Informatica®

Informatica® Multidomain MDM
10.5

Sample ORS Guide

Informatica Multidomain MDM Sample ORS Guide
10.5
September 2022

© Copyright Informatica LLC 2001, 2022

This software and documentation are provided only under a separate license agreement containing restrictions on use and disclosure. No part of this document may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording or otherwise) without prior consent of Informatica LLC.

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation is subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License.

Informatica and the Informatica logo are trademarks or registered trademarks of Informatica LLC in the United States and many jurisdictions throughout the world. A current list of Informatica trademarks is available on the web at https://www.informatica.com/trademarks.html. Other company and product names may be trade names or trademarks of their respective owners.

Portions of this software and/or documentation are subject to copyright held by third parties. Required third party notices are included with the product.

The information in this documentation is subject to change without notice. If you find any problems in this documentation, report them to us at infa_documentation@informatica.com.

Informatica products are warranted according to the terms and conditions of the agreements under which they are provided. INFORMATICA PROVIDES THE INFORMATION IN THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.

Publication Date: 2022-09-29

Table of Contents

Preface 5

Informatica Resources 5

- Informatica Network 5
- Informatica Knowledge Base 5
- Informatica Documentation 5
- Informatica Product Availability Matrices 6
- Informatica Velocity 6
- Informatica Marketplace 6
- Informatica Global Customer Support 6

Chapter 1: Introduction 7

Overview 7

Chapter 2: Data Source Configuration 9

Source Systems 9

Data Origin by Source System 10

Chapter 3: Data Model and Schema Configuration 12

Overview of Data Model and Schema Configuration 12

Data Model Entity-Relationship diagram 12

Base Objects 13

Landing Tables 13

Staging Tables 14

Chapter 4: Data Cleansing Configuration 15

Source System Cleanse Maps 15

Cleanse Functions 17

Chapter 5: Match and Merge Configuration 20

Trust Settings 20

Trust Validation Rules 20

Match Rule Configuration 21

- Match Settings for Party Base Object 22
- Match Settings for Address Base Object 22
- Match Setting for the Party Group Object 22
- Match Settings for Party Group Rel 22
- Match Settings for Product Base Object 23

Chapter 6: Hierarchy Manager (HM) Configuration 24

Entity Objects and Types 24

Hierarchies. 25
Relationship Objects and Types. 25
Profiles. 26
Hierarchy Manager Console. 27

Preface

Refer to the Informatica® *Multidomain MDM Sample ORS Guide* to familiarize yourself with the Operational Reference Store (ORS) configuration. The sample ORS is a repository that is used with the sample applications included in the Multidomain MDM Resource Kit. Understand the ORS configurations related to the data model, data cleansing, and match rules.

Informatica Resources

Informatica provides you with a range of product resources through the Informatica Network and other online portals. Use the resources to get the most from your Informatica products and solutions and to learn from other Informatica users and subject matter experts.

Informatica Network

The Informatica Network is the gateway to many resources, including the Informatica Knowledge Base and Informatica Global Customer Support. To enter the Informatica Network, visit https://network.informatica.com.

As an Informatica Network member, you have the following options:

- Search the Knowledge Base for product resources.
- View product availability information.
- Create and review your support cases.
- Find your local Informatica User Group Network and collaborate with your peers.

Informatica Knowledge Base

Use the Informatica Knowledge Base to find product resources such as how-to articles, best practices, video tutorials, and answers to frequently asked questions.

To search the Knowledge Base, visit https://search.informatica.com. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team at KB_Feedback@informatica.com.

Informatica Documentation

Use the Informatica Documentation Portal to explore an extensive library of documentation for current and recent product releases. To explore the Documentation Portal, visit https://docs.informatica.com.

If you have questions, comments, or ideas about the product documentation, contact the Informatica Documentation team at infa_documentation@informatica.com.

Informatica Product Availability Matrices

Product Availability Matrices (PAMs) indicate the versions of the operating systems, databases, and types of data sources and targets that a product release supports. You can browse the Informatica PAMs at https://network.informatica.com/community/informatica-network/product-availability-matrices.

Informatica Velocity

Informatica Velocity is a collection of tips and best practices developed by Informatica Professional Services and based on real-world experiences from hundreds of data management projects. Informatica Velocity represents the collective knowledge of Informatica consultants who work with organizations around the world to plan, develop, deploy, and maintain successful data management solutions.

You can find Informatica Velocity resources at http://velocity.informatica.com. If you have questions, comments, or ideas about Informatica Velocity, contact Informatica Professional Services at ips@informatica.com.

