

Data Quality   
Components  
for SSIS

# **Data Quality Components** for SSIS

Quick Start Guide

**Melissa Data Corporation**

## Copyright

Companies, names, and data used in examples herein are fictitious unless otherwise noted. No part of this document may be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without the express written permission of Melissa Data Corporation. This document and the software it describes are furnished under a license agreement, and may be used or copied only in accordance with the terms of the license agreement.

Copyright © 2016 by Melissa Data Corporation. All rights reserved.

Information in this document is subject to change without notice. Melissa Data Corporation assumes no responsibility or liability for any errors, omissions, or inaccuracies that may appear in this document.

## Trademarks

Address Object, Phone Object, Name Object, Email Object, GeoCoder Object, MatchUp Object, and SmartMover are registered trademarks of Melissa Data Corp.

Windows is a registered trademark of Microsoft Corp.

The following are registered trademarks of the United States Postal Service®: United States Postal Service; USPS; Suite<sup>Link</sup>; ZIP Code.

All other brands and products are trademarks of their respective holder(s).

## Melissa Data Corporation

22382 Avenida Empresa  
Rancho Santa Margarita; CA 92688-2112  
Phone: 1-800-MELISSA (1-800-635-4772)  
Fax: 949-589-5211

E-mail: [info@MelissaData.com](mailto:info@MelissaData.com)  
Internet: [www.MelissaData.com](http://www.MelissaData.com)

For the most recent version of this document, visit  
<http://www.melissadata.com/>

Document Code: DQTENTSSISDQCQSG  
Revision Number: 17102012.15

---

# Table of Contents

---

<b>Overview</b> .....	<b>1</b>
Data Quality Components for SQL Server	
Integration Services (SSIS) .....	1
About SSIS .....	2
System Requirements .....	2
License and Demo Information .....	2
<b>Functionality Overview</b> .....	<b>3</b>
Contact Verify Component .....	3
Global Verify Component: .....	4
SmartMover Component (Move Updates): .....	4
Fuzzy Match Component: .....	4
MatchUp Component (Duplicate Matching): ...	4
<b>Advanced Capabilities</b> .....	<b>5</b>
<b>Installation</b> .....	<b>6</b>

# Overview

---

### **Data Quality Components for SQL Server Integration Services (SSIS)**

Melissa Data's Data Quality Components for SSIS are a suite of custom Data Transformation components for Microsoft's SQL Server SSIS that encapsulate the six levels of complete Data Quality as defined by Gartner®. High quality data is the basis for solid business decisions. Having an accurate view of your organization, your customers, and your business environment allows organizations to maximize profitability and reduce costly operational inefficiencies. Melissa Data's Data Quality Components for SSIS are the most affordable, full featured enterprise data quality tools to help IT professionals integrate increasingly complex enterprise data - in different formats, structures, and sources, and deliver a single view of critical information assets.

#### **Generalized Cleansing**

Corrects data values to meet specific business standards, customer business rules, or relationship constraints.

#### **Parsing, Standardization, And Verification**

Parses and restructures data into a common format to build more consistent data, such as standardizing addresses to USPS® specifications, or to custom-defined values and patterns specific to a particular business need. Also verifies addresses actually exist.

#### **Matching**

Consolidates customer records into identifiable groups using sophisticated matching technology, and links or merges related records within or across disparate datasets.

#### **Enrichment**

Adds value to customer data by attaching additional bits of data from other sources including; latitude/longitude coordinates; demographic data; full name parsing; phone number verification; and email validation.

## Features and Benefits

- Delivers a single view of the customer
- Supports data quality and MDM initiatives
- Leverages MS SSIS to profile, cleanse and integrate critical information assets
- Lowest cost of ownership

## About SSIS

Microsoft's SSIS is a scalable data integration platform bundled with Microsoft's flagship SQL Server database that features exceptional ETL capabilities with stock connectors to source and target most data formats such as Excel, RDBMS systems, OLE DB connections, and flat files. Out-of-the-box Microsoft provides SSIS with some unique data transformations for sorting, aggregating, and other powerful data manipulations.

