentation.

Considerations

- The Find Similar call employs the matching functionality to produce the potential duplicate results returned in the response. For this reason, certain aspects of the solution must be appropriately paired for the call to function.
- The matching configuration may need to be altered to support the Find Similar call including creating additional Match Codes and Match Algorithms. It is important to examine the current matching landscape to determine if alterations are needed.
- The match algorithm being used to create and manage golden records may not be the best algorithm for the Find Similar call. Consider if external system's business needs align with the business needs of the existing governing match algorithm.
- Determine if a person or a machine will make the decision when reviewing potential duplicates and create an appropriate match algorithm:
  - If a person (such as a Customer Service Representative) in an external system asks customers confirming questions about the data that is found, a looser match algorithm may be desired.
  - If a machine makes the decision, there is no need for a looser match algorithm because the decision is made solely on the score.

Find Similar Use Cases

The following topics show how Find Similar can be used to create unique records:

- Business-to-Business (B2B) - **Find Similar Acquisitions**
- Business to Customer (B2C) - **Find Similar Preventing Duplicates**

Find Similar Acquisitions

Benny, from the sales team, is responsible for retailers and distributors that have contracts to resell ACME's products.

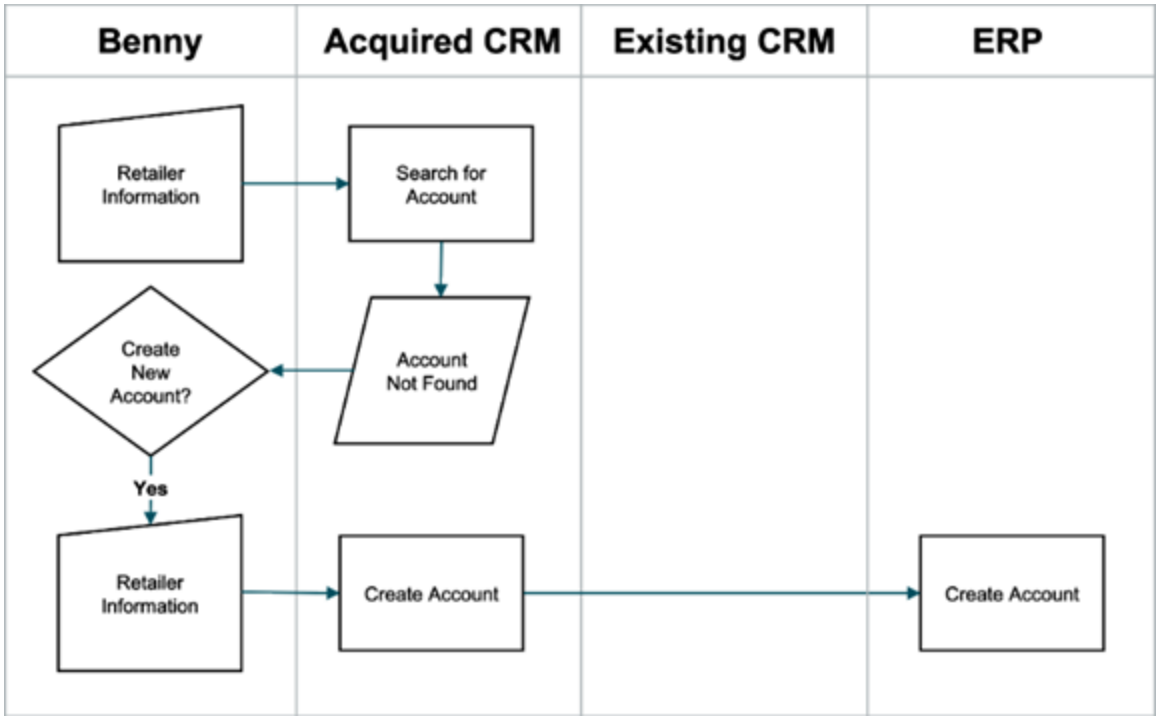
Benny needs to be able to check if a retailer or distributor already exists because ACME recently acquired a new company with a separate CRM.

Since Benny doesn't have clear visibility into other systems, he cannot identify the record since the record does not exist in his CRM. Additionally, the other company's CRM cannot identify the record based on a fuzzy match if Benny cannot provide information for an exact match like a unique identifier.

Benny has difficulty merging records in the ERP system once they've already been created because they may have orders associated with them. It is also common that existing records from another system contain information that is valuable to Benny, in the process where he is creating the new record. i.e., the customer may have made a purchase recently, have an open support issue, or in the process of paying debts. It is therefore of high value to know of the existing record as well as to prevent creation of duplicate records.

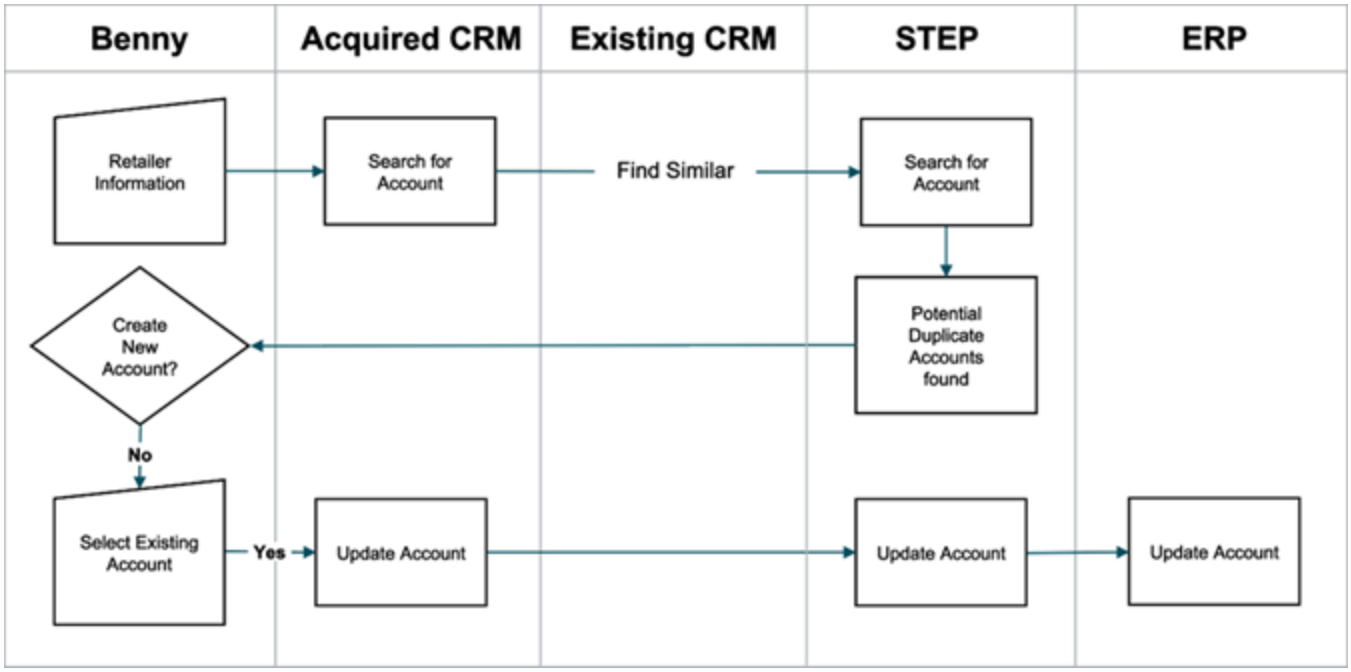
Due to his lack of knowledge of other engagements with the same customer, Benny cannot provide a positive customer experience. Additionally, he is increasing the risk of poor customer experiences in the future as he causes multiple copies of the same account with potential inconsistent data and increased complexity in merging the records later.

Current Duplicate Search



With the current system landscape, Benny has a difficult time determining if the retailer already exists. This leads Benny to create a duplicate retailer in his ERP.

Future Duplicate Search



With a Customer MDM solution in place, all data is ingested by STEP. This allows other systems to make Find Similar calls to STEP to determine if potential duplicates already exist. Benny can then make an informed decision to either update an existing record accordingly or create a new record.

Find Similar Preventing Duplicates

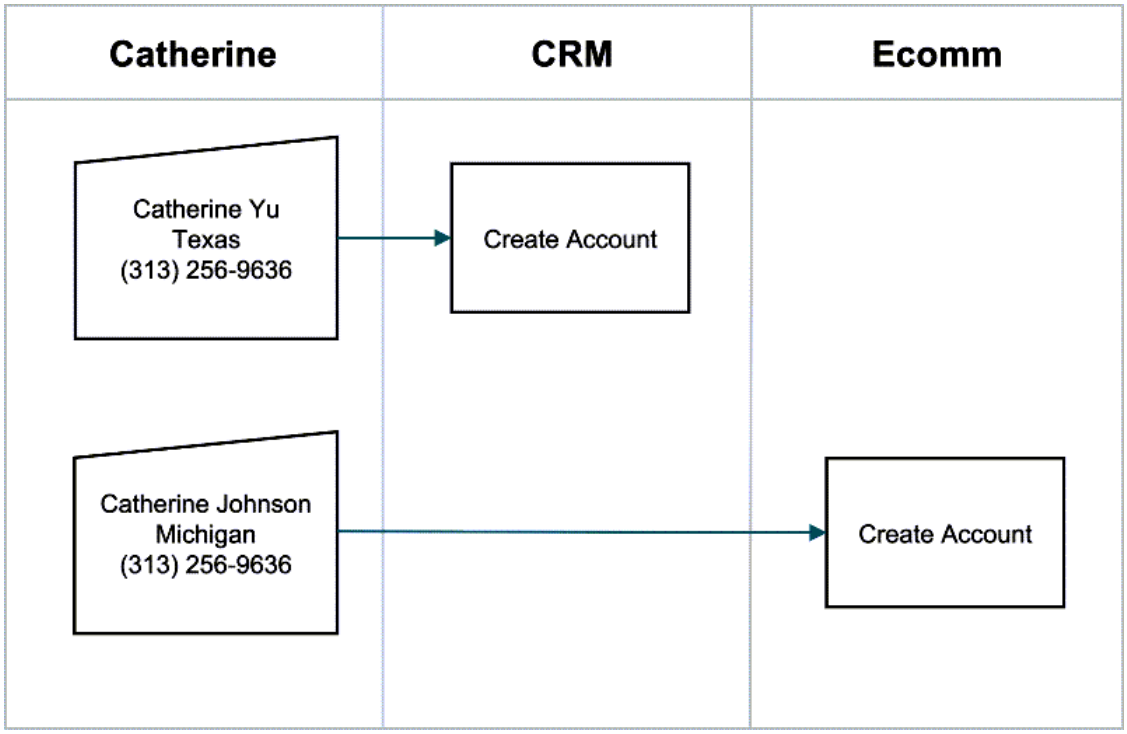
Catherine Yu signed up in-store for an ACME loyalty account to take advantage of an in-store promotion. Catherine is engaged and living in Texas. She is across the country from her fiancé who is in Michigan. Catherine signed up for the loyalty program using her maiden name, her Texas address, and her lifelong cell phone number.

After she's married, Catherine changed her name and moved to Michigan with her new husband. She goes shopping online and forgets she had signed up for the loyalty program and attempts to sign up again. Catherine uses her new last name, her Michigan address, and her lifelong cell phone number to sign up.

Without the exact same data, it is impossible for the CRM to find Catherine's loyalty account and previous purchases.

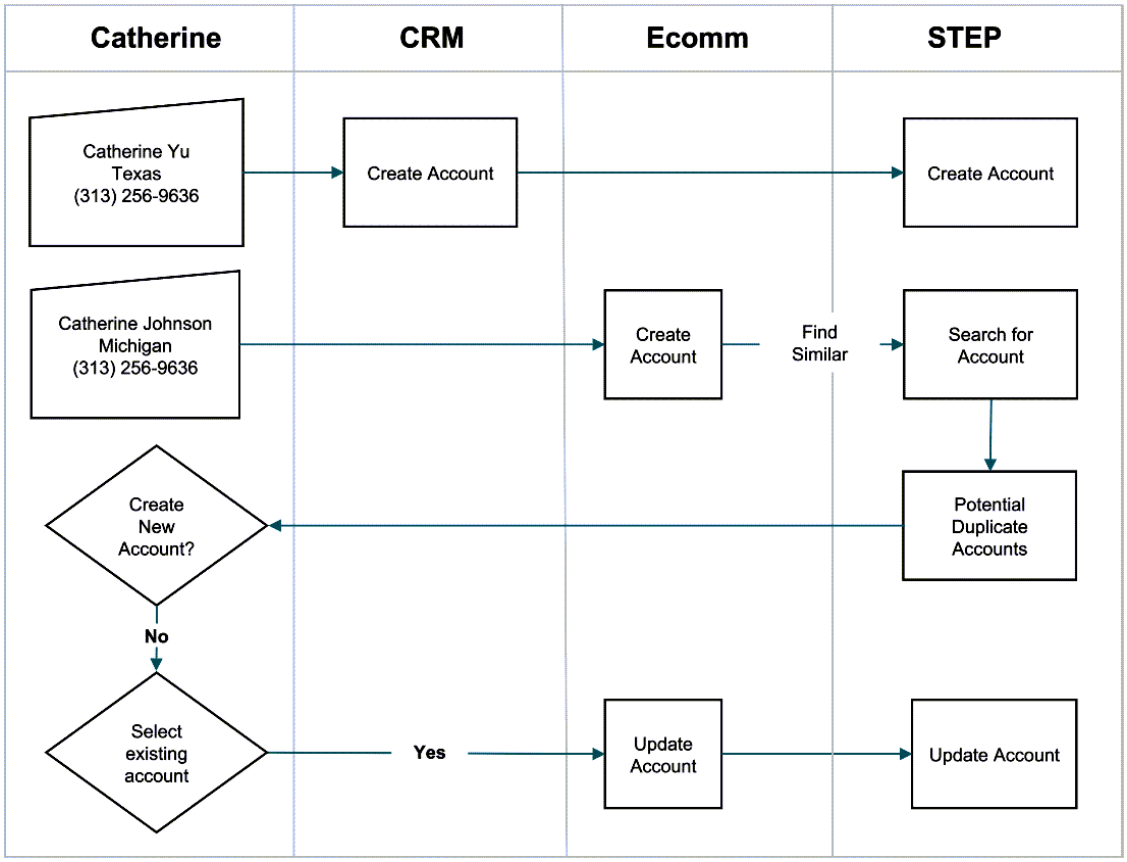
Catherine creates a duplicate loyalty account. Flyers and promotions are now being sent to two addresses. ACME missed cross sell and upsell opportunities because disjointed records make it impossible to identify Catherine's purchase history.

Current Duplicate Search



With the current system landscape, Catherine cannot see that a loyalty account exists under her maiden name. ACME has a disjointed view of who Catherine is and the products she's purchased.

To-Be Duplicate Search



The Customer MDM system can identify Catherine's existing loyalty account for other systems through the Find Similar call. Catherine can identify and update her existing loyalty account preventing a duplicate loyalty account from being created.

Find Similar Examples

Below are examples for an Organizational Customer, a Contact Person, and an Individual Customer, which are included as part of the initial configuration. Households have been excluded as they are generally not created in systems external to the Customer MDM system.

- Organizational Customer
- Contact Person
- Individual Customer
- Supplier

Organizational Customer

The Organizational Customer Find Similar call uses the Organization Matching Algorithm to identify potential duplicate Organization Customers.

Required values to identify match:

- Name
- Address

Request

See the online version of this topic to view the web service request.

Response

See the online version of this topic to view the web service response.

Contact Person

The Contact Person Find Similar call uses the Contact Person Matching Algorithm to identify potential duplicate records.

Required values to identify match:

- Name
- Phone number

The GetSimilarContactOrgID must be bound to the match code for Organization IDs to be compared.

Request

See the online version of this topic to view the web service request.

Response

See the online version of this topic to view the web service response.

Individual Customer

The Individual Customer Find Similar call uses the IndividualMatchingAlgorithm to identify potential duplicate Individual Customers.

Required values for successful validation

- Email

Request

See the online version of this topic to view the web service request.

Response

See the online version of this topic to view the web service response.

Supplier

The Supplier Find Similar call uses the Supplier Matching Algorithm to identify potential duplicate supplier records.

Required values to identify match:

- Legal name
- Address
- Source Record ID

Request

See the online version of this topic to view the web service response.

Response

See the online version of this topic to view the web service response.

Data Flow for Centralized MDM in B2B Solutions

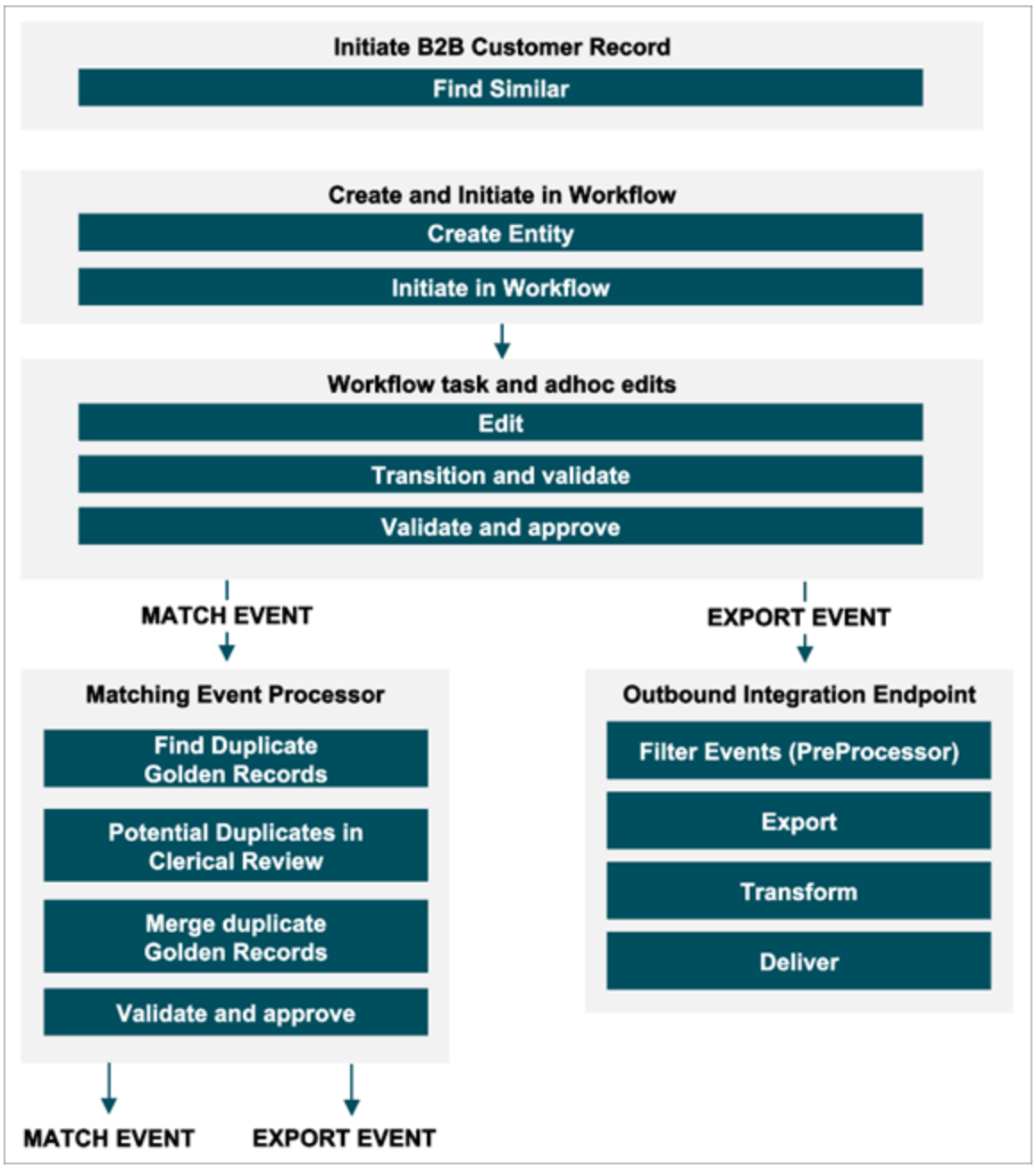
In a centralized implementation methodology, origination and stewardship / maintenance of entity records reside within MDM. While many different types of data are often associated with entity records, it is important to note that the role of MDM is to capture the customer master data that drives operational processes within an organization. Within this section, we will focus primarily on centralized methodology for Business-to-Business (B2B) organizations within the Consumer Packaged Goods (CPG) industry.