Informatica Marketplace

The Informatica Marketplace is a forum where you can find solutions that extend and enhance your Informatica implementations. Leverage any of the hundreds of solutions from Informatica developers and partners on the Marketplace to improve your productivity and speed up time to implementation on your projects. You can find the Informatica Marketplace at https://marketplace.informatica.com.

Informatica Global Customer Support

You can contact a Global Support Center by telephone or through the Informatica Network.

To find your local Informatica Global Customer Support telephone number, visit the Informatica website at the following link:
https://www.informatica.com/services-and-training/customer-success-services/contact-us.html.

To find online support resources on the Informatica Network, visit https://network.informatica.com and select the eSupport option.

CHAPTER 1

Introduction

This chapter includes the following topic:

- Overview, 7

Overview

The following configuration areas of a Informatica MDM Hub implementation are described in this guide:

- Data source configuration
  - Source systems
  - Data origin by source system
- Data model and schema configuration
  - Data model entity-relationship diagram
  - Base objects
  - Landing tables
  - Stage tables
- Data cleansing configuration
  - Source system cleanse maps
  - Cleanse functions
- Match and merge configuration
  - Trust settings
  - Trust validation rules
  - Match rule configuration with match examples
- Hierarchy Manager configuration
  - Entity objects and types
  - Hierarchies
  - Relationship objects and types

The following Informatica MDM Hub Resource Kit applications rely on the sample ORS and configuration. For more information on these applications, refer to the documentation in the Resource Kit sample folder for the application.

- Informatica Data Director demo application
  - Customer and product Master Data Management
  - Informatica MDM workflow example
  - Organization, household and product Hierarchy Management
- Business Process Management integration with Lombardi Teamworks
  - Integration with SIF APIs using SOAP
  - Integration with business entity services using EJBs
- SIF SDK usage and SOA enablement (Web Service examples)
  - Business entity services creation from SIF APIs
  - Lookup code management services
  - Cross-reference services

CHAPTER 2

Data Source Configuration

This chapter includes the following topics:

- Source Systems, 9
- Data Origin by Source System, 10

Source Systems

This sample implementation of the Informatica MDM Hub integrates primary source applications, one data augmentation source, and a reference system. These sources provide the data for Party, Address, Household, and Product information.

The primary data sources are:

Informatica Data Director

A data governance application that provides an interface to data that is stored and managed by the Informatica MDM Hub.

SFA

A Sales Force Automation system providing party, address, household and account data.

Legacy

A legacy system providing party, address, household and relationship data

Product

A PIM system providing product and product relationship data.

The augmentation data source is:

Org Data Inc

Org Data Inc (ODI) source for additional party relationship and organization data.

The reference system is:

Lookups

A source system providing reference table data for lookups.

Data Origin by Source System

The data from the five source systems are associated with each system as follows:

SFA

- Party
- Address
- Account
- Electronic Address
- Country
- State
- Party Phone
- Household
- Party Group relationship
- Party to Product relationship
- Party Phone relationship
- Party to Address
- Party to Party Relationship

Legacy

- Party
- Shipping Address
- Billing Address
- Household
- Household to Party relationship
- Party to Billing Address relationship
- Party to Shipping Address Relationship

Product

- Product
- Product relationship
- Product to Product relationship

ODI

- Party
- Address
- Party Phone
- Party Organization detail
- Party Hierarchy data
- Party Address Relationship

Lookup

- Type codes for Phone, Product, Address and Electronic Address

Lookup codes for Country, State, Gender, Marital Status, Product Status, Name Prefix, Generation Suffix etc.

CHAPTER 3

Data Model and Schema Configuration

This chapter includes the following topics:

- Overview of Data Model and Schema Configuration, 12
- Data Model Entity-Relationship diagram, 12
- Base Objects, 13
- Landing Tables, 13
- Staging Tables, 14

