Informatica®
10.2.1

New Features Guide

Informatica New Features Guide
10.2.1
May 2018

Publication Date: 2019-04-09

Table of Contents

Preface

The *Informatica New Features and Enhancements Guide* is written for all Informatica software users. This guide lists the new features and enhancements in Informatica products.

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.

CHAPTER 1

New Features (10.2.1)

This chapter includes the following topics:

Application Services

This section describes new application service features in version 10.2.1.

Content Management Service

Effective in version 10.2.1, you can optionally specify a schema to identify reference tables in the reference data database as a property on the Content Management Service.

To specify the schema, use the Reference Data Location Schema property on the Content Management Service in Informatica Administrator. Or, run the infacmd cms updateServiceOptions command with the DataServiceOptions.RefDataLocationSchema option.

If you do not specify a schema for reference tables on the Content Management Service, the service uses the schema that the database connection specifies. If you do not explicitly set a schema on the database connection, the Content Management Service uses the default database schema.

Note: Establish the database and the schema that the Content Management Service will use for reference data before you create a managed reference table.

For more information, see the "Content Management Service" chapter in the *Informatica 10.2.1 Application Service Guide* and the "infacmd cms Command Reference" chapter in the *Informatica 10.2.1 Command Reference*.

Data Integration Service

Effective in version 10.2.1, the Data Integration Service properties include a new execution option.

JDK Home Directory

The JDK installation directory on the machine that runs the Data Integration Service. Required to run Sqoop mappings or mass ingestion specifications that use a Sqoop connection on the Spark engine, or to process a Java transformation on the Spark engine. Default is blank.

Mass Ingestion Service

Effective in version 10.2.1, you can create a Mass Ingestion Service. The Mass Ingestion Service is an application service in the Informatica domain that manages mass ingestion specifications. You configure the mass ingestion specifications in the Mass Ingestion tool to ingest large amounts of data from a relational source to a Hive or HDFS target.

To manage mass ingestion specifications, the Mass Ingestion Service performs the following tasks:

- Manages and validates a mass ingestion specification.
- Schedules a mass ingestion job to run on a Data integration Service.
- Monitors the results and statistics of a mass ingestion job.
- Restarts a mass ingestion job.

For more information on the Mass Ingestion Service, see the "Mass Ingestion Service" chapter in the *Informatica 10.2.1 Application Service Guide*.

Metadata Access Service

Effective in version 10.2.1, you can create a Metadata Access Service. The Metadata Access Service is an application service that allows the Developer tool to access Hadoop connection information to import and preview metadata. When you import an object from a Hadoop cluster, the following adapters use Metadata Access Service to extract the object metadata at design time:

- PowerExchange for HBase
- PowerExchange for HDFS
- PowerExchange for Hive
- PowerExchange for MapR-DB

For more information, see the "Metadata Access Service" chapter in the *Informatica 10.2.1 Application Service Guide*.

Model Repository Service

Azure SQL Database as Model Repository

Effective in version 10.2.1, you can use the Azure SQL database as the Model repository.

For more information, see the "Model Repository Service" chapter in the *Informatica 10.2.1 Application Service Guide*.

Git Version Control System

Effective in version 10.2.1, you can integrate the Model repository with the Git version control system. Git is a distributed version control system. When you check out and check in an object, a copy of the version is saved to the local repository and to the Git server. If the Git server goes down, the local repository retains all the versions of the object. To use the Git version control system, enter the URL of the global repository for Git in the **URL** field, login credentials for the global repository in the **Username** and **Password** fields, and the path of the local repository for the Model Repository Service in the **VCS Local Repository Path** field.

For more information, see the "Model Repository Service" chapter in the *Informatica 10.2.1 Application Service Guide*.

Big Data Management

This section describes new Big Data Management features in version 10.2.1.

Blaze Engine Resource Conservation

Effective in version 10.2.1, you can preserve the resources that the Blaze engine infrastructure uses.

Set the infagrid.blaze.service.idle.timeout property to specify the number of minutes that the Blaze engine remains idle before releasing resources. Set the infagrid.orchestrator.svc.sunset.time property to specify the maximum number of hours for the Blaze orchestrator service. You can use the infacmd isp createConnection command, or set the property in the Blaze Advanced properties in the Hadoop connection in the Administrator tool or the Developer tool.

For more information about these properties, see the *Informatica Big Data Management 10.2.1 Administrator Guide*.

Cluster Workflows

You can use new workflow tasks to create a cluster workflow.

A cluster workflow creates a cluster on a cloud platform and runs Mapping and other workflow tasks on the cluster. You can choose to terminate and delete the cluster when workflow tasks are complete to save cluster resources.

Two new workflow tasks enable you to create and delete a Hadoop cluster as part of a cluster workflow:

Create Cluster Task

The Create Cluster task enables you to create, configure and start a Hadoop cluster on the following cloud platforms:

- Amazon Web Services (AWS). You can create an Amazon EMR cluster.
- Microsoft Azure. You can create an HDInsight cluster.

Delete Cluster Task

The optional Delete Cluster task enables you to delete a cluster after Mapping tasks and any other tasks in the workflow are complete. You might want to do this to save costs.

Previously, you could use Command tasks in a workflow to create clusters on a cloud platform. For more information about cluster workflows and workflow tasks, see the *Informatica 10.2.1 Developer Workflow Guide*.

Note: In 10.2.1, the Command task method of creating and deleting clusters now supports Cloudera Altus clusters on AWS. For more information, see the article "How to Create Cloudera Altus Clusters with a Cluster Workflow on Big Data Management" on the Informatica Network.

Mapping Task

Mapping task advanced properties include a new ClusterIdentifier property. The ClusterIdentifier identifies the cluster to use to run the Mapping task.

For more information about cluster workflows, see the *Informatica 10.2.1 Developer Workflow Guide*.

Cloud Provisioning Configuration

A cloud provisioning configuration is an object that contains information about connecting to a Hadoop cluster.

The cloud provisioning configuration includes information about how to integrate the domain with Hadoop account authentication and storage resources. A cluster workflow uses the information in the cloud provisioning configuration to connect to and create a cluster on a cloud platform such as Amazon Web Services or Microsoft Azure.

For more information about cloud provisioning, see the "Cloud Provisioning Configuration" chapter in the *Informatica Big Data Management 10.2.1 Administrator Guide*.

High Availability

Effective in version 10.2.1, you can enable high availability for the following services and security systems in the Hadoop environment on Cloudera CDH, Hortonworks HDP, and MapR Hadoop distributions:

- Apache Ranger
- Apache Ranger KMS
- Apache Sentry
- Cloudera Navigator Encrypt
- HBase
- Hive Metastore
- HiveServer2
- Name node
- Resource Manager

Hive Functionality in the Hadoop Environment

This section describes new features for Hive functionality in the Hadoop environment in version 10.2.1.

Hive Table Truncation

Effective in version 10.2.1, you can truncate external partitioned Hive tables on all run-time engines.

You can truncate tables in the following Hive storage formats:

- Avro
- ORC
- Parquet

- RCFile
- Sequence
- Text

You can truncate tables in the following Hive external table formats:

- Hive on HDFS
- Hive on Amazon S3
- Hive on Azure Blob
- Hive on WASB
- Hive on ADLS

For more information on truncating Hive targets, see the "Mapping Targets in the Hadoop Environment" chapter in the *Informatica Big Data Management 10.2.1 User Guide*.

Pre- and Post-Mapping SQL Commands

Effective in version 10.2.1, you can configure PreSQL and PostSQL commands against Hive sources and targets in mappings that run on the Spark engine.

For more information, see the *Informatica Big Data Management 10.2.1 User Guide*.

Importing from PowerCenter

This section describes new import from PowerCenter features in version 10.2.1.

Import Session Properties from PowerCenter

Effective in version 10.2.1, you can import session properties, such as SQL-based overrides in relational sources and targets and overrides for the Lookup transformation from the PowerCenter repository to the Model repository.

For more information about the import from PowerCenter functionality, see the "Import from PowerCenter" chapter in the *Informatica 10.2.1 Developer Mapping Guide*.

SQL Parameters

Effective in version 10.2.1, you can specify an SQL parameter type to import all SQL-based overrides into the Model repository. The remaining session override properties map to String or a corresponding parameter type.

For more information, see the "Import from PowerCenter" chapter in the *Informatica 10.2.1 Developer Mapping Guide*.

Import a Command Task from PowerCenter

Effective in version 10.2.1, you can import a Command task from PowerCenter into the Model repository.

For more information, see the "Workflows" chapter in the *Informatica 10.2.1 Developer Workflow Guide*.

Intelligent Structure Model

Effective in version 10.2.1, you can use the intelligent structure model in Big Data Management.

Spark Engine Support for Data Objects with Intelligent Structure Model

You can incorporate an intelligent structure model in an Amazon S3, Microsoft Azure Blob, or complex file data object. When you add the data object to a mapping that runs on the Spark engine, you can process any input type that the model can parse.

The data object can accept input and parse PDF forms, JSON, Microsoft Excel, Microsoft Word tables, CSV, text, or XML input files, based on the file which you used to create the model.

Intelligent structure model in the complex file, Amazon S3, and Microsoft Azure Blob data objects is available for technical preview. Technical preview functionality is supported but is unwarranted and is not production-ready. Informatica recommends that you use these features in non-production environments only.

For more information, see the *Informatica Big Data Management 10.2.1 User Guide*.

Mass Ingestion

Effective in version 10.2.1, you can perform mass ingestion jobs to ingest or replicate large amounts of data for use or storage in a database or a repository. To perform mass ingestion jobs, you use the Mass Ingestion tool to create a mass ingestion specification. You configure the mass ingestion specification to ingest data from a relational database to a Hive or HDFS target. You can also specify parameters to cleanse the data that you ingest.

A mass ingestion specification replaces the need to manually create and run mappings. You can create one mass ingestion specification that ingests all of the data at once.

For more information on mass ingestion, see the *Informatica Big Data Management 10.2.1 Mass Ingestion Guide*.

Monitoring

This section describes the new features related to monitoring in Big Data Management in version 10.2.1.

Hadoop Cluster Monitoring

Effective in version 10.2.1, you can configure the amount of information that appears in the application logs that you monitor for a Hadoop cluster.

The amount of information in the application logs depends on the tracing level that you configure for a mapping in the Developer tool. The following table describes the amount of information that appears in the application logs for each tracing level:

Tracing Level	Messages
None	The log displays FATAL messages. FATAL messages include non-recoverable system failures that cause the service to shut down or become unavailable.
Terse	The log displays FATAL and ERROR code messages. ERROR messages include connection failures, failures to save or retrieve metadata, service errors.
Normal	The log displays FATAL, ERROR, and WARNING messages. WARNING errors include recoverable system failures or warnings.

Tracing Level	Messages
Verbose initialization	The log displays FATAL, ERROR, WARNING, and INFO messages. INFO messages include system and service change messages.
Verbose data	The log displays FATAL, ERROR, WARNING, INFO, and DEBUG messages. DEBUG messages are user request logs.

For more information, see the "Monitoring Mappings in the Hadoop Environment" chapter in the *Informatica Big Data Management 10.2.1 User Guide*.

Spark Monitoring

Effective in version 10.2.1, the Spark executor listens on a port for Spark events as part of Spark monitoring support and it is not required to configure the SparkMonitoringPort.

The Data Integration Service has a range of available ports, and the Spark executor selects a port from the available range. During failure, the port connection remains available and you do not need to restart the Data Integration Service before running the mapping.

The custom property for the monitoring port is retained. If you configure the property, the Data Integration Service uses the specified port to listen to Spark events.

Previously, the Data Integration Service custom property, the Spark monitoring port could configure the Spark listening port. If you did not configure the property, Spark Monitoring was disabled by default.

Tez Monitoring

Effective in 10.2.1, you can view Tez engine monitoring support related properties. You can use the Hive engine to run the mapping on MapReduce or Tez. The Tez engine can process jobs on Hortonworks HDP, Azure HDInsight, and Amazon Elastic MapReduce. To run a Spark mapping on Tez, you can use any of the supported clusters for Tez.

In the Administrator tool, you can also review the Hive query properties for Tez when you monitor the Hive engine. In the Hive session log and in Tez, you can view information related to Tez statistics, such as DAG tracking URL, total vertex count, and DAG progress.

You can monitor any Hive query on the Tez engine. When you enable logging for verbose data or verbose initialization, you can view the Tez engine information in the Administrator tool or in the session log. You can also monitor the status of the mapping on the Tez engine on the Monitoring tab in the Administrator tool.

For more information about Tez monitoring, see the *Informatica Big Data Management 10.2.1 User Guide* and the *Informatica Big Data Management 10.2.1 Hadoop Integration Guide*.

Processing Hierarchical Data on the Spark Engine

Effective in version 10.2.1, the Spark engine includes the following additional functionality to process hierarchical data:

Map data type

You can use map data type to generate and process map data in complex files.

Complex files on Amazon S3

You can use complex data types to read and write hierarchical data in Avro and Parquet files on Amazon S3. You project columns as complex data type in the data object read and write operations.

For more information, see the "Processing Hierarchical Data on the Spark Engine" chapter in the *Informatica Big Data Management 10.2.1 User Guide*.

Rule Specification Support on the Spark Engine

Effective in version 10.2.1, you can run a mapping that contains a rule specification on the Spark engine in addition to the Blaze and Hive engines.

You can also run a mapping that contains a mapplet that you generate from a rule specification on the Spark engine in addition to the Blaze and Hive engines.

For more information about rule specifications, see the *Informatica 10.2.1 Rule Specification Guide*.

Security

This section describes the new features related to security in Big Data Management in version 10.2.1.

Cloudera Navigator Encrypt

Effective in version 10.2.1, you can use Cloudera Navigator Encrypt to secure the data and implement transparent encryption of data at rest.

EMR File System Authorization

Effective in version 10.2.1, you can use EMR File System (EMRFS) authorization to access data in Amazon S3 on Spark engine.

IAM Roles

Effective in version 10.2.1, you can use IAM roles for EMR File System to read and write data from the cluster to Amazon S3 in Amazon EMR cluster version 5.10.

Kerberos Authentication

Effective in version 10.2.1, you can enable Kerberos authentication for the following clusters:

- Amazon EMR
- Azure HDInsight with WASB as storage

LDAP Authentication

Effective in version 10.2.1, you can configure Lightweight Directory Access Protocol (LDAP) authentication for Amazon EMR cluster version 5.10.

Sqoop

Effective in version 10.2.1, you can use the following new Sqoop features:

Support for MapR Connector for Teradata

You can use MapR Connector for Teradata to read data from or write data to Teradata on the Spark engine. MapR Connector for Teradata is a Teradata Connector for Hadoop (TDCH) specialized connector for Sqoop. When you run Sqoop mappings on the Spark engine, the Data Integration Service invokes the connector by default.

For more information, see the *Informatica Big Data Management 10.2.1 User Guide*.

Spark engine optimization for Sqoop pass-through mappings

When you run a Sqoop pass-through mapping on the Spark engine, the Data Integration Service optimizes mapping performance in the following scenarios:

- You read data from a Sqoop source and write data to a Hive target that uses the Text format.
- You read data from a Sqoop source and write data to an HDFS target that uses the Flat, Avro, or Parquet format.

For more information, see the *Informatica Big Data Management 10.2.1 User Guide*.

Spark engine support for high availability and security features

Sqoop honors all the high availability and security features such as Kerberos keytab login and KMS encryption that the Spark engine supports.

For more information, see the "Data Integration Service" chapter in the *Informatica 10.2.1 Application Services Guide* and "infacmd dis Command Reference" chapter in the *Informatica 10.2.1 Command Reference Guide*.

Spark engine support for Teradata data objects

If you use a Teradata data object and you run a mapping on the Spark engine and on a Hortonworks or Cloudera cluster, the Data Integration Service runs the mapping through Sqoop.

If you use a Hortonworks cluster, the Data Integration Service invokes Hortonworks Connector for Teradata at run time. If you use a Cloudera cluster, the Data Integration Service invokes Cloudera Connector Powered by Teradata at run time.

For more information, see the *Informatica PowerExchange for Teradata Parallel Transporter API 10.2.1 User Guide*.

Transformation Support in the Hadoop Environment

This section describes new transformation features in the Hadoop environment in version 10.2.1.

Transformation Support on the Spark Engine

This section describes new transformation features on the Spark engine in version 10.2.1.

Transformation Support

Effective in version 10.2.1, the following transformations are supported on the Spark engine:

- Case Converter
- Classifier
- Comparison
- Key Generator
- Labeler
- Merge
- Parser
- Python
- Standardizer
- Weighted Average

Effective in version 10.2.1, the following transformations are supported with restrictions on the Spark engine:

- Address Validator
- Consolidation
- Decision
- Match
- Sequence Generator

Effective in version 10.2.1, the following transformation has additional support on the Spark engine:

- Java. Supports complex data types such as array, map, and struct to process hierarchical data.

For more information on transformation support, see the "Mapping Transformations in the Hadoop Environment" chapter in the *Informatica Big Data Management 10.2.1 User Guide*.

For more information about transformation operations, see the *Informatica 10.2.1 Developer Transformation Guide*.

Python Transformation

Effective in version 10.2.1, you can create a Python transformation in the Developer tool. Use the Python transformation to execute Python code in a mapping that runs on the Spark engine.

You can use a Python transformation to implement a machine model on the data that you pass through the transformation. For example, use the Python transformation to write Python code that loads a pre-trained model. You can use the pre-trained model to classify input data or create predictions.

Note: The Python transformation is available for technical preview. Technical preview functionality is supported but is not production-ready. Informatica recommends that you use in non-production environments only.

For more information, see the "Python Transformation" chapter in the *Informatica 10.2.1 Developer Transformation Guide*.

Update Strategy Transformation

Effective in version 10.2.1, you can use Hive MERGE statements for mappings that run on the Spark engine to perform update strategy tasks. Using MERGE in queries is usually more efficient and helps increase performance.

Hive MERGE statements are supported for the following Hadoop distributions:

- Amazon EMR 5.10
- Azure HDInsight 3.6
- Hortonworks HDP 2.6

To use Hive MERGE, select the option in the advanced properties of the Update Strategy transformation.

Previously, the Data Integration Service used INSERT, UPDATE and DELETE statements to perform this task using any run-time engine. The Update Strategy transformation still uses these statements in the following scenarios:

- You do not select the Hive MERGE option.
- Mappings run on the Hive or Blaze engine.
- If the Hadoop distribution does not support Hive MERGE.

For more information about using a MERGE statement in Update Strategy transformations, see the chapter on Update Strategy transformation in the *Informatica Big Data Management 10.2.1 User Guide*.

Transformation Support on the Blaze Engine

This section describes new transformation features on the Blaze engine in version 10.2.1.

Aggregator Transformation

Effective in version 10.2.1, the data cache for the Aggregator transformation uses variable length to store binary and string data types on the Blaze engine. Variable length reduces the amount of data that the data cache stores when the Aggregator transformation runs.

When data that passes through the Aggregator transformation is stored in the data cache using variable length, the Aggregator transformation is optimized to use sorted input and a Sorter transformation is inserted before the Aggregator transformation in the run-time mapping.

For more information, see the "Mapping Transformations in the Hadoop Environment" chapter in the *Informatica Big Data Management 10.2.1 User Guide*.

Match Transformation

Effective in version 10.2.1, you can run a mapping that contains a Match transformation that you configure for identity analysis on the Blaze engine.

Configure the Match transformation to write the identity index data to cache files. The mapping fails validation if you configure the Match transformation to write the index data to database tables.

For more information on transformation support, see the "Mapping Transformations in the Hadoop Environment" chapter in the *Informatica Big Data Management 10.2.1 User Guide*.

Rank Transformation

Effective in version 10.2.1, the data cache for the Rank transformation uses variable length to store binary and string data types on the Blaze engine. Variable length reduces the amount of data that the data cache stores when the Rank transformation runs.

When data that passes through the Rank transformation is stored in the data cache using variable length, the Rank transformation is optimized to use sorted input and a Sorter transformation is inserted before the Rank transformation in the run-time mapping.

For more information, see the "Mapping Transformations in the Hadoop Environment" chapter in the *Informatica Big Data Management 10.2.1 User Guide*.

For more information about transformation operations, see the *Informatica 10.2.1 Developer Transformation Guide*.

Big Data Streaming

This section describes new Big Data Streaming features in version 10.2.1.

Sources and Targets

Effective in version 10.2.1, you can read from or write to the following sources and targets in streaming mappings:

- Azure Event Hubs. Create an Azure EventHub data object to read from or write to Event Hub events. You can use an Azure EventHub connection to access Microsoft Azure Event Hubs as source or target. You can create and manage an Azure Eventhub connection in the Developer tool or through infacmd.

- Microsoft Azure Data Lake Store. Create an Azure Data Lake store data object to write to Azure Data Lake Store. You can use an Azure Data Lake Store connection to access Microsoft Azure Data Lake Store tables as targets. You can create and manage a Microsoft Azure Data Lake Store connection in the Developer tool.
- JDBC-compliant database. Create a relational data object with a JDBC connection.

For more information, see the "Sources in a Streaming Mapping" and "Targets in a Streaming Mapping" chapters in the *Informatica Big Data Streaming 10.2.1 User Guide*.

Stateful Computing in Streaming Mappings

Effective in 10.2.1, you can use window functions in an Expression transformation to perform stateful calculations in streaming mappings.

For more information, see the "Streaming Mappings" chapter in the *Informatica Big Data Streaming 10.2.1 User Guide*.

Transformation Support

Effective in version 10.2.1, you can use the following transformations in streaming mappings:

- Data Masking
- Normalizer
- Python

You can perform an uncached lookup on HBase data in streaming mappings with a Lookup transformation.

For more information, see the "Streaming Mappings" chapter in the *Informatica Big Data Streaming 10.2.1 User Guide*.

Truncate Partitioned Hive Target Tables

Effective in version 10.2.1, you can truncate an external or managed Hive table with or without partitions.

For more information about truncating Hive targets, see the "Targets in a Streaming Mapping" chapter in the *Informatica Big Data Streaming 10.2.1 User Guide*.

Command Line Programs

This section describes new commands in version 10.2.1.

infacmd autotune Commands

autotune is a new infacmd plugin that tunes services and connections in the Informatica domain.

The following table describes new infacmd autotune commands:

Command	Description
Autotune	Configures services and connections in the Informatica domain with recommended settings based on the size description.

For more information, see the "infacmd autotune Command Reference" chapter in the *Informatica 10.2.1 Command Reference*.

infacmd ccps Commands

ccps is a new infacmd plugin that performs operations on cloud platform clusters.

The following table describes new infacmd ccps commands:

Command	Description
deleteClusters	Deletes clusters on the cloud platform that a cluster workflow created.
listClusters	Lists clusters on the cloud platform that a cluster workflow created.
updateADLSCertifcate	Updates the Azure Data Lake Service Principal certificate.

For more information, see the "infacmd ccps Command Reference" chapter in the *Informatica 10.2.1 Command Reference*.

infacmd cluster Commands

The following table describes new infacmd cluster commands:

Command	Description
updateConfiguration	Updates the Hadoop distribution version of a cluster configuration. Use the -dv option to change the distribution version of the Hadoop distribution of a cluster configuration.

The following table describes changes to infacmd cluster commands:

Command	Change Description
listConfigurationProperties	Effective in 10.2.1, you can specify the general configuration set when you use the -cs option to return the property values in the general configuration set. Previously, the -cs option accepted only .xml file names.
createConfiguration	Effective in 10.2.1, you can optionally use the -dv option to specify a Hadoop distribution version when you create a cluster configuration. If you do not specify a version, the command creates a cluster configuration with the default version for the specified Hadoop distribution. Previously, the createConfiguration command did not contain the option to specify the Hadoop version.

For more information, see the "infacmd cluster Command Reference" chapter in the *Informatica 10.2.1 Command Reference*.

infacmd cms Commands

The following table describes new Content Management Service options for infacmd cms updateServiceOptions:

Command	Description
DataServiceOptions.RefDataLocationSchema	Identifies the schema that specifies the reference data tables in the reference data database.

For more information, see the "infacmd cms Command Reference" chapter in the *Informatica 10.2.1 Command Reference*.

infacmd dis Commands

The following table describes new infacmd dis commands:

Command	Description
listMappingEngines	Lists the execution engines of the deployed mappings on a Data Integration Service.

For more information, see the "infacmd dis Command Reference" chapter in the *Informatica 10.2.1 Command Reference*.

infacmd ihs Commands

The following table describes new infacmd ihs commands:

Command	Description
ListServiceProcessOptions	Lists process options for the Informatica Cluster Service.
UpdateServiceProcessOptions	Updates service options for the Informatica Cluster Service.

For more information, see the "infacmd ihs Command Reference" chapter in the *Informatica 10.2.1 Command Reference*.

infacmd isp Commands

The following table describes new infacmd isp commands:

Command	Description
PingDomain	Pings a domain, service, domain gateway host, or node.
GetPasswordComplexityConfig	Returns the password complexity configuration for the domain users.
ListWeakPasswordUsers	Lists the users with passwords that do not meet the password policy.

For more information, see the "infacmd isp Command Reference" chapter in the *Informatica 10.2.1 Command Reference*.

infacmd ldm Commands

The following table describes new infacmd ldm commands:

Command	Description
ListServiceProcessOptions	Lists options for the Catalog Administrator process.
UpdateServiceProcessOptions	Updates process options for the Catalog Service.

For more information, see the "infacmd ldm Command Reference" chapter in the *Informatica 10.2.1 Command Reference*.

infacmd mi Commands

mi is a new infacmd plugin that performs mass ingestion operations.

The following table describes new infacmd mi commands:

Command	Description
abortRun	Aborts the ingestion mapping jobs in a run instance of a mass ingestion specification.
createService	Creates a Mass Ingestion Service. Disabled by default. To enable the Mass Ingestion Service, use infacmd isp enableService.
deploySpec	Deploys a mass ingestion specification.
exportSpec	Exports the mass ingestion specification to an application archive file.
extendedRunStats	Gets the extended statistics for a mapping in the deployed mass ingestion specification.
getSpecRunStats	Gets the detailed run statistics for a deployed mass ingestion specification.
listSpecRuns	Lists the run instances of a deployed mass ingestion specification.
listSpecs	Lists the mass ingestion specifications.

Command	Description
restartMapping	Restarts the ingestion mapping jobs in a mass ingestion specification.
runSpec	Runs a mass ingestion specification that is deployed to a Data Integration Service.

For more information, see the "infacmd mi Command Reference" chapter in the *Informatica 10.2.1 Command Reference*.

infacmd mrs Commands

The following table describes new infacmd mrs commands:

Command	Description
listMappingEngines	Lists the execution engines of the mappings that are stored in a Model repository.
listPermissionOnProject	Lists all the permissions on multiple projects for groups and users.
updateStatistics	Updates the statistics for the monitoring Model repository on Microsoft SQL Server.

For more information, see the "infacmd mrs Command Reference" chapter in the *Informatica 10.2.1 Command Reference*.

infacmd wfs Commands

The following table describes new infacmd wfs commands:

Command	Description
pruneOldInstances	Deletes workflow process data from the workflow database.

To delete the process data, you must have the Manage Service privilege on the domain.

For more information, see the "infacmd wfs Command Reference" chapter in the *Informatica 10.2.1 Command Reference*.

infasetup Commands

The following table describes new infasetup commands:

Command	Description
UpdatePasswordComplexityConfig	Enables or disables the password complexity configuration for the domain.

For more information, see the "infasetup Command Reference" chapter in the *Informatica 10.2.1 Command Reference*.

Enterprise Data Catalog

This section describes new Enterprise Data Catalog features in version 10.2.1.

Adding a Business Title to an Asset

Effective in version 10.2.1, you can add a business title to any asset in the catalog except for Business Glossary and Axon glossary assets. You can either associate a business term or provide a display name to add a business title to an asset.

For more information about adding a business title, see the *Informatica 10.2 .1 Enterprise Data Catalog User Guide*.

Cluster Validation Utility in Installer

Effective in version 10.2.1, when you install Enterprise Data Catalog, the installer provides an option to run the cluster-validation utility. The utility helps you validate the prerequisites to install Enterprise Data Catalog in an embedded cluster and existing cluster. The utility also validates the configuration settings for Informatica domain, cluster hosts, and the Hadoop cluster services.

For more information about the utility, see the *Informatica Enterprise Data Catalog 10.2 .1 Installation and Configuration Guide* and the following knowledge base articles:

- HOW TO: Validate Embedded Cluster Prerequisites with Validation Utility in Enterprise Information Catalog
- HOW TO: Validate Informatica Domain, Cluster Hosts, and Cluster Services Configuration

Data Domain Discovery Types

Effective in version 10.2.1, when you configure the data domain discovery profile settings, you can choose one of the following data domain discovery types:

- Run Discovery on Source Data. Scanner runs data domain discovery on source data.
- Run Discovery on Source Metadata. Scanner runs data domain discovery on source metadata.
- Run Discovery on both Source Metadata and Data. Scanner runs data domain discovery on source data and source metadata.
- Run Discovery on Source Data Where Metadata Matches. Scanner runs data domain discovery on the source metadata to identify the columns with inferred data domains. The scanner then runs discovery on the source data for the columns that have inferred data domains.

For more information about data domain discovery types, see the *Informatica 10.2 .1 Catalog Administrator Guide*.

Filter Settings

Effective in version 10.2.1, you can use the filter settings in the Application Configuration page to customize the search filters that you view in the **Filter By** panel of the search results page.

For more information about search filters, see the *Informatica Enterprise Data Catalog 10.2.1 User Guide*.

Missing Links Report

Effective in version 10.2.1, you can now generate a missing links report to identify the connection links that are missing after you assign schemas from a resource to connections.

For more information about the missing links report, see the *Informatica 10.2.1 Catalog Administrator Guide*.

New Resource Types

Effective in version 10.2.1, Informatica Enterprise Data Catalog extracts metadata from several new data sources.

You can create resources in Informatica Catalog Administrator to extract metadata from the following data sources:

Azure Data Lake Store

Online cloud file storage platform.

Database Scripts

Database scripts to extract lineage information. The Database Scripts resource is available for technical preview. Technical preview functionality is supported but is unwarranted and is not production-ready. Informatica recommends that you use these features in non-production environments only.

Microsoft Azure Blob Storage

Cloud-based file storage web service.

QlikView

Business Intelligence tool that allows you to extract metadata from the QlikView source system.

SharePoint

Import metadata from files in SharePoint.

OneDrive

Import metadata from files in OneDrive.

For more information about the new resources, see the *Informatica 10.2 .1 Catalog Administrator Guide*.

REST APIs

Effective in version 10.2.1, you can use Informatica Enterprise Data Catalog REST APIs to load and monitor resources.

For more information about the REST APIs, see the *Informatica 10.2 .1 Enterprise Data Catalog REST API Reference*.

SAML Authentication for Enterprise Data Catalog Applications

Effective in version 10.2.1, you can enable Single Sign-on using SAML authentication for Enterprise Data Catalog applications. You can either use SAML authentication using OKTA with Active Directory or Active Directory Federation Services with Active Directory.

For more information, see the *Informatica Enterprise Data Catalog 10.2 .1 Installation and Configuration Guide*.

SAP Resource

Effective in version 10.2.1, you can choose the **Enable Streaming for Data Access** option for SAP R/3 resources to extract data by using the HTTP protocol.

For more information about the option, see the *Informatica 10.2 .1 Catalog Administrator Guide*.

Import from ServiceNow

Effective in version 10.2.1, Catalog Administrator now connects to ServiceNow to import connections and extract the configuration metadata into the catalog.

The Import from ServiceNow feature is available for technical preview. Technical preview functionality is supported but is unwarranted and is not production-ready. Informatica recommends that you use these features in non-production environments only.

For more information about importing metadata from ServiceNow, see the *Informatica 10.2 .1 Catalog Administrator Guide*.

Similar Columns

Effective in version 10.2.1, you can view the Similar Columns section that displays all the columns that are similar to the column you are viewing. Enterprise Data Catalog discovers similar columns based on column names, column patterns, unique values, and value frequencies.

For more information about column similarity, see the *Informatica 10.2 .1 Enterprise Data Catalog User Guide*.

Specify Load Types for Catalog Service

Effective in version 10.2.1, when you create a Catalog Service, you can choose the option to specify the data size that you want to deploy.

Previously, you had to create the Catalog Service and use the custom properties for the Catalog Service to specify the data size.

For more information, see the *Informatica Enterprise Data Catalog 10.2 .1 Installation and Configuration Guide*.

Supported Resource Types for Data Discovery

Effective in version 10.2.1, you can enable data discovery for the following resources to extract profiling metadata:

- Unstructured file types:
  - Apple Files. Supported extension types include .key, .pages, .numbers, .ibooks, and .ipa.
  - Open Office Files. Supported extension types include .odt, .ott, .odm, .ods, .ots, .odp, .odg, .otp, .odg, .otg, and .odf.
- Structured file types:
  - Avro. Supported extension type is .avro.

    This file type is available for HDFS resource and File System resource. For the File System resource, you can choose only the Local File protocol.

- Parquet. Supported extension type is .parquet.

  This file type is available for HDFS resource and File System resource. For the File System resource, you can choose only the Local File protocol.

- Other resources:
  - Azure Data Lake Store
  - File System. Supported protocols include Local File, SFTP, and SMB/CIFS protocol.
  - HDFS. Supported distribution includes MapR FS.
  - Microsoft Azure Blob Storage
  - OneDrive
  - SharePoint

For more information about new resources, see the *Informatica 10.2 .1 Catalog Administrator Guide*.

Enterprise Data Lake

This section describes new Enterprise Data Lake features in version 10.2.1.

Column Data

Effective in version 10.2.1, you can use the following features when you work with columns in worksheets:

- You can categorize or group related values in a column into categories to make analysis easier.
- You can view the source of the data for a selected column in a worksheet. You might want to view the source of the data in a column to help you troubleshoot an issue.
- You can revert types or data domains inferred during sampling on columns to the source type. You might want to revert an inferred type or data domain to the source type if you want to use the column data in a formula.

For more information, see the "Prepare Data" chapter in the *Informatica 10.2.1 Enterprise Data Lake User Guide*.

Manage Data Lake Resources

Effective in version 10.2.1, you can use the Enterprise Data Lake application to add and delete Enterprise Data Catalog resources. Catalog resources represent the external data sources and metadata repositories from which scanners extract metadata that can be used in the data lake.

For more information, see the "Managing the Data Lake" chapter in the *Informatica 10.2.1 Enterprise Data Lake Administrator Guide*.

Data Preparation Operations

Effective in version 10.2.1, you can perform the following operations during data preparation:

Pivot Data

You can use the pivot operation to reshape the data in selected columns in a worksheet into a summarized format. The pivot operation enables you to group and aggregate data for analysis, such as

summarizing the average price of single family homes sold in each city for the first six months of the year.

Unpivot Data

You can use the unpivot operation to transform columns in a worksheet into rows containing the column data in key value format. The unpivot operation is useful when you want to aggregate data in a worksheet into rows based on keys and corresponding values.

Apply One Hot Encoding

You can use the one hot encoding operation to determine the existence of a string value in a selected column within each row in a worksheet. You might use the one hot encoding operation to convert categorical values in a worksheet to numeric values required by machine learning algorithms.

For more information, see the "Prepare Data" chapter in the *Informatica 10.2.1 Enterprise Data Lake User Guide*.

Prepare JSON Files

Effective in version 10.2.1, you can sample the hierarchal data in JavaScript Object Notation Lines (JSONL) files you add to your project as the first step in data preparation. Enterprise Data Lake converts the JSON file structure into a flat structure, and presents the data in a worksheet that you use to sample the data.

For more information, see the "Prepare Data" chapter in the *Informatica 10.2.1 Enterprise Data Lake User Guide*.

Recipe Steps

Effective in version 10.2.1, you can use the following features when you work with recipes in worksheets:

- You can reuse recipe steps created in a worksheet, including steps that contain complex formulas or rule definitions. You can reuse recipe steps within the same worksheet or in a different worksheet, including a worksheet in another project. You can copy and reuse selected steps from a recipe, or you can reuse the entire recipe.
- You can insert a step at any position in a recipe.
- You can add a filter or modify a filter applied to a recipe step.

For more information, see the "Prepare Data" chapter in the *Informatica 10.2.1 Enterprise Data Lake User Guide*.

Schedule Export, Import, and Publish Activities

Effective in version 10.2.1, you can schedule the exporting, importing, and publishing of data assets. Scheduling an activity enables you to import, export or publish updated data assets on a recurring basis.

When you schedule an activity, you can create a new schedule, or you can select an existing schedule. You can use schedules created by other users, and other users can use schedules that you create.

For more information, see the "Scheduling Export, Import, and Publish Activities" chapter in the *Informatica 10.2.1 Enterprise Data Lake User Guide*.

Security Assertion Markup Language Authentication

Effective in version 10.2.1, the Enterprise Data Lake application supports Security Assertion Markup Language (SAML) authentication.

For more information on configuring SAML authentication, see the *Informatica 10.2.1 Security Guide*.

View Project Flows and Project History

Effective in version 10.2.1, you can view project flow diagrams and review the activities performed within a project.

You can view a flow diagram that shows you how worksheets in a project are related and how they are derived. The diagram is especially useful when you work on a complex project that contains numerous worksheets and includes numerous assets.

You can also review the complete history of the activities performed within a project, including activities performed on worksheets within the project. Viewing the project history might help you determine the root cause of issues within the project.

For more information, see the "Create and Manage Projects" chapter in the *Informatica 10.2.1 Enterprise Data Lake User Guide*.

Informatica Developer

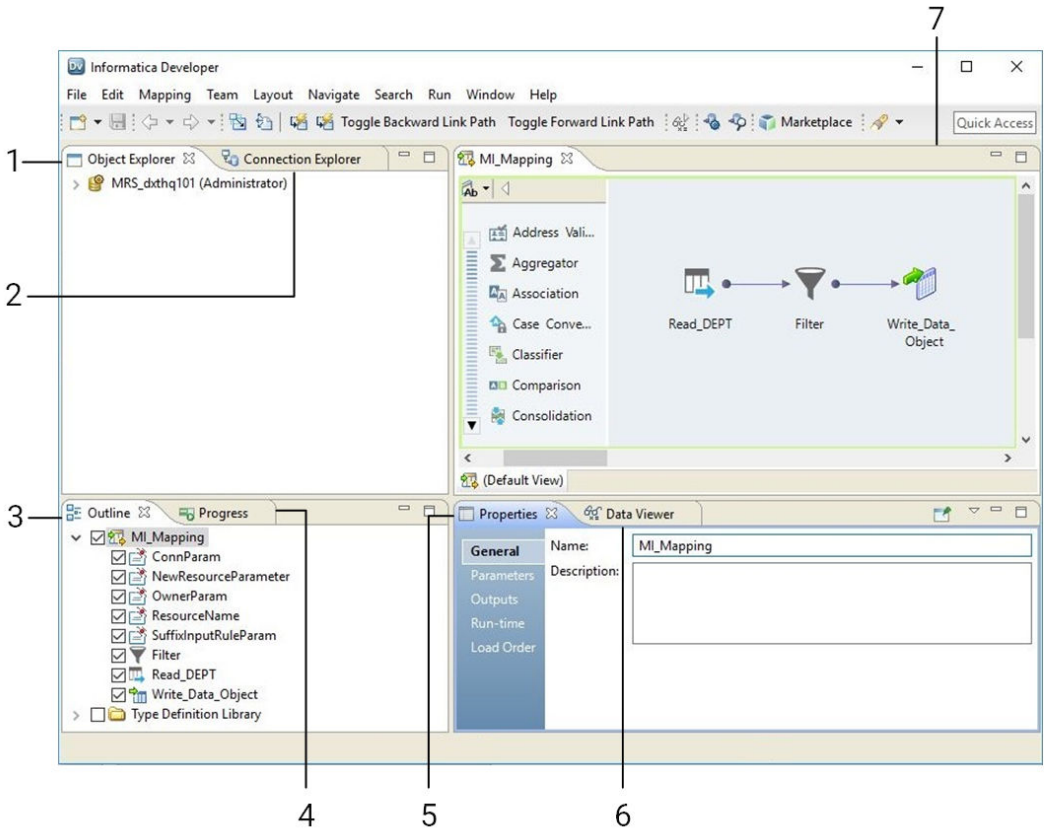
This section describes new Developer tool features in version 10.2.1.

Default Layout

Effective in version 10.2.1, the following additional views appear by default in the Developer tool workbench:

- Connection Explorer view
- Progress view

The following image shows the default Developer tool workbench in version 10.2.1:



1. Object Explorer view
2. Connection Explorer view
3. Outline view
4. Progress view
5. Properties view
6. Data Viewer view
7. Editor

For more information, see the "Informatica Developer" chapter in the *Informatica 10.2.1 Developer Tool Guide*.

Editor Search

Effective in version 10.2.1, you can search for a complex data type definition in mappings and mapplets in the Editor view. You can also show link paths using a complex data type definition.

For more information, see the "Searches in Informatica Developer" chapter in the *Informatica 10.2.1 Developer Tool Guide*.

Import Session Properties from PowerCenter

Effective in version 10.2.1, you can import session properties, such as SQL-based overrides in relational sources and targets and overrides for the Lookup transformation from the PowerCenter repository to the Model repository.

For more information about the import from PowerCenter functionality, see the "Import from PowerCenter" chapter in the *Informatica 10.2.1 Developer Mapping Guide*.

Views

Effective in version 10.2.1, you can expand complex data types to view the complex data type definition in the following views:

- Editor view
- Outline view
- Properties view

For more information, see the "Informatica Developer" chapter in the *Informatica 10.2.1 Developer Tool Guide*.

Informatica Mappings

This section describes new Informatica mapping features in version 10.2.1.

Dynamic Mappings

This section describes new dynamic mapping features in version 10.2.1.

Input Rules

Effective in version 10.2.1, you can perform the following tasks when you create an input rule:

- Create an input rule by complex data type definition.
- Restore source port names when you rename generated ports.
- Select ports by source name when you create an input rule by column name or a pattern.
- View source names and complex data type definitions in the port preview.

For more information, see the "Dynamic Mappings" chapter in the *Informatica 10.2.1 Developer Mapping Guide*.

Port Selectors

Effective in version 10.2.1, you can configure a port selector to select ports by complex data type definition.

For more information, see the "Dynamic Mappings" chapter in the *Informatica 10.2.1 Developer Mapping Guide*.

Validate Dynamic Sources and Targets

Effective in version 10.2.1, you can validate dynamic sources and targets. To validate dynamic sources and targets, resolve the mapping parameters to view a run-time instance of the mapping. Validate the run-time instance of the mapping.

For more information, see the "Dynamic Mappings" chapter in the *Informatica 10.2.1 Developer Mapping Guide*.

Mapping Parameters

This section describes new mapping parameters features in version 10.2.1

Assign Parameters

Effective in version 10.2.1, you can assign parameters to the following mapping objects and object fields:

Object	Field
Customized data object read operation	Custom query Filter condition Join condition PreSQL PostSQL
Customized data object write operation	PreSQL PostSQL Update override
Flat file data object	Compression codec Compression format
Lookup transformation	Custom query. Relational only.
Read transformation	Custom query. Relational only. Filter condition. Relational only. Join condition. Relational only. PreSQL. Relational only. PostSQL. Relational only.
Write transformation	PreSQL. Relational only. PostSQL. Relational only. Update override. Relational only.

For more information, see the "Mapping Parameters" chapter in the *Informatica 10.2.1 Developer Mapping Guide*.

Resolve Mapping Parameters

Effective in version 10.2.1, you can resolve mapping parameters in the Developer tool. When you resolve mapping parameters, the Developer tool generates a run-time instance of the mapping that shows how the Data Integration Service resolves the parameters at run time. You can run the instance of the mapping where the parameters are resolved to run the mapping with the selected parameters.

The following table describes the options that you can use to resolve mapping parameters:

Mapping Parameters	Description
Apply the default values in the mapping	Resolves the mapping parameters based on the default values configured for the parameters in the mapping. If parameters are not configured for the mapping, no parameters are resolved in the mapping.
Apply a parameter set	Resolves the mapping parameters based on the parameter values defined in the specified parameter set.
Apply a parameter file	Resolves the mapping parameters based on the parameter values defined in the specified parameter file.

To quickly resolve mapping parameters based on a parameter set. Drag the parameter set from the Object Explorer view to the mapping editor to view the resolved parameters in the run-time instance of the mapping.

For more information, see the "Mapping Parameters" chapter in the *Informatica 10.2.1 Developer Mapping Guide*.

Validate Mapping Parameters

Effective in version 10.2.1, you can validate mapping parameters in the Developer tool. To validate mapping parameters, first resolve the mapping parameters. When you resolve mapping parameters, the Developer tool generates a run-time instance of the mapping that shows the resolved parameters. Validate the run-time instance of the mapping to validate the mapping parameters.

For more information, see the "Mapping Parameters" chapter in the *Informatica 10.2.1 Developer Mapping Guide*.

Running Mappings

This section describes new run mapping features in version 10.2.1.

Run a Mapping from the Object Explorer View

Effective in version 10.2.1, you can run a mapping from the Object Explorer view. You do not have to open the mapping in the mapping editor. Right-click the mapping in the Object Explorer view and click Run.

For more information, see the *Informatica 10.2.1 Developer Tool Guide*.

Run a Mapping Using Advanced Options

Effective in version 10.2.1, you can run a mapping in the Developer tool using advanced options. In the advanced options, you can specify a mapping configuration and mapping parameters. Specify the mapping configuration and mapping parameters each time that you run the mapping.

The following table describes the options that you can use to specify a mapping configuration:

Option	Description
Select a mapping configuration	Select a mapping configuration from the drop-down menu. To create a new mapping configuration, select New Configuration.
Specify a custom mapping configuration	Create a custom mapping configuration that persists for the current mapping run.