CPG companies source from a multitude of suppliers to manage a wide array of products, brands, categories, and channels, while serving diverse customer segments across multiple geographic regions worldwide. Their end consumers of products are either individual consumers or other businesses such as hospitals, schools, restaurants, etc.

While a major business cost is the procurement of raw materials, goods, and services from various suppliers, most of the revenue is generated by selling through a network of retail chains, distributors, and master wholesalers. These parties are who the CPG company considers its customers. Any CPG company has the following operational functions as part of the sales side of their business:

- Sales
- Distribution
- Accounting and finance

To meet the needs of such operational functions, MDM is implemented to enforce and guide the organization in onboarding suppliers and/or customers by providing a means to enrich, govern, and maintain their entity data in a centralized manner.



Initiate B2B Customer Record

A new supplier entity record request is initiated within MDM through the Web UI interface. Initial minimal data provided may include the legal name of the entity being onboarded. At this point, two options are provided in the Web UI:

- Find Similar – to ensure no other record with a similar legal name already exists, the user is presented with matching candidates if any are found.
- Save – If no matching candidate is found, the user may save as a new record.

For details on Find Similar, see the **Find Similar** topic in the **Web User Interfaces / Web UI Setup and User Guide** documentation.

Create and Initiate in Workflow

Once a record is created, it is automatically initiated into a workflow.

Workflow Task and Ad Hoc Edits

By leveraging STEP Workflows, the operational processes such as onboarding, enrichment, and maintenance of master data may be closely owned by various departmental stakeholders across the organization.

The onboarding workflow allows for detailed enrichment of the new entity record as well as various business conditions to ensure data integrity is upheld.

Users may be presented with the following options:

- Save – Edits made to the new record may be saved. It is not uncommon that the save button executes data verification business action(s), such as address verification.
- Submit – Saved edits to a new record may then be submitted to the next state of the workflow. Data validation may also take place on submission to the next state.
- Delete – User has the option to delete the new record if necessary.

For more information, see the **Workflows in Web UI** topic in the **Web User Interfaces / Web UI Setup and User Guide** documentation.

Note: All user actions available are dependent upon the user privileges afforded to said user. It is recommended that privileges are afforded to users based on their role. For example, a logistics specialist cannot delete a record. Only the MDM Specialist or some other system administrative role can be granted all privileges. For more information, see the Privilege Rules topic in the System Setup / Super User Guide documentation.

For ad hoc edits, like edits to an existing record outside of a workflow, the user is presented with the following options:

- Save – Edits made to the new record may be saved. It is not uncommon that the save button executes data verification business action(s), such as address verification. Data validation may also take place at this point to ensure data integrity is upheld.
- Request Approval – Approval of changes to the entity record should be the responsibility of the MDM Specialist. This action will initiate the record into an Approval workflow for the approval. Further data validations may take place within this workflow to ensure data integrity.

Upon final approval of the entity record, a match event is generated for the matching event processor to execute matching with the newly created record.

Approved Workspace

Data promoted to the approved workspace should be considered as having been validated and completed. Approved entity data may be synchronized to a downstream system to drive external business processes and/or analytics.

Matching Event Processor

Whenever entity records are approved, this will trigger a match event. The event will trigger the match algorithm to execute and compare the new or existing golden record with the entire dataset of entity records.

Prior to matching, the entity record under treatment will have its match codes updated. During matching, data within the Approved workspace is to be evaluated. At this juncture, there are two solution paths that may be taken; the decision of which should be determined using the business requirements gathered by the implementation team.

Merge and Auto-Approve Matches

In the event that a resulting match score is above the auto-merge threshold, candidates are merged by the system. Survivorship rules will update the merged record within the Main workspace and auto-approve the changes to promote to the Approved workspace.

There may still be a clerical review threshold should it be dictated by the business requirements.

Review all Matches

All match scores above a certain threshold are considered as potential matches and candidates are manually reviewed. Survivorship rules will auto-initiate the potential candidates into an Approval / Clerical Review workflow to be reviewed by an MDM specialist.

Match scores that fall beneath a certain threshold may be automatically considered as non-matches.

Outbound Integration Endpoint

Whenever entity records are approved, an export event will trigger. This allows the mastered data to be synchronized with various downstream systems.

Data Flow for Supplier Self-Service

As a natural evolution of centralized onboarding of master data, providing the means for suppliers to create and enrich their own profile, or 'self-service,' should be treated as an extension of centralized Supplier MDM.

For large enterprise corporations, there are various benefits in providing their suppliers with the means to self-service their master data. These benefits may include streamlining the maintenance of large amounts of supplier data while also maintaining a high degree of governance through business and compliance validations. Another benefit may be to provide a collaborative interface where internal operation teams may interact with suppliers to ensure data is accurate and up to date.

Important: See the **Solution Enablement Customer MDM** online help version of the **Data Flow for Supplier Self-Service** topic for all of the swimlane diagrams.

Procurement Manager: Self-Service

As the main user persona that interfaces with establishing and maintain relationships with new and existing suppliers, respectively, the Procurement Manager is the first internal user to conduct supplier on-boarding-related activities.

See the **Solution Enablement Customer MDM** online help version of the **Data Flow for Supplier Self-Service** topic for the swimlane diagram.

Procurement Manager Creates Supplier Entity for New Supplier

To create a new entity for a new supplier in the Web UI, the procurement manager clicks on the ‘New Supplier’ link on the home page widget for ‘Supplier Onboarding’. On the Initiate New Supplier screen, the procurement manager adds the legal name of the entity that is to be created.

MDM conducts a ‘Find Similar’ action to search for existing supplier entities that match the provided name. The ‘Find Similar’ component uses a basic match algorithm to search for existing entities with the same or similar legal name. If no match is found, the procurement manager is directed to the ‘Procurement Enrichment’ workflow state and presented with all supplier attribute fields categorized across multiple tabs in the Web UI.

At this point, the procurement manager must provide a contact email in the ‘Initial Admin Email’ field on ‘Procurement Enrichment’ tab.

Once the user fills out the necessary information, by clicking on 'Create New Supplier' a series of business logic is executed by MDM.

First, a supplier classification structure that is specific to the new supplier is created. This structure is two levels deep – a single parent classification with sub-classifications which are intended to be referenced by different node types. The ‘Supplier Entities Classification’ is only referenced by supplier entities. All entities onboarded by a supplier are referenced into that supplier’s entity classification. This ensures that only users of this supplier can see their supplier entities.

For more information, see the **Supplier Self-Service Privileges** section under the **Supplier and Classification Object Types** topic of this document.

Second, credentials (username and password) for the first Supplier Admin User are created. Only one credential is automatically created. CreatesSubsequent users by the supplier admin user through the Web UI.

Third, the credentials are emailed to the contact person that was just created by leveraging the 'Initial Admin Email' address provided. The supplier admin user may then log into the Web UI with the new credentials and review / enrich the new supplier entity or create additional entities.

Once the new entity has been enriched by the supplier user, clicking on the 'Submit' button transitions the task to subsequent workflow states for internal user personas to review, further enrich, and validate the data. These internal personas include the procurement manager, finance specialist, logistics specialist, and the MDM specialist.

The MDM specialist reviews all data and ensures business validations are adhered to. If any details require further attention, the MDM specialist may provide a question / comment and submit the task back to any of the preceding workflow states. Once the MDM specialist approves the new supplier entity, the MDM system re-executes all business validations and the new entity exits the workflow.

By using a workflow which caters to different various personas, MDM provides a collaborative user experience in which proper governance may be attained to present an efficient supplier onboarding user experience.

Procurement Manager Creates Supplier Entity for Existing Supplier

To create a new entity for an existing supplier in the Web UI, the procurement manager clicks the 'New Supplier' link on the home page widget for 'Supplier Onboarding.' The procurement manager is then presented with the Initiate New Supplier screen which requires the legal name of the entity that is to be created.

MDM conducts a 'Find Similar' action to search for existing supplier entities that match the provided name. The 'Find Similar' component utilizes a basic match algorithm to search for existing entities with the same or similar legal name. If no match is found, then the procurement manager is directed to the 'Procurement Enrichment' workflow state and presented with all supplier attribute fields categorized across multiple tabs in the Web UI.

Once the user fills out the necessary information, by clicking on 'Link to Existing Supplier' a node-picker screen appears for the user to select the Supplier classification this new entity should be referenced to.

Once the new entity has been linked to an existing supplier, clicking on the 'Submit' button transitions the task to subsequent workflow states for internal user personas to review, further enrich, and validate the data. These internal personas include the procurement manager, finance specialist, logistics specialist, and the MDM specialist.

The MDM specialist reviews all data and ensures business validations are adhered to. If any details require further attention, the MDM specialist may provide a question / comment and submit the task back to any of the preceding workflow states. Once the MDM specialist approves the new supplier entity, the MDM system re-executes all business validations and the new entity exits the workflow.

By using a workflow which caters to different various personas, MDM provides a collaborative user experience in which proper governance may be attained to present an efficient supplier onboarding user experience.

Supplier User: Self-Service

See the **Solution Enablement Customer MDM** online help version of the **Data Flow for Supplier Self-Service** topic for the swimlane diagram.

Supplier Creates New Supplier Entity

Once the supplier admin user receives their user credential, they may log onto the Web UI.

To create a new entity, the supplier user navigates to the 'Supplier Onboarding' widget and clicks on 'New Supplier'. From there, the user is taken to the Initiate screen and be required to provide the legal name of the entity to create.

MDM conducts a 'Find Similar' action to search for existing supplier entities that match the provided name. The 'Find Similar' component utilizes a basic match algorithm to search for existing entities with the same or similar legal name. If no match is found, then the user is directed to the 'Initial Enrichment' workflow state and presented with all supplier attribute fields categorized across multiple tabs in the Web UI.

Once the new entity has been enriched by the supplier user, clicking on the 'Submit' button transitions the task to subsequent workflow states for internal user personas to review, further enrich, and validate the data. These internal personas include the procurement manager, finance specialist, logistics specialist, and the MDM specialist.

The MDM specialist reviews all data and ensures business validations are adhered to. If any details require further attention, the MDM specialist may provide a question / comment and submit the task back to any of the preceding workflow states. Once the MDM specialist approves the new supplier entity, the MDM system re-executes all business validations and the new entity exits the workflow.

By using a workflow which caters to different various personas, MDM provides a collaborative user experience in which proper governance may be attained to present an efficient supplier-initiated onboarding process.

Supplier Creates New Contact Person

When creating new suppliers, it is important to associate contacts which may also need to be created. Creating a new contact for a newly created or existing supplier is accomplished in a similar fashion as creating new supplier entities. However, less internal users are required to review and approve new contacts.

Supplier objects must first be created prior to referencing a new or existing contact person.

See the **Solution Enablement Customer MDM** online help version of the **Data Flow for Supplier Self-Service** topic for the swimlane diagram.

Supplier Creates New Users

As Supplier Admin Users receives their MDM credentials, other users within their supplier organization may require access to MDM to carry out tasks. To accommodate this, the initial configurations provide a way for Supplier Admin Users to create additional users within the Web UI.

Supplier Admins can not only create new users but also may grant administrative privileges to both new and existing users, as well as revoke administrative privileges to existing users.

See the **Solution Enablement Customer MDM** online help version of the **Data Flow for Supplier Self-Service** topic for the swimlane diagram.

Supplier Updates Existing Supplier Entity

Upon logging into the Web UI, to maintain / make updates to an existing entity, the supplier user searches for the entity within MDM. Once found, the entity displays as read-only. This intended behavior is to ensure supplier-initiated updates are reviewed and validated by the respective internal persona and business rules. To make edits, the user submits the entity into the 'Initial Enrichment' state of the workflow. To do so, a button on the supplier details page is configured to initiate the current entity into a pre-configured workflow.

Once all necessary updates are made to the entity, the supplier user must submit the revised entity to transition the task to subsequent workflow states for internal user personas to review, further enrich, and validate the data. These internal personas include the procurement manager, finance specialist, logistics specialist, and the MDM specialist.

The MDM specialist reviews all data and ensures business validations are adhered to. If any details require further attention, the MDM specialist may provide a question / comment and submit the task back to any of the preceding workflow states. Once the MDM specialist approves the new supplier entity, the MDM system re-executes all business validations and the new entity exits the workflow.

By using a workflow which caters to different various personas, MDM provides a collaborative user experience in which proper governance may be attained to present an efficient supplier-initiated maintenance process.

Supplier Reviews Updates to Existing Entities

Upon logging into the Web UI, the supplier admin user may navigate to the 'Review' state of the Supplier Onboarding widget. Tasks within this state are changes (payment terms, etc.) made by procurement which require review by the supplier.

Once the supplier user reviews and approves the review tasks, the task must be submitted to subsequent workflow states for internal user personas to review, further enrich, and validate the data. These internal personas include the procurement manager, finance specialist, logistics specialist, and the MDM specialist.

Procurement Manager: No Self-Service

In a non-self-service scenario, to create a new supplier entity in the Web UI the procurement manager clicks on 'New Supplier' on the home page widget for 'Supplier Onboarding'. The procurement manager is then presented with the Initiate New Supplier screen which requires the legal name of the entity that is to be created.

See the **Solution Enablement Customer MDM** online help version of the **Data Flow for Supplier Self-Service** topic for the swimlane diagram.

MDM conducts a ‘Find Similar’ action to search for existing supplier entities that match the provided name. The ‘Find Similar’ component utilizes a basic match algorithm to search for existing entities with the same or similar legal name. If no match is found, then the procurement manager is directed to the ‘Procurement Enrichment’ workflow state and presented with all supplier attribute fields categorized across multiple tabs in the Web UI.

Once the new entity has been enriched by the procurement manager, clicking on the ‘Submit’ button transitions the task to subsequent workflow states for internal user personas to review, further enrich, and validate the data. These internal personas include the finance specialist, logistics specialist, and the MDM specialist.

The MDM specialist reviews all data and ensures business validations are adhered to. If any details require further attention, the MDM specialist may provide a question / comment and submit the task back to any of the preceding workflow states. Once the MDM specialist approves the new supplier entity, the MDM system re-executes all business validations and the new entity exits the workflow.

By using a workflow which caters to different various personas, MDM provides a collaborative user experience in which proper governance may be attained to present an efficient supplier onboarding user experience.

Integrations

Through third-party tools, STEP expands the potential level of data quality. The following topics discuss use cases for these integrations.

- Address Validation
- Dun & Bradstreet
- Experian Data Quality Integration
- SAP Business Partner Integration

Address Validation

Loqate validations can occur at three different points when dealing with customer records:

- During Import of Source Records
- After surviving attribute values have been determined for a newly created Golden Record
- When updating existing Golden Records (ad hoc or within a workflow)

Note: It is recommended to validate address values upon Import and after Survivorship.

For more general information on Loqate, see the **Loqate Integration** section of the **Data Integration** documentation.

Validate on Import

By calling Loqate during Import, each Source Record will have its address values standardized before matching takes place. Raw values taken from Source Systems commonly include typos or inconsistent formatting. By standardizing these values beforehand, the matching algorithm can match records more accurately.

Validate on Survivorship

Validating addresses after survivorship ensures that the surviving address values are standardized. However, because address matching is less reliable when matching raw values taken from Source Systems, it is recommended to validate on Import as well.

Validate Updates to Existing Records

A Data Steward can trigger Loqate for single records or when running a bulk update for a collection. By using either of the maintenance methods listed below, address values can be standardized when records are updated.

Maintenance Validation Methods:

- A 'Run Business Rule' button on a Node Details screen
- Running a Bulk Update configuration (an Address Validation business rule can be configured for bulk updates)
- An Event processor can monitor for changes to the address inputs. When changes are detected, an event is generated to have Loqate re-standardize the updated address.

By using Bulk Update, existing records can have their address values standardized whenever their data is updated.

Validate via Web Service Request

Address standardization can be triggered from an external system which will generate a request and receive a response containing the standardized address information. This requires no direct interaction with STEP and provides a better user experience in allowing external systems to use Loqate's address standardization feature.

In exposing address validation to external applications, address data points may be standardized prior to flowing into Customer MDM, which further ensures that data is trusted and verified.

Note: All customers that have purchased a Loqate license, either cloud or local, will have access to this service.

Validate in Workflow

In addition to the aforementioned areas where address validation may take place, it is common to also require address validation within the context of a defined business or operational workflow. Depending on the nature and complexity of the workflow, the Standardize Address business action may be executed on exit of a particular workflow state in which address information may be altered.

Validate for Match Tuning

The match tuning process uses input address values while the import allows for matching to run against standardized values.

Unexpected match results can occur since an algorithm that was tuned using the input address comparison scores is then used to evaluate standardized values. Instead, use the Loqate Command Line Tool to output standardized address via Excel, which can then be used within match tuning.

For more information, see the **Loqate Local Command Line Tool** topic in the **Data Integration** documentation and see the **Match Tuning** topic of the **Matching, Linking, and Merging** documentation.

Dun & Bradstreet Integration

Dun & Bradstreet is a leading authority in augmenting and enhancing customer and supplier data with industry leading company information within MDM.

Stibo Systems' MDM includes integration with D&B Match- and D&B Company Profile services.

The D&B Match service integration matches organization master data with D&B's database of over 265 million organizational records world-wide and return relevant match candidates. Information such as DUNS number, address and contact information and more are provided using the Basic match.

The D&B Company Profile service provides over 220 attributes of business insight related to marketing and sales information, company structures, financial information, key executive contact information and more.