## System Requirements

### Minimum Requirements

- 32-Bit Microsoft Windows™ Server 2003, 2008, Vista, 7
- Microsoft SQL Server 2008/R2/2012 (32-bit)
- 6 GB hard-disk space

### Recommended Requirements

- 64-Bit Microsoft Windows™ Server 2003, 2008, Vista, 7
- Microsoft SQL Server 2008/R2/2012 (64-bit)
- 6 GB hard-disk space

## License and Demo Information

Customers who have purchased the Component or any trials of the product will be issued a License String by a Melissa Data sales representative. To contact a sales representative, please call 1-800-MELISSA (3).

## Functionality Overview

---

### **Contact Verify Component**

#### **Address Verification:**

- Validate and Correct US and Canadian addresses
- Append missing suites with Suite<sup>Link®</sup> and exclusive AddressPlus technology
- Determine delivery type (business vs. residential)
- Validate and Correct non-USPS addresses

#### **Geocoding (Spatial Processing):**

- Assign Latitude and Longitude to the rooftop
- Census Tract and Block
- CBSA information
- County Name and FIPS
- Place Name and Place Code

#### **Name Processing:**

- Parse and genderize names and multiple names
- Detect vulgarities and suspicious words in name fields
- Detect and standardize companies in the name field

#### **Phone Processing:**

- Verify telephone numbers to 7 - 10 digit level
- Determine Telco Switch location
- Detect landlines vs. cellular or VOIP
- Detect residential, business, small office/home office phones

#### **Email Processing:**

- Validate and correct email domains
- Detect mobile domains (restricted from email by FCC)
- Detect bad mailbox names (spam, noreply...)
- Correct common misspellings and syntax errors

### **Global Verify Component:**

- Validate and Correct International Addresses
- Supports 240 countries in the world

### **SmartMover Component (Move Updates):**

- Replace old addresses with new-move addresses
- Identify undeliverable and incomplete addresses
- Cut your mailing costs with postal discounts
- Increase Postal processing speed and delivery

### **Fuzzy Match Component:**

- Identify duplicate records from multiple sources
- Miscellaneous data of any type, such as departments, titles, companies, products, can be matched and duplicates eliminated
- Large toolbox of state-of-the-art fuzzy matching algorithms
- Granular control on match thresholds and fine tuning of the algorithms
- Based off of the percentage score between records, the component will direct the match results to the appropriate stream (Matches, Possible Matches and Non-Matches)

### **MatchUp Component (Duplicate Matching):**

- Identify and eliminate obvious and not-so-obvious duplicate contact records
- Create your own MatchCodes using any combination of over 35 components, from common fields like ZIP Code™, address, and last name – to less common elements like email address and Social Security Number



# Advanced Capabilities

---

### **Field Mapping**

Map output to existing fields or automatically create new output field names.

### **Save configuration settings**

All SSIS components allow the saving of configuration settings of the current project for future use. This allows power users to run packages developed in a Development environment in Test or Production without any configuration changes. This feature allows the user to provide saved configuration data for deployment with multiple packages to distribute verification processing.

### **Filtering**

Advanced filtering allows decision-makers to decide what constitutes a good or bad record. Choose from common or recommended filters or build your own using the powerful rules builder.

### **Character sets and data types**

The Melissa Data SSIS components allow changes to the character sets and data types for custom database design.

# Installation

To install, simply place the installation disc into the DVD drive and run setup.exe on the root of the DVD. The license string may be inserted during the Setup process for all purchased components, or individually, upon first use of each component. If you are having trouble installing the SSIS Components or want to install the Components manually please see:

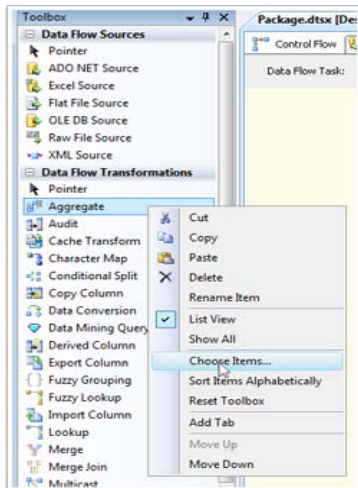
<http://wiki.melissadata.com/index.php?title=SSIS%3AInstallation>

## General Usage

The Melissa Data SSIS components can be easily added to your SQL Server Integration Service project.