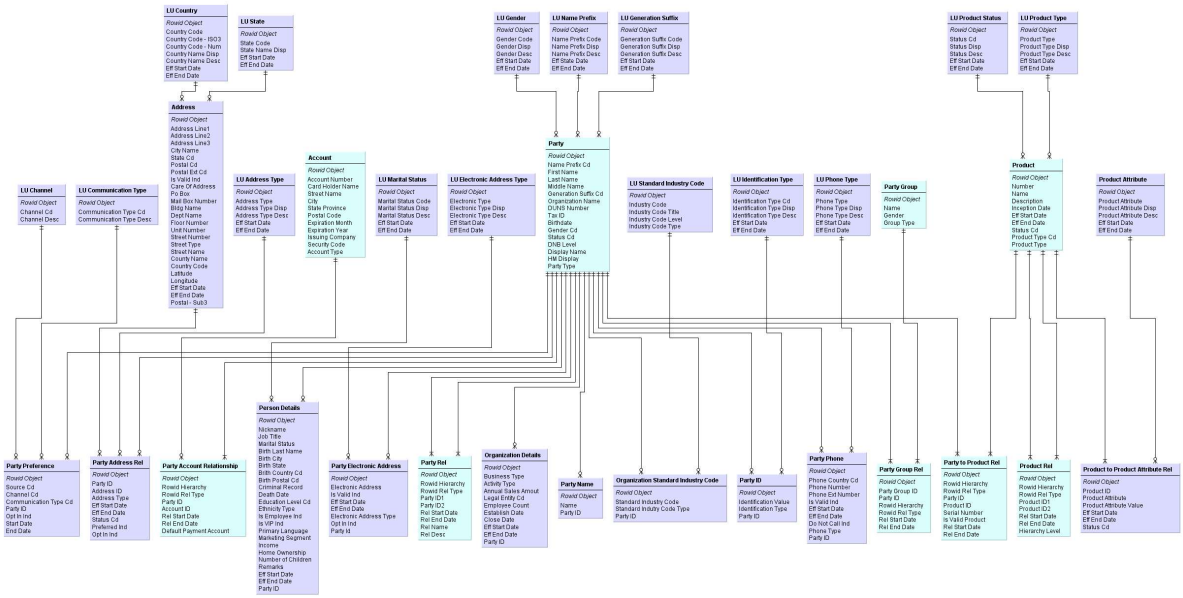
Overview of Data Model and Schema Configuration

This section describes the data model for the sample implementation, and defines the base objects, landing tables, and staging tables that make up the sample ORS.

Data Model Entity-Relationship diagram

A Party data model has been defined to address the Master Data needs for Individual, Organization, and their relationship to Account data. The data model also contains a Product data model to address product Master Data.

The following diagram, generated by the Informatica MDM Schema Viewer, depicts the data model configured in the sample ORS.



Base Objects

The sample Operational Reference Store has the following data models with base objects defined:

- Party
- Product
- Lookup
- Hierarchy Manager

Landing Tables

The MDM Hub sample Operational Reference Store has landing tables defined for data from the following source systems:

- Legacy System Landing Tables
- SFA System Landing Tables
- Lookup System Landing Tables
- ODI System Landing Tables

Staging Tables

The sample Operational Reference Store has the following data models that have staging tables defined:

- Party Data Model Stage Tables
- Product Data Model Stage Tables
- Lookup Stage Tables

CHAPTER 4

Data Cleansing Configuration

This chapter includes the following topics:

- Source System Cleanse Maps, 15
- Cleanse Functions, 17

Source System Cleanse Maps

One of the primary methods to bring data into the Informatica MDM Hub from source systems is through the land, stage, and load processes. During the stage process, the data can be modified by data standardization routines, such as routines that remove erroneous characters, and other data quality procedures. This process is referred to as data cleansing.

Cleanse maps define the rules for how to move data from landing tables to staging tables. These maps are based on one landing table and one staging table. Each field in the landing table can be directly mapped to a field in the staging table or the field can be modified by mapping it through one or more cleanse functions before mapping it to the staging table field. A typical cleanse routine found on a cleanse map might remove excess spaces (trim) from the data, or change all of the characters in a string field to uppercase.