Use Case for Customer MDM

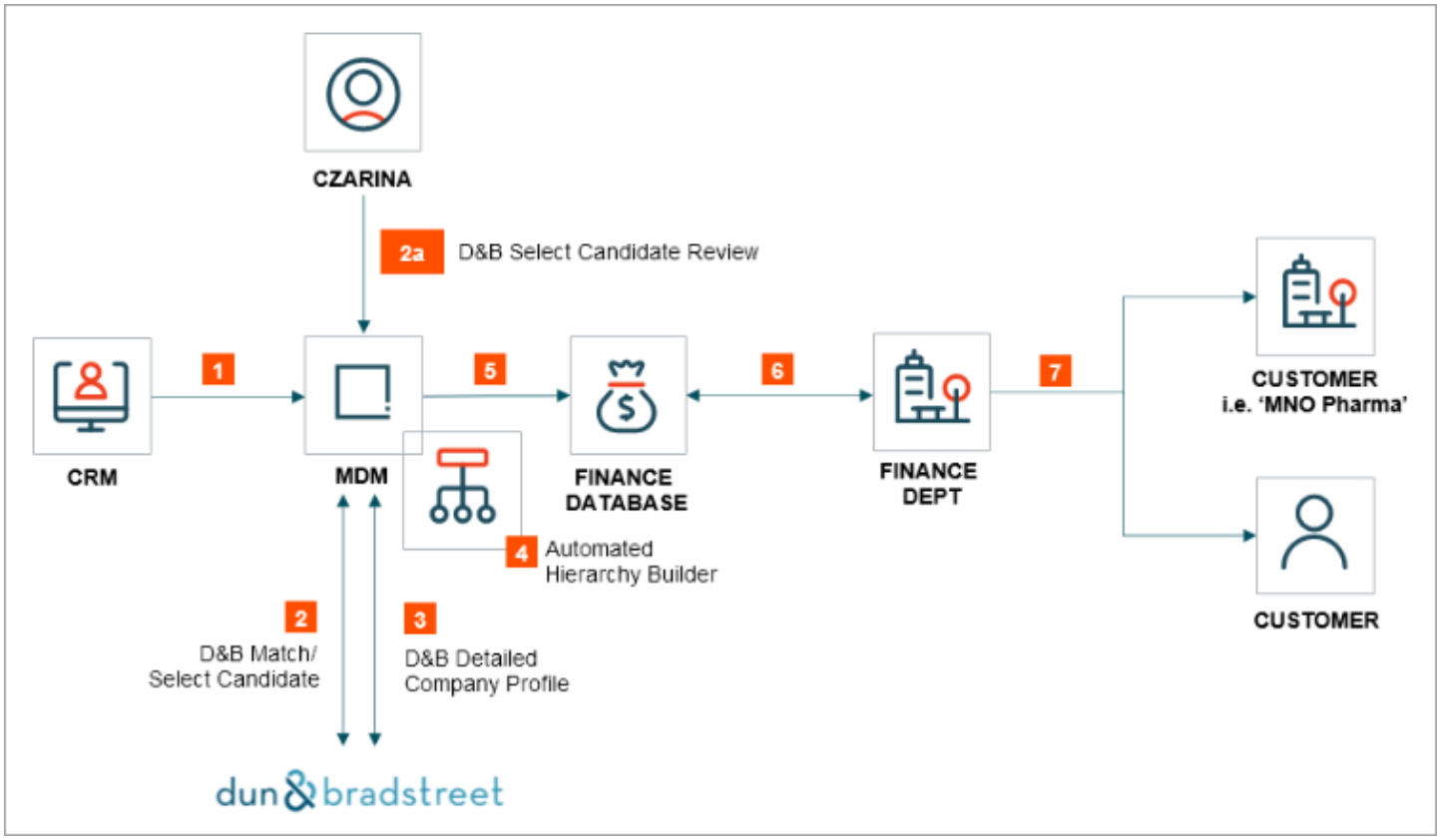
ACME Company is a large, multi-national holding company with subsidiaries that spans across various continents. ACME's customer portfolio consists of many business-to-business (B2B) accounts which operate across various industries.

In an initiative to improve overall reporting to benefit various departments such as legal, marketing, and finance, ACME realizes a need for a better understanding of their customers. To accomplish this, ACME has commissioned MDM to be the central repository for B2B customer data by sending their customer information from various source systems to MDM for deduplication and further enrichment.

Czarina is a data steward for ACME company, whose responsibilities include ensuring ACME's master customer data, which drives all reporting and analytics, is accurate and up to date.

Using the Dun & Bradstreet integration offering, MDM firstly obtains the DUNS number for ACME customers using D&B Match by providing basic information such as legal entity name of the organization or a main organization address.

ACME Company's Stibo Systems' MDM implementation will automatically use the information provided by D&B, to identify duplicate customers and to improve various customer hierarchy structures.



Using the Dun & Bradstreet integration offering, the diagram above depicts the flow of information for this particular use case:

1. Source record information from ACME source system(s) (i.e., CRM) are sent to MDM.
2. As Organization Customer records are created, MDM automatically obtains a DUNS number for the customer(s) by running a D&B Match. The request includes basic information such as legal entity name of the organization, address, phone number, tax identifiers, etc. In response, if D&B returns a match, then a DUNS number is provided and associated with the customer record.

   If multiple potential matches are found, D&B will return the list of candidates and require a manual clerical review. It is part of Czarina's responsibility to review and select the most accurate match; once this is completed the correct DUNS number is associated with the customer record.

   Note: The DUNS number is a D&B unique identifier which directly correlates to other government identification protocols such as EIN or Tax Identification Number. This ensures the uniqueness of the customer entity in question.

3. Having obtained the DUNS number, ACME will further enrich their B2B customer data by automatically leveraging the D&B Detailed Company Profile integration.

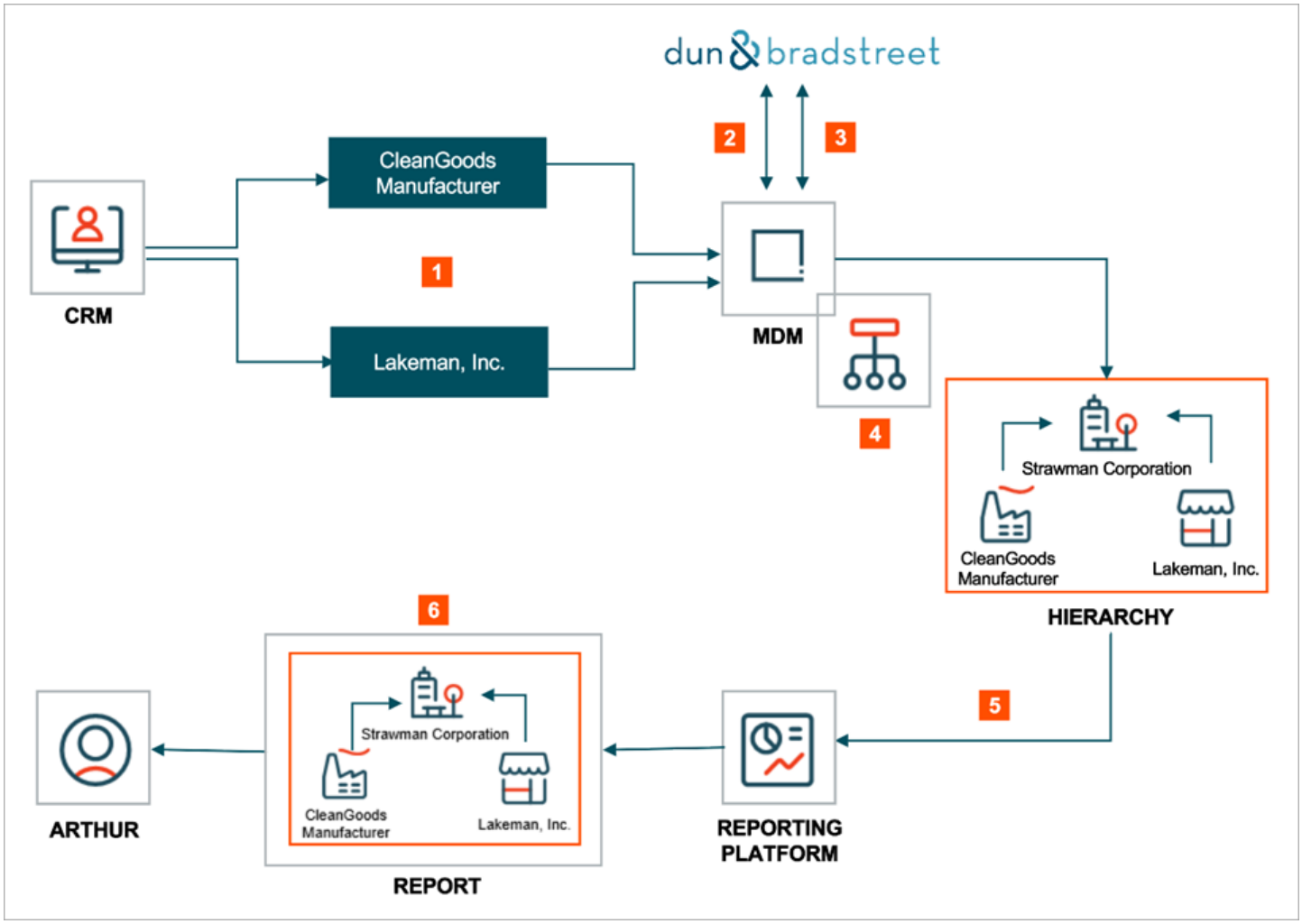
As a global company that operates across multiple industries, ACME plans to utilize this detailed information to build an industry segmentation hierarchy for their customers, which allows the finance department to properly report on, invoice, as well as be made aware of any relevant tax breaks associated to customers within industries.

4. A separate process within MDM is created to automate the creation of industry hierarchies and categorization of customers within such hierarchies. This logic is based on D&B profile data and must execute after D&B Detailed Company Profile enrichment.
5. Once the relevant hierarchy(s) is created, data is then synchronized with external systems (i.e., Finance database) (see details below)
6. From there, the Finance department now has the complete view of their customers with respect to their particular industry and available metadata.
7. The Finance department now can confidently send appropriate invoices and other business transactions.

Use Case for Supplier MDM

ACME Company is a large manufacturer who relies on a large network of suppliers for their raw materials and finished goods.

As part of a company-wide initial to reduce cost both from ACME buy-side and sell-side operations, Arthur, who is on the procurement team and is responsible for managing supplier relations, must be able to leverage third-party Dun & Bradstreet data to identify cost-cutting opportunities.



Using the Dun & Bradstreet integration offering, the diagram above depicts the flow of information:

1. Supplier record information from ACME source system(s) (i.e., CRM) are sent to MDM.
2. As Supplier records are created, MDM automatically obtains a DUNS number for the suppliers by running a D&B Match. The request includes basic information such as legal entity name of the supplier, address, phone number, tax identifiers, etc. In response, if D&B returns a match, then a DUNS number is provided and associated with the supplier record.

   If multiple potential matches are found, D&B will return the list of candidates and require a manual clerical review. It is part of an MDM Specialist's responsibility to review and select the most accurate match; once this is completed the correct DUNS number is associated with the supplier record.

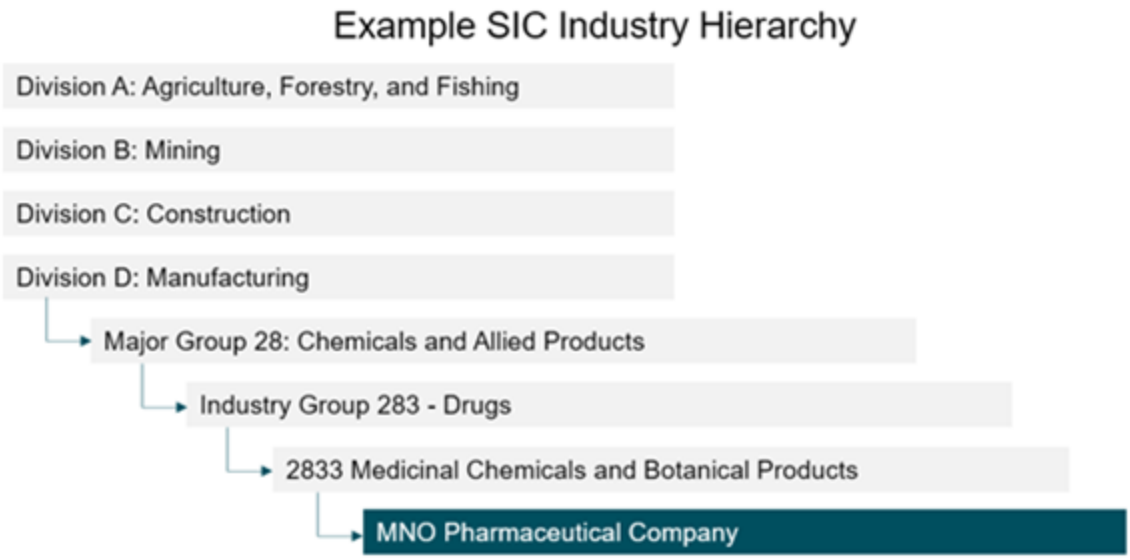
   Note: The DUNS number is a D&B unique identifier which directly correlates to other government identification protocols such as EIN or Tax Identification Number. This ensures the uniqueness of the customer entity in question.

3. Having obtained the DUNS number, ACME will further enrich their supplier data by leveraging the D&B Detailed Company Profile integration.
4. A separate process within MDM is created to automate the creation of legal hierarchies. This logic is based on various DUNS numbers within D&B profile data and must execute after D&B Detailed Company Profile enrichment.
5. Once the legal hierarchy(s) is created, data is then synchronized with external systems (i.e., reporting platform).
6. From there, Arthur may now obtain reports that contain the relevant supplier master data and all the relationships provided by Dun & Bradstreet.

As a global company that operates across multiple industries, ACME plans to utilize this detailed information to build legal hierarchical representations of their existing suppliers' relationships. This allows Arthur to ascertain new suppliers to source and leverage existing relationships to negotiate for lower pricing.

Creation of Hierarchies

Once organization records have been enriched with D&B data, a separate process is initiated to evaluate SIC and NAICS industry codes to automate the building of an industry hierarchy as well as categorization of customers into this hierarchy.



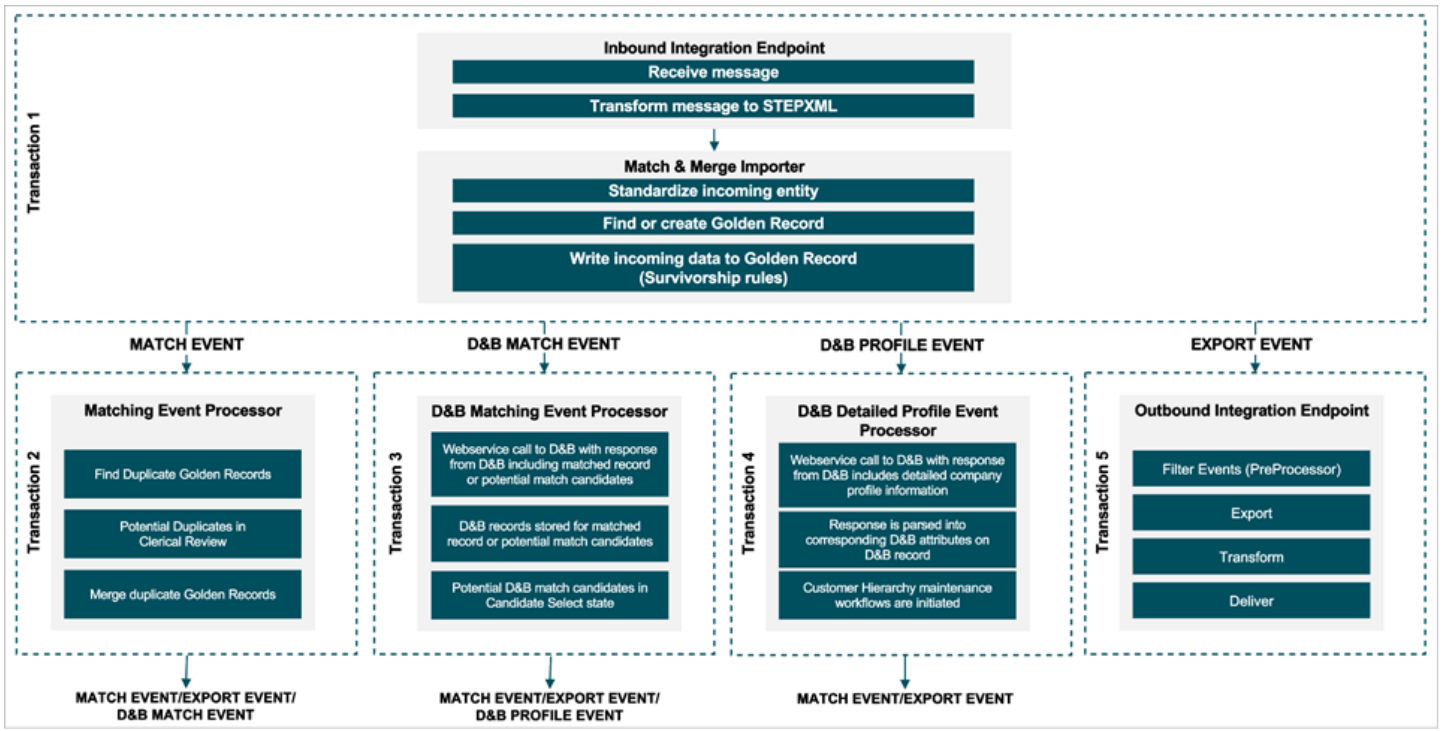
Identifying More Duplicates

As ACME's MDM solution is configured to match customer records on various data provided by D&B, whenever it is available, event triggering of matching algorithms, will automatically identify and merge duplicate entity records that could not previously be identified, due to a lack of common identifiers on the duplicate records.

For further information on MDM's matching capabilities, see the **Matching, Linking, and Merging** documentation.

D&B Data Flow within MDM Application

The progression of data flow for B2B entity records as it is de-duplicated within MDM and also further enriched via Dun & Bradstreet is shown below.



Transaction 1: Inbound Integration Endpoint

An inbound integration endpoint will receive a message and transform it into STEPXML.

Once transformed, the Match & Merge Importer will:

- Execute standardizations and other associated business rules to cleanse the data. Cleansed data allows for the Match algorithm to more accurately identify duplicate records.
- Determine the target Golden Record. For more information, see the **IIEP - Configure Match and Merge Importer Processing Engine** topic of the **Data Exchange** documentation.
- Survivorship rules will write any updates from the incoming data to the golden record. For more information, see the **Golden Records Survivorship Rules** section of the **Matching, Linking, and Merging** documentation.
- To the extent that a Golden Record was created, or an existing Golden Record was updated, events are written to event queues triggering on those changes. In this case, a Matching event processor as well as an Outbound Integration Endpoint.

Transaction 2: Matching Event Processor

When new golden records are created or an existing golden record is updated, this will trigger a matching event, with the purpose of comparing golden records to each other. Prior to matching, the golden record under treatment will have its match codes updated.

Matching golden records may result in a clerical review workflow initiation or a merge, depending on the score. Additionally, survivorship rules should trigger a D&B Match event

For more information, see the **Match and Merge Clerical Review - Merge** topic of the **Matching, Linking, and Merging** documentation.

Transaction 3: D&B Match Event Processor

Golden record events are then triggered for the D&B Match event processor. This triggers a web service call to D&B to obtain a corresponding match to existing entities within D&B database. The D&B response will invoke one of two types of behaviors within MDM.

If a direct match is found, a corresponding D&B entity record is created within MDM and referenced to the golden organization record. Attribute information from the D&B response are then parsed into D&B attributes.

Once a DUNS number is obtained, a D&B Profile event should be triggered for the D&B Detailed Profile Event Processor.

