



Informatica® Data Integration - Free & PayGo

Microsoft CDM Folders V2 Connector

© Copyright Informatica LLC 2021, 2023

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, the Informatica logo, Informatica Cloud, and PowerCenter 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.

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.

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: 2023-04-04

Table of Contents

Preface	4
Informatica Resources.	4
Informatica Documentation.	4
Informatica Intelligent Cloud Services web site.	4
Informatica Intelligent Cloud Services Communities.	4
Informatica Intelligent Cloud Services Marketplace.	4
Simple Data Integration connector documentation.	5
Informatica Knowledge Base.	5
Informatica Intelligent Cloud Services Trust Center.	5
Informatica Global Customer Support.	5
 Chapter 1: Introduction to Microsoft CDM Folders V2 Connector.....	6
Microsoft CDM Folders V2 Connector assets.	6
Administration of Microsoft CDM Folders V2 Connector.	7
 Chapter 2: Connections for Microsoft CDM Folders V2.....	8
Microsoft CDM Folders V2 connection properties.	8
 Chapter 3: Mappings for Microsoft CDM Folders V2.....	10
Microsoft CDM Folders V2 sources in mappings.	10
Overriding the source advanced properties.	11
 Chapter 4: Data type reference.....	12
Microsoft CDM Folders V2 and transformation data types.	12
 Chapter 5: Troubleshooting.....	14
SDK exception when entity name has Unicode (UTF-8) characters.	14
 Index.....	15

Preface

Use *Microsoft CDM Folders V2 Connector* to learn how to read from CDM folders on Microsoft Azure Data Lake Storage Gen2. Learn to create a connection, develop and run mappings, mapping tasks, and data transfer tasks in Data Integration.

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 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 Intelligent Cloud Services web site

You can access the Informatica Intelligent Cloud Services web site at <http://www.informatica.com/cloud>. This site contains information about Informatica Cloud integration services.

Informatica Intelligent Cloud Services Communities

Use the Informatica Intelligent Cloud Services Community to discuss and resolve technical issues. You can also find technical tips, documentation updates, and answers to frequently asked questions.

Access the Informatica Intelligent Cloud Services Community at:

<https://network.informatica.com/community/informatica-network/products/cloud-integration>

Developers can learn more and share tips at the Cloud Developer community:

<https://network.informatica.com/community/informatica-network/products/cloud-integration/cloud-developers>

Informatica Intelligent Cloud Services Marketplace

Visit the Informatica Marketplace to try and buy Data Integration Connectors, templates, and mapplets:

<https://marketplace.informatica.com/>

Simple Data Integration connector documentation

You can access documentation for Simple Data Integration Connectors at the Documentation Portal. To explore the Documentation Portal, visit <https://docs.informatica.com>.

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 Intelligent Cloud Services Trust Center

The Informatica Intelligent Cloud Services Trust Center provides information about Informatica security policies and real-time system availability.

You can access the trust center at <https://www.informatica.com/trust-center.html>.

Subscribe to the Informatica Intelligent Cloud Services Trust Center to receive upgrade, maintenance, and incident notifications. The [Informatica Intelligent Cloud Services Status](#) page displays the production status of all the Informatica cloud products. All maintenance updates are posted to this page, and during an outage, it will have the most current information. To ensure you are notified of updates and outages, you can subscribe to receive updates for a single component or all Informatica Intelligent Cloud Services components. Subscribing to all components is the best way to be certain you never miss an update.

To subscribe, go to <https://status.informatica.com/> and click **SUBSCRIBE TO UPDATES**. You can then choose to receive notifications sent as emails, SMS text messages, webhooks, RSS feeds, or any combination of the four.

Informatica Global Customer Support

You can contact a Customer Support Center by telephone or online.

For online support, click **Submit Support Request** in Informatica Intelligent Cloud Services. You can also use Online Support to log a case. Online Support requires a login. You can request a login at <https://network.informatica.com/welcome>.

