



Informatica®

10.4.0

Guía de nuevas características

© Copyright Informatica LLC 2009, 2019

Este software y la documentación se proporcionan exclusivamente en virtud de un acuerdo de licencia independiente que contiene restricciones de uso y divulgación. Ninguna parte de este documento puede ser reproducida o transmitida de cualquier forma o manera (electrónica, fotocopia, grabación o mediante otros métodos) sin el consentimiento previo de 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, Big Data Management, PowerCenter, and PowerExchange 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.

La información contenida en esta documentación está sujeta a cambios sin previo aviso. Si encuentra algún problema en esta documentación, escribanos a infa_documentation@informatica.com para notificárnoslo.

Los productos de Informatica gozan de garantía en función de los términos y condiciones de los acuerdos conforme a los cuales se proporcionen. INFORMATICA PROPORCIONA LA INFORMACIÓN DE ESTE DOCUMENTO "TAL CUAL" SIN GARANTÍA DE NINGÚN TIPO, EXPRESA O IMPLÍCITA, INCLUIDAS LAS GARANTÍAS DE COMERCIALIZACIÓN, ADAPTACIÓN A UN FIN PARTICULAR Y CUALQUIER GARANTÍA O CONDICIÓN DE NO INCUMPLIMIENTO.

Fecha de publicación: 2019-12-20

Tabla de contenido

Prefacio	8
Recursos de Informatica	8
Informatica Network.	8
Base de conocimiento de Informatica.	8
Documentación de Informatica	8
Matrices de disponibilidad de producto de Informatica.	9
Informatica Velocity.	9
Catálogo de soluciones de Informatica.	9
Servicio internacional de atención al cliente de Informatica.	9
 Capítulo 1: Nuevas características (10.4.0).....	10
CI/CD.	10
API de REST del servicio de integración de datos.	10
Comandos infacmd dis.	11
Servidor proxy inverso.	12
Comandos infacmd roh.	13
Implementación de revisiones de la aplicación.	13
Conexión a una aplicación en tiempo de ejecución.	14
Vista Explorador de objetos.	14
Etiquetas.	15
Programas de la línea de comandos.	15
Comandos infacmd isp (nuevas características de la versión 10.4.0).	16
Data Engineering Integration.	16
Compatibilidad con nuevos tipos de datos	16
Integración de AWS Databricks.	17
Flujos de trabajo de clúster para el acceso de HDInsight a recursos de ALDS Gen2.	17
Acceso al almacenamiento de Databricks Delta Lake.	17
Visualización de los nodos utilizados en una asignación.	18
Agregación de registros	18
Análisis de datos jerárquicos en el motor Spark.	18
Creación de perfiles y opciones de muestreo en el motor Spark.	18
Transformación de Python.	19
Sqoop.	19
Data Engineering Streaming	20
Confluent Schema Registry en asignaciones de transmisión	20
Transformaciones de Data Quality en asignaciones de transmisión.	20
Clúster efímero en asignaciones de transmisión.	20
Puerto FileName en Amazon S3	21
Microsoft Azure Data Lake Storage Gen2.	21
Asignaciones de transmisión en Azure Databricks.	21

Asignaciones dinámicas en Data Engineering Streaming	22
Enterprise Data Catalog.	22
Asignación de atributos personalizados a recursos y clases.	22
Nuevos recursos.	23
Recursos de referencia y activos de referencia.	23
Exportación de activos desde resultados de búsqueda.	23
Filtros de linaje e impacto.	24
Resumen de control de activos.	24
Reglas y cuadros de mandos.	24
Inferencia de clave única.	24
Detección del dominio de datos en el tipo de archivo CLOB.	24
Obtención de datos y opciones de muestreo en el motor Spark.	24
Seguimiento de la vista previa técnica.	25
Vista previa de datos y aprovisionamiento.	25
Tipos de recursos compatibles en la utilidad de escáner independiente.	26
API de REST.	26
Enterprise Data Preparation.	27
Administración de acceso a Data Lake.	27
Microsoft Azure Data Lake Storage como origen de datos.	27
Publicación de archivos en Data Lake.	27
Carga de archivos en Data Lake.	28
Asignaciones de Informatica	28
Vinculación de salidas de asignación con parámetros de asignación.	28
Recomendaciones y conclusiones de CLAIRE.	29
Actualización del nivel del optimizador de asignaciones.	29
Transformaciones de Informatica.	29
Transformación del validador de direcciones.	30
Flujos de trabajo de Informatica.	30
Propiedades avanzadas de la tarea Crear clúster en Amazon EMR.	30
Instalación de Informatica.	30
PostgreSQL.	30
Pre-installation (i10Pi) System Check Tool en modo silencioso.	31
Cifrado de contraseñas en el archivo de propiedades de la instalación silenciosa.	31
Modelo de estructura inteligente.	31
Tipos de entrada adicionales.	31
Creación de modelos a partir de muestras en tiempo de diseño.	31
Administración de datos sin identificar.	31
PowerCenter.	32
Conectividad.	32
Adaptadores de PowerExchange.	32
Adaptadores de PowerExchange para Informatica.	33
Adaptadores de PowerExchange para PowerCenter.	36

Seguridad.	38
Configuración de aplicaciones web para utilizar proveedores de identidades SAML diferentes.	38
Capítulo 2: Nuevas características (10.2.2 HotFix 1).....	39
Programas de la línea de comandos	39
Comandos infacmd ldm.	39
Enterprise Data Catalog.	39
Azure Data Lake Storage Gen2.	40
Vinculación sin distinción entre mayúsculas y minúsculas	40
Extensión Tableau de Enterprise Data Catalog	40
Nuevos recursos.	40
Escáner de almacén de creación de perfiles.	40
API de REST.	41
Búsqueda Mejoras	41
Búsqueda Tabuladores.	41
Tipos de recursos compatibles en la utilidad de escáner independiente.	42
Versión preliminar técnica.	42
Capítulo 3: Nuevas características (10.2.2 Service Pack 1).....	44
Big Data Management.	44
Sqoop.	44
Big Data Transmisión.	45
Destino de Amazon S3.	45
Función TIME_RANGE.	45
Escritura de datos en varios archivos HDFS.	45
Enterprise Data Catalog.	46
Tipos de recursos compatibles en la utilidad de escáner independiente.	46
Enterprise Data Preparation.	46
Revertir todos los tipos de datos inferidos.	46
Preparar archivos Avro y Parquet.	46
Adaptadores de PowerExchange para Informatica.	47
PowerExchange for Hive.	47
Capítulo 4: Nuevas características (10.2.2).....	48
servicios de aplicación.	48
Servicio de ingesta masiva.	48
Servicio de acceso a metadatos.	48
Servicio del concentrador de operaciones REST.	49
Big Data Management.	49
Integración de Azure Databricks.	49
Datos jerárquicos.	50
Alta disponibilidad.	51
Modelo de estructura inteligente.	52

Ingesta masiva.	52
Supervisión.	53
Seguridad.	53
Destinos.	53
Big Data Transmisión	53
Objetos de datos de centros de eventos de Azure.	54
Roles IAM de cuentas cruzadas en Amazon Kinesis Connection.	54
Modelo de estructura inteligente.	54
Puertos de encabezado para objetos de datos de Big Data Streaming.	54
Perfil de credenciales de AWS en la conexión de Amazon Kinesis.	54
Transmisión estructurada de Spark.	55
Transformación de ventana.	55
Programas de la línea de comandos.	55
Comandos infacmd dis.	55
Comandos infacmd ihs.	56
Comandos infacmd ipc.	57
Comandos infacmd ldm.	57
Comandos infacmd mi.	58
Comandos infacmd ms.	58
Comandos infacmd oie.	59
Comandos infacmd tools.	59
Comandos de infasetup.	59
Enterprise Data Catalog.	60
Asignar automáticamente un título empresarial a una columna.	60
Colaboración del usuario en activos.	60
Crear servicios de aplicación de Enterprise Data Catalog usando el programa de instalación.	60
Utilidad de validación de metadatos personalizada.	61
Notificaciones de cambio.	61
Business Glossary Informe de asignación.	61
Perfiles de sistema operativo.	61
API de REST.	61
Filtro de perfiles de datos y metadatos de origen.	62
Utilidad de analizador.	62
Tipos de recursos.	62
Enterprise Data Lake.	63
Aplicar reglas activas.	63
Eliminar filas duplicadas.	63
Agrupar y categorizar datos de columna	63
Recomendaciones basadas en CLAIRE.	63
Agregación condicional.	64
Enmascaramiento de datos.	64
Localización.	64

Orígenes y destinos de la partición.	64
Agregar comentarios a pasos de la receta.	64
Guardar una receta como asignación.	65
Amazon S3, ADLS, WASB, MapR-FS como orígenes de datos.	65
Funciones estadísticas.	65
Funciones de fecha y hora.	66
Funciones matemáticas.	66
Funciones de texto.	66
Funciones de ventana.	67
Purgar eventos de auditoría.	67
Motor de ejecución de Spark.	67
Informatica Developer.	68
Aplicaciones.	68
Asignaciones de Informatica.	68
Tipos de datos.	68
Salidas de asignación.	68
Parámetros de asignación.	69
Niveles del optimizador.	69
Sqoop.	69
Transformaciones de Informatica.	70
Transformación del validador de direcciones.	70
Transformación de estrategia de actualización.	73
Adaptadores de PowerExchange para Informatica.	73
PowerExchange for Amazon Redshift.	73
PowerExchange for Amazon S3.	74
PowerExchange for Google BigQuery.	75
PowerExchange for HBase.	75
PowerExchange for HDFS.	75
PowerExchange for Hive.	76
PowerExchange for MapR-DB.	77
PowerExchange for Microsoft Azure Blob Storage.	77
PowerExchange for Microsoft Azure Cosmos DB SQL API.	77
PowerExchange for Microsoft Azure Data Lake Store.	78
PowerExchange for Microsoft Azure SQL Data Warehouse.	78
PowerExchange for Salesforce.	79
PowerExchange for Snowflake.	79
PowerExchange para la API del transportador paralelo de Teradata.	80

Prefacio

Consulte la *Guía de nuevas características y mejoras de Informatica®* para conocer las nuevas características y mejoras en los productos de Informatica en esta versión.

Recursos de Informatica

Informatica proporciona una variedad de recursos de productos a través de Informatica Network y otros portales en línea. Use los recursos para sacar el mayor provecho de los productos y las soluciones de Informatica y aprender de otros expertos en la materia y usuarios de Informatica.

Informatica Network

Informatica Network es la puerta de entrada a muchos recursos, entre ellos, la base de conocimientos de Informatica y el servicio internacional de atención al cliente de Informatica. Para entrar en Informatica Network, visite <https://network.informatica.com>.

Como miembro de Informatica Network, tiene las siguientes opciones:

- Buscar recursos de productos en la base de conocimientos
- Ver la información de disponibilidad del producto
- Crear y revisar casos de soporte
- Buscar su red de grupos de usuarios de Informatica locales y colaborar con sus pares

Base de conocimiento de Informatica

Use la base de conocimientos de Informatica para encontrar recursos de productos como artículos prácticos, procedimientos recomendados, tutoriales de video y respuestas a preguntas frecuentes.

Para buscar en la base de conocimiento, visite <https://search.informatica.com>. Si tiene preguntas, comentarios o ideas relacionadas con la base de conocimiento de Informatica, póngase en contacto con el equipo de la base de conocimiento de Informatica en KB_Feedback@informatica.com.

Documentación de Informatica

Use el portal de documentación de Informatica para recorrer una extensa biblioteca de documentación para las versiones de productos actuales y recientes. Para recorrer el portal de documentación, visite <https://docs.informatica.com>.

Si tiene preguntas, comentarios o ideas acerca de la documentación de los productos, póngase en contacto con el equipo de la documentación de Informatica en infa_documentation@informatica.com.

Matrices de disponibilidad de producto de Informatica

Las matrices de disponibilidad de producto (PAM, Product Availability Matrixes) indican las versiones de sistemas operativos, bases de datos y otros tipos de orígenes y destinos de datos admitidos por la versión de un producto. Puede recorrer las PAM de Informatica en

<https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Velocity

Informatica Velocity es una colección de consejos y procedimientos recomendados desarrollados por los servicios profesionales de Informatica que se basan en experiencias reales de cientos de proyectos de administración de datos. Informatica Velocity representa el conocimiento colectivo de los consultores de Informatica que trabajan con organizaciones de todo el mundo para planificar, desarrollar, implementar y dar mantenimiento a soluciones de administración de datos exitosas.

Puede encontrar recursos de Informatica Velocity en <http://velocity.informatica.com>. Si tiene alguna pregunta, comentario o idea acerca de Informatica Velocity, póngase en contacto con los servicios profesionales de Informatica en ips@informatica.com.

Catálogo de soluciones de Informatica

El catálogo de soluciones de Informatica es un foro donde puede buscar soluciones que aumenten, amplíen o mejoren sus implementaciones de Informatica. Aproveche cualquiera de los cientos de soluciones de socios y desarrolladores de Informatica que se encuentran en el catálogo para mejorar su productividad y acelerar la implementación de los proyectos. Puede encontrar el catálogo de soluciones de Informatica en

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

Servicio internacional de atención al cliente de Informatica

Puede ponerse en contacto con un centro de atención global por teléfono o a través del Informatica Network.

Para encontrar el número de teléfono local del servicio internacional de atención al cliente de Informatica, visite el sitio web de Informatica en el siguiente vínculo:

<https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

Para encontrar recursos de soporte en línea en Informatica Network, visite <https://network.informatica.com> y seleccione la opción eSupport.

CAPÍTULO 1

Nuevas características (10.4.0)

Este capítulo incluye los siguientes temas:

- [CI/CD, 10](#)
- [Programas de la línea de comandos, 15](#)
- [Data Engineering Integration, 16](#)
- [Data Engineering Streaming , 20](#)
- [Enterprise Data Catalog, 22](#)
- [Enterprise Data Preparation, 27](#)
- [Asignaciones de Informatica , 28](#)
- [Transformaciones de Informatica, 29](#)
- [Flujos de trabajo de Informatica, 30](#)
- [Instalación de Informatica, 30](#)
- [Modelo de estructura inteligente, 31](#)
- [PowerCenter, 32](#)
- [Adaptadores de PowerExchange, 32](#)
- [Seguridad, 38](#)

CI/CD

En esta sección se describen las mejoras realizadas en CI/CD en la versión 10.4.0.

CI/CD, o integración continua y entrega continua, es una práctica que automatiza las operaciones de integración y entrega en un canal de CI/CD. En la versión 10.4.0, puede incorporar las mejoras en su canal de CI/CD para mejorar la forma en que implementa, prueba y proporciona objetos en el entorno de producción.

API de REST del servicio de integración de datos

A partir de la versión 10.4.0, puede utilizar la API de REST del servicio de integración de datos para automatizar tareas en un canal de CI/CD.

Algunas de las tareas que puede automatizar la API de REST incluyen las siguientes:

Consultar objetos.

Consultar objetos, incluidos los objetos en tiempo de diseño en un repositorio de modelos y objetos en tiempo de ejecución que se implementan en un servicio de integración de datos.

Puede pasar la consulta a otras solicitudes de la API de REST. Por ejemplo, puede pasar una consulta a una operación de control de versiones para realizar un control de versiones en un conjunto de objetos específico. También puede pasar una consulta para implementar objetos en tiempo de diseño específicos en un archivo de almacenamiento de revisiones de aplicación.

Realizar operaciones de control de versiones.

Realice operaciones de control de versiones para proteger, desproteger, deshacer una desprotección o reasignar un objeto en tiempo de diseño desprotegido a otro desarrollador.

Administrar etiquetas.

Administre las etiquetas asignadas a objetos en tiempo de diseño. Puede asignar una nueva etiqueta o sustituir las etiquetas de un objeto. También puede eliminar las etiquetas de un objeto.

Actualizar aplicaciones.

Implemente objetos en tiempo de diseño en un archivo de almacenamiento de revisiones de aplicación e implemente el archivo en un servicio de integración de datos para actualizar una aplicación incremental implementada.

Administrar aplicaciones.

Inicie, anule la implementación o detenga una aplicación en función de los requisitos del proyecto o de la organización.

Ejecutar asignaciones implementadas.

Ejecute una asignación implementada para probar la salida de la aplicación.

Comparar asignaciones.

Compare dos asignaciones en el mismo dominio.

Por ejemplo, puede comparar dos asignaciones en tiempo de diseño o una asignación en tiempo de diseño y otra en tiempo de ejecución.

Para ver las solicitudes de API de REST que puede utilizar y los parámetros de cada solicitud, acceda a la documentación de API de REST a través de las propiedades de proceso del servicio de integración de datos o las propiedades del servicio del concentrador de operaciones REST en la Herramienta del administrador.

En comparación con los programas de línea de comandos infacmd, la API de REST no tiene ningún requisito de configuración y puede ejecutarla en entornos sin servicios de Informática instalados en el equipo cliente.

Para obtener información sobre la API de REST, consulte el capítulo "API de REST del servicio de integración de datos" en la *Guía del servicio de aplicación de Informática 10.4.0*.

Comandos infacmd dis

A partir de la versión 10.4.0, puede utilizar los comandos infacmd dis para realizar operaciones en objetos en tiempo de diseño, en objetos en tiempo de ejecución en una aplicación y en asignaciones implementadas.

También puede utilizar los comandos infacmd dis para ejecutar utilidades del servicio de integración de datos.

La siguiente tabla describe los nuevos comandos infacmd dis:

Comando	Descripción
compareMapping	Compara dos asignaciones consultadas. Consulte las asignaciones para comparar propiedades de asignación, propiedades de transformación y puertos en transformaciones. Para consultar asignaciones de tiempo de diseño, especifique el repositorio de modelos de tiempo de diseño. Para consultar asignaciones de tiempo de ejecución, no especifique un repositorio de modelos. La consulta utilizará el servicio de integración de datos que especifique para ejecutar el comando.
deployObjectsToFile	Implementa los objetos en tiempo de diseño para un archivo de almacenamiento de revisión de aplicación.
queryDesignTimeObjects	Consulta los objetos en tiempo de diseño de un servicio de repositorio de modelos.
queryRunTimeObjects	Consulta los objetos en tiempo de ejecución que están implementados en un servicio de integración de datos y devuelve una lista de objetos.
replaceAllTag	Reemplaza las etiquetas con las etiquetas especificadas en los objetos consultados en el servicio de repositorio de modelos.
tag	Asigna una etiqueta en los objetos consultados en el servicio de repositorio de modelos.
untag	Elimina una etiqueta en los objetos consultados en el servicio de repositorio de modelos.
listPatchNames	Enumera todas las revisiones que se han realizado a una aplicación incremental.

Para obtener más información, consulte el capítulo "Referencia del comando infacmd dis" en la *Referencia de comandos de Informatica® 10.4.0*.

Servidor proxy inverso

A partir de la versión 10.4.0, se puede utilizar el concentrador de operaciones REST para iniciar el servidor proxy inverso que realiza el equilibrio de carga para las solicitudes de API del servicio de integración de datos en los nodos de una cuadrícula del servicio de integración de datos.

El concentrador de operaciones REST está habilitado para cuadrículas.

Para obtener información acerca del servidor proxy inverso, consulte el capítulo "Servicios del sistema" en la Guía del servicio de aplicación de Informatica 10.4.0.

Comandos infacmd roh

A partir de la versión 10.4.0, puede utilizar los siguientes comandos infacmd roh para actualizar el proceso del servicio del concentrador de operaciones REST.

En la siguiente tabla se describen los nuevos comandos infacmd roh:

Comandos	Descripción
listReverseProxyServerOptions	Enumera las propiedades del servidor proxy inverso.
listServiceProcessOptions	Enumera las propiedades del proceso del servicio del concentrador de operaciones REST.
updateServiceOptions	Actualiza las propiedades del servicio del concentrador de operaciones REST.
updateReverseProxyServerOptions	Actualiza las propiedades del servidor proxy inverso en el dominio.

A partir de la versión 10.4.0, se ha cambiado el nombre de los siguientes comandos infacmd roh:

- listROHProperties a listProcessProperties.
- updateROHService a updateServiceProcessOptions.

Nota: Actualice todos los scripts que utilice con el nombre de comando anterior.

Para obtener más información, consulte el capítulo "Referencia del comando infacmd roh" en la *Referencia de comandos de Informatica 10.4.0*.

Implementación de revisiones de la aplicación

A partir de la versión 10.4.0, la implementación de revisiones de la aplicación tiene nuevas funciones en las siguientes áreas:

Asistente de implementación incremental

A partir de la versión 10.4.0, se pueden realizar las siguientes tareas en el asistente de implementación incremental:

- De forma opcional, puede introducir una descripción de la revisión.
- De forma opcional, puede optar por conservar o descartar la información de estado.
La información de estado hace referencia a las propiedades de la asignación y las propiedades de los objetos en tiempo de ejecución, tales como las salidas de asignación o la transformación de generador de secuencia.

Para obtener más información sobre el asistente de implementación incremental, consulte el capítulo "Implementación de revisiones de la aplicación" en la *Guía de Informatica 10.4.0 Developer Tool*.

Para obtener más información sobre la información de estado, consulte el capítulo "Implementación de aplicaciones" en la *Guía de Informatica 10.4.0 Developer Tool*.

Historial de revisiones

A partir de la versión 10.4.0, el historial de revisiones en el asistente de implementación incremental muestra tanto el nombre como la descripción de las revisiones que se implementaron para actualizar la aplicación incremental. La hora de creación de la revisión se anexa al principio de la descripción de la revisión.

Además, puede utilizar la Herramienta del administrador para ver el historial de revisiones de una aplicación incremental implementada.

Para obtener más información sobre el historial de revisiones, consulte el capítulo "Implementación de revisiones de la aplicación" en la *Guía de Informatica 10.4.0 Developer Tool*.

Para obtener más información sobre las aplicaciones implementadas, consulte el capítulo "Aplicaciones del servicio de integración de datos" en la *Guía del servicio de aplicación de Informatica 10.4.0*.

Archivos de almacenamiento de revisiones de aplicación

A partir de la versión 10.4.0, los archivos de almacenamiento de revisiones de aplicación también almacenan la descripción de la revisión e indican si la información de estado se conserva.

Para obtener más información sobre el historial de revisiones, consulte el capítulo "Implementación de revisiones de la aplicación" en la *Guía de Informatica 10.4.0 Developer Tool*.

Comandos infacmd tools

A partir de la versión 10.4.0, el comando `infacmd tools patchApplication` tiene la siguiente opción nueva:

Opción	Argumento	Descripción
<code>-RetainStateInformation</code> <code>-rsi</code>	True False	Opcional. Indica si la información de estado se conserva o se descarta. La información de estado hace referencia a las propiedades de la asignación y las propiedades de los objetos en tiempo de ejecución, tales como las salidas de asignación o la transformación de generador de secuencia.

Para obtener más información, consulte el capítulo "Referencia del comando `infacmd tools`" en la *Referencia de comandos de Informatica 10.4.0*.

Conexión a una aplicación en tiempo de ejecución

A partir de la versión 10.4.0, puede conectarse a una aplicación en tiempo de ejecución en Developer tool. Después de conectarse a una aplicación en tiempo de ejecución, puede expandir la aplicación en la vista **Explorador de objetos** y abrir copias de solo lectura de los objetos en tiempo de ejecución en el **Editor**.

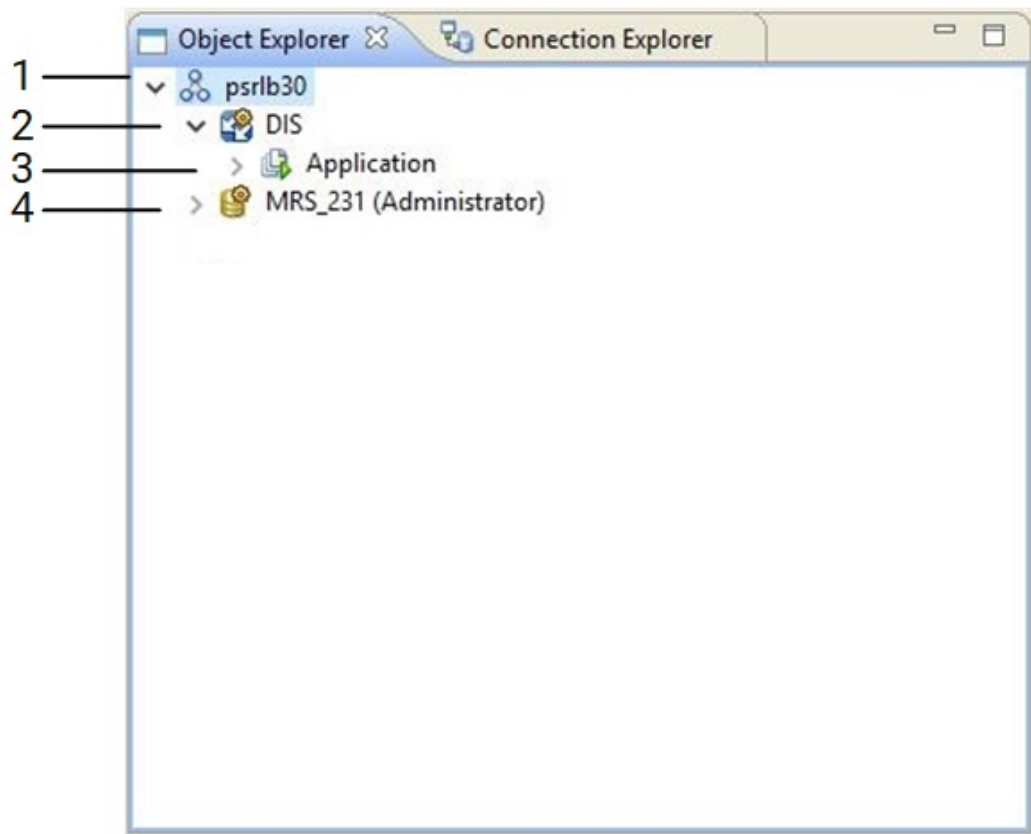
Después de conectarse a una aplicación en tiempo de ejecución, las búsquedas que realice en Developer tool pueden ubicar objetos en tiempo de ejecución en la aplicación.

Para obtener más información sobre la conexión a una aplicación en tiempo de ejecución y la visualización de los objetos en tiempo de ejecución, consulte el capítulo "Implementación de aplicaciones" en la *Guía de Informatica 10.4.0 Developer Tool*.

Vista Explorador de objetos

A partir de la versión 10.4.0, la vista **Explorador de objetos** de Developer tool muestra el dominio después de que se conecte a un repositorio de modelos o a una aplicación en tiempo de ejecución. Puede expandir el dominio para ver los objetos en tiempo de diseño en el repositorio de modelos o los objetos en tiempo de ejecución en la aplicación en tiempo de ejecución.

La siguiente imagen muestra la vista **Explorador de objetos**:



1. Dominio
2. Servicio de integración de datos
3. Aplicación en tiempo de ejecución
4. Repositorio de modelos

Para obtener más información sobre la interfaz de usuario de Developer tool, consulte el capítulo "Informatica Developer" en la *Guía de Informatica 10.4.0 Developer Tool*.

Etiquetas

A partir de la versión 10.4.0, las etiquetas tienen la siguiente funcionalidad:

- Cuando implementa una asignación que está asociada a una etiqueta, la etiqueta se propaga a la versión en tiempo de ejecución de la asignación en el servicio de integración de datos.
- Si actualiza la asignación implementada mediante una revisión de la aplicación, el nombre de la revisión se asocia como una etiqueta a la versión en tiempo de ejecución de la asignación.

Para obtener más información, consulte el capítulo "Informatica Developer" de la *Guía de Informatica 10.4.0 Developer Tool*.

Programas de la línea de comandos

En esta sección se describen los nuevos comandos de la versión 10.4.0.

Comandos infacmd isp (nuevas características de la versión 10.4.0)

En la siguiente tabla se describen los nuevos comandos infacmd isp:

Comando	Descripción
addCustomLDAPType	Agrega un tipo LDAP personalizado que define un servicio de directorio de LDAP.
listAllCustomLDAPTypes	Enumera la información de configuración de todos los tipos LDAP personalizados utilizados por el dominio especificado.
listAllLDAPConnectivity	Enumera la información de configuración de todas las configuraciones de LDAP utilizadas por el dominio especificado.
listCustomLDAPType	Enumera la información de configuración de un tipo LDAP personalizado.
removeCustomLDAPType	Elimina el tipo LDAP personalizado especificado del dominio especificado.
removeLDAPConnectivity	Elimina la configuración de LDAP especificada del dominio especificado.
updateCustomLDAPType	Actualiza el tipo LDAP personalizado especificado.
updateLDAPConnectivity	Actualiza la configuración de LDAP especificada.

Para obtener más información, consulte la *Referencia de comandos de Informatica 10.4.0*.

Data Engineering Integration

En esta sección se describen las nuevas características de Data Engineering Integration en la versión 10.4.0.

Compatibilidad con nuevos tipos de datos

A partir de la versión 10.4.0, puede utilizar los siguientes tipos de datos nuevos para los archivos complejos:

- Cuando ejecuta una asignación que lee o escribe en objetos de archivos complejos de Avro y Parquet en el entorno nativo o en el entorno Hadoop, puede utilizar los siguientes tipos de datos:
 - Date
 - Decimal
 - Timestamp
- Puede utilizar el tipo de datos Time para leer y escribir en objetos de archivos complejos de Avro o Parquet en el entorno nativo o en el motor de Blaze.
- Puede utilizar los tipos de datos Date, Time, Timestamp y Decimal cuando ejecuta una asignación en el motor de Databricks Spark.

Los nuevos tipos de datos se aplican a los siguientes adaptadores:

- PowerExchange for HDFS
- PowerExchange for Amazon S3

- PowerExchange for Google Cloud Storage
- PowerExchange for Microsoft Azure Blob Storage
- PowerExchange for Microsoft Azure Data Lake Storage Gen1
- PowerExchange for Microsoft Azure Data Lake Storage Gen2

Para obtener más información sobre los tipos de datos, consulte el capítulo "Referencia de tipos de datos" en la *Guía del usuario de Data Engineering Integration 10.4.0*.

Integración de AWS Databricks

A partir de la versión 10.4.0, se puede integrar el dominio de Informatica con Databricks on AWS.

Puede utilizar AWS Databricks para ejecutar asignaciones con la siguiente funcionalidad:

- Puede ejecutar asignaciones en orígenes y destinos de Amazon Simple Storage Service (S3) y Amazon Redshift dentro del entorno de Databricks.
- Puede desarrollar flujos de trabajo de clúster para crear clústeres efímeros mediante Databricks on AWS.
- Puede agregar la transformación de Python a una asignación configurada para ejecutarse en el motor de Databricks Spark.

La transformación de Python solo es compatible para la vista previa técnica.

AWS Databricks admite los mismos tipos de datos que Azure Databricks.

Para obtener más información, consulte las siguientes guías:

Guía de integración de Data Engineering 10.4.0

Guía del administrador de Data Engineering 10.4.0

Guía del usuario de Data Engineering Integration 10.4.0

Guía de flujo de trabajo de Informatica 10.4.0 Developer

Flujos de trabajo de clúster para el acceso de HDInsight a recursos de ALDS Gen2

A partir de la versión 10.4.0, puede crear un flujo de trabajo de clúster para ejecutar en un clúster de Azure HDInsight y acceder a recursos de ADLS Gen2.

Para obtener más información sobre los flujos de trabajo de clúster, consulte la *Guía del usuario de Informatica Data Engineering Integration 10.4.0*.

Acceso al almacenamiento de Databricks Delta Lake

A partir de la versión 10.4.0, puede acceder al almacenamiento de Databricks Delta Lake como orígenes y destinos.

Las asignaciones pueden acceder a los recursos de Delta Lake en las plataformas AWS y Azure.

Para obtener información sobre la configuración del acceso a las tablas de Delta Lake, consulte la *Data Engineering Integration Guide*. Para obtener información sobre la creación de asignaciones para acceder a tablas de Delta Lake, consulte la *Data Engineering Integration User Guide*.

Visualización de los nodos utilizados en una asignación

A partir de la versión 10.4.0, puede ver el número máximo de nodos de clúster utilizados por una asignación en un tiempo determinado.

Puede utilizar la API del concentrador de operaciones REST `ClusterStats(startTimeInmillis=[value], endTimeInmillis=[value])` para ver el número máximo de nodos de Hadoop en una configuración de clústeres utilizada por una asignación en un tiempo determinado.

Para obtener más información sobre la API de REST, consulte el capítulo "Referencia de la API de REST de supervisión" en la *Guía del administrador de Data Engineering 10.4.0*.

Agregación de registros

A partir de la versión 10.4.0, se pueden obtener registros agregados para las asignaciones implementadas que se ejecutan en el entorno Hadoop.

Puede recopilar los registros de clústeres agregados de una asignación basándose en el ID del trabajo en la Herramienta de supervisión o mediante el comando `infacmd ms fetchAggregatedClusterLogs`. Puede obtener un archivo .zip o tar.gz de los registros de clústeres agregados de una asignación basándose en el ID del trabajo y escribir el archivo de registros agregados en un directorio de destino.

Para obtener más información, consulte la *Guía del administrador de Informatica 10.4.0*.

Análisis de datos jerárquicos en el motor Spark

A partir de la versión 10.4.0, se pueden utilizar funciones complejas para analizar hasta 5 MB de datos intermedios en una asignación.

El motor Spark puede analizar datos de origen de cadenas sin formato mediante las siguientes funciones complejas:

- PARSE_JSON
- PARSE_XML

Las funciones complejas analizan los datos JSON o XML en la cadena de origen y generan datos de destino estructurados.

Para obtener más información, consulte el capítulo "Procesamiento de datos jerárquicos" en la *Guía del usuario de Informatica Data Engineering Integration 10.4.0*.

Para obtener más información acerca de las funciones complejas, consulte el capítulo "Funciones" de la *Referencia del lenguaje de transformación para Informatica 10.4.0 Developer*.

Creación de perfiles y opciones de muestreo en el motor Spark

A partir de la versión 10.4.0, puede ejecutar perfiles y seleccionar opciones de muestreo en el motor Spark.

Creación de perfiles en el motor Spark

Puede crear y ejecutar perfiles en el motor Spark en las herramientas Informatica Developer e Informatica Analyst. Puede realizar la detección de dominios de datos y la creación de cuadros de mandos en el motor Spark.

Opciones de muestreo en el motor Spark

Puede seleccionar las siguientes opciones de muestreo para ejecutar perfiles en el motor Spark:

- La opción de muestreo **Limit n** ejecuta un perfil basado en el número de filas en el objeto de datos. Cuando decide ejecutar un perfil en el entorno Hadoop, el motor Spark recopila muestras de varias particiones del objeto de datos y las envía a un único nodo para calcular el tamaño de la muestra. No puede aplicar las opciones de muestreo Limit n en los perfiles con un filtro avanzado. Admitido en bases de datos Oracle, SQL Server y DB2.
- La opción de muestreo **Random percentage** ejecuta un perfil en un porcentaje de las filas del objeto de datos.

Para obtener información sobre los perfiles y las opciones de muestreo en el motor Spark, consulte la *Guía de Informatica 10.4.0 Data Discovery*.

Transformación de Python

A partir de la versión 10.4.0, la transformación de Python tiene la siguiente funcionalidad:

Modo activo

Puede crear una transformación de Python activa. Como transformación activa, la transformación de Python puede cambiar el número de filas que pasan por ella. Por ejemplo, la transformación de Python puede generar múltiples filas de salida a partir de una única fila de entrada, o la transformación puede generar una única fila de salida a partir de múltiples filas de entrada.

Para obtener más información, consulte el capítulo "Transformación de Python" en la *Guía del usuario de Informatica Data Engineering Integration 10.4.0*.

Datos particionados

Puede ejecutar código Python para procesar datos entrantes basándose en el esquema de particionamiento predeterminado de los datos, o puede volver a particionar los datos antes de ejecutar el código Python. Para volver a particionar los datos antes de ejecutar el código Python, seleccione uno o varios puertos de entrada como clave de la partición.

Para obtener más información, consulte el capítulo "Transformación de Python" en la *Guía del usuario de Informatica Data Engineering Integration 10.4.0*.

Sqoop

A partir de la versión 10.4.0, puede configurar los siguientes argumentos de Sqoop en la conexión JDBC:

- --update-key
- --update-mode
- --validate
- --validation-failurehandler
- --validation-threshold
- --validator
- --mapreduce-job-name
- --bindir
- --class-name
- --jar-file
- --outdir

- --package-name

Para obtener más información sobre cómo configurar estos argumentos de Sqoop, consulte la documentación de Sqoop.

Data Engineering Streaming

En esta sección se describen las nuevas características de Data Engineering Streaming en la versión 10.4.0.

Confluent Schema Registry en asignaciones de transmisión

A partir de la versión 10.4.0, se puede utilizar Confluent Kafka como orígenes y destinos en las asignaciones de transmisión que utilizan Schema Registry.

Puede utilizar Confluent Kafka para almacenar y recuperar esquemas de Apache Avro en asignaciones de transmisión. Schema Registry utiliza Kafka como mecanismo de almacenamiento subyacente.

Para obtener más información, consulte la *Guía del usuario de Data Engineering Streaming 10.4.0*.

Transformaciones de Data Quality en asignaciones de transmisión

A partir de la versión 10.4.0, se pueden utilizar transformaciones de Data Quality en las asignaciones de transmisión.

Puede utilizar las siguientes transformaciones de Data Quality en las asignaciones de transmisión para aplicar el proceso de calidad de datos en los datos de transmisión:

- Transformación del validador de direcciones
- transformación de clasificador
- Transformación de analizador
- Transformación de estandarizador

Para obtener más información, consulte la *Guía del usuario de Data Engineering Streaming 10.4.0*.

Clúster efímero en asignaciones de transmisión

A partir de la versión 10.4.0, se puede ejecutar un flujo de trabajo para crear un clúster efímero que ejecute asignaciones y otras tareas en un clúster de plataforma en la nube.

Para reanudar el procesamiento de los datos desde el punto en que se elimina un clúster, puede ejecutar asignaciones de transmisión en un clúster efímero especificando un almacenamiento externo y un directorio de puntos de comprobación.

Para obtener más información, consulte la *Guía del usuario de Data Engineering Streaming 10.4.0*.

Puerto FileName en Amazon S3

A partir de la versión 10.4.0, cuando se crea una operación de escritura en objeto de datos para archivos de Amazon S3, el puerto FileName aparece de forma predeterminada.

En tiempo de ejecución, el servicio de integración de datos crea directorios independientes para cada valor del puerto FileName y agrega los archivos de destino en los directorios.

Para obtener más información, consulte la *Guía del usuario de Data Engineering Streaming 10.4.0*.

Microsoft Azure Data Lake Storage Gen2

A partir de la versión 10.4.0, se puede utilizar Microsoft Azure Data Lake Storage Gen2 como destino en las asignaciones de transmisión.

Azure Data Lake Storage Gen2 está basado en Azure Blob Storage. Azure Data Lake Storage Gen2 tiene las capacidades tanto de Azure Data Lake Storage Gen1 como de Azure Blob Storage. Puede utilizar Azure Databricks 5.4 o Azure HDInsight 4.0 para acceder a los datos almacenados en Azure Data Lake Storage Gen2.

Para obtener más información, consulte la *Guía del usuario de Data Engineering Streaming 10.4.0*.

Asignaciones de transmisión en Azure Databricks

A partir de la versión 10.4.0, puede ejecutar asignaciones de transmisión en el servicio Azure Databricks en los servicios en la nube de Microsoft Azure.

Orígenes y destinos

Puede ejecutar asignaciones de transmisión en los siguientes orígenes y destinos dentro del entorno de Databricks:

- Microsoft Azure Event Hubs

- Azure Data Lake Storage Gen2 (ADLS Gen2)

Transformaciones

Puede agregar las siguientes transformaciones a una asignación de transmisión de Databricks:

- Agregación

- Expresión

- Filtro

- Unión

- Normalizador

- Rango

- Enrutador

- Unión

- Ventana

Tipos de datos

Se admiten los siguientes tipos de datos:

- Matriz

- Bigint

Fecha/Hora
Decimal
Doble
Entero
Asignar
Struct
Texto
Cadena

Flujos de trabajo

Puede desarrollar flujos de trabajo de clúster para crear clústeres efímeros en el entorno de Databricks. Utilice Azure Data Lake Storage Gen1 (ADLS Gen1) y Azure Data Lake Storage Gen2 (ADLS Gen2) para crear clústeres efímeros en el entorno de Databricks.

Para obtener más información sobre las asignaciones de transmisión en Azure Databricks, consulte la *Guía del usuario de Data Engineering Streaming 10.4.0*.

Asignaciones dinámicas en Data Engineering Streaming

A partir de la versión 10.4.0, la compatibilidad con las asignaciones dinámicas en Data Engineering Streaming está disponible para la vista previa técnica.

Puede usar objetos de datos de Confluent Kafka como orígenes y destinos dinámicos en una asignación de transmisión.

La funcionalidad de vista previa técnica se puede utilizar con fines de evaluación, pero no se garantiza ni está preparada para usarla en producción. Informatica recomienda utilizarla solo en entornos que no sean de producción. Informatica tiene la intención de incluir la funcionalidad de vista previa en una próxima versión para poder usarla en producción, pero podría decidir no hacerlo en función de cuáles sean los requisitos del mercado o las circunstancias técnicas en ese momento. Para obtener más información, póngase en contacto con el servicio internacional de atención al cliente de Informatica.

Enterprise Data Catalog

En esta sección se describen las nuevas características de Enterprise Data Catalog en la versión 10.4.0.

Asignación de atributos personalizados a recursos y clases

A partir de la versión 10.4.0, se pueden asignar atributos personalizados a recursos y clases específicos del catálogo o a todos los recursos y clases del catálogo. Puede realizar esta tarea cuando cree o modifique un atributo personalizado.

Para obtener más información, consulte la *Guía del administrador de Informatica 10.4.0 Catalog*.

Nuevos recursos

A partir de la versión 10.4.0, los siguientes recursos nuevos se añaden a Enterprise Data Catalog:

- AWS Glue
- Microsoft Power BI
- Apache Cassandra
- Snowflake
- Google Cloud Storage

Puede extraer información de metadatos, relaciones y linaje de todos estos recursos. Para obtener más información, consulte la *Guía de configuración del escáner de Informatica 10.4.0 Enterprise Data Catalog*.

Recursos de referencia y activos de referencia

A partir de la versión 10.4.0, puede configurar recursos para extraer metadatos de orígenes de datos u otros recursos del catálogo al que hace referencia el recurso. Algunos ejemplos son las tablas de origen y de destino en las asignaciones de PowerCenter, y las tablas y los archivos de origen de los informes de Tableau. Estos orígenes de datos referenciados se conocen como "recursos de referencia" y los activos incluidos en los orígenes de datos, como "activos de referencia". Puede ver los recursos y los activos de referencia en las aplicaciones de Enterprise Data Catalog. Para ver el linaje completo en Enterprise Data Catalog, puede realizar asignaciones de conexión para los recursos de referencia. Puede realizar asignaciones de conexión entre los recursos de referencia y los recursos que configure en Enterprise Data Catalog para los orígenes de datos de los recursos de referencia.

Puede configurar los siguientes recursos para extraer metadatos de orígenes de datos u otros recursos del catálogo al que hace referencia el recurso:

- PowerCenter
- AWS Glue
- Tableau Server
- Coudera Navigator
- Apache Atlas
- Informatica Intelligent Cloud Services
- Informatica Platform
- SQL Server Integration Service

Para obtener más información, consulte la *Guía del administrador de Informatica 10.4.0 Catalog* y la *Guía del usuario de Informatica 10.4.0 Enterprise Data Catalog*.

Exportación de activos desde resultados de búsqueda

A partir de la versión 10.4.0, se pueden exportar todos los activos que aparecen en los resultados de búsqueda a un archivo de valores separados por comas (CSV) y volver a importar el archivo CSV en Enterprise Data Catalog. Exporte los activos después de refinar los resultados de búsqueda en cualquier ficha de búsqueda.

Para obtener más información, consulte el capítulo sobre *tareas de activos* en la *Guía del usuario de Informatica 10.4.0 Enterprise Data Catalog*.

Filtros de linaje e impacto

A partir de la versión 10.4.0, puede crear y aplicar filtros a la vista Linaje e impacto. Utilice los filtros para restringir los resultados en la vista Linaje e impacto. Cree filtros personalizados utilizando distintas combinaciones de opciones de filtro basadas en los tipos de activos disponibles en la vista Linaje e impacto.

Para obtener más información, consulte el capítulo sobre la *visualización del linaje y su impacto* en la *Guía del usuario de Informatica 10.4.0 Enterprise Data Catalog*.

Resumen de control de activos

A partir de la versión 10.4.0, puede ver activos del flujo de control en la ficha **Resumen de control de activos** para los tipos de activos de tabla y archivo. Un flujo de control coloca restricciones o condiciones en el flujo de datos. Por ejemplo, una consulta SQL puede contener restricciones en la cláusula WHERE o las asignaciones pueden contener transformaciones.

Para obtener más información, consulte el capítulo sobre la *visualización del linaje y su impacto* en la *Guía del usuario de Informatica 10.4.0 Enterprise Data Catalog*.

Reglas y cuadros de mandos

A partir de la versión 10.4.0, se pueden extraer resultados de reglas y cuadros de mandos del almacén de creación de perfiles y verlos en Enterprise Data Catalog. Puede ver los resultados de columna de una regla y un cuadro de mando para medir la calidad de los datos de un origen de datos.

Para obtener más información, consulte el capítulo sobre *visualización de activos* en la *Guía del usuario de Informatica 10.4.0 Enterprise Data Catalog*.

Inferencia de clave única

A partir de la versión 10.4.0, se pueden generar candidatos de clave única a partir de las columnas en el objeto de datos de origen. Puede ver los resultados de la inferencia de clave única para comprender los problemas de calidad de los datos identificando las columnas que tienen valores duplicados.

Puede aceptar o rechazar los resultados de inferencia de claves únicas deducidas. Después de aceptar o rechazar una inferencia de clave única deducida, puede restablecer la inferencia de clave única para restaurar el estado deducido.

Para obtener más información, consulte el capítulo sobre *visualización de activos* en la *Guía del usuario de Informatica 10.4.0 Enterprise Data Catalog*.

Detección del dominio de datos en el tipo de archivo CLOB

A partir de la versión 10.4.0, se puede realizar la detección del dominio de datos en el tipo de archivo CLOB.

Para obtener más información, consulte el capítulo *Conceptos de Enterprise Data Catalog* en la *Guía del administrador de Informatica 10.4.0 Enterprise Catalog*.

Obtención de datos y opciones de muestreo en el motor Spark

A partir de la versión 10.4.0, puede ejecutar perfiles para detectar dominios de datos y seleccionar opciones de muestreo en el motor Spark.

Detección de dominios de datos en el motor Spark

Puede realizar la detección de dominios de datos en el motor Spark.

Opciones de muestreo en el motor Spark

Puede seleccionar las siguientes opciones de muestreo para detectar dominios de datos en el motor Spark:

- La opción de muestreo **Limit n** ejecuta un perfil basado en el número de filas en el objeto de datos. Cuando decide detectar dominios de datos en el entorno Hadoop, el motor Spark recopila muestras de varias particiones del objeto de datos y las envía a un único nodo para calcular el tamaño de la muestra.
- La opción de muestreo **Random percentage** ejecuta un perfil en un porcentaje de las filas del objeto de datos.

Para obtener más información, consulte el capítulo *Conceptos de Enterprise Data Catalog* en la *Guía del administrador de Informatica 10.4.0 Enterprise Catalog*.

Seguimiento de la vista previa técnica

La versión 10.4.0 de Enterprise Data Catalog incluye funcionalidad que está disponible para la vista previa técnica.

La funcionalidad de vista previa técnica es compatible, pero no se garantiza ni está preparada para usarla en producción. Informatica recomienda utilizarla solo en entornos que no sean de producción. Informatica tiene la intención de incluir la funcionalidad de vista previa en una próxima versión de disponibilidad general (GA) para poder usarla en producción, pero podría decidir no hacerlo en función de cuáles sean los requisitos del mercado o las circunstancias técnicas en ese momento. Para obtener más información, póngase en contacto con el servicio internacional de atención al cliente de Informatica.

- A partir de la versión 10.4.0, puede optar por mostrar la vista compacta de la vista Linaje e impacto. La vista Linaje e impacto compacta muestra el diagrama de linaje e impacto resumido en el nivel de recursos.

Para obtener más información, consulte el capítulo sobre la *visualización del linaje y su impacto* en la *Guía del usuario de Informatica 10.4.0 Enterprise Data Catalog*.

- A partir de la versión 10.4.0, puede extraer metadatos de orígenes de SAP Business Warehouse, SAP BW/4HANA, IBM InfoSphere DataStage y Oracle Data Integrator cuando no se puede acceder a ellos en tiempo de ejecución o sin conexión.

Para obtener más información, consulte la *Guía del administrador de Informatica 10.4.0 Catalog*.

- A partir de la versión 10.4.0, puede extraer metadatos de los orígenes de datos de SAP Business Warehouse y SAP BW/4HANA.

Para obtener más información, consulte la *Guía de configuración del escáner de Informatica 10.4.0 Enterprise Data Catalog*.

Vista previa de datos y aprovisionamiento

A partir de la versión 10.4.0, se puede realizar el aprovisionamiento de datos después de completar la obtención de datos en el catálogo. Gracias al aprovisionamiento de datos, podrá mover los datos a un destino para realizar otros análisis. Puede obtener una vista previa de los datos para que los orígenes relacionales evalúen los datos antes de transferirlos al destino.

Para obtener más información sobre la obtención de una vista previa y el aprovisionamiento de los datos, consulte la *Guía del administrador de Informatica 10.4.0 Catalog* y la *Guía del usuario de Informatica 10.4.0 Enterprise Data Catalog*.

Tipos de recursos compatibles en la utilidad de escáner independiente

A partir de la versión 10.4, se pueden extraer metadatos de los siguientes orígenes externos cuando no se pueda acceder a ellos en tiempo de ejecución o cuando estén desconectados:

- Amazon Redshift
- Amazon S3
- Apache Cassandra
- Axon
- Azure Data Lake Store
- Azure Microsoft SQL Data Warehouse
- Azure Microsoft SQL Server
- Business Glossary
- Linaje personalizado
- Scripts de la base de datos
- Erwin
- Glue
- Google BigQuery
- Google Cloud Storage
- Informatica Cloud Service
- IBM Cognos
- Microsoft Azure Blob Storage
- Microsoft SQL Server Integration Services
- PowerBI
- QlikView Business Intelligence
- Salesforce
- SAP HANA
- SAP PowerDesigner
- Snowflake
- Workday

Para obtener más información, consulte el capítulo sobre extracción de metadatos de recursos inaccesibles o sin conexión en la *Guía del administrador de Informatica Enterprise Data Catalog 10.4*.

API de REST

A partir de la versión 10.4, puede usar las siguientes API de REST de Informatica Enterprise Data Catalog:

- API de REST de provisión de datos. Además de las API de REST existentes, puede ver si el aprovisionamiento de datos está disponible para el usuario y enumerar los recursos que admiten aprovisionamiento de datos.
- API de REST de filtro de linaje. Puede crear, actualizar, enumerar o eliminar un filtro de linaje.

- API de REST de información de modelo. Además de las API de REST existentes, puede enumerar las facetas predefinidas del control deslizante, la definición de la faceta del control deslizante y las definiciones del filtro de linaje.
- API de REST de modificación de modelo. Además de las API de REST existentes, puede crear, actualizar y eliminar una definición de faceta del control deslizante.
- API de REST de información de supervisión. Puede enviar o enumerar trabajos que incluyan trabajos de tipo de exportación de objeto, tipo de importación de objeto, tipo de exportación de recurso y tipo de exportación de búsqueda.
- API de REST de recuento de activos secundarios de un objeto. Puede enumerar el total de activos secundarios de un objeto.
- API de REST de información del producto. Puede enumerar los detalles de Enterprise Data Catalog, que incluyen la versión de lanzamiento, la versión de compilación y la fecha de compilación.

Para obtener más información sobre las API de REST, consulte la *Referencia de las API de REST de Informatica Enterprise Data Catalog 10.4*.

Enterprise Data Preparation

En esta sección se describen las nuevas características de Enterprise Data Preparation en la versión 10.4.0.

Administración de acceso a Data Lake

A partir de la versión 10.4.0, puede administrar el acceso a Data Lake mediante la asignación de permisos en los recursos de Enterprise Data Catalog que representan esquemas de Hive y ubicaciones de HDFS en Data Lake a los usuarios y grupos de usuarios de Enterprise Data Preparation.

Cuando se conceden permisos en esquemas o ubicaciones específicos a un usuario o grupo de usuarios, la aplicación solo muestra aquellos esquemas o ubicaciones en los que el usuario tiene permiso para realizar operaciones de importación, publicación o carga.

Para obtener más información, consulte la *Guía del administrador de Enterprise Data Preparation 10.4.0*.

Microsoft Azure Data Lake Storage como origen de datos

A partir de la versión 10.4.0, se puede utilizar una ubicación de Microsoft Azure Data Lake Storage Gen2 como origen de datos.

Publicación de archivos en Data Lake

A partir de la versión 10.4.0, puede publicar datos preparados en un archivo en una ubicación de HDFS en Data Lake.

Cuando publica datos, puede seleccionar el tipo de archivo en el que escribir los datos en Data Lake. Por ejemplo, si decide publicar datos como un archivo de valores separados por comas, la aplicación escribe los datos en Data Lake como un archivo .csv.

Para obtener más información, consulte la *Guía del usuario de Enterprise Data Preparation 10.4.0*.

Carga de archivos en Data Lake

A partir de la versión 10.4.0, se pueden cargar varios formatos de archivo desde la unidad local a Data Lake.

Puede elegir una de las siguientes opciones de carga:

Cargar un archivo directamente en Data Lake.

Puede cargar en Data Lake un archivo delimitado por comas, un archivo Avro, un archivo JSON o un archivo Parquet en formato UTF-8 directamente desde la unidad local, sin obtener una vista previa de los datos. Podría optar por esta opción si desea cargar un archivo sin obtener una vista previa de los datos.

Permitir a CLAIRE determinar la estructura del archivo y luego cargar el archivo en Data Lake.

Puede cargar los datos en un archivo delimitado por comas o una hoja de cálculo de Microsoft Excel en Data Lake. Cuando se carga el archivo, Enterprise Data Preparation utiliza el motor de detección integrado de CLAIRE para determinar la estructura del archivo y muestra una vista previa de los datos.

Cuando utiliza esta opción para cargar una hoja de cálculo de Excel, el motor de CLAIRE detecta las hojas y las tablas en la hoja de cálculo. Puede seleccionar la hoja y la tabla de las que quiere obtener una vista previa.

Nota: Cargar una hoja de cálculo de Excel es una característica de vista previa.

Definir la estructura del archivo y luego cargar el archivo en Data Lake.

Puede cargar los datos en un archivo delimitado por comas de la unidad local en Data Lake. Cuando se carga el archivo, puede obtener una vista previa de los datos, especificar la estructura del archivo y configurar atributos de columna que cumplan sus requisitos. Puede optar por esta opción si necesita modificar atributos de columna antes de cargar el archivo.

Para obtener más información, consulte la *Guía del usuario de Enterprise Data Preparation 10.4.0*.

Asignaciones de Informatica

En esta sección se describen las nuevas características para las asignaciones de Informatica en la versión 10.4.0.

Vinculación de salidas de asignación con parámetros de asignación

A partir de la versión 10.4.0, puede conservar y enlazar salidas de asignación en asignaciones que estén implementadas para ejecutarse en el entorno nativo o en el motor Spark.

Cree una salida de asignación. Enlace la salida con un parámetro de asignación para utilizar el valor en ejecuciones subsiguientes de la asignación. Cuando ejecute la asignación, el servicio de integración de datos pasará el valor de la salida de asignación al parámetro de asignación. Para conservar salidas de asignación, debe especificar un nombre de instancia en tiempo de ejecución mediante la opción `-RuntimeInstanceName` para el comando `infacmd ms runMapping`.

Developer tool incluye ahora una columna **Enlace** en la vista **Propiedades** de las asignaciones para enlazar una salida de asignación con un parámetro.

Para obtener información sobre las salidas de asignación en asignaciones implementadas, consulte el capítulo "Salidas de asignación" en la *Guía de asignaciones de Informatica 10.4.0 Developer*.

Comandos infacmd ms

En la siguiente tabla se describen los nuevos y actualizados comandos infacmd ms:

Comando	Descripción
deleteMappingPersistedOutputs	Nuevo comando que elimina todas las salidas de asignación persistentes de una asignación implementada. Especifique las salidas que se deban eliminar con el nombre de la aplicación y el nombre de instancia en tiempo de ejecución de la asignación. Para eliminar salidas específicas, use la opción - OutputNamesToDelete.
getMappingStatus	Comando actualizado que ahora devuelve el nombre del trabajo. Si ha definido un nombre de instancia en tiempo de ejecución en infacmd ms runMapping, el nombre del trabajo es el nombre de la instancia en tiempo de ejecución.
listMappingPersistedOutputs	Nuevo comando que enumera las salidas de asignación persistentes de una asignación implementada. Las salidas se enumeran de acuerdo con el nombre de la aplicación y del nombre de instancia en tiempo de ejecución de la asignación.

Para obtener más información, consulte el capítulo "Referencia del comando infacmed ms" de la *Referencia de comandos de Informatica 10.4.0*.

Recomendaciones y conclusiones de CLAIRE

A partir de la versión 10.4.0, puede utilizar la inteligencia artificial de CLAIRE para ajustar y corregir las asignaciones durante el desarrollo.

Cuando habilita las recomendaciones, CLAIRE se ejecuta automáticamente en las asignaciones conforme las desarrolla, y muestra recomendaciones que le permiten corregir o ajustar la asignación.

También puede ejecutar análisis de CLAIRE en las asignaciones de un proyecto o en la carpeta del proyecto. Cuando analiza un grupo de asignaciones, CLAIRE muestra conclusiones acerca de las similitudes entre las asignaciones.

Para obtener más información sobre las recomendaciones y las conclusiones, consulte la *Data Engineering Integration User Guide*.

Actualización del nivel del optimizador de asignaciones

A partir de la versión 10.4.0, se puede utilizar el comando infacmd ms UpdateOptimizationDefaultLevel para establecer el nivel del optimizador de las asignaciones en una aplicación en el nivel predeterminado: Auto.

Cuando ejecute el comando, debe especificar un nombre de aplicación. UpdateOptimizationDefaultLevel establece el nivel del optimizador para todas las asignaciones en la aplicación.

Para obtener más información, consulte la *Referencia de comandos de Informatica 10.4.0* y la *Guía de asignaciones de Informatica 10.4.0 Developer*.

Transformaciones de Informatica

En esta sección se describen las nuevas características de transformación de Informatica en la versión 10.4.0.

Transformación del validador de direcciones

En esta sección se describen las nuevas características de la transformación del validador de direcciones.

La transformación del validador de direcciones contiene funciones de dirección adicionales para los siguientes países:

Estados Unidos

A partir de la versión 10.4, el validador de direcciones reconoce MC como una versión alternativa de MSC (Mail Stop Code) en las direcciones de Estados Unidos.

Para obtener información completa acerca de las características y las operaciones de la versión del motor de software de verificación de direcciones de la versión 10.4, consulte la *Guía de desarrollador de Informatica Address Verification 5.15.0*.

Flujos de trabajo de Informatica

En esta sección se describen las nuevas características de los flujos de trabajo de Informatica en la versión 10.4.0.

Propiedades avanzadas de la tarea Crear clúster en Amazon EMR

A partir de la versión 10.4.0, hay nuevas propiedades avanzadas disponibles cuando se configura una tarea Crear clúster con una conexión Amazon EMR en un flujo de trabajo.

Puede configurar la siguiente funcionalidad en una tarea Crear clúster:

- Tamaño de volumen de EBS de dispositivo raíz. El número de GB del volumen del dispositivo raíz de EBS.
- ID de AMI personalizada. ID de una imagen de máquina de Amazon (Amazon Machine Image, AMI) personalizada de Amazon Linux.
- Configuración de seguridad. El nombre de una configuración de seguridad para la autenticación y el cifrado en el clúster.

Para obtener más información, consulte la *Guía del usuario de Data Engineering Integration 10.4.0* y la *Guía de flujo de trabajo de Informatica® 10.4.0 Developer*.

Instalación de Informatica

En esta sección se describen las nuevas características de instalación de la versión 10.4.0.

PostgreSQL

A partir de la versión 10.4.0, se puede utilizar la base de datos de PostgreSQL para el repositorio de configuración del dominio, el repositorio de modelos y el repositorio de PowerCenter. Para Enterprise Data Preparation, solo puede utilizar la base de datos de PostgreSQL en el servicio de repositorio de modelos adicional.

También puede instalar la versión 10.6 de la aplicación cliente psql para que PostgreSQL funcione en Linux o en Windows.

Para obtener más información sobre PostgreSQL, consulte las guías de instalación de *Informatica 10.4.0*.

Pre-installation (i10Pi) System Check Tool en modo silencioso

A partir de la versión 10.4.0, es posible ejecutar Pre-installation (i10Pi) System Check Tool en modo silencioso. Puede ejecutar i10pi en modo silencioso para comprobar si la máquina cumple los requisitos del sistema para realizar la instalación sin interacción por parte del usuario.

Para obtener más información sobre cómo ejecutar i10Pi en modo silencioso, consulte la Guía de instalación de Informatica 10.4.0.

Cifrado de contraseñas en el archivo de propiedades de la instalación silenciosa

A partir de la versión 10.4.0, el programa de instalación de Informatica incluye una utilidad que se puede usar para cifrar contraseñas que se establecen en el archivo de propiedades que contiene las opciones necesarias para instalar los servicios en modo silencioso.

Cuando el programa de instalación se ejecuta en modo silencioso, el marco de la instalación descifra las contraseñas cifradas.

Para obtener más información, consulte la *Guía de instalación y configuración de Informatica*.

Modelo de estructura inteligente

En esta sección se describen las nuevas características del modelo de estructura inteligente en la versión 10.4.0.

Tipos de entrada adicionales

A partir de la versión 10.4.0, Intelligent Structure Discovery puede procesar entradas de ORC, Avro y Parquet.

Para obtener más información, consulte la *Guía del usuario de Data Engineering Integration 10.4.0*.

Creación de modelos a partir de muestras en tiempo de diseño

A partir de la versión 10.4.0, Intelligent Structure Discovery puede crear un modelo basado en un archivo de ejemplo que seleccione al crear un objeto de datos complejo. Puede refinar el modelo más adelante en la integración de datos de Informatica Intelligent Cloud Services.

Esta funcionalidad es compatible con archivos de ejemplo XML, JSON, ORC, AVRO y Parquet.

Para obtener más información, consulte la *Guía del usuario de Data Engineering Integration 10.4.0*.

Administración de datos sin identificar

A partir de la versión 10.4.0, Intelligent Structure Discovery ordena la salida de los datos sin identificar en el modelo con formato JSON estructurado.

Para obtener más información, consulte la *Guía del usuario de Data Engineering Integration 10.4.0*.

PowerCenter

En esta sección se describen las nuevas características de PowerCenter de la versión 10.4.0.

Transformación de HTTP

A partir de la versión 10.4.0, la transformación de HTTP también incluye los siguientes métodos para la construcción de URL final: SIMPLE PATCH, SIMPLE PUT y SIMPLE DELETE.

Puede realizar una actualización parcial y los datos de entrada no necesitan formar un cuerpo completo con el método SIMPLE PATCH. Puede utilizarlo para actualizar los datos del puerto de entrada como una revisión del recurso.

Puede realizar un reemplazo completo del documento con el método SIMPLE PUT. Puede crear datos desde un puerto de entrada como un único bloque de datos para un servidor HTTP. Si los datos ya existen, puede actualizar los datos de un puerto de entrada como un único bloque de datos para un servidor HTTP.

Puede eliminar datos del servidor HTTP con el método SIMPLE DELETE.

También puede parametrizar la URL base de la transformación HTTP.

Anteriormente, solo se podía especificar la construcción de URL final para estos dos métodos: SIMPLE GET y SIMPLE POST. Tampoco se podía parametrizar la URL final para la transformación HTTP.

Para obtener más información, consulte el capítulo "Transformación HTTP" de la *Guía de transformación de PowerCenter 10.4.0*.

Conectividad

En esta sección se describen las nuevas características de conectividad de la versión 10.4.0.

Suplantación de usuario para conexiones de Oracle

A partir de la versión 10.4.0, puede especificar el nombre del usuario suplantado en la conexión de Oracle en PowerCenter para conectar con orígenes y destinos de Oracle. Cuando especifica un usuario suplantado, puede conectarse a Oracle en nombre de ese usuario y realizar operaciones de lectura o escritura.

Para obtener más información, consulte la *Guía de conceptos básicos de flujos de trabajo de Informatica 10.4.0 PowerCenter*.

Compatibilidad con las vistas de modelos de base de datos de SAP HANA

A partir de la versión 10.4.0, puede leer datos de los siguientes tipos de vistas de modelos de base de datos de SAP HANA:

- Vistas de análisis
- Vistas de atributos
- Vistas calculadas

Para obtener más información, consulte la *Guía de Informatica 10.4.0 PowerCenter Designer*.

Adaptadores de PowerExchange

En esta sección se describen las nuevas características del adaptador de PowerExchange en la versión 10.4.0.

Adaptadores de PowerExchange para Informatica

En esta sección se describen las nuevas características del adaptador de Informatica de la versión 10.4.0.

PowerExchange for Amazon Redshift

A partir de la versión 10.4.0, PowerExchange for Amazon Redshift incluye las siguientes características:

- Puede ejecutar asignaciones en el entorno de AWS Databricks.
- Puede seleccionar un nombre de región de clúster en la propiedad de conexión Región del clúster, incluso aunque haya especificado el nombre de la región de clúster en la propiedad de conexión de URL de JDBC.
- Puede conservar valores nulos cuando lee datos de Amazon Redshift.
- Puede especificar el número de archivos de transferencia provisional por lote, cuando escribe datos en Amazon Redshift.
- Puede conservar el orden de los registros, cuando escribe datos desde un origen CDC a un destino de Amazon Redshift.

Para obtener más información, consulte la *Guía del usuario de Informatica 10.4.0 PowerExchange for Amazon Redshift*.

PowerExchange for Amazon S3

A partir de la versión 10.4.0, PowerExchange for Amazon S3 incluye las siguientes características:

- Puede ejecutar asignaciones en el entorno de AWS Databricks.
- Puede utilizar las credenciales de seguridad temporales que emplean AssumeRole para acceder a los recursos de AWS.
- Puede parametrizar el tipo de formato y el esquema de datos en las propiedades de las operaciones de lectura y escritura en tiempo de ejecución.

Para obtener más información, consulte la *Guía del usuario de Informatica 10.4.0 PowerExchange for Amazon S3*.

PowerExchange for Google BigQuery

A partir de la versión 10.4.0, PowerExchange for Google BigQuery incluye las siguientes características:

- Puede utilizar un clúster de Google Dataproc para ejecutar asignaciones en el motor Spark.
- Puede aumentar el rendimiento de la asignación ejecutándola en el modo Spark optimizado. Cuando utiliza el modo Spark optimizado para leer datos, puede especificar el número de particiones que se utilizarán. Puede especificar si la asignación se ejecutará en modo **Genérico** u **Optimizado** en las propiedades avanzadas de las operaciones de lectura y escritura. El modo Spark optimizado aumenta el rendimiento de las asignaciones.
- También puede configurar un reemplazo de SQL para reemplazar la consulta SQL predeterminada que se utiliza para extraer datos del origen de Google BigQuery.
- Puede leer o escribir datos de tipo de datos NUMERIC en Google BigQuery. El tipo de datos NUMERIC es un valor numérico exacto con 38 dígitos de precisión y 9 dígitos de escala. Cuando lee o escribe el tipo de datos NUMERIC, el servicio de integración de datos asigna el tipo de datos NUMERIC al tipo de datos Decimal transformation y la precisión permitida es hasta 38 y la escala, hasta 9.

Para obtener más información, consulte la *Guía del usuario de Informatica 10.4.0 PowerExchange for Google BigQuery*.

PowerExchange for Google Cloud Storage

A partir de la versión 10.4.0, PowerExchange for Google Cloud Storage incluye las siguientes características:

- Puede utilizar un clúster de Google Dataproc para ejecutar asignaciones en el motor Spark.
- Puede configurar las siguientes propiedades avanzadas de operación de lectura de objeto de datos de Google Cloud Storage cuando lee datos de un origen de Google Cloud Storage:

Ruta de Google Cloud Storage

Reemplaza la ruta de Google Cloud Storage al archivo que seleccionó en el objeto de datos de Google Cloud Storage.

Use el siguiente formato:

```
gs://<bucket name> o gs://<bucket name>/<folder name>
```

Nombre del archivo de origen

Reemplaza el nombre del archivo de origen de Google Cloud Storage especificado en el objeto de datos de Google Cloud Storage.

Es un directorio

Lee todos los archivos disponibles en la carpeta especificada en la propiedad avanzada de operación de lectura de objeto de datos de **Ruta de Google Cloud Storage**.

Para obtener más información, consulte la *Guía del usuario de Informatica 10.4.0 PowerExchange for Google Cloud Storage*.

PowerExchange for Microsoft Azure Blob Storage

A partir de la versión 10.4.0, PowerExchange for Microsoft Azure Blob Storage incluye las siguientes características:

- Puede parametrizar el tipo de formato y el esquema de datos en las propiedades de las operaciones de lectura y escritura en tiempo de ejecución.
- Puede utilizar la autenticación con firmas de acceso compartido mientras crea una conexión con Microsoft Azure Blob Storage.

Para obtener más información, consulte la *Guía del usuario de Informatica 10.4.0 PowerExchange for Microsoft Azure Blob Storage*.

PowerExchange for Microsoft Azure SQL Data Warehouse

A partir de la versión 10.4.0, se pueden leer y escribir datos en un endpoint de Microsoft Azure SQL Data Warehouse que resida en una red virtual (VNet).

Para obtener más información, consulte la *Guía del usuario de Informatica 10.4.0 PowerExchange for Microsoft Azure SQL Data Warehouse*.

PowerExchange for Salesforce

A partir de la versión 10.4.0, PowerExchange for Salesforce incluye las siguientes características:

- Puede utilizar la versión 45.0, 46.0 y 47.0 de la API de Salesforce para crear una conexión de Salesforce y acceder a los objetos de Salesforce.
- Puede habilitar la fragmentación de claves principales para consultas sobre un objeto compartido que representa una entrada que se comparte en el objeto de destino. La fragmentación de claves principales solo es compatible con los objetos compartidos si el objeto principal es compatible. Por ejemplo, si

quiere consultar sobre CaseHistory, la fragmentación de claves principales debe ser compatible con el objeto principal Case.

- Puede crear reglas de asignaciones para volver a asignar atributos en registros cuando se insertan, actualizan o se actualizan e insertan registros para los objetos de destino Lead y Case que utilizan la API estándar.

Para obtener más información, consulte la *Guía del usuario de Informatica 10.4.0 PowerExchange for Salesforce*.