If potential match candidates are returned by D&B, then corresponding D&B entities are created and submitted to the 'Select Candidate' workflow state.

Transaction 4: D&B Profile Event Processor

Once a D&B match has been found, a second D&B event processor sends the D&B record with its corresponding DUNS number back to D&B for retrieval of detailed company profile information.

The D&B response from this call is then parsed into corresponding D&B attributes on the D&B record.

Once a D&B response is received and parsed, two events should be triggered:

- Match Event – in the case that the detailed profile provides more information that may affect matching results, the Matching Event Processor is invoked to re-run the match algorithm.
- Export Event – to synchronize the golden record information with external systems, an export event is triggered for an Outbound Integration Endpoint(s)

Transaction 5: Outbound Integration Endpoint

Whenever Golden Records are created or changed, an export event will trigger. This allows the mastered data to be synced back to the various source systems.

For more information, see the **Outbound Integration Endpoints** topic of the **Data Exchange** documentation.

Experian Data Quality Integration

Experian Data Quality's industry-leading Clean Web Service is an asynchronous validation tool which provides bulk verification of email syntax, format, domain and account for customers' email addresses.

In an age when digital communication is at the forefront, keeping customers' email information accurate and up to date is vital for a variety of reasons including, but not limited to:

- Email is the primary means of communication.
- Maintaining brand awareness among existing and prospective customers.
- Ensures marketing campaigns adequately and efficiently reaches target audiences.
- Protects your brand reputation.
- Maintain good standing status with email providers by reducing the amount of bounce-backs.

Stibo Systems' Customer MDM offers an integration with Experian Clean Web Service to validate emails in bulk for customers, contacts, and prospects.

Use Case

Pearl is a data steward who works closely with the marketing department to build seasonal campaigns and corresponding collateral. Today, the use of digital marketing platform to drive sales, increase brand awareness, and promote customer loyalty is widespread. This implies that for marketing strategies to succeed, having the proper contact information for customers and prospective customers is paramount.

ACME's marketing department relies on email contacts provided by existing customers within loyalty programs as well as other marketing initiatives. To assist with ensuring success of the campaign, Pearl is tasked with making sure that email addresses are valid within the customer master (Customer MDM). By having an analysis of the accuracy of email addresses, ACME may then determine the cause of any discrepancies with the existing workflow of gathering email contacts.

In leveraging Customer MDM's email validation solution using Experian, Pearl can validate email addresses in bulk. Validated email data may then be synchronized with the marketing team's CRM application. Having validated their email contacts, Pearl can now search for and identify the various reasons email address may be bad or unreachable. For more information, see the **Experian Email Validation Integration** topic of the **Data Integration** documentation.

SAP Business Partner Integration

This section outlines an integration with the SAP Business Partner API implemented as SOAP Services: Services for Business Partner (SOAP) - SAP Help Portal.

The communication pattern is as follows:

- Send request
- Receive asynchronous confirmation with status of processing of request

The API is comprised of the following services:

- BP-Replication from STEP to SAP
- BP-Replication from SAP to STEP
- BP-Replication Confirmation from SAP to STEP
- BP-Replication Confirmation from STEP to SAP
- RL-Replication from STEP to SAP
- RL-Replication from SAP to STEP
- RL-Replication Confirmation from SAP to STEP
- RL-Replication Confirmation from STEP to SAP

For more information on an SAP Business Partner Integration, see the following topics in this documentation:

- **SAP Integration Supported MDM Implementation Methodologies**
- **SAP Publishing From STEP**
- **SAP Publishing To STEP**
- **SAP Survivorship Rules**

SAP Integration Supported MDM Implementation Methodologies

The following implementation methodologies are supported for use with SAP. For the complete list of MDM implementation methodologies, see the **MDM Overview** topic.

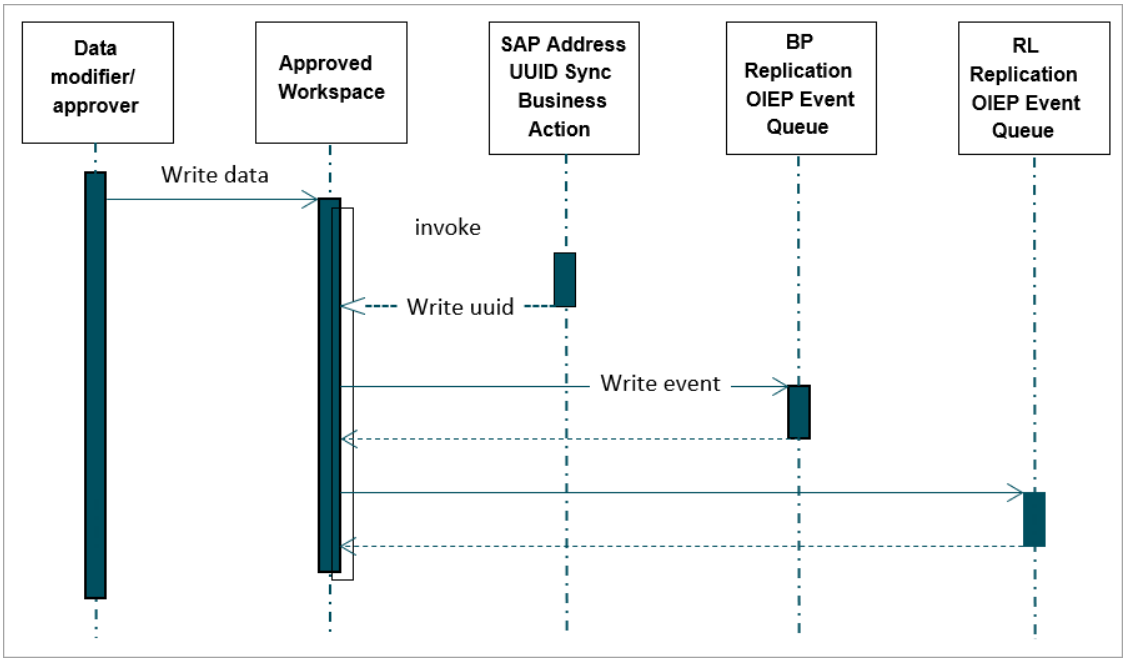
- **Consolidation** - BP records are authored in upstream SAP systems and then published to STEP, where they are matched and merged into golden records with preservation of source record identifiers. After the consolidation, the golden records may be enriched with additional information and published to downstream systems.
- **Centralized Onboarding** - BP records are created and maintained in STEP through workflows. As records are approved, new records and updates to existing records are published to multiple downstream SAP systems. As BP records are created and updated, duplicates may be identified and merged automatically or through clerical review. As BP records merge, the deactivation of the non-survivor is communicated the downstream SAP systems and references to the deactivated records from other master data records are moved to the surviving record, which is also communicated to the downstream systems.

SAP Publishing From STEP

Publishing of data to SAP is handled with an event-based Outbound Integration Endpoint (OIEP).

Inbound Integration Endpoints (IIEPs) with Match and Merge Processing Engine are used to process acknowledgment messages from SAP to STEP.

A separate OIEP is configured for BP replication and RL (Relationship Link) replication.

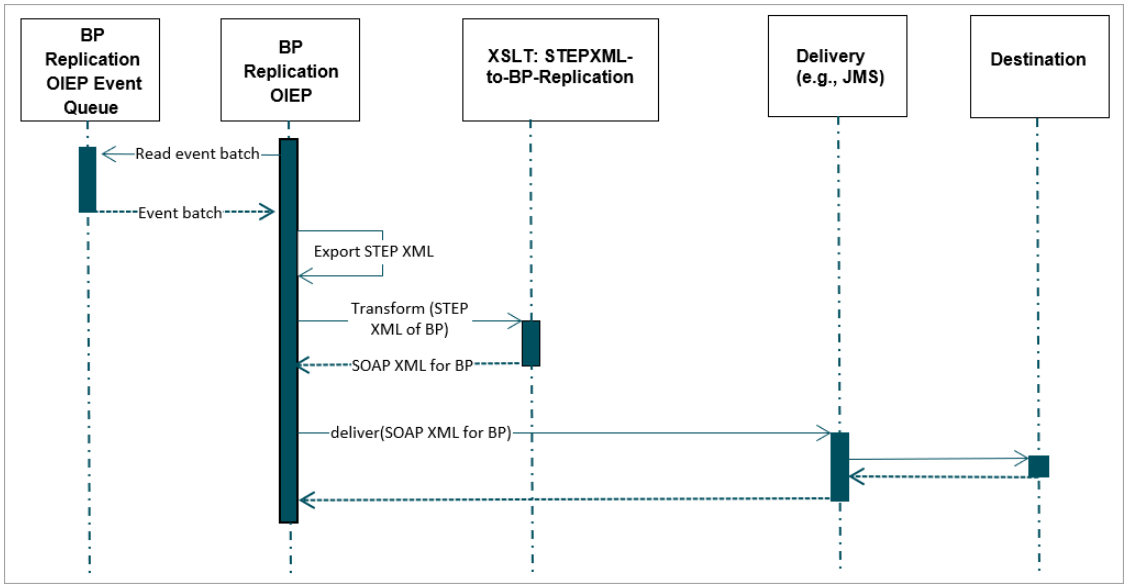


As the BP replication reads the events in batches, it exports the batch of entities and uses XSLT to transform this into a SOAP message. The SAP BP / RL replication APIs support receiving and acknowledging multiple BPs in one message.

The event batch size of the OIEP should be limited so that the message size does not grow too big. For example, the STEP XSLT processors are limited to handle files not bigger than a certain size. See the **OIEP - Post-Processor - Transformation by XSLT** topic in the **Data Exchange** documentation.

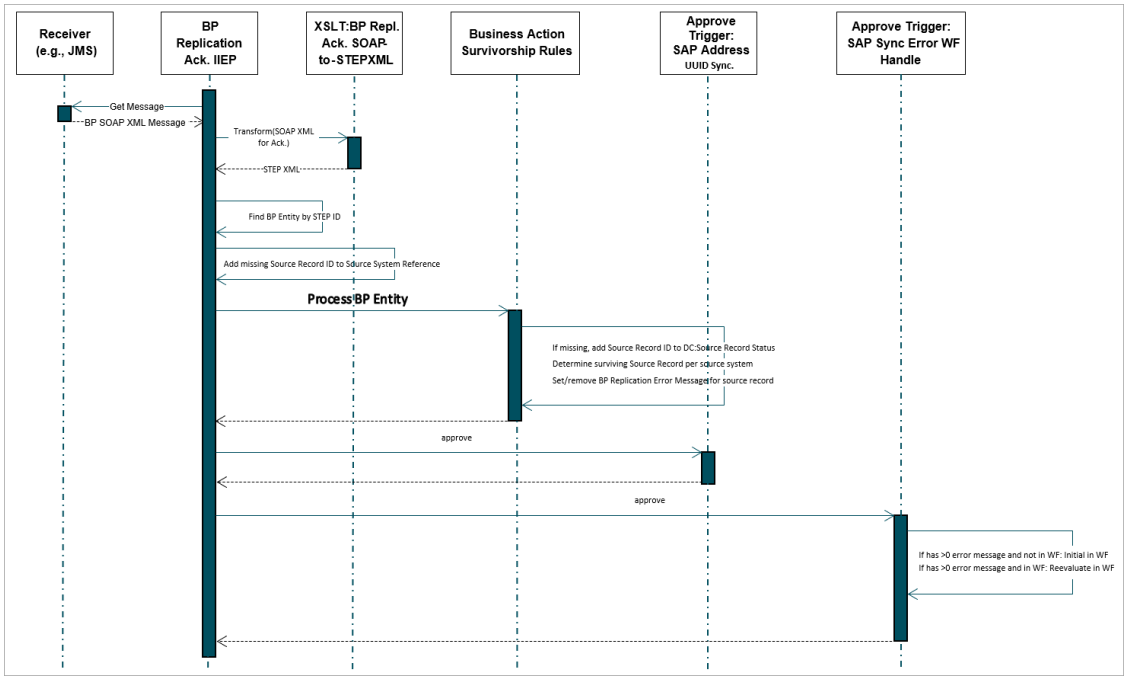
The SOAP messages are delivered to SAP via middleware using, for example, REST Upload, JMS delivery, FTP, etc.

This diagram is best viewed in online help.



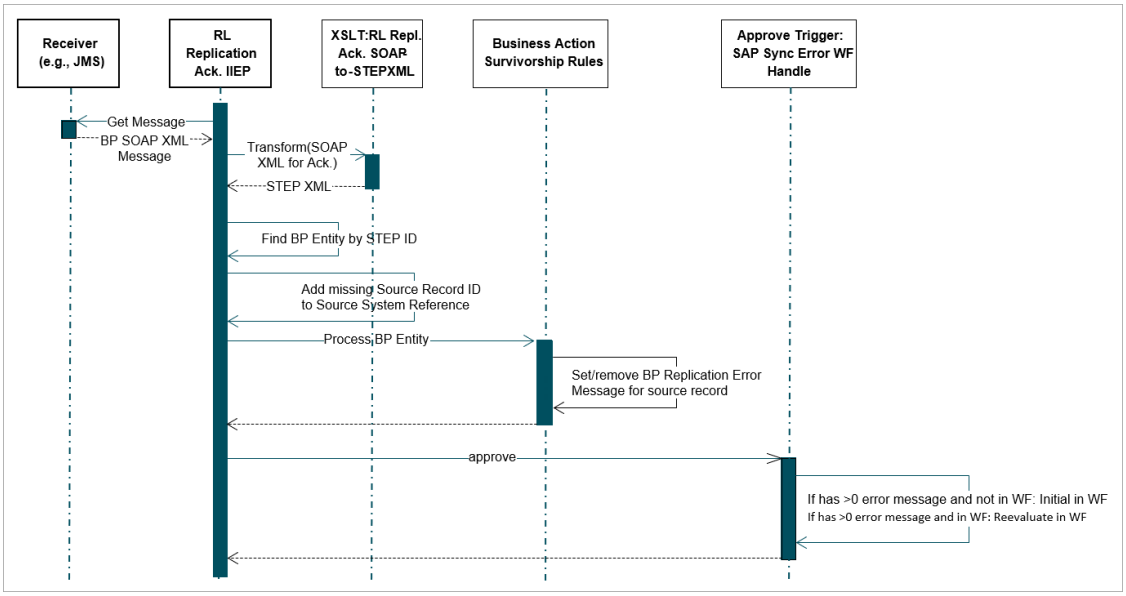
As BP replication acknowledgment comes back from SAP, missing identifiers and status information is updated on the BP entity.

This diagram is best viewed in online help.



The RL replication uses the same pattern to export, transform, and deliver messages to SAP as the BP replication.

This diagram is best viewed in online help.



As the RL replication message comes back from STEP, the status is written to STEP.

Correlation of Messages and Data Objects

The async confirmation message can be correlated to the request by a message ID provided by the sender in the request.

The async confirmation message also contains the identifiers of the BP object. Both the sender provider and the receiver created identifiers are included.

Address UUIDs generated by the receiver are not part of the async confirmation.

STEP Solution

Since the STEP Integration Endpoints do not support async message id correlation patterns, the sequence diagrams above outline a simpler solution based on correlating async acknowledgment messages to BP entities. Consequently, the STEP ID of a BP entity published from STEP to SAP must be stored in SAP and must be part of the async acknowledgment message. This is possible since SAP allows BP objects to have an ID that is unique for a specific sender system (STEP).

It is also recommended that each SAP system generate its own BP IDs and that the MDM system store those IDs as Source Record IDs according to the Merge Golden Record Component model.

Since IDs of addresses are not part of the async confirmation message, to refer to the same addresses in later requests, you must send to SAP the ID given to an address by STEP. The ID of an address in SAP is a UUID, which is a generated identifier that can and must be considered globally unique. Therefore, STEP can use the [uuid] STEP ID Pattern for the address, email and phone data container type and the STEP ID of the address can be used in SAP. However, as multiple records in the same SAP system merge into the same Golden Record, the same surviving addresses, emails and phones must be represented by multiple SAP UUID. See the details of handling this in SAP Address Source Record Relations.

Email and Phone Number Approve Trigger

For information on the modeling of Address, Email and Phone Number, see the **Emails and Phone Numbers and the Approve Trigger** section of the **SAP Business Partners and Enterprise Structure Definitions** topic.

This approval trigger is invoked by the Match and Merge Importer, by the Matching Event Processor, and after manual merge in the Clerical Review Workflow.

Use the following specific logic:

- Group addresses, emails and phones by usages and validity dates
  - Each address is a group
  - Allocate the zero-to-many addresses for each email and phone by usage and validity period
  - Make a separate group for each email and phone that are not allocated to at least one address
- Determine one STEP address ID for each group
  - Priority 1: use STEP ID of address
  - Priority 2: use lowest STEP ID of emails and phones
- On each email and phone, replace existing STEP address IDs with the set of STEP address ID of each group it is member of.

SAP Address Record relations must be maintained as addresses, emails and phones are approved and after the STEP address ID on email and phone has been maintained.

Validate that the union of all STEP IDs address data containers and STEP address IDs on email and phone data containers each has different SAP UUID for each SAP BP record. This ensures that there are no conflicts of address UUIDs in SAP, even when one BP entity in STEP represents multiple BP records within one SAP system.

Atomic Scopes of Content to Update

When sending data from STEP to SAP, the SAP BP / RL replication web services support different action codes for adding, removing, and replacing different parts of the content of the BP objects.

To keep things simple, it is recommended to always use 'Action Code 04' which:

- makes modifications to existing information in SAP
- creates new information present in the message but not present in SAP
- removes information present in SAP but not present in the message

Message Sequencing

The SAP BP / RL replication web services support ignoring messages that has been superseded by newer messages by including a 'changeOrdinalNumberValue' on the BP XML element.

In STEP, this can be set to system time in milliseconds as the message is generated and exported. This is implemented directly in the XSLT that transforms the exported STEP XML to SOAP messages.

For the inbound integration of acknowledgments from SAP to STEP, it is expected that the transport of messages will preserve the sequence of messages and therefore the latest received message can always be expected to contain the latest information about a given BP. For this reason, STEP ignores the 'changeOrdinalNumberValue' provided by SAP.

SAP Publishing To STEP

Receiving data from SAP is handled with Inbound Integration Endpoints (IIEPs) using the Match and Merge Importer processing engine.

Object Identification

To ensure consecutive updates to the same source record are automatically routed to the existing golden record when BPs are received from SAP, the BP internal ID must be mapped to the Source Record ID attribute on the Source System Reference in the Merge Golden Record Component Model.

To ensure survivorship rules for addresses, emails, and phone numbers are handled, address UUIDs from SAP must be mapped to an incoming SAP address UUID attribute on address, email, and phone data containers. See the **SAP Address Source Record Relations** section of the **SAP Business Partners and Enterprise Structure Definitions** topic.

Error Handling

Since errors that must be addressed by business users are expected to be managed in STEP, there is no communication back to the sending SAP application. This also means that for challenges with value harmonization of LOVs, etc., the golden records data model in STEP must be designed to store data that is not harmonized in STEP and to use workflows and business rules to overcome such challenges within STEP.

For more information on error handling, see the **Replication Error Messages** section below.