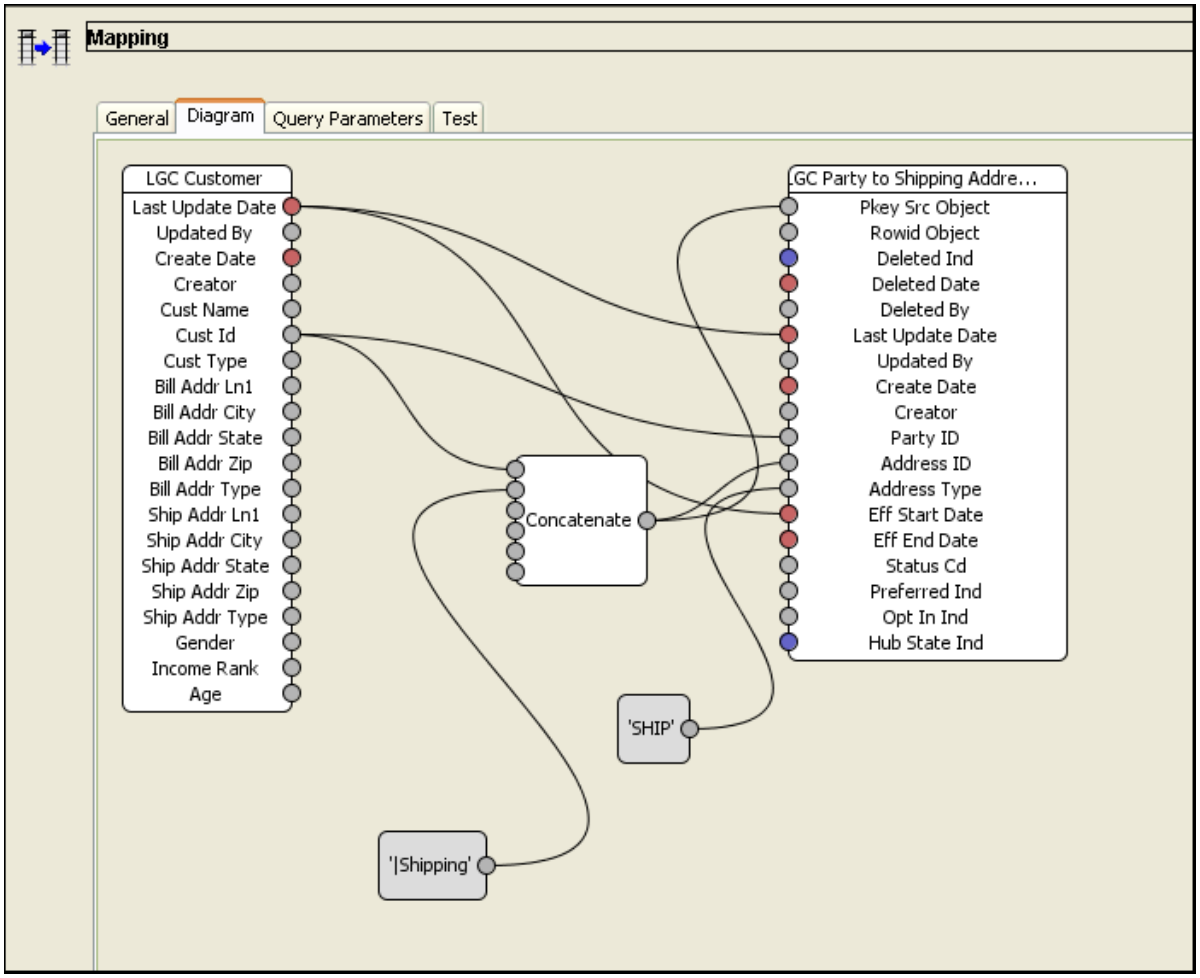
These maps are also called stage maps and are executed using a batch process by running the stage job associated with the specific staging table used by the map. The maps can also be called in real-time through calls to the SIF API.

The cleanse maps defined in the MDM Hub sample Operational Reference Store are classified based on the following systems:

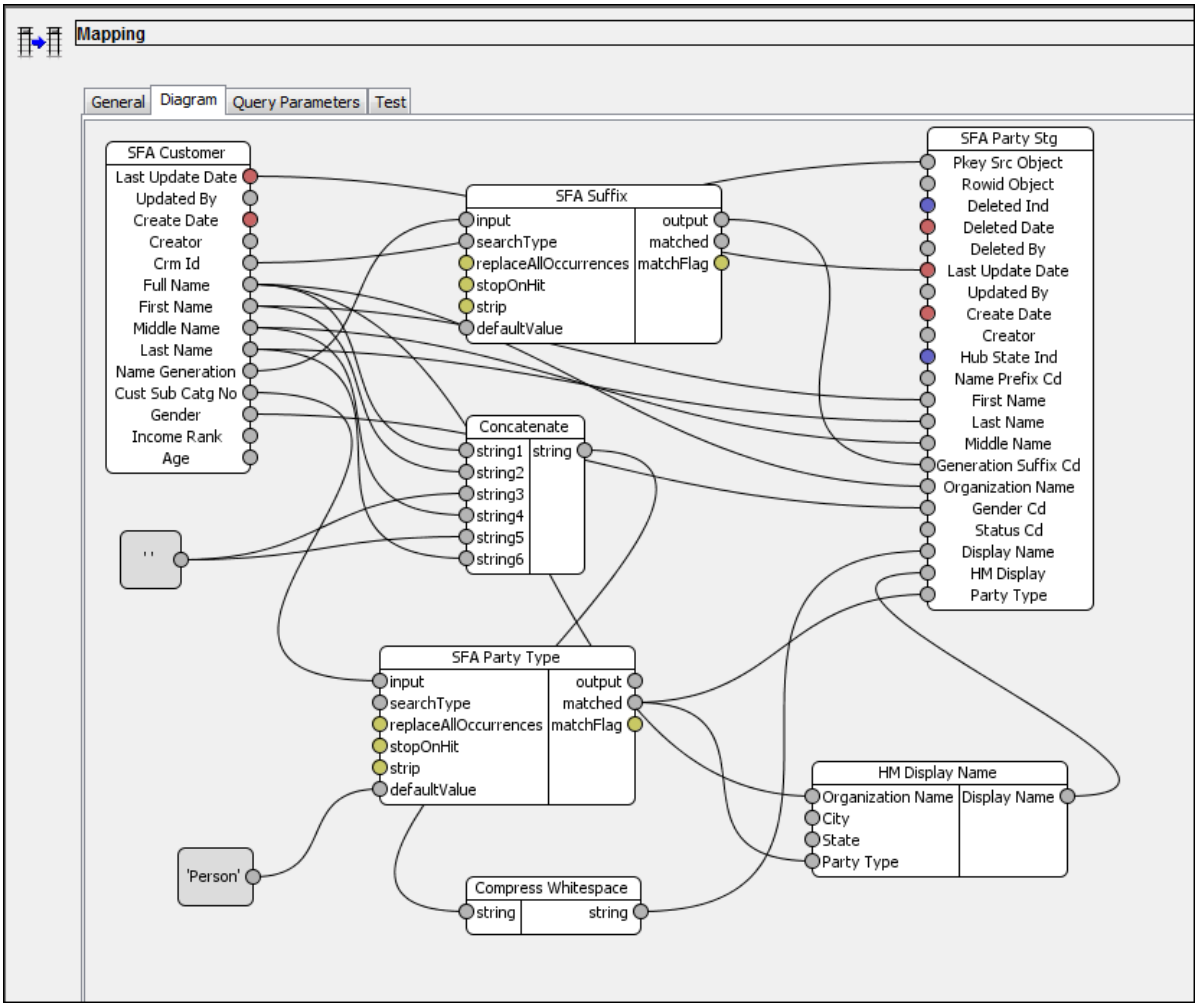
- Legacy
- SFA
- Lookup
- Product
- ODI

Some cleanse maps are simple while others are complex, using multiple custom cleanse functions.

The following image shows the sample of a simple cleanse map:



The following image shows the sample of a complex cleanse map:



A sample of the custom cleanse functions used on the cleanse maps is defined in the section "Cleanse Functions."

Cleanse Functions

Informatica MDM Hub comes with a standard set of cleanse functions that consist of common string manipulation functions, logical operations, data conversion functions, and pre-built cleanse lists (a specific type of cleanse function). You can combine these standard functions into custom functions that perform data manipulation specific to the data cleanse requirements of a particular source system.

A number of custom cleanse functions have been created as part of the sample ORS configuration. These functions can be found in the Cleanse Function tool. They are stored in the Custom, CustomUtilityLibrary, General Processing, and IDD Cleanse and Validation Library folders. There are also some sample cleanse lists in the Noise Filters folder.

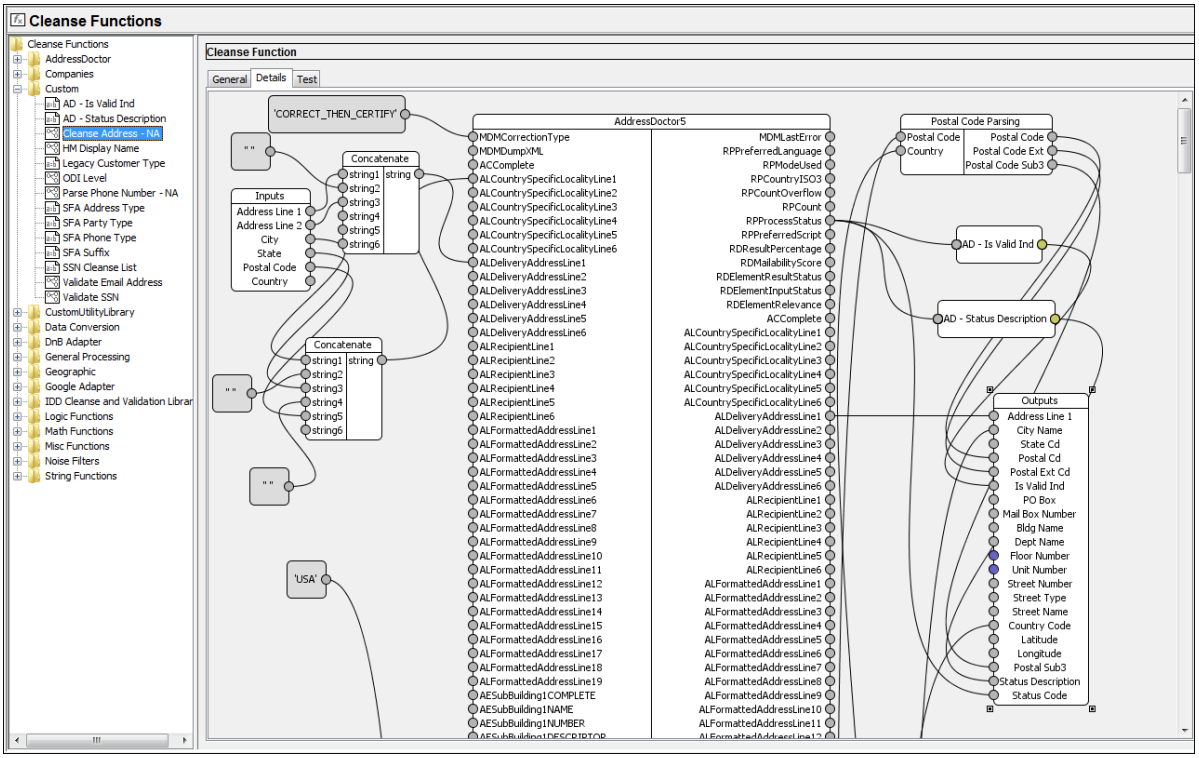
In addition to these custom functions, the sample ORS contains cleanse function libraries (folders) for third-party data quality tools (for example, Informatica Address Verification) and third-party data service providers. You can access the functionality of these products using special adapters developed on the Informatica MDM Open Cleanse architecture that allows for plugging-in third-party data quality tools.

