



Informatica® MDM Big Data Relationship
Management
10.0 HotFix 6

Release Guide

© Copyright Informatica LLC 2014, 2018

This software and documentation contain proprietary information of Informatica LLC and are provided under a license agreement containing restrictions on use and disclosure and are also protected by copyright law. Reverse engineering of the software is prohibited. 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. This Software may be protected by U.S. and/or international Patents and other Patents Pending.

Use, duplication, or disclosure of the Software by the U.S. Government is subject to the restrictions set forth in the applicable software license agreement and as provided in DFARS 227.7202-1(a) and 227.7702-3(a) (1995), DFARS 252.227-7013(1)(ii) (OCT 1988), FAR 12.212(a) (1995), FAR 52.227-19, or FAR 52.227-14 (ALT III), as applicable.

The information in this product or documentation is subject to change without notice. If you find any problems in this product or documentation, please report them to us in writing.

Informatica, Informatica Platform, Informatica Data Services, PowerCenter, PowerCenterRT, PowerCenter Connect, PowerCenter Data Analyzer, PowerExchange, PowerMart, Metadata Manager, Informatica Data Quality, Informatica Data Explorer, Informatica B2B Data Transformation, Informatica B2B Data Exchange Informatica On Demand, Informatica Identity Resolution, Informatica Application Information Lifecycle Management, Informatica Complex Event Processing, Ultra Messaging, Informatica Master Data Management, and Live Data Map are trademarks or registered trademarks of Informatica LLC in the United States and in jurisdictions throughout the world. All 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, including without limitation: Copyright DataDirect Technologies. All rights reserved. Copyright © Sun Microsystems. All rights reserved. Copyright © RSA Security Inc. All Rights Reserved. Copyright © Ordinal Technology Corp. All rights reserved. Copyright © Aandacht c.v. All rights reserved. Copyright Genivia, Inc. All rights reserved. Copyright Isomorphic Software. All rights reserved. Copyright © Meta Integration Technology, Inc. All rights reserved. Copyright © Intalio. All rights reserved. Copyright © Oracle. All rights reserved. Copyright © Adobe Systems Incorporated. All rights reserved. Copyright © DataArt, Inc. All rights reserved. Copyright © ComponentSource. All rights reserved. Copyright © Microsoft Corporation. All rights reserved. Copyright © Rogue Wave Software, Inc. All rights reserved. Copyright © Teradata Corporation. All rights reserved. Copyright © Yahoo! Inc. All rights reserved. Copyright © Glyph & Cog, LLC. All rights reserved. Copyright © Thinkmap, Inc. All rights reserved. Copyright © Clearpace Software Limited. All rights reserved. Copyright © Information Builders, Inc. All rights reserved. Copyright © OSS Nokalva, Inc. All rights reserved. Copyright Edifecs, Inc. All rights reserved. Copyright Cleo Communications, Inc. All rights reserved. Copyright © International Organization for Standardization 1986. All rights reserved. Copyright © ej-technologies GmbH. All rights reserved. Copyright © Jaspersoft Corporation. All rights reserved. Copyright © International Business Machines Corporation. All rights reserved. Copyright © yWorks GmbH. All rights reserved. Copyright © Lucent Technologies. All rights reserved. Copyright (c) University of Toronto. All rights reserved. Copyright © Daniel Veillard. All rights reserved. Copyright © Unicode, Inc. Copyright IBM Corp. All rights reserved. Copyright © MicroQuill Software Publishing, Inc. All rights reserved. Copyright © PassMark Software Pty Ltd. All rights reserved. Copyright © LogiXML, Inc. All rights reserved. Copyright © 2003-2010 Lorenzi Davide, All rights reserved. Copyright © Red Hat, Inc. All rights reserved. Copyright © The Board of Trustees of the Leland Stanford Junior University. All rights reserved. Copyright © EMC Corporation. All rights reserved. Copyright © Flexera Software. All rights reserved. Copyright © Jinfonet Software. All rights reserved. Copyright © Apple Inc. All rights reserved. Copyright © Telerik Inc. All rights reserved. Copyright © BEA Systems. All rights reserved. Copyright © PDFlib GmbH. All rights reserved. Copyright © Orientation in Objects GmbH. All rights reserved. Copyright © Tanuki Software, Ltd. All rights reserved. Copyright © Ricebridge. All rights reserved. Copyright © Sencha, Inc. All rights reserved. Copyright © Scalable Systems, Inc. All rights reserved. Copyright © jqWidgets. All rights reserved. Copyright © Tableau Software, Inc. All rights reserved. Copyright © MaxMind, Inc. All Rights Reserved. Copyright © TMat Software s.r.o. All rights reserved. Copyright © MapR Technologies Inc. All rights reserved. Copyright © Amazon Corporate LLC. All rights reserved. Copyright © Highsoft. All rights reserved. Copyright © Python Software Foundation. All rights reserved. Copyright © BeOpen.com. All rights reserved. Copyright © CNRI. All rights reserved.