### Adding a component to your SSIS toolbox

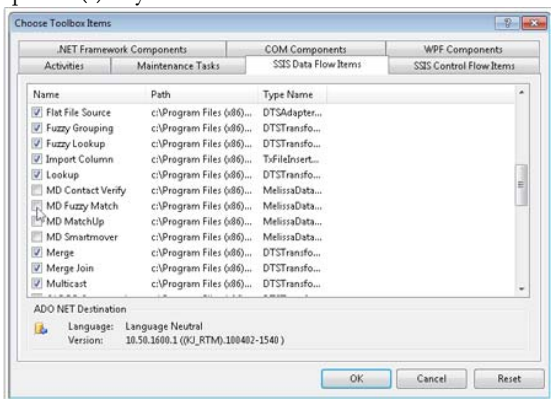
For SQL Server 2005 and 2008 only: The first time you install the component, you will need to manually add a component to your toolbox. Right-click on your toolbox and select **Choose Items**.



## Data Quality Components for SSIS

### Choose Items

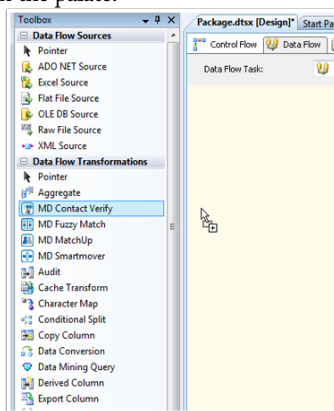
Go to the **SSIS Data Flow Items** tab and find the desired components. Select the check box(es) and click **OK** to add the component(s) to your toolbox.



Toolbox Items

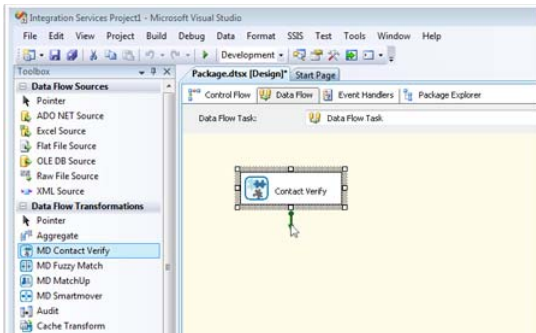
### Drag a Component from the toolbox and snap it to the data flow

To add a Melissa Data component to your project, locate the desired item on the palette.



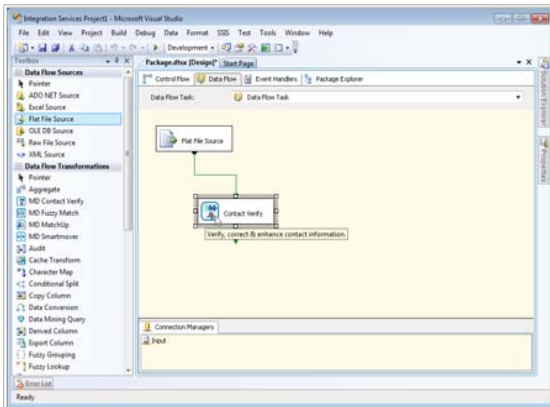
Melissa Data Components in the Palette.

Drag the component onto the Data Flow screen. This will snap the component into your workflow space.



### Connect the input data source

Select a data flow source to be your input data. SSIS is able to use many formats as Sources, including SQL tables, Excel files, flat files, or miscellaneous OLE DB data sources. Connect this data source to the component by dragging the arrow from your data flow source to the component box.



# Data Quality Components for SSIS

Connected to a Flat File Source

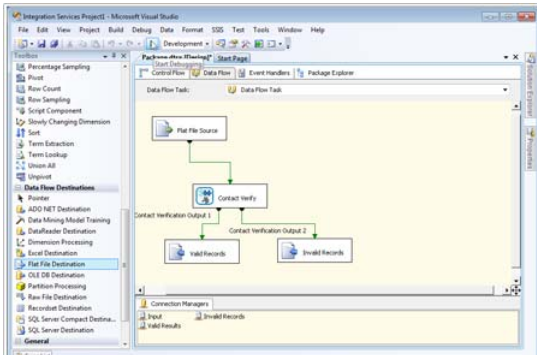
## Configure the component settings

Double-click the component to bring up the interface.

For detailed information on configuring your component, see the reference guide for the specific product in use.

## Connect to Output Data Source

Select the output data sources to direct output into. The number of output data tables will vary according to the component.



Connected to Output Tables

## Run Project

Now, the project is ready to run. It is possible to observe in real time as records flow from your input source through the component and pipe to the output sources.