These third-party cleanse adapters can be purchased separately, as needed, from Informatica. For more details, please contact your Informatica Account Manager.

Note: The cleanse functions based on these products will not function unless the underlying third-party software and Informatica MDM cleanse adapter is correctly installed and configured on your system.

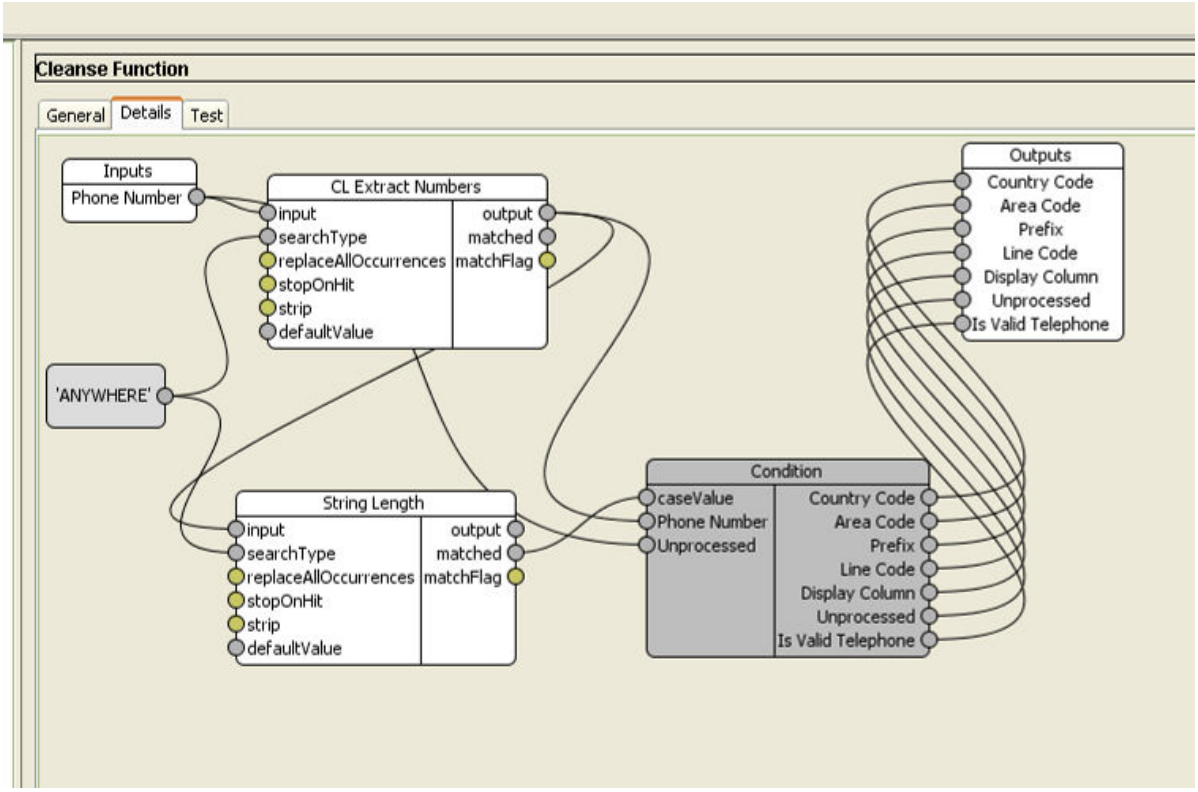
Cleanse Address - NA

This cleanse graph function is used to cleanse North American (NA) addresses. It combines the Informatica Address Verification cleanse function with other cleanse functions to create a complex function that is used as a component of the address cleanse maps. The function can also be called as a stand alone service through the SIF API.