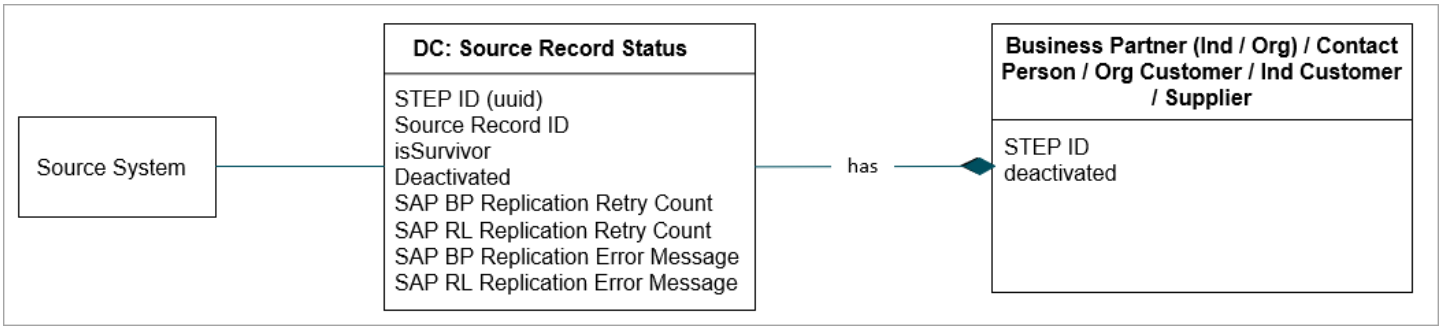
Message Sequencing

To preserve the sequence of BP (Business Partner) and RL (Relationship Link) replication messages, so that, for example, an RL message for a new BP does not arrive prior to the new BP, one IIEP is used to receive both BP and RL replication messages. This means that the transformation from both BP- and RL- Replication messages must be written in the same XSLT.

Since the inbound integration from SAP to STEP expects that the transport of messages preserves the sequence of messages, the latest received message is expected to contain the latest information about a given BP. For this reason, STEP ignores the 'changeOrdinalNumberValue' provided by SAP.

Source Record Status

Source Record Status is a data container (ID= SAPBPSourceRecordStatusData).



- **Source Record ID** - As inbound integrations or merges of existing golden records add source records to the source system relation, survivorship rules must also add those source record IDs as source record status objects, so that 'Is Survivor' and 'Replication Error Messages' can be managed.

- **Is Survivor** - As multiple BPs from the same source (SAP) system merge into the same golden record in STEP, a surviving source record, per source system, is determined by survivorship rules.

  When a golden record is a reference target of another golden record and the reference is exported to a given source system, it is the surviving source record that must be target of the reference in the message sent to the source system.

- **Handling in Partner Function and RL Replication (BP-to-BP References)**

  When a reference source or target does not have a surviving source record ID for a given SAP system, the STEP ID must be mapped to the BusinessPartnerInternalID and/or RelationshipBusinessPartnerInternalID in the RL replication message and to the PartnerFunction/PartyInternalID in the BP replication message.

  The missing source record ID may be because the record is missing in the target system. Since the integration is asynchronous, the message that creates the target may be in transit and arrive earlier than this message. If not, this replication must fail and the async response should reveal the error, leading to the error workflow where the entity replication is retried exactly once. If the next attempt fails, the entity is placed in a manual fallout task where it is up to a user to resolve the inconsistency in data and resubmit the entity to the target system. See the **Error handling workflow** information below.

  Use the following exact logic for the survivorship rule:

  - Find source record IDs in the source record / non-surviving golden record
  - Add source record IDs in the source record status if they do not already exist
  - Add all source record statuses of non-surviving golden records to the survivor (to cover merge of existing golden records)
  - Determine exactly one surviving source record per source system. The default logic is:
    - Priority 1: survivor is the existing surviving source record of the target golden record
    - Priority 2: survivor is the existing surviving source record of non-surviving golden record (merge of existing golden records)
    - Priority 3: survivor is the new source record ID (update / merge on import)

- **Replication Error Messages** - As 'BP-Replication Confirmation-' and 'RL-Replication Confirmation' messages are received from SAP containing error messages, error messages are written to the replication error messages by survivorship rules.

  Separate replication error messages must exist for each atomic scope of data that is replicated, so that errors related to one problem cannot overwrite errors correlated to another problem.

  Specifically, BP replication and Relationship Link (RL) replication are two atomic scopes of data that is replicated.

  When transforming the SAP response to STEPXML, the error messages should be mapped to an SAP Source Record Status Data Container object.

  Use survivorship rules to write error messages to the Source Record Status Data Container objects with the following specific logic:

  - Find target DC by Source System + Source Record ID.
  - Set relevant error message, overwriting any existing error message.

    Replication Error Messages Approve Trigger

    If a golden record has at least one error message, it must be initiated into the error handling workflow. This determination should be done with an approve trigger running for these scenarios:

    - post survivorship rules on import
    - on auto-merge in the event processor
    - post merge in the clerical review workflow

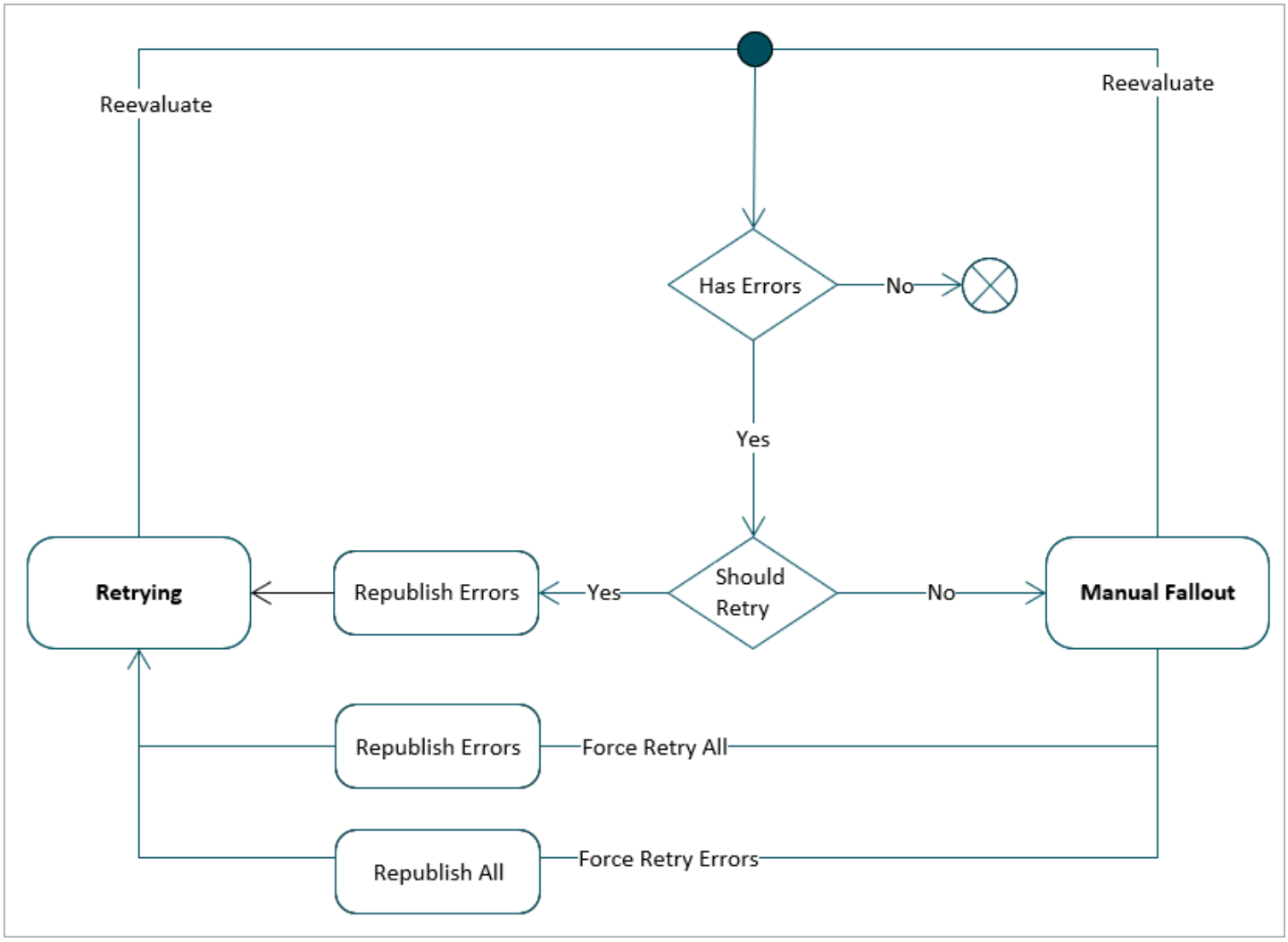
      If the golden record has zero error messages in the Source Record Status data container but it is already in the workflow, it must be removed from the workflow, by letting the workflow transition back to the initial state and reevaluate.

    Error Handling Workflow

    The error handling workflow must determine the appropriate action for a given error. The recommended logic is:

    - retry the replication once (see the **Handling in Partner Function and RL Replication** section above)
    - if it keeps failing, send it to a manual queue

When republishing events to the relevant OIEPs, the retry count for the relevant source records must be incremented.


Reevaluate
Reevaluate
Has Errors
No
Yes
Retrying
Republish Errors
Yes
Should Retry
No
Manual Fallout
Republish Errors
Force Retry All
Republish All
Force Retry Errors

SAP Survivorship Rules

Survivorship rules are used to aggregate information from source records on to the consolidated golden record. In particular, because the business logic for merging data that has validity periods and for addresses, emails, and phones is complicated, the survivorship rules are also complicated. Survivorship rules for each of these structures are defined in the following sections.

Survivorship of Validity Periods

The complexity of validity periods is relevant to Business Partner Roles, Business Partner Relations, Addresses, Emails, and Phones. The following are valid strategies for merging validity periods:

- **Winning set of periods, specifically 'Most Recent' and 'Trusted Source' patterns** - As two sets of periods must be merged, one set is selected as the surviving set and the period that is not in the surviving set, does not survive. The winning set can be determined, for example, by a most recent date and time or a prioritized list of trusted sources.
- **Aggregate periods** - As two sets of periods are merged, keep all periods, but merge periods with the same 'from' date and the same 'to' date. This pattern does not allow removing a validity period once it has been added, which may be a problem in a consolidation solution. This method is not a problem in a centralized onboarding solution, where survivorship rules are only applied as golden records merge.
- **Union periods aggregation** - As two sets of periods are merged, keep all periods but merge overlapping periods to the lowest 'from' date and the highest 'to' date. This pattern does not allow reducing validity periods once they are added, which may be a problem in a consolidation solution. This method is not a problem in a centralized onboarding solution where survivorship rules are only applied as golden records merge.
- **Common periods aggregation** - As two sets of periods are merged, only keep overlapping periods and reduce overlapping periods to the highest 'from' date and the lowest 'to' date.

Survivorship of Data Container: BP Role

Survivorship of BP Roles should union all Business Partner Roles across source records. In addition, it must be considered how to aggregate validity periods. See the **Survivorship of Validity Periods** section above.

Survivorship of Business Partner Relations

Business Partner Relations vary in complexity and must be handled differently as described below.

- **BP Relation Types without Validity Periods (simple reference types)**
  - If only a single reference is allowed - use a simple survivorship pattern like 'Most Recent' or 'Trusted Source'.
  - If multiple references are allowed - union all relations across source records.
- **BP Relation Types with Validity Periods (Data Container: BP Relation)**

- If only a single reference is allowed - sort all relations by 'from' date and as periods overlap, use a most recent pattern to determine which relation is shortened in period or removed, so that there are no overlaps in periods.
- If multiple references are allowed - union all relations across source records, but merge overlapping periods to the lowest 'from' date and the highest 'to' date.

Survivorship of Addresses, Emails, and Phones

Use the following to create survivorship rules for addresses, emails, and phones.

- **Merging addresses, emails, and phones** - union all objects across source records and union validity periods per address usage. This means that one data container may merge into multiple because the union on validity periods is different for different address usages of the same address, email, or phone.
- **Adding incoming SAP Address IDs to SAP Address Source Record Relations** - As an incoming address is associated with a target data container object, this relationship must be added to SAP Address Source Record Relations.
- **Merging SAP Address Source Record Relations** - When survivorship rules complete the merge of addresses, emails, and phone numbers, the SAP Address Record relations must be maintained. First, however, the STEP address ID on emails and phones must be maintained. See the **Email and Phone Number Approve Trigger** section in the **SAP Publishing From STEP** topic.

Survivorship of Specific Data

Use the following to create survivorship rules for **company code data**, **sales area data**, **purchasing organization data**, and **bank data**.

- Survivorship should union all data containers and merge those with the same key.
- Values and references on the data containers that are single valued should be merged following a single survivor pattern like the 'Most Recent' or 'Trusted Source' pattern.
- Values and references of the data containers that are multivalued may union all values from all source records or may follow a 'Most Recent' or 'Trusted Source' pattern.

User Authentication within a Self-Service

Within a self-service solution, users will constantly need to be created and linked to the right user groups to ensure they have the correct privileges. In addition, these user credentials must be validated upon each log in via an authentication setup. A couple of options exist for the best approach to this issue, each with their own advantages and disadvantages. The following describes these approaches as well as how to use the provided example to this within the initial configurations.

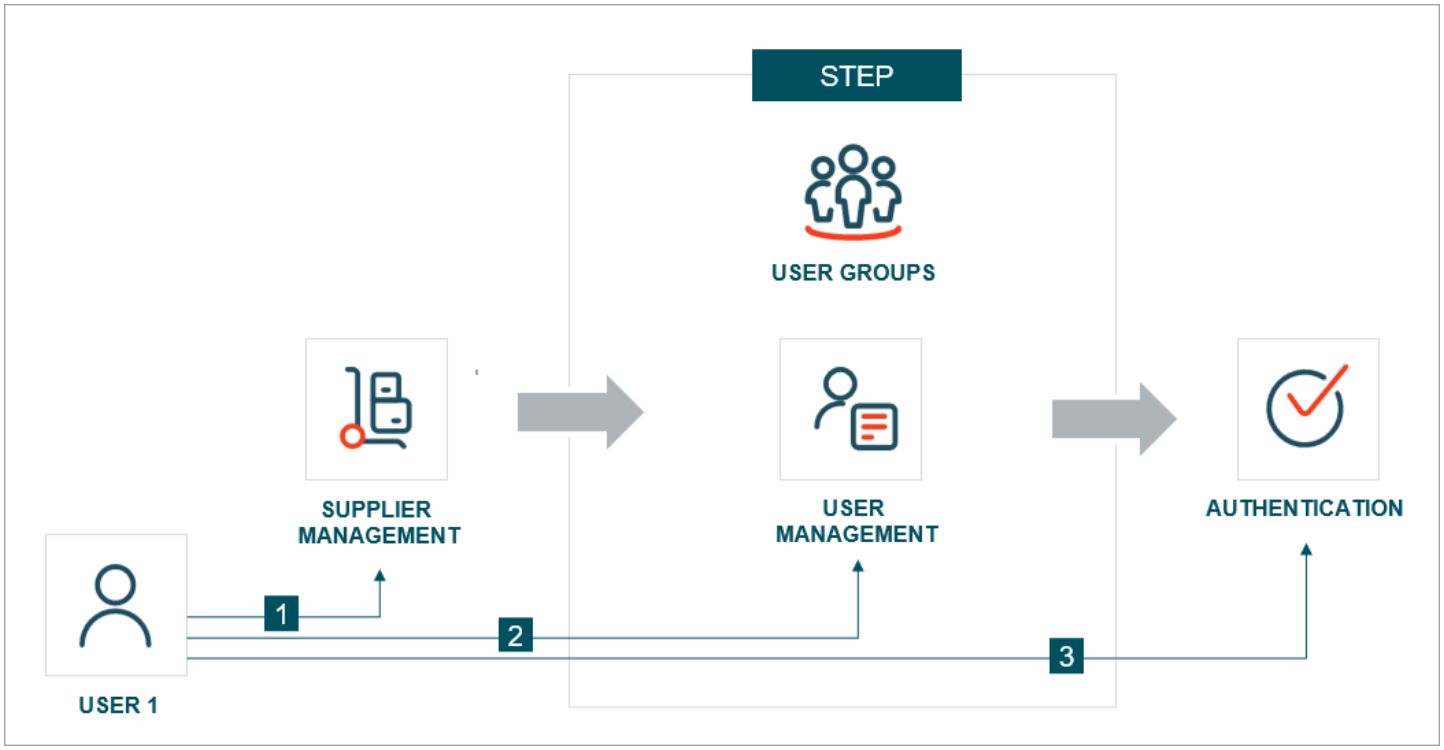
Identity Providers

STEP as an identity provider for external users

One approach to supplier user management & authentication is to allow the initial user creation and communication to the external supplier to be managed via STEP workflows and business rules. This approach allows the supplier admin user to create and manage his own subset of users within the STEP Web UI. Authentication mechanism, password policies, etc. are limited to what STEP offers. Additional information can be found in the **Security Policy** topic of the **System Setup / Super User Guide** documentation.

A disadvantages to this approach is that authentication monitoring is not offered.

In the following image we can see the process involved for user management and authentication where STEP is utilized as the idenitity provider.



1. A new user is created under the relevant supplier user group. This enables view restrictions to only see objects related to this particular supplier group.

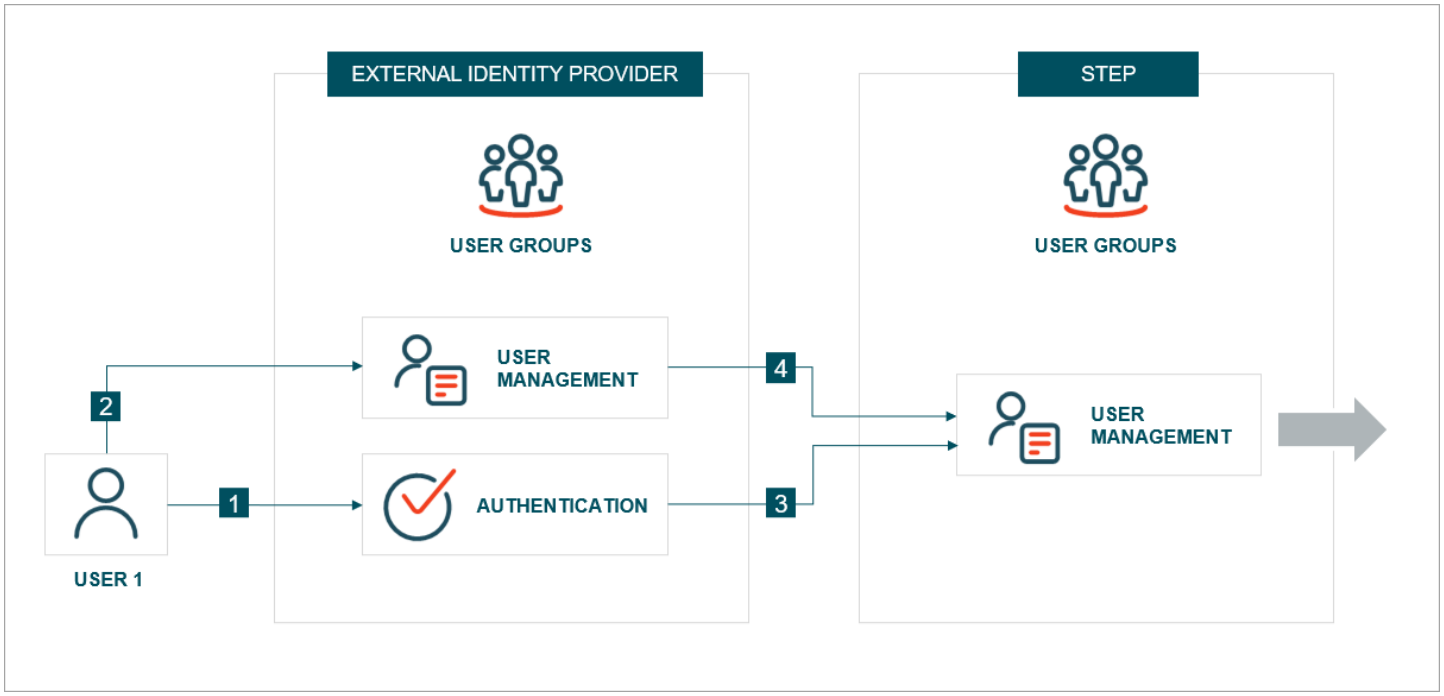
2. The new user may be linked into additional user groups. This enables privilege customization through the user gaining cumulative privileges of all groups they are a part of.
3. This user must have a password that is validated upon each login. This is user authentication.

Integrate with customer's identity provider

It is likely customers may have their own user management & authentication management solutions in place. Either automatically or through a manual process the identify provider must be integrated with STEP to allow for the initial creation and authentication of supplier users. In these scenarios, the customer's identity provider must manage the authentication and monitoring of users. Stibo Systems technical services can provide additional guidance and support for the specific integration between STEP and an external identity provider. It is possible to support multiple identity providers.

A disadvantage to this approach is there is no user management in Web UI, so onboarding of additional users for an existing supplier, must happen through a different process outside of STEP.

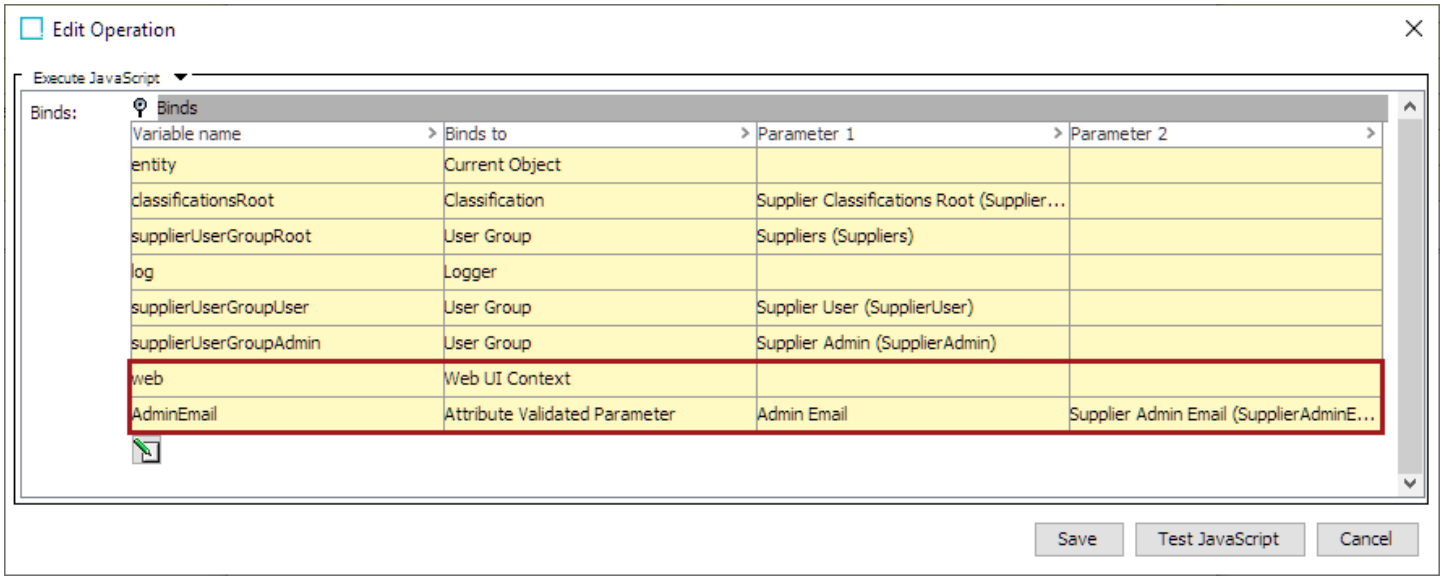
In the following image we can see the process involved for user management and authentication where an outside identity provider is used.



1. User credentials are authenticated via an outside system
2. Users are managed in the outside system
3. Authentication information is passed to STEP to allow supplier users access
4. User group information is synced to STEP and users are placed into the appropriate groups.

Initial Configuration Solution

The initial configurations for Supplier Onboarding with self-service use the first approach discussed above by managing the creation of supplier users via a workflow with contributing business rules. As part of the onboarding process of a self-service supplier, the user is asked to provide the email address of the initial supplier admin. The required Web UI bind for passing this value to the business rule is 'Attribute validated parameter.'



This email address is used to provide the external user with a username and password. This initial user is an admin and will have privileges to create additional users. They accomplish this by having additional user groups that contain various privileges. Supplier admin users can then link users to multiple groups to gain the cumulative privilege set.

Data Stewardship

Data stewardship in a Customer or Supplier MDM application is performed by two distinct roles: data stewards and business users. Each role requires the configuration of different Web UI screens to perform certain tasks. In this topic, the roles (and the tasks performed by each) are defined and Web UI configuration guidelines are provided.

Note: When configuring any Web UI, consult the **Web User Interfaces / Web UI Setup and User Guide** documentation. This documentation contains details on how to add various components of the Web UI.

Role of the Data Stewards

Experienced data stewards must be provided with the proper tools to ensure data quality.

These tasks include:

- Clerical Review - Data Stewards can review potential duplicates.
- Manually maintaining records - Data Stewards update and edit records as needed.
- Monitoring / Data Quality - Data Stewards should review customer data quality in the Web UI, and monitor Source system performance.

Data Steward Tasks

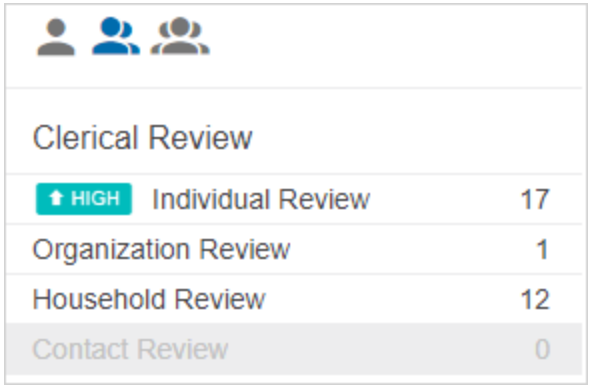
Typically, Data Stewards will need to perform the following tasks: clerical review, ad hoc data stewardship, data governance, hierarchy maintenance, and attribute maintenance.

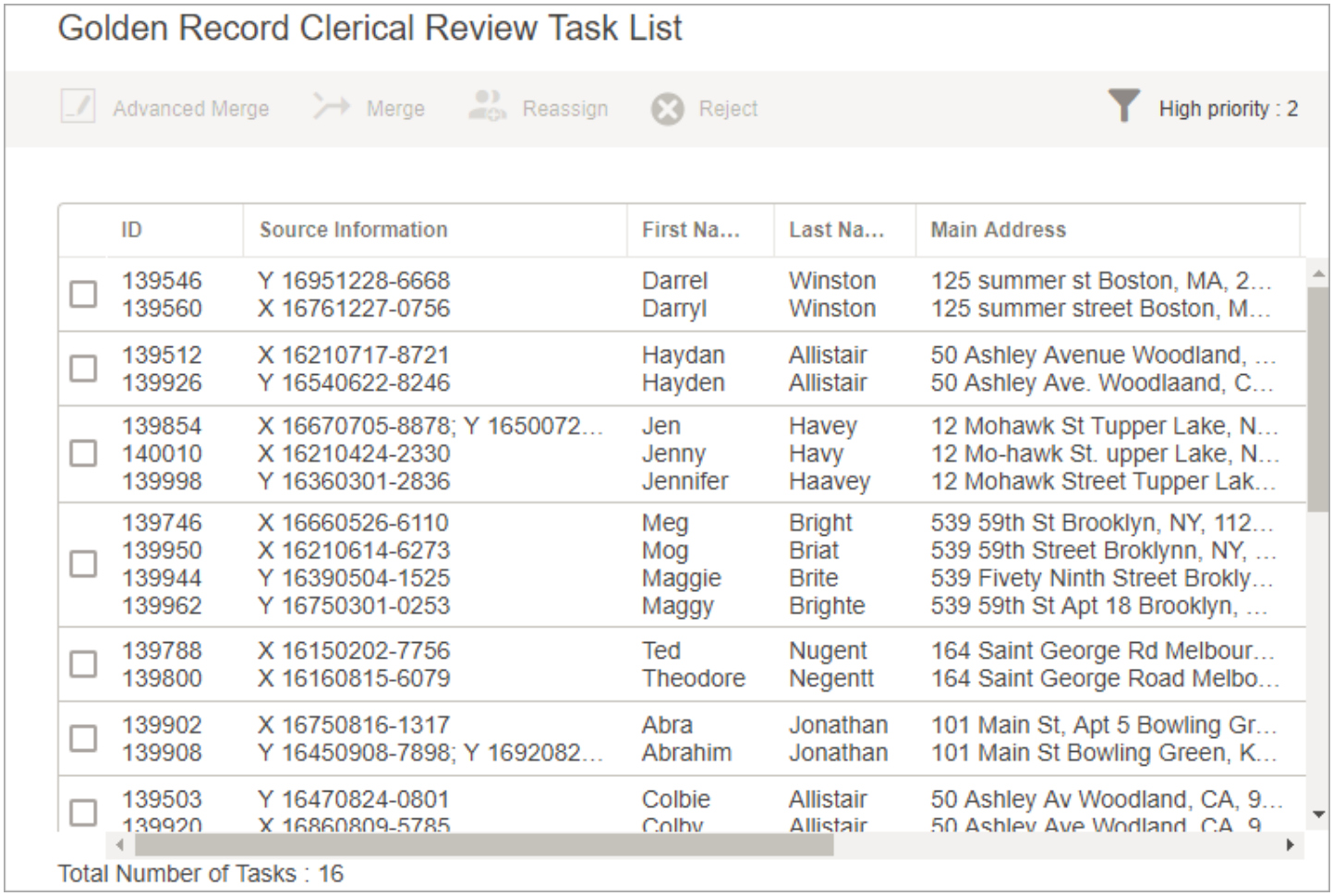
Note: The online help topics for these tasks are extensive and cover configuration and typical use.

Clerical Review

Clerical reviews allow a data steward to approve matched duplicates, reject erroneously matched records, and reassign tasks that are better suited for other users.

To assist data stewards in identifying critical issues, Clerical Review task lists can be filtered to display only high priority tasks. To flag tasks as high priority, a workflow status flag and accompanying business condition must be configured on the relevant matching algorithm.





For more information on clerical reviews, see the **Golden Record Clerical Review Task List** topic in the **Matching, Linking, and Merging** documentation.

Use Cases

Arthur is a senior data steward responsible for the carrying out several data management tasks and must interact with many areas of the Web UI to do so.

When Arthur begins his day, he will first handle any clerical review tasks that may have been generated from the previous day's activity. Some new customers coming into MDM have a risk flag set to 'yes', this requires Arthur's immediate attention. Arthur is made aware of these high priority tasks through a visual cue (see below). Arthur is then able to enter the task list and filter based on priority. This allows Arthur to immediately address the high priority clerical reviews, before continuing with his other tasks.

Components

Arthur utilizes the following components to address the story presented above:

1. The Global Navigation Panel - Allows Arthur to see all clerical review tasks grouped in one central location. For more information, see the **Global Navigation Panel Component** section of the **Web UI** documentation.
2. Golden Record Clerical Review Task List – Allows Arthur to address deduplication tasks. More specifically, this screen allows for merging duplicate records, task reassignment, and rejecting potential duplicates.
3. Advanced Merge – Allows Arthur to manually dictate survivorship by selecting individual attribute values.

Hierarchy Maintenance

Maintaining the company hierarchy is key to ensuring that data can be correctly represented across the enterprise. Using a hierarchy visualization implementation helps represent the flow of companies and how their ownership functions.

For more information on setting up and using a company hierarchy, see the **Company Hierarchy Visualization and Maintenance** topic of the **Web UI** documentation.

Use Cases

A volcanic eruption has destroyed an ACME subsidiary. Arthur has been asked to remove the subsidiary from ACME's corporate hierarchy.

Arthur navigates the subsidiary in question and selects the hierarchy tab. From this screen, Arthur can remove the link to its parental organization using built in functions.

Components

Arthur utilizes the following components to address the story above:

1. Global Search – Arthur can navigate to the subsidiary based on source record ID.
2. Company Data Hierarchy Screen – This screen allows Arthur to add, edit, and remove links throughout the hierarchy. Several different views are also available for each hierarchy.

Attribute Maintenance

Data stewards must create and maintain various attributes and attribute groups.

For more information on attributes, see the **Attribute and LOV Creation and Maintenance in Web UI** topic of the **Web UI** documentation.

Use Cases

Arthur has been notified by management that an upstream SAP system is now tracking additional financial attributes (Income Verified Flag). Arthur would like to create this attribute in MDM, so he will have access to the most complete view of the customers.

To do so, Arthur will utilize the attribute maintenance screens. First, Arthur will need to identify the correct party data attribute group. Since this is new financial data, Arthur decides it belongs best in the IncomeData group. Arthur selects this group and can create the attribute using the buttons on the screen.

Components

Arthur utilizes the following components to address the story above:

1. Attribute Group Management Screen – Using this screen, Arthur can view all attributes within the attribute group and judge whether the new attribute he wants to create belongs there. For more information, see the **Attribute Group Management Screen** section of the **Web UI** documentation.
2. Attribute Link Editor Screen - Arthur can establish or remove links between attributes via this screen. For more information, see the **Attribute Link Editor Screen** section of the **Web UI** documentation.
3. Attribute Management Screen – Arthur can create new attributes or edit existing ones via the Attribute Management Screen. He can edit the attribute details, change the attribute validity, and maintain attribute links via various tabs configurable on this screen. For more information, see the **Attribute Management Screen** section of the **Web UI** documentation.

Additional Considerations when Performing Attribute Maintenance

Udate import configurations with the appropriate mappings, otherwise the new incoming data is not written to any attribute.

If the attribute does not fit into any existing attribute groups, a new group should be created first. If the attribute fits into a current display group, no Web UI configuration is needed. Standalone attributes must be manually added to the appropriate Details screens.

If the attribute should be considered during matching, adjust the algorithms and/or match codes. If the intention is for the attribute to 'pool' similar customers together, it should be added as a match code function. If the attribute is intended to be considered by the algorithm, a new corresponding normalizer and matcher must be added. Additionally, the rules must be updated to include how the final score provided by the algorithm is affected. In order for survivorship to properly execute, the attribute must be placed in the relevant survivorship groups.

An onboarding process may be considered to verify that all considerations have been accounted for prior to finalizing new attribute(s).

If the attribute should be monitored via data governance, update the required policies and widgets to account for it.

Role of Business User

Experienced business users need to be able to accurately search for customers to verify information, and export customer data.

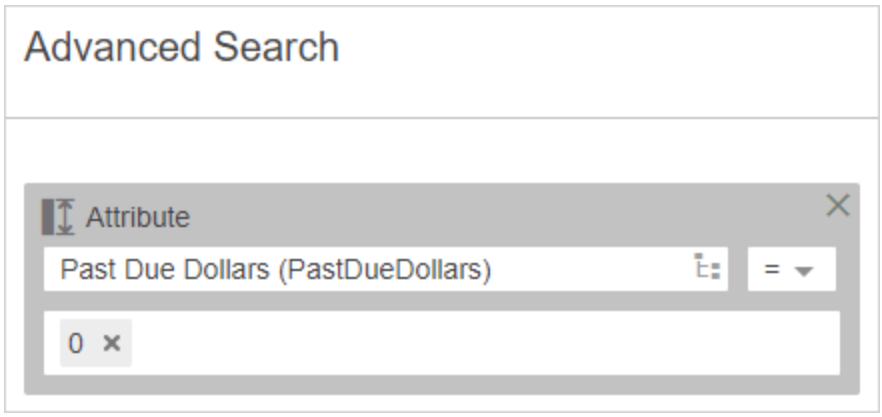
Business User Tasks

Business users need to be able to maintain small areas of the database, mostly involving searching for customers to find records and exporting this data.

Advanced Search and Data Export

Business users need to be able to navigate the large amount of data in the system. To do this, they need to create customized, granular searches using a myriad of search criteria.

Nikki, a business user, has been asked to export a list of customers who have outstanding payments. To do so, Nikki utilizes an advanced search. Nikki then saves the results to a collection and uses the built-in buttons to export.



To build these search queries, see the **Using Advanced Search** topic in the **Web UI** documentation.

Export the data once the data is isolated for external support. For more information, see the **Export Manager** topic of the **Data Exchange** documentation.

Manual Edits

Business users may also manually edit customer information within the Web UI. One example is a business user working within a call center may receive an updated contact number from a customer. The business user may edit an existing phone number or add an additional number to the customer's record.

While working in a call center, Nikki receives a phone call from a customer. The customer has notified Nikki that they are moving and would like to update their address to continue receiving promotional mailings. Nikki can look this customer up based on a provided loyalty card number and manually update the address. Once Nikki saves the address update, Loqate will trigger and standardize the new address.

Manual edits within the Web UI is generally carried out within the Node Details Screen.

Hierarchy Maintenance

Business users may also be responsible for maintaining company hierarchies.

Maintaining the company hierarchy is key to ensuring that data can be correctly represented across the enterprise. Using a hierarchy visualization implementation helps represent the flow of companies and how their ownership functions.

For more information on setting up and using a company hierarchy, see the **Company Hierarchy Visualization and Maintenance** topic of the **Web UI** documentation.

Ad Hoc Data Stewardship

When dealing with large volumes of data it is often difficult to identify trends and outliers within a dataset. MDM's data profiling functionality allows data stewards to observe data in its current state and identify patterns and trends. A few examples of when a data steward would utilize data profiling include:

- Profiling and analyzing data
  - Investigate a decline in quality of data as reported by data quality policies.
  - A potential upstream problem may have been identified.
- Ensuring address quality of master data
  - For a marketing campaign targeting a group of customers or specific geographical region.
  - Increase in return mail or bounced emails.
- Running a search and seeing the result of the search
  - Determine number of resulting records from the search.
  - See examples of a search that returns a large subset of entity records (ex. more than 1 million).
  - See details of the various source systems that serve as input for organization records.
  - List all data that came in from Sweden on a Saturday 2 weeks ago.
- Investigating what an existing collection contains
  - See if an old collection can be deleted.
  - Identify who created the collection. When? Why?
- Exporting a list of customers
  - 1000 records with phone numbers for marketing research.
  - 42 million records with address and anonymized data points for a BI platform.

Use Case 1: Ad Hoc Work on Lists of Records

Bruce is a data steward whose primary role is to support and ensure the quality of customer data for the ACME's marketing team. The Marketing team is promoting the upcoming season and is running multiple campaigns which target existing customers residing in specific geographical areas within the United States. All campaign collateral us addressed to each individual customer and sent via both email and direct mail.

Existing data policy monitoring metrics are in place and show that a sizable percentage of marketing collateral is either returned by the United States Postal Service (USPS), emails have bounced back, or contact information is not present. In the last two campaign seasons the statistical breakdown is as follows:

- 9% of emails have bounced back
- 12% of direct mails were undeliverable and returned

Because the upcoming campaign is targeted towards individuals in a particular region of the US, the marketing team would like to understand the quality of contact data for those individuals before rolling out the campaign. As such, the team has compiled a mailing list comprised of customers and prospects they would like to mail collateral to.

Bruce's task is to run this campaign list against the current MDM database to determine the quality of contact data currently on file. This determination can be made by seeing if the following requirements are met:

- Customer has sufficient contact information
  - Main Address is validated at least once within the last 365 days
    - If address has a low quality, no flyer is sent
  - Email is validated within the last 120 days
    - If email is no longer valid, a process is started that notifies the customer and requests that they update their email address.

To accomplish this, Bruce can create a collection of customer records based on the mailing list and generate a data profile on the Collection. He then evaluates the data profile against the requirements.

Note: It is assumed that the profiling configuration Bruce uses includes the necessary data elements to be analyzed, i.e., Address and Email attributes.

With the data profiles in place, the MDM system can highlight the data quality problems for Bruce, allowing him to report the following back to the marketing team:

- 4% of the customers did not have mailing addresses.
- 18% of the addresses that are present are incomplete (i.e., missing state or zip code).
- 8% of the emails are marked by Experian as 'disposable', indicating the address is associated with a disposable email provider. Usually they are unreliable and could potentially be spam accounts.
- 4% of the emails are 'illegitimate', indicating it is highly likely they are spam accounts or inactive domains.

Based on this information, marketing may then determine the best course of action, i.e., to either proceed with the mailing (if the data quality is acceptable) or place a larger emphasis on an e-campaign if a large percentage of the target customers only have an email address.

Use Case 2: Investigate a Decrease in Completeness of Contact Information

Chris is a data steward whose primary task is to continuously monitor the quality of data within MDM. A focal data point of Chris' responsibilities is to monitor overall quality of contact information. Data quality policies monitor contact information by utilizing a completeness metric which evaluates completeness of attributes such as Address, Email, Phone, Last Contact Date, etc.

In the last two weeks, Chris has noticed there has been an increase in frequency of policy breaches for quality of customer contact information. Since ACME's marketing strategy includes heavy reliance upon informing existing customers of the latest products and promotions, having dependable contact data is a top priority.

Chris would like to use MDM to profile the records that contain incomplete contact information to determine what aspects within contact information are incomplete / missing. Once these customers are identified, Chris will then:

1. Provide the results to the marketing team, so they may best determine how to proceed with their marketing activities.
2. Collaborate with marketing in identifying and changing existing business processes that caused the capture of incomplete contact data so they may address the issue at the source.

By using the profiling capabilities of MDM, Chris was able to support the marketing team and their campaign efforts. In doing so, it was identified that ACME's CRM system, which manages all customer loyalty programs and sign-ups, was lacking validations to ensure complete and accurate contact information is captured at the point of entry. Specifically, the CRM system experienced an issue where city and zip code information were no longer mandatory as part of the new customer sign-up process.

Furthermore, the marketing team was able to strategically accommodate for this in their marketing activities as they awaited resolution of the issue within the CRM system. Once the issue was addressed, marketing then proceeded to carry out their campaign with confidence in their data

Executing Users

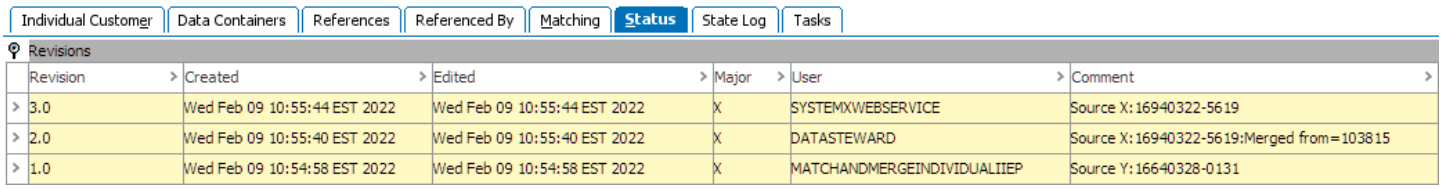
It is important to consider the executing user when setting up several system configurations. These include event processors, inbound / outbound integration endpoints, and web service endpoints. Having dedicated user groups for each allows for easy identification of where a data change originated. Stibo Systems recommends that these user groups are configured with super user privileges. This is to avoid any privilege limitations, as the purpose of these users is to provide traceability.

For example, in the following screen shot, we can see the different system configurations that contributed to the revisions on this entity.

Revision 1: Entity was created via a match & merge inbound integration endpoint.

Revision 2: A data steward actions a clerical review task that results in a merge which causes survivorship rules to be reevaluated.

Revision 3: A web service request causes a data update based on a source record ID match.

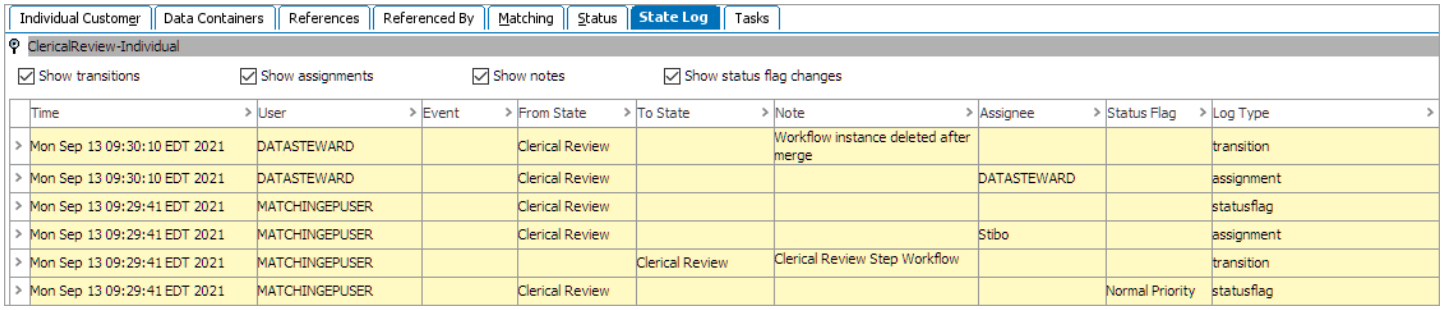

Individual Customer | Data Containers | References | Referenced By | Matching | Status | State Log | Tasks

Revisions

Revision	Created	Edited	Major	User	Comment
3.0	Wed Feb 09 10:55:44 EST 2022	Wed Feb 09 10:55:44 EST 2022	X	SYSTEMXWEBSERVICE	Source X:16940322-5619
2.0	Wed Feb 09 10:55:40 EST 2022	Wed Feb 09 10:55:40 EST 2022	X	DATASTEWARD	Source X:16940322-5619:Merged from=103815
1.0	Wed Feb 09 10:54:58 EST 2022	Wed Feb 09 10:54:58 EST 2022	X	MATCHANDMERGEINDIVIDUALIIEP	Source Y:16640328-0131

Additionally, in the following screenshot we can see:

- The matching event processor user has initiated the entity into clerical review.
- A data steward actions a task resulting in a merge, and the entity is then removed from clerical review.


Individual Customer | Data Containers | References | Referenced By | Matching | Status | State Log | Tasks

ClericalReview-Individual

Show transitions | Show assignments | Show notes | Show status flag changes

Time	User	Event	From State	To State	Note	Assignee	Status Flag	Log Type
Mon Sep 13 09:30:10 EDT 2021	DATASTEWARD		Clerical Review		Workflow instance deleted after merge			transition
Mon Sep 13 09:30:10 EDT 2021	DATASTEWARD		Clerical Review			DATASTEWARD		assignment
Mon Sep 13 09:29:41 EDT 2021	MATCHINGEPUSER		Clerical Review					statusflag
Mon Sep 13 09:29:41 EDT 2021	MATCHINGEPUSER		Clerical Review			Stibo		assignment
Mon Sep 13 09:29:41 EDT 2021	MATCHINGEPUSER			Clerical Review	Clerical Review Step Workflow			transition
Mon Sep 13 09:29:41 EDT 2021	MATCHINGEPUSER		Clerical Review				Normal Priority	statusflag

Best Practices

Administrators should manage background process (BGP) users within a single user group configured with super user equivalent privileges. Additional considerations:

- Event processors should have separate unique users.
- Each Match & Merge Importer should have separate unique users.

- Every Source System should have unique web service users.
- BGP users should be process specific. For example: Policy Monitoring & Hotfolder Imports.

Data Governance

Data governance is the overall management of the availability, usability, integrity, and security of data used within an enterprise. A sound data governance program is driven by a governing body or council, which encompasses stakeholders from across the business, including data stewards. It includes a defined set of procedures, policies, and a plan to execute on those procedures.

Realization of corporate data governance is largely motivated by a desire to improve business operations and performance by gaining better oversight and management of corporate information. While a data governance program institutes policies and processes designed to produce more accurate and consistent data throughout an organization, it primarily becomes the job of the data steward to put those policies and processes into practice by ensuring compliance. It's through governance and enforcement of said policies where you are ultimately supporting business process integrity, which in turn drives positive business outcomes.

Data Quality Policies

Data policies allow users like data stewards to define thresholds, and monitor breaches and deviations in the quality of the existing data as well as incoming data.

Data quality policies apply Metrics on Datasets to measure the quality of data. Thresholds define when users must be notified.

Policies enable data stewards to proactively monitor data. Data stewards can define policies to ensure data completeness, uniqueness, accuracy and more.

For more information, see the **Data Policies** topic in the **Data Governance** documentation.

Existing Data vs Incoming Data

Existing Data Policies evaluate data that exists in STEP each night. Incoming Data Policies only evaluates the incoming data from an Inbound Integration Endpoint of the Merge Golden Record type, allowing early warnings if the source system starts sending bad data.

Data Quality Dimensions

With Data Quality there are dimensions that need to be considered which are key cogs in driving the definition of Data Quality Policies.

> **Note:** Data Quality dimensions are not to be confused with language dimensions, country dimensions, etc. that are platform-specific concepts.

Data Quality policies help organizations to ensure that data quality complies with the business' expectations.

Data quality policies use logical metrics on entity data to test the quality threshold. These thresholds show a simplified view of the metric performance. If data quality does not comply with the policy, a data policy breach is recorded. These policies update when the data quality returns to normal expectations.

With these policies, data stewards are able to proactively monitor, control, and maintain entity data from within MDM. Data stewards can build policies to ensure data completeness, uniqueness, accuracy, and more.

Examples of data quality dimensions are:

- Accuracy– Is the data verified, accurate and up to date? Our 3rd party integrations can now verify aspects of party data using the very latest trusted reference sources.
- Completeness– What data is missing or unusable? Is all the necessary data present?
- Timeliness - The degree to which data represents reality from a required point in time - Is data available at the time needed?

Metrics define the specifics of how to measure data quality.

For use cases with data governance concepts, see the **ACME Holding Group Example Case** topic in this documentation.

ACME Holding Group Example Case

The parent company Acme Group operates across three business units; ACME Healthcare, ACME Inc., and Obits Technologies. Each business unit spans across multiple countries.

The three business units operate independently but have a large overlap in customer and supplier base. It is a company strategy to align the business units both in a customer and supplier centric way around one central brand that services the customers end-to-end and aligns supplier procurement. As part of this journey, master data must be aligned between the three business units across the various countries. For that purpose, ACME Group is turning to Stibo Systems' MDM solution to help address their needs.

There is not much overlap in the customer base across countries. Because of this, information stewardship is divided between each country, as there are many country-specific quirks in the data that local data stewards can more easily deal with.

Suppliers can potentially present a larger overlap across countries and business units, however managing supplier master data may also be relegated to local data stewards as certain region and line of business-specific data points may need to be governed separately.

To facilitate this organizational stewardship, Acme Group has established a governance organization with members from each business unit in each country. This organization will define a set of local policies for many different things. They expect that the MDM platform can monitor and report on these policies.

While master data across various countries may adhere to different standards of completeness, it is widely accepted by ACME that the data being considered while evaluating completeness of a record shall remain the same. As such, ACME has determined that they require scoring metrics to be agreed upon across the organization. From there, individual policies may be built on top of the metrics to better define the country-specific thresholds that are appropriate.

Furthermore, since the focus of each business unit and country may be different, it is not easy for the team to understand what they should monitor. Therefore, some metrics may not be shared and must be considered as separate metrics for each business unit-country combination. In other instances, some metrics may be shared, however, the policy thresholds may be different.

Use Cases

See the **Email Completeness Policy** topic for an example of a policy shared between entities.

See the **Individual Completeness** topic for a locally stored policy.

See the **Source System Performance** topic for a stream data policy.

Email Completeness Policy

Dylan and Sonya are both data stewards within ACME Group whose team is responsible for analyzing and reporting on organizational governance across the globe. Dylan is responsible for ACME Healthcare in the US, whereas Sonya is responsible for ACME Inc. in Brazil.

To assist in planning for an upcoming promotional campaign, ACME Group would like to analyze the current quality of contact information of their customers across the two business units. Since the digital age means email is the engine that drives all digital campaigns, having reliable email addresses of their customers is key to achieving their quarterly fiscal goals.

To accomplish this goal, ACME Group would like to evaluate email completeness over a period to identify trends in data quality and devise internal process improvements to improve email quality.

Individual Completeness

Dylan and Sonya are both data stewards within ACME Group whose team is responsible for analyzing and reporting on organizational governance across the globe. Dylan is responsible for ACME Healthcare in the US, whereas Sonya is responsible for ACME Inc. in Brazil.

Having trusted customer data is critical to an organization's ability to confidently identify who their customers are in driving revenue and strategy. To achieve this level of maturity, ACME Group would like to obtain an accurate view of completeness of the data they have on their customers based on business unit and country.

The challenge they face is that there is no way for ACME to have meaningful visibility in assessing the quality of their customer data based even on a small set of core attributes over a period. The limitations stem from the siloed structure of the various CRM platforms for business units, countries, and subsidiaries operating today.

ACME Group has chosen to leverage Stibo Systems' Customer MDM data quality offering to master their customer data. Typically, performance metrics will vary from country to country due to differences in consumer behavior, culture, standards, regulations, etc.

To achieve a level of uniformity in performance indicators across business units and countries, it therefore makes sense for Dylan and Sonya to define a common set of attributes which defines a completeness metric from which to gauge customer data. These core attributes consist of:

- First name
- Last name
- Address Line
- City
- State/Province
- Postal Code
- Country

Since the quality of customer information gathered may vary from country to country, Dylan and Sonya have realized that while the US and Brazil may share the same metrics, there is a need to track the quality of data within each country by using separate policies. This provides both Dylan and Sonya visibility into how their respective countries are "performing" over time.

Since ACME is interested in the overall quality of data on their customers, the data quality policy that Dylan and Sonya are responsible for will execute and profile against the existing customer data set as opposed to monitoring incoming data from specific source systems.

The result is a single completeness metric that is utilized by multiple policies between the US and Brazil (and potentially other countries) with thresholds unique to each country. As such, Dylan and Sonya can view and edit policy thresholds for their respective countries. They are also able to view the policy's performance via the data quality dashboard. Should a policy breach occur for their respective countries, Dylan and Sonya are notified via email.

Source System Performance

Bobby is a senior data analyst who is responsible for monitoring and evaluating the overall performance and quality of customer data for the many source systems (CRM platforms) within ACME Group subsidiaries.

To help support the upcoming season, Bobby is assisting the marketing department in streamlining the promotional mailings process and reduce overhead costs. To do so, Bobby requires the monitoring of address quality ACME Group possesses for their customers, particularly to identify quality within source systems to better identify and address the root problems.

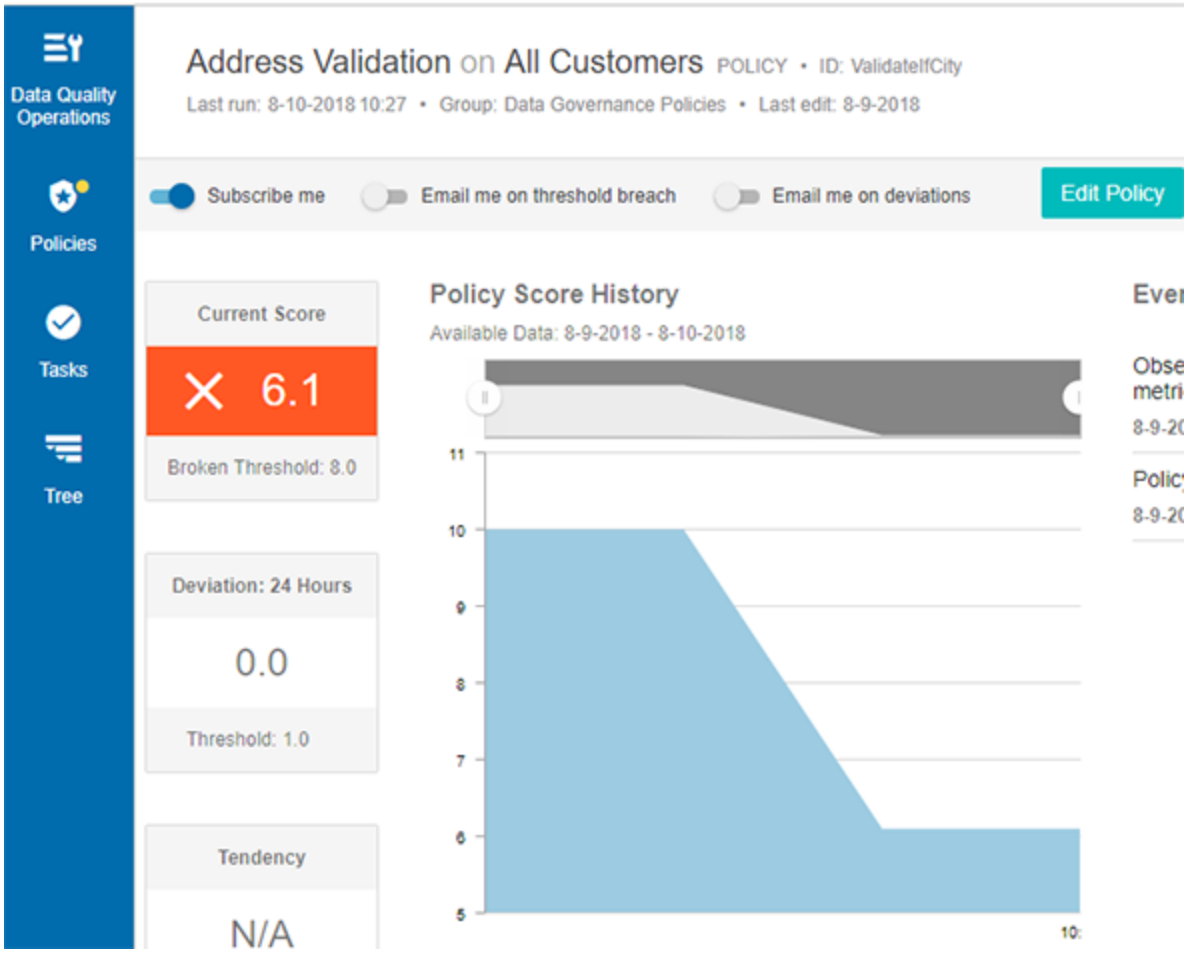
While an existing data quality policy allows for Bobby to gauge the quality of addresses among the existing data set within ACME Group, should there ever be a change in how data is captured in one of ACME Group's many source systems, there is not a good way for Bobby to realize this in a timely fashion. As such, a stream policy which monitors data as they come in from individual source systems would allow Bobby to more easily monitor the behavior / performance of each source system.

With Customer MDM, a value metric allows for evaluating of attribute values and mapping to specific scores. In the scenario of ACME Group's address quality policy, all customer addresses are verified and standardized via the Loqate integration which returns an Address Quality Index (AQI). The AQI is a letter score which of the quality of input address that was validated against Loqate's database. This letter score is evaluated and resolved as a corresponding numeric score for the policy to profile and display against.

The following table shows the evaluation of the AQI:

Score	Quality	Description	Policy Score
A	EXCELLENT	Verifiable to at least Premise level without changes	10
B	GOOD	Verifiable to at least Thoroughfare level with minor changes	8
C	AVERAGE	Verifiable to at least Locality level with moderate changes	6
D	POOR	Only verifiable to at least Locality level with more than moderate changes	4
E	BAD	N/A	2
[null]	N/A	No Address Validation / Quality Index	0

With this value metric defined, Bobby can configure individual policies to evaluate source system performance of customer addresses. He is also able to be notified of breaches and view the performance of a particular source system via the policy dashboard:


Data Quality Operations
Policies
Tasks
Tree
Address Validation on All Customers POLICY • ID: ValidateIfCity
Last run: 8-10-2018 10:27 • Group: Data Governance Policies • Last edit: 8-9-2018
Subscribe me
Email me on threshold breach
Email me on deviations
Edit Policy
Current Score
6.1
Broken Threshold: 8.0
Deviation: 24 Hours
0.0
Threshold: 1.0
Tendency
N/A
Policy Score History
Available Data: 8-9-2018 - 8-10-2018
11
10
9
8
7
6
5

Customer & Supplier MDM Configuration Guide

The initial configurations for the enablement solution provide a generic set of configurations for a basic Customer & Supplier MDM system for the purpose of providing quicker time to value.

This document provides an in-depth overview of the initial configurations as a compliment to the Customer & Supplier MDM Enablement Guidelines. The target audience for this guide is any party interested in taking the initial configurations as a starting point to a new implementation.

Enablement Initial Configuration Overview

The initial configurations include basic data models, i.e., object types, attributes, references, etc., component models, workflows for clerical review of entity deduplication, inbound integrations for data consumption, and web user interfaces (Web UI) for data stewardship.

The initial configurations are largely based on the Merge Golden Record solution and addresses various implementation styles. The implementation styles that are addressed in this guide are:

- Consolidation
- Coexistence
- Centralized

Based on the implementation style that is most relevant to your requirements, the following configuration files may be provided:

Base Configuration

- Mandatory (select one of):
  - Consolidation & Coexistence
  - Centralized
  - SAP Data Model
- Optional Extension for B2B: D&B Web UI Configuration

It is recommended that the initial configurations are loaded into a STEP environment with the appropriate STEP Base. STEP Base is a clean, fresh install of the STEP application. For details on which STEP Base to take which corresponds to the purchased STEP version, contact Stibo Systems' Technical Services team.

STEP System Components

The following components are required to be activated in your STEP system prior to loading the initial configurations:

- Party Data Matching (partydatamatching)
- Company Data Visualization (companydata-visualization)
- Customer MDM Monitoring (cmdm-monitoring)
- Profiling (profiling)

Deployment of STEP components require adequate working knowledge of patching STEP environments. For deploying STEP components, it is recommended to solicit the assistance of Stibo Systems Support or Technical Services.

Third-party services

Stibo Systems provides native integrations with the following third-party services to enhance the data quality and reliability of master data.

- Dun & Bradstreet – for entity matching and company profiling in B2B implementations
- Loqate – for address verification
- Experian – for email validations

If any of these third-party integrations are to be part of a solution, contact your Stibo Systems account manager or partner manager to ensure properly licensing is procured.

Installation of third-party services that require dedicated system components are not included in this guide. Information regarding the Customer and Supplier MDM third-party integrations is available in the **Data Integration** documentation. Further detailed deployment instructions may also be available upon request.

Customer MDM Base Configuration

The Customer MDM Base Configurations includes a basic data model and other configurations that are shared across the subsequent configurations. Hence, the Base Configuration file should be loaded first.

Object types and corresponding attribute definitions that are included are Organization Customer, Supplier, and Contact Person.

For more information on data modeling, see the **Data Modeling** section of the this guide.

Data Model

Organization Customer

For B2B implementations, the initial configurations provide attributes, references, a match algorithm, and clerical review workflow for Organization Customer entity type.

Supplier

For B2B implementations, the initial configurations provide attributes, references, a match algorithm, and clerical review workflow for Supplier entity type.

Contact Persons

Contact Person are also relevant for B2B implementations. From an attribution standpoint, Contact Person is similar to Individual Customer with the exception being that a Contact Person has its own algorithm and clerical review workflow.

A Contact Person is associated to an Organization Customer or a Supplier.

Hierarchy Object Types

In company hierarchies, aggregate types may be used to represent aspects of an entity that do not directly represent customer or supplier entities. Aggregates may be used to model categorizations to accommodate specific organizational hierarchy structures. For example, a legal hierarchy structure may include aggregates representing a parent holding company, business units, subsidiaries, and divisions. A sales hierarchy structure may include aggregates representing regions and territories.

Aggregates are applicable to both internal and external hierarchy types, such as legal hierarchies, financial hierarchies, etc.

Address Standardization

Included with the Base configuration is the Address Standardization business rule using the Loqate integration. For details on configuring and setting up the integration, see the **Loqate Integration** topic in the **Data Integration** documentation.

Lookup Tables

Lookup tables may be used to compliment the match algorithms. A Word Alias Table is included for organizations Organization (Word Alias Table). An Unmatched Word Factor Table is also included for organizations (Organization Unmatched Word Factor Table).

Consolidation & Coexistence MDM Style Configuration

The system setup of the Consolidated and Coexistence styles is rather similar, hence, only a single configuration file is required for either implementation style. However, one major difference between the two styles are with regards to the synchronization of master data back to external systems. Data synchronization requirements must be carefully vetted and designed for during the implementation.

Data Model

Individual Customer

For B2C implementations, the initial configurations provide a match algorithm and clerical review workflow for Individual Customer entity type.

Household

The Household entity type is relevant to the Individual Customer use case for B2C implementations. The initial configurations provide a match algorithm and clerical review workflow for Household.

Workflows

Four clerical review workflows are provided with this configuration.

- Individual Customer Clerical Review
- Organization Customer Clerical Review
- Contact Person Clerical Review
- Household Clerical Review
- Supplier Clerical Review

Clerical review workflows provide the data steward with an interface to view and resolve potential duplicate findings by the respective algorithms of each object type.

For more information on configuring Match Algorithms and Clerical Review Workflows, see the **Matching, Linking, and Merging** documentation.

User Privileges

The Consolidated and Coexistence configurations provide a generic user group for Data Stewards. No special privileges are configured for this group and all users have system super user privileges.

Event processor

The preconfigured event processor collects events for Individual Customer, Organization Customer, Supplier Contact Person, and Household entities to run through their respective match algorithms

Web UI

A Web UI configuration is provided for the purpose of meeting the operational needs of the MDM Data Steward. The Web UI affords the data steward the ability to carry out their daily tasks such as clerical review remediation, ad hoc stewardship responsibilities, maintaining internal hierarchies, monitoring of data governance, and running data profiles.

For more information on configuring a Web UI, see the **Creating a New Web UI** topic in the **Web User Interfaces / Web UI Setup and User Guide** documentation. For more information on Data Stewardship, see the **Data Stewardship** topic in this documentation.

Inbound Integration Endpoint

An Inbound Integration Endpoint (IIEP) is provided and used in conjunction with the Match and Merge Importer. For more information, see the **Inbound Integration Endpoints** topic and the **IIEP - Configure Match and Merge Importer** topic in the **Data Exchange** documentation.

Component Models

The following component models are included in the Consolidated and Coexistence Configuration file.

- Matching, configured for:
    - Individual Customer
    - Household
    - Organization Customer
    - Contact Person
    - Supplier

- Matching – Merge Golden Record, configured for:
  - Individual Customer
  - Organization Customer
  - Contact Person
  - Supplier
- Matching – Link Golden Record, configured for:
  - Household

For more information, see the **Component Models** topic in the **System Setup / Super User** documentation.

Centralized MDM Style Configuration for B2B Customer & Supplier

In a centralized MDM style, all master data originates within MDM and may be synchronized downstream / throughout the enterprise. In such a scenario, operational oversight and data governance is of utmost importance to an organization. The MDM must be able to accommodate complexities of data structures as well as business processes.

Data Model

As the centralized style deals primarily with the B2B use case, companies must be able to model B2B customers and suppliers as well as their respective of lines of businesses. As such, the centralized configuration includes Organization Customer, Supplier and Contact Person with line of business data containers. These data containers contain attribution that are specific to the line of business the customer does business with.

Organization customers, Suppliers, Contact Persons, and line of business data containers are provided in the Base Configuration file.

Workspaces

The Approved workspace is utilized as part of the centralized onboarding story. Workspaces allow for distinguishing the final state of a new entity which denotes active and publishing status to surrounding applications.

Attributes

Corresponding attribution and match algorithms are also provided in the configuration.

Workflows

As data origination / authorship resides within MDM, an onboarding, mechanism must be present to allow all departments and personas that are part of the process a means to interact with the data and even collaborate with one another.

Organization Customer Onboarding Workflow

The Centralized configurations include an **Organization Customer Onboarding Workflow** that is designed for contribution by the following personas:

- Sales Manager
- Finance Specialist
- Logistics Specialist
- MDM Specialist

The workflow allows each persona to fulfill their respective enrichment activities, while providing a means for intra-departmental collaboration.

Supplier Onboarding Workflow without Self-Service

The Centralized configurations include a Supplier Onboarding Workflow that is designed for contribution by the following personas:

- Procurement Manager
- Finance Specialist
- Logistics Specialist
- MDM Specialist

The workflow allows each persona to fulfill their respective supplier onboarding and enrichment activities, while providing a means for intra-departmental collaboration.

Organization Customer & Supplier Approval Workflow

Complimentary to the Onboarding workflows, an Approval workflow is also in place which is owned by the MDM Specialist. The purpose of this workflow is to accommodate ad hoc stewardship of data as it relates to changes to customer and supplier hierarchy relationships and providing a means for the MDM Specialist to review such changes and execute relevant business validations.

Complimentary to the Onboarding workflow, an **Approval workflow** is also in place which is owned by the MDM Specialist. The purpose of this workflow is to accommodate ad hoc stewardship of customer data and providing a means for the MDM Specialist to review changes and execute any business validation on the changes.

User Privileges

The Centralized configurations provide distinct system user privileges for each user persona that interfaces with customer and supplier data. User privileges are important to governance and provides are tighter control in ensuring the right people deals with the right data.

User Privileges for Supplier MDM

Currently, the MDM Specialist may be considered as the *super user* of the Centralized MDM and has privileges to access, edit, and approve all data in the system, whereas Sales, Finance, and Logistics do not have permission to approve customer data.

When it comes to Supplier onboarding activities, the Procurement manager will initiate the onboarding of new supplier entities. The Procurement manager, along with other internal users such as Finance, and Logistics also has privileges to edit and enrich, whereas the MDM Specialist may access, edit, and approve the data.

Event Processor

The preconfigured event processor collects events for Organization Customer, Supplier, and Contact Person entities to run against their respective match algorithm.

Web UI

Organization Customer Onboarding Web UI

A Web UI configuration is provided which allows direct user interface to onboard, maintain, and approve customer data as it pertains to the company's operational processes. In the Web UI, users from Sales, Finance, Logistics, Procurement, Merchandiser as well as MDM Specialist are able to carry out their day to day activities.

Supplier Onboarding Web UI

A Web UI configuration for Supplier Onboarding is provided which allows users to onboard, maintain, and approve supplier data. In the Web UI, users from Procurement, Finance, Logistics, and the MDM Specialist are able to collaborate with each other and interface with their supplier data.

The Web UI configuration also provides the means to create new contacts to be associated with supplier entities.

For more information on configuring a Web UI, see the **Creating a New Web UI** topic in the **Web User Interfaces / Web UI Setup and User Guide** documentation.

Inbound Integration Endpoint

No inbound integration endpoint is provided with the Centralized Configuration.

A Find Similar algorithm is provided to support the Onboarding workflow user experience. This allows users to determine whether a given party they intend to onboard, already exists in the system.

Component Models

The following component models are included in the Centralized Configuration file:

- Matching, configured for:
  - Organization Customer
  - Supplier
  - Contact Person
- Matching – Merge Golden Record, configured for:
  - Organization Customer

- Supplier
- Contact Person

For more information, see the **Component Models** topic in the **System Setup / Super User** documentation.

Centralized MDM Style Configuration for Supplier with Self-Service

As a compliment to the Centralized Configuration, the initial configurations include an extension that allows supplier users to actively participate in the onboarding and maintenance of master data.

The configuration includes:

- Supplier Privileges and User Groups
- Supplier Classification Data Model
- Supplier- and Contact Person onboarding and maintenance workflows that includes the supplier persona
- Workflows for setting up suppliers for self service
- A Supplier Web UI

Data Model

Supplier Entities Classification

Supplier Entities Classification folder structures are included for Supplier Onboarding with Self-Service. Entity objects for a particular supplier are referenced to a corresponding classification folder for the purpose of governing and ensuring that supplier users may only view and edit entities associated to their respective supplier.

Workflows

As an extension to the support of data origination / authorship within MDM, a self-service onboarding mechanism is made available to external supplier users who must interface with MDM to manage their supplier information.

Supplier Onboarding Workflow with Self-Service

The Supplier Onboarding Workflow with Self-Service is intended for contribution by the following personas:

- Procurement Manager
- Supplier Admin User
- Supplier User
- Finance Specialist
- Logistics Specialist
- MDM Specialist

This workflow not only allows for the onboarding and enrichment of supplier entities, but also facilitates the use case of supplier admin users self-servicing their relevant supplier entities by directly interfacing with MDM.

Contact Person Onboarding Workflow

As Supplier Admin Users and Supplier Users onboard new supplier location entities, they may also need to create new contact persons to associate to the new supplier location entities. The initial configurations provide a Contact Onboarding workflow for this very purpose.

User Privileges

When it comes to Supplier Onboarding with Self-Service activities, Supplier User Groups are included to provided delineation of users for different suppliers. The Supplier Admin User can create additional users called Supplier Users who may optionally have administrative privileges. Both the Supplier Admin User and Supplier Users can view and edit only the entities associated with their respective Supplier.

Web UI

Supplier Onboarding with Self-Service Web UI

A Web UI configuration for Supplier Onboarding with Self-Service is provided which allows internal users to onboard, maintain, and approve supplier data. Additionally, external supplier users may also interface directly with MDM to enrich and maintain their supplier data. This provides suppliers as well as internal users a collaborative interface to ensure supplier data is sustained and governed.

The Web UI configuration also provides Supplier Admin Users the means to create new contacts to be associated with supplier entities as well as creating additional Supplier Users and assigning or revoking admin privileges for those users.

For more information on configuring a Web UI, see the **Creating a New Web UI** topic in the **Web User Interfaces / Web UI Setup and User Guide** documentation.

Inbound Integration Endpoint

No inbound integration endpoint is provided with the Centralized Configuration.

Customer MDM SAP Data Model Configuration

The initial configurations include data modeling SAP business processes in a B2B scenario. As many companies utilize SAP as their primary ERP solution, the goal of this configuration is to allow for a more seamless bridge between customer master data and SAP data structures.

The SAP configuration will include the SAP data model and corresponding business conditions to enforce SAP data modeling constraints.

Data Model

The SAP data model configuration uses the same Organization Customer entity types as the other configuration sets. The intent of this configuration is to be able to model complex SAP data structures to provide a more seamless user experience for companies that utilize SAP as their primary ERP platform. As such, the usage of relationships in between data structures is highly emphasized.

Data Model Constraints

The support of complex SAP data structures denotes the supporting of data integrity complexities and maintain proper constraints between the data structures. This configuration includes basic SAP data integrity validation

rules to sure data integrity is maintained in MDM.

For more information on SAP data model constraints, it is recommended to research SAP documentation for the Sales & Distribution module for customers and vendors, as well as the relevant sections in this enablement guideline as a complimentary resource.

Extension: Dun & Bradstreet Web UI Configuration

A separate Dun & Bradstreet Web UI configuration is available for those who are interested in implementing the third-party integration. The Web UI configuration provides an interface for data stewards to run D&B Match to obtain the DUNS number and other relevant match information. Furthermore, data stewards may also level the Web UI configuration to run the D&B Company Profile service to obtain a more detailed overview of an organization.

Component Model

For security reasons, D&B subscription credentials may not be loaded via STEPXML, so users are recommended to utilize the 'Easy Set-up' feature of the D&B component model within the Customer MDM application. Deployment of 'Easy Set-up' will automatically create the relevant D&B object types, attribution, and references in the system.

Extension : Supplier Self Service

As a compliment to the Centralized Configuration, the initial configurations include an extension that allows supplier users to actively participate in the onboarding and maintenance of master data.

For more information, see the **Dun & Bradstreet Integration** topic in this documentation.

The configuration includes:

- Supplier Privileges and User Groups
- Supplier Classification Data Model
- Supplier- and Contact Person onboarding and maintenance workflows that includes the supplier persona
- Workflows for setting up suppliers for self service
- A Supplier Web UI

Sample Data Load Files

For the Consolidated, Coexistence, and Centralized configurations, three sample data Excel files are available for each of the three customer entity types to be imported via the match and merge import configuration (i.e., Individuals, Organizations, and Contact Persons).

The names of these sample data files are:

Extension: Supplier Self Service

As a compliment to the Centralized Configuration, the initial configurations include an extension that allows supplier users to actively participate in the onboarding and maintenance of master data.

The configuration includes:

- SampleIndividuals.xlsx
- SampleOrganizationCustomers.xlsx
- SampleSuppliers.xlsx
- SampleCustomerContacts.xlsx
- SampleSupplierContacts.xlsx

Note: Users should import SampleOrganizationCustomers.xlsx before importing SampleCustomerContacts.xlsx and import SampleSuppliers.xlsx before importing SampleSupplierContacts.xlsx.

Sample data files must be imported with the Match and Merge Importer via the Asset Importer Widgets available within the Web UI.