The following table describes the options that you can use to specify mapping parameters:

Mapping Parameters	Description
Apply the default values in the mapping	Resolves the mapping parameters based on the default values configured for the parameters in the mapping. If parameters are not configured for the mapping, no parameters are resolved in the mapping.
Apply a parameter set	Resolves the mapping parameters based on the parameter values defined in the specified parameter set.
Apply a parameter file	Resolves the mapping parameters based on the parameter values defined in the specified parameter file.

For more information, see the *Informatica 10.2.1 Developer Mapping Guide*.

Truncate Partitioned Hive Target Tables

Effective in version 10.2.1, you can truncate an external or managed Hive table with or without partitions.

Previously, you could design a mapping to truncate a Hive target table, but not an external, partitioned Hive target table.

For more information on truncating Hive targets, see the "Mapping Targets in the Hadoop Environment" chapter in the *Informatica Big Data Management 10.2.1 User Guide*.

Informatica Transformation Language

This section describes Informatica Transformation Language new features in 10.2.1.

Complex Functions for Map Data Type

Effective in version 10.2.1, the transformation language introduces complex functions for map data type. Use complex functions for map data type to generate or process map data on the Spark engine.

The transformation language includes the following complex functions for map data type:

- COLLECT_MAP
- MAP
- MAP_FROM_ARRAYS
- MAP_KEYS
- MAP_VALUES

Effective in version 10.2.1, you can use the SIZE function to determine the size of map data.

For more information about complex functions, see the "Functions" chapter in the *Informatica 10.2.1 Developer Transformation Language Reference*.

Complex Operator for Map Data Type

Effective in version 10.2.1, you can use a complex operator in mappings that run on the Spark engine to access elements in a map data type.

Map data type contains an unordered collection of key-value pair elements. Use the subscript operator [ ] to access the value corresponding to a given key in the map data type.

For more information about complex operators, see the "Operators" chapter in the *Informatica 10.2.1 Developer Transformation Language Reference*.

Informatica Transformations

This section describes new Informatica transformation features in version 10.2.1.

Address Validator Transformation

This section describes the new Address Validator transformation features.

The Address Validator transformation contains additional address functionality for the following countries:

Argentina

Effective in version 10.2.1, you can configure Informatica to return valid suggestions for an Argentina address that you enter on a single line.

Enter an Argentina address in the following format:

```
[Street] [House Number] [Dependent Locality] [Post Code] [Locality]
```

To verify single-line addresses, enter the addresses in the `Complete Address` port.

Brazil

Effective in version 10.2.1, you can configure Informatica to return valid suggestions for a Brazil address that you enter on a single line.

Enter a Brazil address in the following format:

```
[Street] [House Number] [Locality] [State Code] [Post Code]
```

To verify single-line addresses, enter the addresses in the `Complete Address` port.

Colombia

Effective in version 10.2.1, Informatica validates an address in Colombia to house number level.

Hong Kong

Effective in version 10.2.1, Informatica supports rooftop geocoding for Hong Kong addresses. Informatica can return rooftop geocoordinates for a Hong Kong address that you submit in the Chinese language or the English language.

Informatica can consider all three levels of building information when it generates the geocoordinates. It delivers rooftop geocoordinates to the lowest level available in the verified address.

To retrieve rooftop geocoordinates for Hong Kong addresses, install the `HKG5GCRT.MD` database.

India

Effective in version 10.2.1, Informatica validates an address in India to house number level.

Mexico

Effective in version 10.2.1, you can configure Informatica to return valid suggestions for a Mexico address that you enter on a single line.

Enter a Mexico address in the following format:

```
[Street] [House Number] [Sub-locality] [Post Code] [Locality] [Province]
```

To verify single-line addresses, enter the addresses in the `Complete Address` port.

South Africa

Effective in version 10.2.1, Informatica improves the parsing and verification of delivery service descriptors in South Africa addresses.

Informatica improves the parsing and verification of the delivery service descriptors in the following ways:

- Address Verification recognizes Private Bag, Cluster Box, Post Office Box, and Postnet Suite as different types of delivery service. Address Verification does not standardize one delivery service descriptor to another. For example, Address Verification does not standardize Postnet Suite to Post Office Box.
- Address Verification parses Postnet Box as a non-standard delivery service descriptor and corrects Postnet Box to the valid descriptor Postnet Suite.
- Address Verification does not standardize the sub-building descriptor Flat to Fl.

South Korea

Effective in version 10.2.1, Informatica introduces the following features and enhancements for South Korea:

- The South Korea address reference data includes building information. Informatica can read, verify, and correct building information in a South Korea address.
- Informatica returns all of the current addresses at a property that an older address represents. The older address might represent a single current address or it might represent multiple addresses, for example if multiple residences occupy the site of the property.

  To return the current addresses, first find the address ID for the older property. When you submit the address ID with the final character A in address code lookup mode, Informatica returns all current addresses that match the address ID.

  Note: The Address Validator transformation uses the `Max Result Count` property to determine the maximum number of addresses to return for the address ID that you enter. The `Count Overflow` property indicates whether the database contains additional addresses for the address ID.

Thailand

Effective in version 10.2.1, Informatica introduces the following features and enhancements for Thailand:

Improvements to Thailand Addresses

Informatica improves the parsing and validation of Thailand addresses in a Latin script.

Additionally, Informatica validates an address to house number level.

Native Support for Thailand Addresses

Informatica can read and write Thailand addresses in native Thai and Latin scripts. Informatica updates the reference data for Thailand and adds reference data in the native Thai script.

Informatica provides separate reference databases for Thailand addresses in each script. To verify addresses in the native Thai script, install the native Thai databases. To verify addresses in a Latin script, install the Latin databases.

Note: If you verify Thailand addresses, do not install both database types. Accept the default option for the `Preferred Script` property.

United Arab Emirates

Effective in version 10.2.1, Informatica verifies street names in United Arab Emirates addresses. To verify street names in United Arab Emirates, install the current reference address databases for the United Arab Emirates.

United Kingdom

Effective in version 10.2.1, Informatica can return a United Kingdom territory name.

Informatica returns the territory name in the Country_2 element. Informatica returns the country name in the Country_1 element. You can configure an output address with both elements, or you can omit the Country_1 element if you post mail within the United Kingdom. The territory name appears above the postcode in a United Kingdom address on an envelope or label.

To return the territory name, install the current United Kingdom reference data.

United States

Effective in version 10.2.1, Informatica can recognize up to three sub-building levels in a United States address.

In compliance with the United States Postal Service requirements, Informatica matches the information in a single sub-building element with the reference data. If the Sub-building_1 information does not match, Informatica compares the Sub-building_2 information. If the Sub-building_2 information does not match, Address Verification compares the Sub-building_3 information. Address Verification copies the unmatched sub-building information from the input address to the output address.

Austria, Germany, and Switzerland

Effective in version 10.2.1, Informatica supports the uppercase character `ß` in Austria, Germany, and Switzerland addresses.

Informatica supports the character `ß` in the following ways:

- If you set the `Casing` property to UPPER, Informatica returns the German character `ß` as `ß`. If you set the `Casing` property to LOWER, Informatica returns the German character `ß` as `ß`.
- Informatica treats `ß` and `ß` as equally valid characters in an address. In reference data matches, Informatica can identify a perfect match when the same values contain either `ß` or `ß`.
- Informatica treats `ß` and `ss` as equally valid characters in an address. In reference data matches, Informatica can identify a standardized match when the same values contain either `ß` or `ss`.
- If you set the `Preferred Script` property to ASCII_SIMPLIFIED, Informatica returns the character `ß` as `S`.
- If you set the `Preferred Script` property to ASCII_EXTENDED, Informatica returns the character `ß` as `SS`.

For comprehensive information about the features and operations of the address verification software engine version that Informatica embeds in version 10.2.1, see the *Informatica Address Verification 5.12.0 Developer Guide*.

Informatica Workflows

This section describes new Informatica workflow features in version 10.2.1.

Import a Command Task from PowerCenter

Effective in version 10.2.1, you can import a Command task from PowerCenter into the Model repository.

For more information, see the "Workflows" chapter in the *Informatica 10.2.1 Developer Workflow Guide*.

PowerExchange Adapters for Informatica

This section describes new Informatica adapter features in version 10.2.1.

PowerExchange for Amazon Redshift

Effective in version 10.2.1, PowerExchange for Amazon Redshift includes the following features:

- You can configure a cached lookup operation to cache the lookup table on the Spark engine and an uncached lookup operation in the native environment.
- For a server-side encryption, you can configure the customer master key ID generated by AWS Key Management Service in the connection in the native environment and Spark engine.

For more information, see the *Informatica PowerExchange for Amazon Redshift 10.2.1 User Guide*.

PowerExchange for Amazon S3

Effective in version 10.2.1, PowerExchange for Amazon S3 includes the following features:

- For a client-side encryption, you can configure the customer master key ID generated by AWS Key Management Service in the connection in the native environment. For a server-side encryption, you can configure the customer master key ID generated by AWS Key Management Service in the connection in the native environment and Spark engine.
- For a server-side encryption, you can configure the Amazon S3-managed encryption key or AWS KMS-managed customer master key to encrypt the data while uploading the files to the buckets.
- You can create an Amazon S3 file data object from the following data source formats in Amazon S3:
  - Intelligent Structure Model
    The intelligent structure model feature for PowerExchange for Amazon S3 is available for technical preview. Technical preview functionality is supported but is not production-ready. Informatica recommends that you use in non-production environments only.
  - JSON
  - ORC
- You can compress an ORC data in the Zlib compression format when you write data to Amazon S3 in the native environment and Spark engine.
- You can create an Amazon S3 target using the **Create Target** option in the target session properties.
- You can use complex data types on the Spark engine to read and write hierarchical data in the Avro and Parquet file formats.
- You can use Amazon S3 sources as dynamic sources in a mapping. Dynamic mapping support for PowerExchange for Amazon S3 sources is available for technical preview. Technical preview functionality is supported but is unwarranted and is not production-ready. Informatica recommends that you use these features in non-production environments only.

For more information, see the *Informatica PowerExchange for Amazon S3 10.2.1 User Guide*.

PowerExchange for Cassandra

Effective in version 10.2.1, the Informatica Cassandra ODBC driver supports asynchronous write.

To enable asynchronous write on a Linux operating system, you must add the **EnableAsynchronousWrites** key name in the `odbc.ini` file and set the value to 1.

To enable asynchronous write on a Windows operating system, you must add the **EnableAsynchronousWrites** property in the Windows registry for the Cassandra ODBC data source name and set the value as 1.

For more information, see the *Informatica PowerExchange for Cassandra 10.2.1 User Guide*.

PowerExchange for HBase

Effective in version 10.2.1, you can use an HBase data object read operation to look up data in an HBase resource. You can enable lookup caching and also parameterize the lookup condition.

For more information, see the *Informatica PowerExchange for HBase 10.2.1 User Guide*.

PowerExchange for HDFS

Effective in version 10.2.1, you can use the following new PowerExchange for HDFS features:

Intelligent structure model support for complex file data objects

You can incorporate an intelligent structure model in a complex file data object. When you add the data object to a mapping that runs on the Spark engine, you can process any input type that the model can parse.

The intelligent structure model feature for PowerExchange for HDFS is available for technical preview. Technical preview functionality is supported but is not production-ready. Informatica recommends that you use in non-production environments only.

For more information, see the *Informatica PowerExchange for HDFS 10.2.1 User Guide*.

Dynamic mapping support for complex file sources

You can use complex file sources as dynamic sources in a mapping.

Dynamic mapping support for complex file sources is available for technical preview. Technical preview functionality is supported but is unwarranted and is not production-ready. Informatica recommends that you use these features in non-production environments only.

For more information about dynamic mappings, see the *Informatica Developer Mapping Guide*.

PowerExchange for Hive

Effective in version 10.2.1, PowerExchange for Hive supports mappings that run PreSQL and PostSQL queries against Hive sources and targets on the Spark engine.

For more information, see the *Informatica PowerExchange for Hive 10.2.1 User Guide*.

PowerExchange for Microsoft Azure Blob Storage

Effective in version 10.2.1, PowerExchange for Microsoft Azure Blob Storage includes the following functionality:

- You can run mappings on the Spark engine.

- You can read and write .csv, Avro, and Parquet files when you run a mapping on the Spark engine and in the native environment.
- You can read and write JSON and intelligent structure files when you run a mapping on the Spark engine.
- You can read a directory when you run a mapping on the Spark engine.
- You can generate or skip header rows when you run a mapping in the native environment. On the Spark engine, the header row is created by default.
- You can append an existing blob. The append operation is applicable to only to the append blob and in the native environment.
- You can override the blob or container name. In the Blob Container Override field, specify the container name or sub-folders in the root container with the absolute path.
- You can read and write .csv files compressed in the gzip format.

All new functionality for PowerExchange for Microsoft Azure Blob Storage is available for technical preview. Technical preview functionality is supported but is not production-ready. Informatica recommends that you use in non-production environments only.

For more information, see the *Informatica PowerExchange for Microsoft Azure Blob Storage 10.2.1 User Guide*.

PowerExchange for Microsoft Azure SQL Data Warehouse

Effective in version 10.2.1, PowerExchange for Microsoft Azure SQL Data Warehouse includes the following features:

- You can run mappings on the Spark engine.
- You can configure key range partitioning when you read data from Microsoft Azure SQL Data Warehouse objects.
- You can override the SQL query and define constraints when you read data from a Microsoft Azure SQL Data Warehouse object.
- You can configure pre-SQL and post-SQL queries for source and target objects in a mapping.
- You can configure the native expression filter for the source data object operation.
- You can perform update, upsert, and delete operations against Microsoft Azure SQL Data Warehouse tables.
- You can configure a cached lookup operation to cache the lookup table on the Spark engine and an uncached lookup operation in the native environment.

For more information, see the *Informatica PowerExchange for Microsoft Azure SQL Data Warehouse 10.2.1 User Guide*.

PowerExchange for Salesforce

Effective in version 10.2.1, you can use version 41 of Salesforce API to create a Salesforce connection and access Salesforce objects. You can use big objects with source and target transformations.

For more information, see the *Informatica PowerExchange for Salesforce 10.2.1 User Guide*.

PowerExchange for SAP NetWeaver

Effective in version 10.2.1, you can run mappings on the Spark engine to read data from SAP tables.

For more information, see the *Informatica PowerExchange for SAP NetWeaver 10.2.1 User Guide*.

PowerExchange for Snowflake

Effective in version 10.2.1, PowerExchange for Snowflake includes the following features:

- You can configure a lookup operation on a Snowflake table. You can also enable lookup caching for a lookup operation to increase the lookup performance. The Data Integration Service caches the lookup source and runs the query on the rows in the cache.
- You can parameterize the Snowflake connection, and data object read and write operation properties.
- You can configure key range partitioning for Snowflake data objects in a read or write operation. The Data Integration Service distributes the data based on the port or set of ports that you define as the partition key.
- You can specify a table name in the advanced target properties to override the table name in the Snowflake connection properties.

For more information, see the *Informatica PowerExchange for Snowflake 10.2.1 User Guide*.

Security

This section describes new security features in version 10.2.1.

Password Complexity

Effective in version 10.2.1, you can enable password complexity to validate the password strength. By default this option is disabled.

For more information, see the "Security Management in Informatica Administrator" chapter in the *Informatica 10.2.1 Security Guide*.