Parse Phone Number - NA

This cleanse graph function parses phone numbers into the different components of a North American phone number.



CHAPTER 5

Match and Merge Configuration

This chapter includes the following topics:

- Trust Settings, 20
- Trust Validation Rules, 20
- Match Rule Configuration, 21

Trust Settings

The MDM Hub sample Operational Reference Store provides trust settings for the following base objects:

- LU Country
- LU Electronic Address Type
- LU Gender
- LU Generation Suffix
- LU Marital Status
- LU Name Prefix
- LU Phone Type
- LU State
- Party
- Party Group

Trust Validation Rules

The "Downgrade Unknown Gender" Trust Validation rule on the Party base object downgrades the name fields and the Gender Code field if the Gender code is not "M" or "F".

Rule Name:

Downgrade Unknown Gender

Rule Type:

Domain Check

Rule Columns	Downgrade Percentage
First Name	50
Last Name	50
Gender Cd	50
Middle Name	50

Validation Rule SQL:

```
WHERE S.GENDER_CD NOT IN( 'M', 'F' )
```

Match Rule Configuration

Match rules can be set up in Informatica MDM to match and merge the duplicate records not only from different source systems but also from within the same source. Match rules can be set up using attributes from any related base object tables in the Informatica MDM target data model. Below are the steps and components involved in defining Match rules.

1. Identify the Base Objects that need data Matched and Merged.
2. Determine if a Base Object requires Fuzzy Match capabilities.
3. Define Path to related base objects.
4. Define Match Columns for the Base Object. Attributes from related base objects can be used only if a Path is setup first.
5. Define Match Rule Sets if different Match Rules need to be run at different times.
6. Create Match Rules using the various combinations of Match Columns.
7. Set Match Rules as Auto-Merge or Manual rules depending on the precision of the match rule.

Here are the base objects that are configured for matching in the Sample Schema:

Party

Match rules for Individual and Organization data

Match data based on name and address

Address

Match based on address information

Party Address Rel

Match based on Party and Address IDs and Address Type.

Party Group

Match based on group type and address

Party Group Rel

Match based on Party and Party Group IDs, Hierarchy and Rel Type

Product

Match based on Product Name and Number

Note: The match configuration for the Party, Party Group, and Address base objects are detailed in the next two sections.

Match Settings for Party Base Object

Configure the following match settings for the Party base object:

- Match paths
- Match columns
- Rule sets
- Match rules
- Fuzzy key settings

Match Settings for Address Base Object

Configure the following match settings for the Address base object:

- Match paths
- Match columns
- Rule sets
- Match rules
- Fuzzy key settings

Match Setting for the Party Group Object

Configure the following match settings for the Party Group base object:

- Match paths
- Match columns
- Rule sets
- Match rules
- Fuzzy key settings

Match Settings for Party Group Rel

Configure the following match settings for the Party Group Rel base object:

- Match paths
- Match columns
- Rule sets
- Match rules
- Fuzzy key settings

Match Settings for Product Base Object

Configure the following match settings for the Product base object:

- Match paths
- Match columns
- Rule sets
- Match rules
- Fuzzy key settings

CHAPTER 6

Hierarchy Manager (HM) Configuration

This chapter includes the following topics:

- Entity Objects and Types, 24
- Hierarchies, 25
- Relationship Objects and Types, 25
- Profiles, 26
- Hierarchy Manager Console, 27

Entity Objects and Types

An entity object is a base object that has been configured to store information about HM entities.

Entity types are defined for the following entity objects in the MDM Hub sample Operational Reference Store:

- Account
- Party
- Party Group
- Product

The following table lists the entity objects and the entity types that are defined in the MDM Hub sample Operational Reference Store:

Entity Object	Entity Types
Account	Credit Card
Party	Organization Person
Party Group	Household
Product	Product Product Group Account

Hierarchies

A hierarchy is logical grouping of relationship types. Hierarchies allow you to group the same data into multiple hierarchy structures.

The following hierarchies have been defined for the sample schema:

- Customer
- ODI
- Product

Relationship Objects and Types

A relationship object is a base object that has been configured to store information about HM relationships.