The telephone numbers for Informatica Global Customer Support are available from the Informatica web site at <https://www.informatica.com/services-and-training/support-services/contact-us.html>.

CHAPTER 1

Introduction to Microsoft CDM Folders V2 Connector

You can use Microsoft CDM Folders V2 Connector to read data from CDM folders on Microsoft Azure Data Lake Storage Gen2.

You can create a common data model folder within a file system in the ADLS Gen2 storage. A common data model folder contains data in the `.csv` file format and the associated metadata in the `model.json` and `manifest.json` file formats. You can read from CDM folders that contain metadata in the `model.json` or `manifest.json` format.

You can use Microsoft CDM Folders V2 Connector to read from CDM format files.

You can use Microsoft CDM Folders V2 objects as sources in mappings and mapping tasks.

For more information about the common data model, see the Microsoft documentation at the following website:

<https://docs.microsoft.com/en-us/common-data-model/>

Microsoft CDM Folders V2 Connector assets

Create assets in Data Integration to integrate data using Microsoft CDM Folders V2 Connector.

When you use Microsoft CDM Folders V2 Connector, you can include the following Data Integration assets:

- Data transfer task
- Mapping
- Mapping task

For more information about configuring assets and transformations, see *Mappings*, *Transformations*, and *Tasks* in the Data Integration documentation.

Administration of Microsoft CDM Folders V2 Connector

Before you use Microsoft CDM Folders V2 objects in tasks, an administrator must perform the following tasks:

- Create an ADLS Gen2 account and provide **Contributor** or **Reader** role to users.
 - Using the contributor role, you can have full access to manage all resources in the storage account, but you cannot assign roles.
 - Using the reader role, you can view all resources in the storage account, but you cannot make any changes.

To add or remove role assignments, you must have write and delete permissions, such as an Owner role.

- Create an Azure Active Directory application to authenticate users to access the ADLS Gen2 account. Provide **Storage Blob Data Contributor** or **Storage Blob Data Reader** role to the application.
 - Using the Storage Blob Data Contributor role, you can read, write, and delete Azure Storage containers and blobs in the storage account.
 - Using the Storage Blob Data Reader role, you can only read and list Azure Storage containers and blobs in the storage account.

To write to or delete Azure Storage containers and blobs, you must have the Contributor role either at the storage account level or the container level.

- Create a file system for ADLS Gen2.
- Create an Azure AD web application for service-to-service authentication with ADLS Gen2.

Note: Ensure that you have superuser privileges to access the folders or files created in the application using the connector.

CHAPTER 2

Connections for Microsoft CDM Folders V2

Create a Microsoft CDM Folders V2 connection to securely read data from CDM folders on Microsoft Azure Data Lake Storage Gen2. You can use a Microsoft CDM Folders V2 connection to specify sources in mappings and mapping tasks.

Microsoft CDM Folders V2 connection properties

When you set up a Microsoft CDM Folders V2 connection, configure the connection properties.

The following table describes the Microsoft CDM Folders V2 connection properties:

Property	Description
Connection Name	Name of the connection. Each connection name must be unique within the organization. Connection names can contain alphanumeric characters, spaces, and the following special characters: _ . + -, Maximum length is 255 characters.
Description	Description of the connection. Maximum length is 4000 characters.
Type	The Microsoft CDM Folders V2 connection type.
Runtime Environment	The name of the runtime environment where you want to run the tasks. You can specify a Secure Agent or a Hosted Agent.
ADLSGen2 Storage Account Name	Name of the ADLS Gen2 storage account.
Azure AD App Client ID	The client ID of the Azure Active Directory account to authenticate user access to the storage account. You can get the application ID from the Microsoft Azure Active Directory administrator.
Azure AD App Client Secret	The client secret key of the Azure Active Directory application to authenticate access to the storage account. You can get the key value from the Microsoft Azure Active Directory administrator.

Property	Description
Azure Tenant ID	The tenant ID of the Azure Active Directory account to authenticate user access to the storage account. You can get the directory ID from the Microsoft Azure Active Directory administrator.
ADLSGen2 File System Name	The name of the file system that you create in the Azure Storage Explorer application. A file system can contain more than one common data model folders.
CDM Folder Path	The path of the common data model folder that you create within the file system. You can use the following values for CDM folder path: <ul style="list-style-type: none"> - / - /folder1 - /folder1/folder2 The recommended CDM folder path is /folder1. Default is empty.
Adls Gen2 End-point	The ADLS Gen2 endpoint core.windows.net.

CHAPTER 3

Mappings for Microsoft CDM Folders V2

When you configure a mapping, you describe the flow of data from the source to the target.

A mapping defines reusable data flow logic that you can use in mapping tasks.

When you create a mapping, you define the Source transformation to represent a Microsoft CDM Folders V2 object. Use the Mapping Designer in Data Integration to add the Source or Target transformations in the mapping canvas and configure the Microsoft CDM Folders V2 source properties.

You can use Monitor to monitor the jobs.

Microsoft CDM Folders V2 sources in mappings

In a mapping, you can configure a source transformation to represent a single Microsoft CDM Folders V2 object.

The following table describes the Microsoft CDM Folders V2 source properties that you can configure in a source transformation:

Property	Description
Connection	Name of the source connection.
Source Type	Type of the CDM folder source objects. Select Single Object or Parameter .
Object	Name of the source object. Select an existing object.

The following table describes the Microsoft CDM Folders V2 source advanced properties that you can configure in a Source transformation:

Property	Description
Concurrent Threads	Number of concurrent connections to read data from CDM Folders V2. When you read from a large file or object, consider spawning multiple threads to process data. Default is 10.
Block Size	Divides a large file or object into smaller parts each of specified block size. When you read from a large file, consider dividing the file into smaller parts and configure concurrent connections to spawn the required number of threads to process data in parallel. Default is 8 MB.
File System Name Override	Overrides the default file system name at run time with the file system name that you specify.
CDM Folder Override	Overrides the default CDM folders. You can use the following paths for CDM folder override: <ul style="list-style-type: none"> - / - /folder1 - /folder1/folder2 - /folder1/filename, for example, /folder1/default.manifest.cdm.json or /folder1/model.json
Entity Name Override	Overrides the default entity name in the <code>manifest.json</code> or <code>model.json</code> file.
Tracing Level	Sets the amount of detail that appears in the log file. You can choose terse, normal, verbose initialization or verbose data. Default is normal.

Note: You cannot use special characters in an entity name.

Overriding the source advanced properties

You can choose to override the default file system name, CDM folder, or entity name in the source advanced properties.

During mapping development you might use connection attributes, and during run time you can change the attributes without changing the connection attributes.

The CDM folder path and entity name must be valid when you override the default file system name, CDM folder, or entity name. Otherwise, the mapping fails with an error.

You can override the file system name when you select the file system name override option. For example, if you configured connection attributes for file system name `cdmsd`, CDM folder path `demo/model`, and the entity name `ff_cdm`, you can change the attributes during run time. You can select the file system name override and specify the system name `cdmsd_prod`. However, `cdmsd_prod` must have the same CDM folder path and the entity name that you specified in Microsoft CDM Folders V2 connection.

You can also override the filename when you select the CDM folder override option. For example, if you specified the CDM folder path `/folder1/default.manifest.cdm.json` in the Microsoft CDM Folders V2 connection to read from `manifest.cdm.json` file, you can choose to read from `model.json` file during run time.

CHAPTER 4

Data type reference

Data Integration uses the following data types in Microsoft CDM Folders V2 mappings and mapping tasks:

- Microsoft CDM Folders V2 native data types appear in the Source transformation when you choose to edit metadata for the fields.
- Transformation data types. Set of data types that appear in the transformations. These are internal data types based on ANSI SQL-92 generic data types, which the Secure Agent uses to move data across platforms. They appear in all transformations in a mapping.