This product includes software developed by the Apache Software Foundation (<http://www.apache.org/>), and/or other software which is licensed under various versions of the Apache License (the "License"). You may obtain a copy of these Licenses at <http://www.apache.org/licenses/>. Unless required by applicable law or agreed to in writing, software distributed under these Licenses is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the Licenses for the specific language governing permissions and limitations under the Licenses.

This product includes software which was developed by Mozilla (<http://www.mozilla.org/>), software copyright The JBoss Group, LLC, all rights reserved; software copyright © 1999-2006 by Bruno Lowagie and Paulo Soares and other software which is licensed under various versions of the GNU Lesser General Public License Agreement, which may be found at <http://www.gnu.org/licenses/lgpl.html>. The materials are provided free of charge by Informatica, "as-is", without warranty of any kind, either express or implied, including but not limited to the implied warranties of merchantability and fitness for a particular purpose.

The product includes ACE(TM) and TAO(TM) software copyrighted by Douglas C. Schmidt and his research group at Washington University, University of California, Irvine, and Vanderbilt University, Copyright (©) 1993-2006, all rights reserved.

This product includes software developed by the OpenSSL Project for use in the OpenSSL Toolkit (copyright The OpenSSL Project. All Rights Reserved) and redistribution of this software is subject to terms available at <http://www.openssl.org> and <http://www.openssl.org/source/license.html>.

This product includes Curl software which is Copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://curl.haxx.se/docs/copyright.html>. Permission to use, copy, modify, and distribute this software for any purpose with or without fee is hereby granted, provided that the above copyright notice and this permission notice appear in all copies.

The product includes software copyright 2001-2005 (©) MetaStuff, Ltd. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://www.dom4j.org/license.html>.

The product includes software copyright © 2004-2007, The Dojo Foundation. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://dojotoolkit.org/license>.

This product includes ICU software which is copyright International Business Machines Corporation and others. All rights reserved. Permissions and limitations regarding this software are subject to terms available at <http://source.icu-project.org/repos/icu/icu/trunk/license.html>.

This product includes software copyright © 1996-2006 Per Bothner. All rights reserved. Your right to use such materials is set forth in the license which may be found at <http://www.gnu.org/software/kawa/Software-License.html>.

This product includes OSSP UUID software which is Copyright © 2002 Ralf S. Engelschall, Copyright © 2002 The OSSP Project Copyright © 2002 Cable & Wireless Deutschland. Permissions and limitations regarding this software are subject to terms available at <http://www.opensource.org/licenses/mit-license.php>.

This product includes software developed by Boost (<http://www.boost.org/>) or under the Boost software license. Permissions and limitations regarding this software are subject to terms available at http://www.boost.org/LICENSE_1_0.txt.

This product includes software copyright © 1997-2007 University of Cambridge. Permissions and limitations regarding this software are subject to terms available at <http://www.pcre.org/license.txt>.

This product includes software copyright © 2007 The Eclipse Foundation. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://www.eclipse.org/org/documents/epl-v10.php> and at <http://www.eclipse.org/org/documents/edl-v10.php>.

This product includes software licensed under the terms at <http://www.tcl.tk/software/tcltk/license.html>, <http://www.bosrup.com/web/overlib/?License>, <http://www.stlport.org/doc/license.html>, <http://asm.ow2.org/license.html>, <http://www.cryptix.org/LICENSE.TXT>, <http://hsqldb.org/web/hsqLicense.html>, <http://httpunit.sourceforge.net/doc/license.html>, <http://jung.sourceforge.net/license.txt>, http://www.gzip.org/zlib/zlib_license.html, <http://www.openldap.org/software/release/license.html>, <http://www.libssh2.org>, <http://slf4j.org/license.html>, <http://www.sente.ch/software/OpenSourceLicense.html>, <http://fusesource.com/downloads/license-agreements/fuse-message-broker-v-5-3-license-agreement>, <http://antlr.org/license.html>, <http://aopalliance.sourceforge.net/>, <http://www.bouncycastle.org/licence.html>, <http://www.jgraph.com/jgraphdownload.html>, <http://www.jcraft.com/jsch/LICENSE.txt>, http://jotm.objectweb.org/bsd_license.html, <http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231>, <http://www.slf4j.org/license.html>, <http://nanoxml.sourceforge.net/orig/copyright.html>, <http://www.json.org/license.html>, <http://forge.ow2.org/projects/javaservice/>, <http://www.postgresql.org/about/licence.html>, <http://www.sqlite.org/copyright.html>, <http://www.tcl.tk/software/tcltk/license.html>, <http://www.jaxen.org/faq.html>, <http://www.jdom.org/docs/faq.html>, <http://www.slf4j.org/license.html>, <http://www.iodbc.org/dataspace/iodbc/wiki/IODBC/License>, <http://www.keplerproject.org/md5/license.html>, <http://www.toedter.com/en/jcalendar/license.html>, <http://www.edankert.com/bounce/index.html>, <http://www.net-snmp.org/about/license.html>, <http://www.openmdx.org/#FAQ>, http://www.php.net/license/3_01.txt, <http://srp.stanford.edu/license.txt>, <http://www.schneider.com/blowfish.html>, <http://www.jmock.org/license.html>, <http://xsom.java.net>, <http://benalman.com/about/license/>, <https://github.com/CreateJS/EaselJS/blob/master/src/easeljs/display/Bitmap.js>, <http://www.h2database.com/html/license.html#summary>, <http://jsoncpp.sourceforge.net/LICENSE>, <http://jdbc.postgresql.org/license.html>, <http://protobuf.googlecode.com/svn/trunk/src/google/protobuf/descriptor.proto>, <https://github.com/rantav/hector/blob/master/LICENSE>, <http://web.mit.edu/Kerberos/krb5-current/doc/mitK5license.html>, <http://jibx.sourceforge.net/jibx-license.html>, <https://github.com/lyokato/libgeohash/blob/master/LICENSE>, <https://github.com/hjiang/jsonxx/blob/master/LICENSE>, <https://code.google.com/p/lz4/>, <https://github.com/jedisct1/libsodium/blob/master/LICENSE>, <http://one-jar.sourceforge.net/index.php?page=documents&file=license>, <https://github.com/EsotericSoftware/kryo/blob/master/license.txt>, <http://www.scala-lang.org/license.html>, <https://github.com/tinkerpop/blueprints/blob/master/LICENSE.txt>, <http://gee.cs.oswego.edu/dl/classes/EDU/oswego/cs/dl/util/concurrent/intro.html>, <https://aws.amazon.com/ssl/>, <https://github.com/twbs/bootstrap/blob/master/LICENSE>, <https://sourceforge.net/p/xmlunit/code/HEAD/tree/trunk/LICENSE.txt>, <https://github.com/documentcloud/underscore-contrib/blob/master/LICENSE>, and <https://github.com/apache/hbase/blob/master/LICENSE.txt>.

This product includes software licensed under the Academic Free License (<http://www.opensource.org/licenses/afl-3.0.php>), the Common Development and Distribution License (<http://www.opensource.org/licenses/cddl1.php>), the Common Public License (<http://www.opensource.org/licenses/cpl1.0.php>), the Sun Binary Code License Agreement Supplemental License Terms, the BSD License (<http://www.opensource.org/licenses/bsd-license.php>), the new BSD License (<http://opensource.org/licenses/BSD-3-Clause>), the MIT License (<http://www.opensource.org/licenses/mit-license.php>), the Artistic License (<http://www.opensource.org/licenses/artistic-license-1.0>) and the Initial Developer's Public License Version 1.0 (<http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/>).

This product includes software copyright © 2003-2006 Joe Walnes, 2006-2007 XStream Committers. All rights reserved. Permissions and limitations regarding this software are subject to terms available at <http://xstream.codehaus.org/license.html>. This product includes software developed by the Indiana University Extreme! Lab. For further information please visit <http://www.extreme.indiana.edu/>.

This product includes software Copyright (c) 2013 Frank Balluffi and Markus Moeller. All rights reserved. Permissions and limitations regarding this software are subject to terms of the MIT license.

See patents at <https://www.informatica.com/legal/patents.html>.

DISCLAIMER: Informatica LLC provides this documentation "as is" without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of noninfringement, merchantability, or use for a particular purpose. Informatica LLC does not warrant that this software or documentation is error free. The information provided in this software or documentation may include technical inaccuracies or typographical errors. The information in this software and documentation is subject to change at any time without notice.

NOTICES

This Informatica product (the "Software") includes certain drivers (the "DataDirect Drivers") from DataDirect Technologies, an operating company of Progress Software Corporation ("DataDirect") which are subject to the following terms and conditions:

1. THE DATADIRECT DRIVERS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT.
2. IN NO EVENT WILL DATADIRECT OR ITS THIRD PARTY SUPPLIERS BE LIABLE TO THE END-USER CUSTOMER FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, CONSEQUENTIAL OR OTHER DAMAGES ARISING OUT OF THE USE OF THE ODBC DRIVERS, WHETHER OR NOT INFORMED OF THE POSSIBILITIES OF DAMAGES IN ADVANCE. THESE LIMITATIONS APPLY TO ALL CAUSES OF ACTION, INCLUDING, WITHOUT LIMITATION, BREACH OF CONTRACT, BREACH OF WARRANTY, NEGLIGENCE, STRICT LIABILITY, MISREPRESENTATION AND OTHER TORTS.

Publication Date: 2018-09-27

Table of Contents

Preface	6
Informatica Resources.	6
Informatica Network.	6
Informatica Knowledge Base.	6
Informatica Documentation.	6
Informatica Product Availability Matrixes.	7
Informatica Velocity.	7
Informatica Marketplace.	7
Informatica Global Customer Support.	7
 Part I: Version 10.0.0	 8
 Chapter 1: New Features, Changes, and Release Tasks (10.0 HotFix 6)	 9
New Features (10.0 HotFix 6).	9
Consolidation Process.	9
Consolidation Rules.	9
Retrieving the Preferred Records.	9
Hive Enabler Job.	10
Changes (10.0 HotFix 6).	10
Version Number of the RESTful Web Services.	10
 Chapter 2: New Features, Changes, and Release Tasks (10.0 HotFix 5)	 11
New Features (10.0 HotFix 5).	11
Pagination Support.	11
Distributed Search.	11
Security for the RESTful Web Services.	11
Processing Data in Motion.	12
Region Splitter Job.	12
Post-Clustering Job.	12
Changes (10.0 HotFix 5).	12
Updating Linked Data in Hive.	12
License File.	13
Version Number of the RESTful Web Services.	13
Deprecated Batch Jobs and Operations.	13
 Chapter 3: New Features, Changes, and Release Tasks (10.0.0 HotFix 4)	 14
New Features (10.0.0 HotFix 4).	14
Support for the Kerberos Authentication in Cloudera.	14
Debug Mode for the MULTISEARCH Operation.	14
Changes (10.0.0 HotFix 4).	14

Search.	15
Chapter 4: New Features, Changes, and Release Tasks (10.0.0 HotFix 3). . . .	16
New Features (10.0.0 HotFix 3).	16
Tokenization.	16
MapReduce Jobs.	16
Search.	17
Changes (10.0.0 HotFix 3).	18
Matching Rules File.	18
Chapter 5: New Features, Changes, and Release Tasks (10.0.0 HotFix 2). . . .	19
New Features (10.0.0 HotFix 2).	19
Advanced Matching.	19
MapReduce Jobs.	19
Search.	20
Changes (10.0.0 HotFix 2).	20
Simple Matching.	21
Updating Indexed and Linked Data.	21

Preface

The *Informatica MDM Big Data Relationship Management Release Guide* lists new features and enhancements, behavior changes between versions, and tasks you might need to perform after you upgrade from a previous version. The *Informatica MDM Big Data Relationship Management Release Guide* is written for users of Informatica MDM Big Data Relationship Management. This guide assumes that you have knowledge of the features for which you are responsible.

Informatica Resources

Informatica Network

Informatica Network hosts Informatica Global Customer Support, the Informatica Knowledge Base, and other product resources. To access Informatica Network, visit <https://network.informatica.com>.

As a member, you can:

- Access all of your Informatica resources in one place.
- Search the Knowledge Base for product resources, including documentation, FAQs, and best practices.
- View product availability information.
- 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 search Informatica Network for product resources such as documentation, how-to articles, best practices, and PAMs.