Relationship types are defined for the following relationship objects in the MDM Hub sample Operational Reference Store:

- Party Account Relationship
- Party Group Rel
- Party Rel
- Party to Product Rel
- Product Rel

The following table lists the relationship objects and relationship types that are defined in the MDM Hub sample Operational Reference Store:

Relationship Object	Relationship Types
Party Account Relationship	Credit Card to Organization
Party Group Rel	Contains member
Party Rel	Employs is ODI parent of
Party to Product Rel	Person owns product Person owns Product Group
Product Rel	Product Group is parent of Product Product Group is parent of Product Group Product is parent of Product Product is parent of Account

Profiles

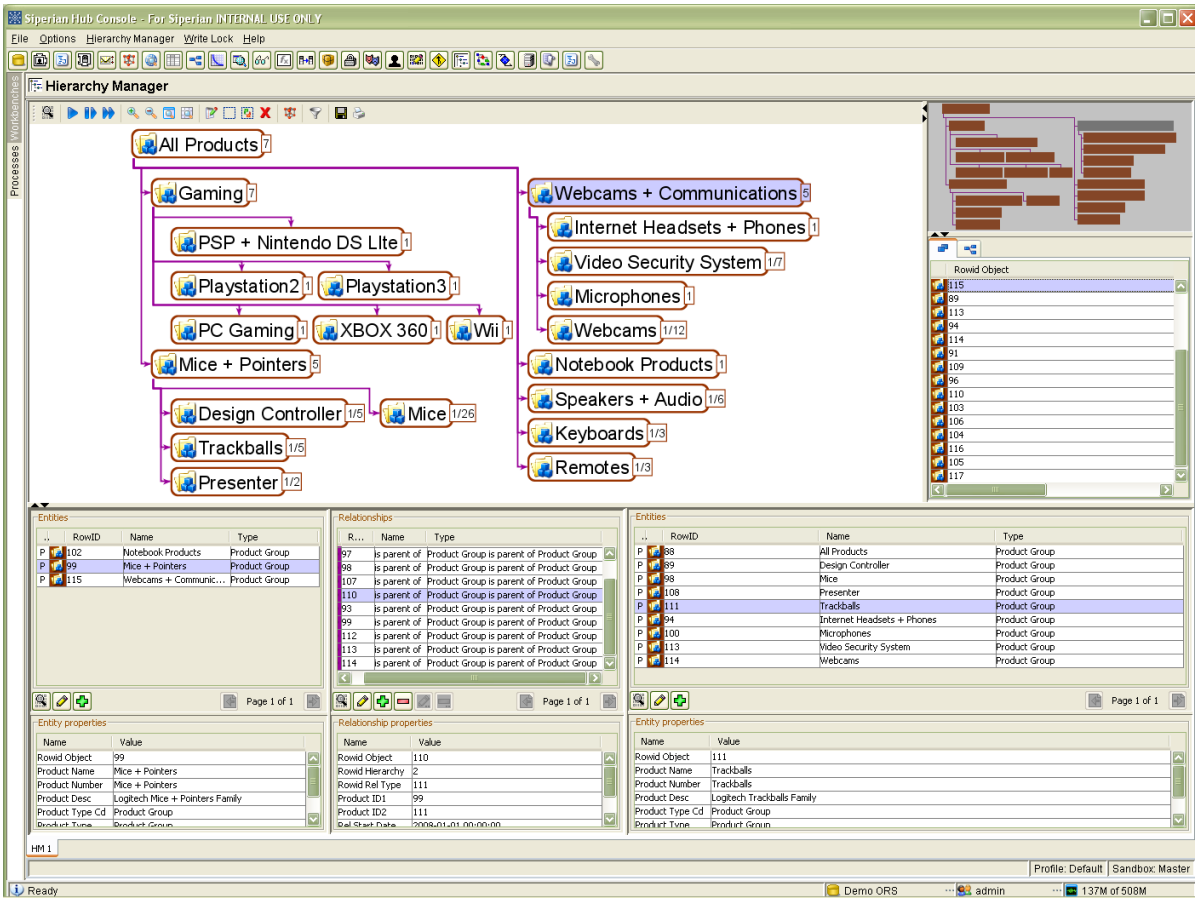
Profiles are used to define user access to Hierarchy Manager objects. A new profile called Default is created before you access the Hierarchy Manager.

The following table lists the entity objects and the entity types of the Default profile:

Entity Object	Entity Type
Party Group	Household
Product	Account Product Product Group
Account	Credit Card
Party	Organization Person

The following table lists the relationship objects and the relationship types of the Default profile:

Relationship Object	Relationship Type
Party to Product Rel	person owns product person prefers product group
Product Rel	Product group is parent of Product. Product group is parent of Product Group Product is parent of Account Product is parent of Parent
Party Group Rel	contains member
Party Rel	employs is ODI parent of
Party Account Relationship	Credit Card to Organization

Hierarchy Manager Console

The following figure shows the Informatica MDM Hub Hierarchy Manager console: