



**DATAFLUX**  
data management

# What's New in DataFlux<sup>®</sup> Data Management Studio 2.3 and 2.4



The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2012. *What's New in the DataFlux® Web Studio 2.3 and 2.4*. Cary, NC: SAS Institute Inc.

### **What's New in the DataFlux® Web Studio 2.3 and 2.4**

Copyright © 2012, SAS Institute Inc., Cary, NC, USA

All rights reserved. Produced in the United States of America.

**For a hard-copy book:** No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

**For a web download or e-book:** Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

**U.S. Government Restricted Rights Notice:** Use, duplication, or disclosure of this software and related documentation by the U.S. government is subject to the Agreement with SAS Institute and the restrictions set forth in FAR 52.227-19, Commercial Computer Software-Restricted Rights (June 1987).

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

SAS provides a complete selection of books and electronic products to help customers use SAS® software to its fullest potential. For more information about our e-books, e-learning products, CDs, and hard-copy books, visit [support.sas.com/bookstore](http://support.sas.com/bookstore) or call 1-800-727-3228.

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

# What's New in DataFlux® Data Management Studio 2.3 and 2.4

## Overview

The main enhancements for DataFlux Data Management Studio 2.4 include the following:

- new Loqate node for international address verification and geocoding.
- new JMS nodes for Java Message Service processing.
- The Java Plugin node is now production.
- improved clustering and sorting.

## New Loqate Node for International Address Verification and Geocoding

The new **Loqate** node enables you to use Loqate global address verification and geocoding technology in a data job. For more information, see the Loqate node topic in the help.

## New JMS Nodes for Java Message Service Processing

The Java Message Service (JMS) is a Java API that enables applications to create, send, receive, and read messages. It defines a common set of interfaces that enable programs written in the Java programming language to communicate with other messaging implementations. For example, a JMS message queue could be used to exchange information between applications from multiple vendors. As long as the vendor applications support JMS message queues, then they can all communicate with each other.

Data Management Studio has four new job nodes that can read from and write to a Java Message Service. For more information, see the Overview of JMS Nodes topic in the help.

## Java Plugin Node Is Now Production

The **Java Plugin** node, which was experimental, is now production. You can use this node to run a Java program in the context of a data job. For an example of how this node can be used, see the Running a Java Program in a Data Job topic in the help.

## Improved Clustering and Sorting

The performance of the **Data Sort** node for data jobs was enhanced. The cluster engine was also enhanced. As a result, the following nodes should see performance gains: the **Cluster** data job node, the **Cluster Update** data job node, the **Frequency Distribution** data job node, and Frequency Distribution within a profile.

The main enhancements for DataFlux Data Management Studio 2.3 include the following:

- new Address Verification (World 2) node for verifying international addresses in data jobs.
- new Execute Business Rule node for executing business rules in data jobs.
- new Execute Custom Metric node for applying custom metrics in data jobs.
- new Document Conversion node for converting documents in the context of data jobs.
- new Domains node and Domain Items node for working with items in a DataFlux Web Studio Reference Data Manager domain.
- new Expression node for adding a DataFlux Expression Engine Language (EEL) expression to a data job.

The Business Rules Manager now enables you to work with items in a DataFlux Web Studio Reference Data Manager domain.

- Master data entity metadata can now be shared between the DataFlux Quality Master Data Management Solution (qMDM) 3.1 and Master Data Management Foundation 2.3.
- new ODBC drivers for Sybase IQ and Salesforce.com.

## New Data Job Nodes for 2.3

You can use the **Address Verification (World 2)** node in a data job to validate, correct, or complete the address format of your input records. The node uses Address Doctor 5 software to process the addresses, so you must install Address Doctor 5 software before you can use this node. For more information about using this node, see the help topic: Working with the Address Verification (World 2) Node.

The **Execute Business Rule** node enables you to select an existing, row-based business rule and apply it to rows of data as they flow through a data job. This node bypasses the need to associate business rules to tasks and associated events to triggered rows. For an example of how to use this node, see the help topic: Using the Execute Business Rule Node.

The **Execute Custom Metric** node enables you to select an existing custom metric and apply it to a set of data in a data job. This node extends the usage of custom metrics beyond group and set rules and data profiling. For an example of how to use this node, see the help topic: Using the Execute Custom Metric Node.

The **Document Conversion** node enables you to work with unstructured data by obtaining UTF-8 encoded text for different document types. In addition to the UTF-8 encoded text, the node can output metadata found in the converted document, such as author, title, and the number of pages. For an example of how to use this node, see the help topic: Converting and Extracting a Document.

The **Domains** node lists Reference Data Manager domains. You can use the List controls in this dialog box to list domains by version and by other attributes. The **Domain Items** node lists the items within one or more Reference Data Manager domains. You can use the List controls in this dialog box to list domains and items by version and by other attributes. For an example of how to use these two nodes, see the help topic: Using Reference Data Manager Domain Items in a Job.

The **Expression** node enables you to add a DataFlux Expression Engine Language (EEL) expression to a data job. For more information about this node, see the help topic for the node. Refer to the *DataFlux Expression Language Reference Guide* for additional information about EEL.

## Other New Features for 2.3

The Business Rules Manager now enables you to work with items in a DataFlux Web Studio Reference Data Manager domain. For more information, see the help topic for the Expression Builder. The Expression Builder enables you to use pre-built DataFlux expressions to create one or more checks for a business rule.

New ODBC drivers for Sybase IQ and Salesforce.com enable you to access these data sources for profiles, data jobs, and other features in DataFlux Data Management Studio.