To access the Knowledge Base, visit <https://kb.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

To get the latest documentation for your product, browse the Informatica Knowledge Base at https://kb.informatica.com/_layouts/ProductDocumentation/Page/ProductDocumentSearch.aspx.

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

Informatica Product Availability Matrixes

Product Availability Matrixes (PAMs) indicate the versions of operating systems, databases, and other types of data sources and targets that a product release supports. If you are an Informatica Network member, you can access 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. Developed from the real-world experience of hundreds of data management projects, Informatica Velocity represents the collective knowledge of our consultants who have worked with organizations from around the world to plan, develop, deploy, and maintain successful data management solutions.

If you are an Informatica Network member, you can access 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 augment, extend, or enhance your Informatica implementations. By leveraging any of the hundreds of solutions from Informatica developers and partners, you can improve your productivity and speed up time to implementation on your projects. You can access Informatica Marketplace at <https://marketplace.informatica.com>.

Informatica Global Customer Support

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

To find your local Informatica Global Customer Support telephone number, visit the Informatica website at the following link:

<http://www.informatica.com/us/services-and-training/support-services/global-support-centers>.

If you are an Informatica Network member, you can use Online Support at <http://network.informatica.com>.

Part I: Version 10.0.0

This part contains the following chapters:

- [New Features, Changes, and Release Tasks \(10.0 HotFix 6\), 9](#)
- [New Features, Changes, and Release Tasks \(10.0 HotFix 5\), 11](#)
- [New Features, Changes, and Release Tasks \(10.0.0 HotFix 4\), 14](#)
- [New Features, Changes, and Release Tasks \(10.0.0 HotFix 3\), 16](#)
- [New Features, Changes, and Release Tasks \(10.0.0 HotFix 2\), 19](#)

CHAPTER 1

New Features, Changes, and Release Tasks (10.0 HotFix 6)

This chapter includes the following topics:

- [New Features \(10.0 HotFix 6\), 9](#)
- [Changes \(10.0 HotFix 6\), 10](#)

New Features (10.0 HotFix 6)

This section describes new features in version 10.0 HotFix 6.

Consolidation Process

Use the consolidation process to merge data from different records in a cluster to create a preferred or golden record based on the consolidation rules. The input data can be batch or streaming data.

For more information about the consolidation process, see the *Informatica MDM Big Data Relationship Management User Guide*.

Consolidation Rules

The consolidation process uses the consolidation rules to create a preferred record for each cluster. You must define the consolidation rules in the consolidation rules file. You can use the GETSTRATEGIES web service or command-line operation to get a list of supported consolidation rules.

For more information about the consolidation rules, see the *Informatica MDM Big Data Relationship Management Installation and Configuration Guide*.

Retrieving the Preferred Records

After you consolidate the input data, use the GETPREFERREDRECORD or PREFERREDRECORDSEARCH web service or command-line operation to retrieve the preferred records of the clusters.

For more information about the GETPREFERREDRECORD and PREFERREDRECORDSEARCH web services or command-line operations, see the *Informatica MDM Big Data Relationship Management User Guide*.

Hive Enabler Job

You can use the Hive enabler job to load the consolidated data from HDFS or repository to a Hive table.

For more information about the Hive enabler job, see the *Informatica MDM Big Data Relationship Management User Guide*.

Changes (10.0 HotFix 6)

This section describes changes in version 10.0 HotFix 6.

Version Number of the RESTful Web Services

Effective in version 10.0 HotFix 6, all the RESTful web services use v3.0 as the version number in the URLs.

The following sample URLs use v3.0 as the version number:

```
[
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/SEARCH",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/GETCLUSTER",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/GETSEARCHLAYOUT",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/GETCLUSTERLAYOUT",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/GETMULTISEARCHLAYOUT",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/MULTISEARCH",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/RULESEARCH",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/INGEST",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/GETINGESTLAYOUT",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/GETRECORDLAYOUT",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/GETRECORD",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/GETMANAGECLUSTERLAYOUT",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/MANAGECLUSTER",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/DELETERECORD",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/AUTHENTICATE",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/GETPREFERREDRECORDLAYOUT",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/GETPREFERREDRECORD",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/GETSTRATEGIES",
  "http://BDRMServer:8080/MDMBDRMSYS/v3.0/SYS/PREFERREDRECORDSEARCH"
]
```

Previously, all the RESTful web services used v2.0 as the version number in the URLs.

CHAPTER 2

New Features, Changes, and Release Tasks (10.0 HotFix 5)

This chapter includes the following topics:

- [New Features \(10.0 HotFix 5\), 11](#)
- [Changes \(10.0 HotFix 5\), 12](#)

New Features (10.0 HotFix 5)

This section describes new features in version 10.0 HotFix 5.

Pagination Support

When you perform a search, you can enable pagination for the search results and specify the maximum number of results to return. If you enable pagination, a search request returns a token with the search results. You can use the token in the subsequent requests to get the search results from cache to avoid performing the search again.

For more information about pagination, see the *Informatica MDM Big Data Relationship Management User Guide*.

Distributed Search

Use distributed search to distribute a search request to multiple region servers of the repository and perform the search on the region servers. Distributed search optimally uses the available resources and results in improved search performance.

For more information about distributed search, see the *Informatica MDM Big Data Relationship Management Installation and Configuration Guide*.

Security for the RESTful Web Services

You can configure security for the RESTful web services to prevent any unauthorized access to the data. MDM Big Data Relationship Management uses the OpenAM-based authentication system to secure the web services.

For more information about securing the RESTful web services, see the *Informatica MDM Big Data Relationship Management Installation and Configuration Guide*.

Processing Data in Motion

You can link or tokenize the input data in motion. The input data in motion indicates streaming data. Use the JSON format to stream the input data. MDM Big Data Relationship Management processes the input data and persists the linked or tokenized data in the repository.

For more information about processing data in motion, see the *Informatica MDM Big Data Relationship Management Installation and Configuration Guide* and *Informatica MDM Big Data Relationship Management User Guide*.

Region Splitter Job

Use the region splitter job to uniformly distribute the linked or tokenized data across all the regions in the repository. The uniform distribution of the linked data optimally uses the resources and improves the search performance. Run the region splitter job before you run the load clustering job for the first time. A load clustering job can use the output files of a region splitter job to distribute the linked data.

For more information about the region splitter job, see the *Informatica MDM Big Data Relationship Management User Guide*.

Post-Clustering Job

Use the post-clustering job to read the output files of an initial clustering job in HDFS and process the input data based on the mode that you configure.

You can run the post-clustering job in the following modes:

- Skip
- Recluster
- Longtail
- Export

For more information about the post-clustering job, see the *Informatica MDM Big Data Relationship Management User Guide*.

Changes (10.0 HotFix 5)

This section describes changes in version 10.0 HotFix 5.

Updating Linked Data in Hive

Effective in version 10.0 HotFix 5, if you persist the linked data in HDFS, you can update the linked data in Hive with incremental data. If you persist the linked data in a repository, you can link a Hive table to the repository table so that the Hive table accesses the repository table for the linked data.

Previously, you could not update the linked data in Hive.

License File

Effective in version 10.0 HotFix 5, you do not require a license file when you install MDM Big Data Relationship Management.

Previously, when you installed MDM Big Data Relationship Management, you had to copy the license file to the installation directory.

Version Number of the RESTful Web Services

Effective in version 10.0 HotFix 5, all the RESTful web services use v2.0 as the version number in the URLs.

The following sample URLs use v2.0 as the version number:

```
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/SEARCH",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/GETCLUSTER",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/GETSEARCHLAYOUT",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/GETCLUSTERLAYOUT",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/GETMULTISEARCHLAYOUT",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/MULTISEARCH",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/RULESEARCH",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/INGEST",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/GETINGESTLAYOUT",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/GETRECORDLAYOUT",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/GETRECORD",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/GETMANAGECLUSTERLAYOUT",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/MANAGECLUSTER",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/DELETERECORD",
"http://BDRMServer:8080/MDMBDRMSYS/v2.0/SYS/AUTHENTICATE"
```

Previously, all the RESTful web services used v1.0 as the version number in the URLs.

Deprecated Batch Jobs and Operations

Effective in version 10.0 HotFix 5, the following batch jobs and search operations that use simple matching are deprecated, and Informatica will drop support for them in the next major release:

- Initial linking job
- Initial loading job
- Incremental linking job
- Update linking job
- SEARCH operation
- RULESEARCH operation

Informatica recommends that you use initial clustering job, load clustering jobs, and MULTISEARCH operation instead.

CHAPTER 3

New Features, Changes, and Release Tasks (10.0.0 HotFix 4)

This chapter includes the following topics:

- [New Features \(10.0.0 HotFix 4\), 14](#)
- [Changes \(10.0.0 HotFix 4\), 14](#)

New Features (10.0.0 HotFix 4)

This section describes new features in version 10.0.0 HotFix 4.

Support for the Kerberos Authentication in Cloudera

In Cloudera, you can use the Kerberos authentication for the MapReduce jobs to access data in HBase and Hive. To use the Kerberos authentication, you must configure the parameters related to Kerberos in the configuration file.

For more information about the parameters related to the Kerberos authentication, see the *Informatica MDM Big Data Relationship Management User Guide*.

Debug Mode for the MULTISEARCH Operation

You can perform the MULTISEARCH operation in the debug mode. The debug mode returns performance metrics that you can use for troubleshooting purposes.

For more information about configuring the debug mode, see the *Informatica MDM Big Data Relationship Management User Guide*.

Changes (10.0.0 HotFix 4)

This section describes changes in version 10.0.0 HotFix 4.

Search

Effective in version 10.0.0 HotFix 4, the MULTISEARCH and RULESEARCH operations use multiple threads to perform parallel processing for matching. The parallel processing improves the search performance.

Previously, the MULTISEARCH and RULESEARCH operations used a single thread to perform matching.

CHAPTER 4

New Features, Changes, and Release Tasks (10.0.0 HotFix 3)

This chapter includes the following topics:

- [New Features \(10.0.0 HotFix 3\), 16](#)
- [Changes \(10.0.0 HotFix 3\), 18](#)

New Features (10.0.0 HotFix 3)

This section describes new features in version 10.0.0 HotFix 3.

Tokenization

Use the tokenization process to create match tokens for the input data based on the columns that you configure as index fields in the matching rules file. You can persist the tokenized data in HDFS or in a repository, and you can search the tokenized data for matching records.

For more information about the tokenization process, see the *Informatica MDM Big Data Relationship Management User Guide*.

MapReduce Jobs

This section describes new features and enhancements to the MapReduce jobs.

HDFS Tokenization Job

The HDFS tokenization job creates match tokens for the input data based on the columns that you configure as index fields in the matching rules file. You can also run the HDFS tokenization job in the incremental mode to update the tokenized data that a HDFS tokenization job creates.

For more information about the HDFS tokenization job, see the *Informatica MDM Big Data Relationship Management User Guide*.

Repository Tokenization Job

The repository tokenization job creates match tokens for the input data in HDFS and loads the tokenized data into the repository. The repository tokenization job uses the columns that you configure as index fields to

generate the match tokens. The repository tokenization job performs the tasks of a HDFS tokenization job and a load clustering job.

For more information about the repository tokenization job, see the *Informatica MDM Big Data Relationship Management User Guide*.

HDFS Batch Search Job

The HDFS batch search job identifies the matching records for the input data in the output files of a HDFS tokenization job based on the match tokens.

For more information about the HDFS batch search job, see the *Informatica MDM Big Data Relationship Management User Guide*.

Repository Batch Search Job

The repository batch search job identifies the matching records for the input data in the repository based on the match tokens.

For more information about the repository batch search job, see the *Informatica MDM Big Data Relationship Management User Guide*.

Repository Update Job

The repository update job updates the tokenized data in the repository with the input data and creates match tokens for the input data in the repository.

For more information about the repository update job, see the *Informatica MDM Big Data Relationship Management User Guide*.

Search

This section describes new features and enhancements to the search process.

RULESEARCH Operation

The RULESEARCH operation generates key ranges, compares the key ranges with the key ranges in the repository data, and gets the matching records based on the rules that you configure in the matching rules file.

You can use the RULESEARCH operation to search the repository data if you define a single index in the matching rules file and use simple matching to link the records. Use the REST API or the `run_client.sh` script in the command line to perform the RULESEARCH operation.

For more information about the RULESEARCH operation, see the *Informatica MDM Big Data Relationship Management User Guide*.

Sorting Search Results

When you perform the MULTISEARCH or RULESEARCH operation in the command line, you can sort the search results based on a column that you define in the `PZMAP` section of the configuration file. You can also sort the search results based on the score of the search results.

For more information about sorting the search results, see the *Informatica MDM Big Data Relationship Management User Guide*.

Changes (10.0.0 HotFix 3)

This section describes changes in version 10.0.0 HotFix 3.

Matching Rules File

This section describes changes to the matching rules file in version 10.0.0 HotFix 3.

Multiple Populations

Effective in version 10.0.0 HotFix 3, you can define matching rules for multiple populations in the matching rules file.

Previously, you could define matching rules only for a single population in the matching rules file.

For more information about configuring matching rules, see the *Informatica MDM Big Data Relationship Management User Guide*.

Index Fields

Effective in version 10.0.0 HotFix 3, you can define multiple columns as index fields for an index.

Previously, you could define only one column for an index. To use multiple columns as index fields, you had to define multiple indexes.

For more information about configuring indexes in the matching rules file, see the *Informatica MDM Big Data Relationship Management User Guide*.

CHAPTER 5

New Features, Changes, and Release Tasks (10.0.0 HotFix 2)

This chapter includes the following topics:

- [New Features \(10.0.0 HotFix 2\), 19](#)
- [Changes \(10.0.0 HotFix 2\), 20](#)

New Features (10.0.0 HotFix 2)

This section describes new features in version 10.0.0 HotFix 2.

Advanced Matching

Use advanced matching to index multiple columns and configure multiple matching rules for each index. You can define indexes and matching rules in the matching rules file. You can use the matching rules file when you run the MapReduce jobs.

For more information about advanced matching, see the *Informatica MDM Big Data Relationship Management User Guide*.

MapReduce Jobs

This section describes new features and enhancements to the MapReduce jobs.

Initial Clustering Job

An initial clustering job indexes and links the input data based on the rules in the matching rules file and the parameters in the configuration file. The initial clustering job reads the input data in HDFS and persists the indexed and linked data to HDFS. You can also run the initial clustering job in the incremental mode to incrementally update the indexed and linked data in HDFS.

For more information about the initial clustering job, see the *Informatica MDM Big Data Relationship Management User Guide*.

Load Clustering Job

A load clustering job loads the integrated and linked data from HDFS into a repository. Before you run the load clustering job, you must run the initial clustering job in the initial or incremental mode.

For more information about the load clustering job, see the *Informatica MDM Big Data Relationship Management User Guide*.

Delete Clustering Job

A delete clustering job matches the input data in HDFS with the indexed and linked data in HDFS that the initial clustering job creates. The job deletes the matching records from the indexed and linked data in HDFS.

For more information about the delete clustering job, see the *Informatica MDM Big Data Relationship Management User Guide*.

Hive Enabler

If you persist the indexed and linked data in HDFS, you can use the Hive enabler job to load the indexed and linked data into the Hive database.

For more information about the Hive enabler job, see the *Informatica MDM Big Data Relationship Management User Guide*.

Search

This section describes new features and enhancements to the search process.

GETMULTISEARCHLAYOUT Operation

The GETMULTISEARCHLAYOUT operation gets the search layout for the matching rules file. The layout contains the required fields that you can specify as input parameters in the JSON format to perform the MULTISEARCH operation. Use the REST API or the `run_client.sh` script in the command line to perform the GETMULTISEARCHLAYOUT operation.

For more information about the GETMULTISEARCHLAYOUT operation, see the *Informatica MDM Big Data Relationship Management User Guide*.

MULTISEARCH Operation

The MULTISEARCH operation generates keys and keys ranges, compares the key ranges with the key ranges in the repository data, and gets the matching records based on the searching and matching criteria that you specify in the matching rules file. Use the REST API or the `run_client.sh` script in the command line to perform the MULTISEARCH operation.

For more information about the MULTISEARCH operation, see the *Informatica MDM Big Data Relationship Management User Guide*.

Changes (10.0.0 HotFix 2)

This section describes changes in version 10.0.0 HotFix 2.

Simple Matching

Effective in version 10.0.0 HotFix 2, you can define multiple matching rules for an index in the matching rules file to perform simple matching. You can use the matching rules file when you run the MapReduce jobs.

Previously, you could define only a single matching rule for an index in the configuration file.

For more information about defining multiple matching rules, see the *Informatica MDM Big Data Relationship Management User Guide*.

Updating Indexed and Linked Data

Effective in version 10.0.0 HotFix 2, you can incrementally update the indexed and linked data in HDFS.

Previously, you could incrementally update the indexed and linked data in a repository but not in HDFS.