When the Secure Agent reads source data, it converts the native data types to the comparable transformation data types before transforming the data.

Microsoft CDM Folders V2 and transformation data types

The following table lists the supporting Microsoft CDM Folders V2 data types and the corresponding transformation data types:

Microsoft CDM Folders V2 Data Type	Transformation Data Type	Description
Boolean	String	TRUE (1) or FALSE (0)
Date	Date	Date values. Microsoft CDM Folders V2 Connector uses the following format: yyyy-MM-dd Maximum precision 29, scale 9. Default precision 29.
DateTime	Date/Time	Date values. Microsoft CDM Folders V2 Connector uses the following format: yyyy-MM-dd'T'HH:mm:ssZ Maximum precision 29, scale 9. Default precision 29.
Decimal	Decimal	Maximum precision 31, scale 0. Default precision 31.

Microsoft CDM Folders V2 Data Type	Transformation Data Type	Description
Double	Double	Maximum precision 15, scale 0. Default precision 15.
Guid	String	Maximum precision 255, scale 0. Default precision 255.
int64	BigInteger	Maximum precision 20, scale 0. Default precision 20.
int32	Integer	Maximum precision 11, scale 0. Default precision 11.
int16	Integer	Maximum precision 10, scale 0. Default precision 10.
Json	String	Maximum precision 255, scale 0. Default precision 255.
String	String	Maximum precision 255, scale 0. Default precision 255.

CHAPTER 5

Troubleshooting

Use the following section to troubleshoot errors in Microsoft CDM Folders V2 Connector.

SDK exception when entity name has Unicode (UTF-8) characters

When you run a mapping to read from a CDM Folders V2 source object where the entity name contains UTF-8 characters, the task fails with the following error:

```
"com.informatica.adapter.sdkadapter.exceptions.AdapterSDKException: [SDK_APP_COM_20000] error [No Data Partitions found indefault.manifest.cdm.json]"
```

To resolve this issue, perform the following tasks to set `-Dfile.encoding=UTF8` in the JVM option in the Secure Agent:

1. Select **Administer > Runtime Environments**.
2. On the **Runtime Environments** page, select the Secure Agent for which you want to increase memory from the list of available Secure Agents.
3. In the upper-right corner, click **Edit**.
4. In the **System Configuration Details** section, select the **Type** as **DTM** for the Data Integration Service.
5. Edit the **JVMOption** as **-Dfile.encoding=UTF8**.
6. Restart the Secure Agent manually.

INDEX

C

Cloud Application Integration community
URL [4](#)
Cloud Developer community
URL [4](#)
connections
Microsoft CDM Folders V2 [8](#)

D

Data Integration community
URL [4](#)
data type reference
overview [12](#)

I

Informatica Global Customer Support
contact information [5](#)
Informatica Intelligent Cloud Services
web site [4](#)

M

maintenance outages [5](#)
mappings
Microsoft CDM Folders V2 properties [10](#)
Microsoft CDM Folders V2
connection properties [8](#)
Source transformation [10](#)
Sources in mappings [10](#)
supported object types [6](#)

Microsoft CDM Folders V2 (*continued*)
supported task types [6](#)
Microsoft CDM Folders V2 Connection
overview [8](#)
Microsoft CDM Folders V2 Connector
administration [7](#)
data types [12](#)
Microsoft Common Data Model V2 Connector
overview [6](#)

S

Source transformation
Microsoft CDM Folders V2 properties [10](#)
Sources
Microsoft CDM Folders V2 in mappings [10](#)
status
Informatica Intelligent Cloud Services [5](#)
system status [5](#)

T

trust site
description [5](#)

U

upgrade notifications [5](#)

W

web site [4](#)