PowerExchange for SAP NetWeaver

A partir de la versión 10.4.0, PowerExchange for SAP NetWeaver incluye las siguientes características:

- Puede configurar la transmisión HTTPS para las asignaciones de lector de tablas de SAP.
- Puede leer datos de las vistas de ABAP CDS utilizando el extractor de tablas de SAP si utiliza la versión 7.50 o posterior del sistema SAP NetWeaver.
- Puede leer datos de tablas de SAP con campos que tengan los siguientes tipos de datos:
 - DF16_DEC
 - DF32_DEC
 - DF16_RAW
 - DF34_RAW
 - INT8
 - RAWSTRING
 - SSTRING
 - STRING

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for SAP NetWeaver 10.4.0*.

PowerExchange for Snowflake

A partir de la versión 10.4.0, PowerExchange for Snowflake incluye las siguientes características:

- Puede ejecutar asignaciones de Snowflake en el entorno de Databricks.
- Puede utilizar objetos de Snowflake como orígenes y destinos dinámicos en una asignación.
- Puede crear un destino de Snowflake mediante la opción **Crear destino**.
- Puede configurar una estrategia de esquema de destino para un destino de Snowflake en una asignación. Puede seleccionar entre las opciones disponibles para conservar el esquema de destino existente o crear un destino si no existe. También puede especificar las opciones de la estrategia de esquema de destino como un valor de parámetro.
- Puede especificar un nombre de archivo y una ruta rechazados en las propiedades de destino avanzadas de Snowflake que el servicio de integración de datos utiliza para escribir registros que se rechazan al escribir en el destino.
- Cuando el tipo de proveedor ODBC en la conexión ODBC de Snowflake es Snowflake, puede configurar la optimización de inserciones para insertar la lógica de transformación en la base de datos de Snowflake.
- Puede leer o escribir datos de tipo de datos Decimal con una precisión y una escala de 38 bits configurando el indicador personalizado `EnableSDKDecimal38` en las propiedades del servicio de integración de datos.

Para obtener más información, consulte la *Guía del usuario de Informatica 10.4.0 PowerExchange for Snowflake*.

PowerExchange for HDFS

A partir de la versión 10.4.0, PowerExchange for HDFS incluye las siguientes características:

- Puede parametrizar el tipo de formato y el esquema de datos en las propiedades de las operaciones de lectura y escritura en tiempo de ejecución.
- Puede dar formato al esquema de un objeto de datos de un archivo complejo para una operación de lectura o escritura.

Para obtener más información, consulte la *Guía del usuario de Informatica 10.4.0 PowerExchange for HDFS*.

Adaptadores de PowerExchange para PowerCenter

En esta sección se describen las nuevas características del adaptador de PowerCenter en la versión 10.4.0.

PowerExchange for Google BigQuery

A partir de la versión 10.4.0, se pueden leer o escribir datos de tipo de datos NUMERIC en Google BigQuery. El tipo de datos NUMERIC es un valor numérico exacto con 38 dígitos de precisión y 9 dígitos de escala. Cuando lee o escribe el tipo de datos NUMERIC, el servicio de integración de PowerCenter asigna el tipo de datos NUMERIC al tipo de datos Decimal transformation y la precisión permitida es hasta 28 y la escala, hasta 9.

Para obtener más información, consulte la *Guía del usuario de Informatica 10.4.0 PowerExchange for Google BigQuery para PowerCenter*.

PowerExchange for Google Cloud Storage

A partir de la versión 10.4.0, puede configurar las siguientes propiedades avanzadas de operación de lectura de objeto de datos de Google Cloud Storage cuando lee datos de un origen de Google Cloud Storage:

Ruta de Google Cloud Storage

Reemplaza la ruta de Google Cloud Storage al archivo que seleccionó en el objeto de datos de Google Cloud Storage.

Use el siguiente formato:

```
gs://<nombre de cubo> o gs://<nombre de cubo>/<nombre de carpeta>
```

Nombre del archivo de origen

Reemplaza el nombre del archivo de origen de Google Cloud Storage especificado en el objeto de datos de Google Cloud Storage.

Es un directorio

Lee todos los archivos disponibles en la carpeta especificada en la propiedad avanzada de operación de lectura de objeto de datos de **Ruta de Google Cloud Storage**.

Para obtener más información, consulte la *Guía del usuario de Informatica 10.4.0 PowerExchange for Google Cloud Storage para PowerCenter*.

PowerExchange for Greenplum

A partir de la versión 10.4.0, puede utilizar PowerExchange for Greenplum para leer datos de Greenplum. Puede configurar propiedades específicas de sesión para los orígenes de Greenplum para determinar cómo extraer datos de Greenplum.

Cuando ejecuta una sesión de Greenplum para leer datos, el servicio de integración de PowerCenter invoca el servidor de archivos paralelos de la base de datos de Greenplum, gpfdist, que es el programa de distribución de archivos, para leer los datos.

Para obtener más información, consulte la *Guía del usuario de Informática PowerExchange for Greenplum 10.4.0 para PowerCenter*.

PowerExchange for JD Edwards EnterpriseOne

A partir de la versión 10.4.0, puede utilizar la versión 9.2 de la API de JD Edwards EnterpriseOne para crear una conexión de JD Edwards EnterpriseOne y acceder a los objetos de JD Edwards EnterpriseOne.

Para obtener más información, consulte la *Guía del usuario de Informática 10.4.0 PowerExchange for JD Edwards EnterpriseOne para PowerCenter*.

PowerExchange for Kafka

A partir de la versión 10.4.0, se pueden utilizar las siguientes propiedades SSL para habilitar una conexión segura a un agente de Kafka:

- Modo SSL
- Ruta de acceso de archivo del TrustStore SSL
- Contraseña del TrustStore SSL
- Ruta de acceso de archivo de almacén de claves SSL
- Contraseña del almacén de claves SSL

Puede configurar el agente de mensajería de Kafka para utilizar la versión 0.10.1.1 o superior del agente de Kafka.

Para obtener más información, consulte la *Guía del usuario de Informática PowerExchange for Kafka 10.4.0 para PowerCenter*.

PowerExchange for Salesforce

A partir de la versión 10.4.0, puede utilizar las versiones 46.0 y 47.0 de la API de Salesforce para crear una conexión de Salesforce y acceder a los objetos de Salesforce.

Para obtener más información, consulte la *Guía del usuario de Informática 10.4.0 PowerExchange for Salesforce para PowerCenter*.

PowerExchange for SAP NetWeaver

A partir de la versión 10.4.0, puede utilizar el extractor de tablas ABAP dinámicas de PowerExchange for SAP para leer datos de tablas de SAP y vistas de ABAP Core Data Services (CDS) a través de transmisiones HTTP/HTTPS. Puede leer datos de las vistas de ABAP CDS utilizando el extractor de tablas ABAP dinámicas de PowerExchange for SAP si utiliza la versión 7.50 o posterior del sistema SAP NetWeaver.

Para obtener más información, consulte la *Guía del usuario de Informática PowerExchange for SAP NetWeaver 10.4.0*.

Seguridad

En esta sección se describen las nuevas características de seguridad de la versión 10.4.0.

Configuración de aplicaciones web para utilizar proveedores de identidades SAML diferentes

A partir de la versión 10.4.0, puede configurar aplicaciones web de Informatica que se ejecuten en un dominio para utilizar proveedores de identidades de lenguaje de marcado de aserción de seguridad (Security Assertion Markup Language, SAML). Por ejemplo, podría configurar Informatica Administrator para utilizar Active Directory Federation Services como el proveedor de identidades, y configurar Informatica Analyst para utilizar PingFederate como el proveedor de identidades.

Para obtener más información, consulte la *Guía de seguridad de Informatica 10.4.0*.

CAPÍTULO 2

Nuevas características (10.2.2 HotFix 1)

Este capítulo incluye los siguientes temas:

- [Programas de la línea de comandos , 39](#)
- [Enterprise Data Catalog, 39](#)

Programas de la línea de comandos

En esta sección se describen los nuevos comandos de la versión 10.2.2 HotFix 1.

Comandos infacmd Idm

En la siguiente tabla se describen las nuevas opciones de comando infacmd Idm:

Nueva opción	Descripción
-Force -fr	Opcional. Si desea forzar la copia de seguridad cuando el modo de copia de seguridad es sin conexión. Realiza a la fuerza la copia de seguridad y sobrescribe la copia de seguridad existente.
-Force -fr	Opcional. Si desea limpiar el contenido existente de HDFS y Apache Zookeeper. Restaura a la fuerza los datos de la copia de seguridad.

Para obtener más información, consulte el capítulo "Referencia del comando infacmd Idm" en la *Referencia de comandos de Informatica 10.2.2 HotFix 1*.

Enterprise Data Catalog

En esta sección se describen las nuevas características de la versión 10.2.2 HotFix 1 de Enterprise Data Catalog.

Azure Data Lake Storage Gen2

A partir de la versión 10.2.2 HotFix 1, se pueden extraer metadatos desde un origen de Azure Data Lake Storage Gen2. Puede acceder a los metadatos de los archivos y las carpetas, así como extraerlos en una cuenta de almacenamiento de Azure Data Lake Storage Gen 2 para Enterprise Data Catalog.

Para obtener más información, consulte el capítulo "Azure Data Lake Store" de la *Guía de configuración del escáner de Informatica Enterprise Data Catalog 10.2.2 HotFix 1*.

Vinculación sin distinción entre mayúsculas y minúsculas

A partir de la versión 10.2.2 HotFix 1, los recursos se pueden configurar para que no distingan entre mayúsculas y minúsculas. Esta opción le permite ver el linaje vinculado entre los orígenes y los destinos independientemente del uso de minúsculas y mayúsculas en los nombres de los recursos.

Para obtener información sobre la vinculación sin distinción entre mayúsculas y minúsculas, consulte el capítulo sobre la administración de recursos en la *Guía del administrador de Informatica Catalog 10.2.2 HotFix 1*.

Extensión Tableau de Enterprise Data Catalog

A partir de la versión 10.2.2 HotFix 1, la extensión Tableau de Enterprise Data Catalog es una extensión nativa del panel de Tableau que se puede utilizar dentro de Tableau.

La extensión Tableau de Enterprise Data Catalog se puede utilizar en Tableau Desktop, Tableau Server y en todos los navegadores web que sean compatibles con Tableau. Descargue la extensión desde la aplicación de Enterprise Data Catalog y añádala a un panel en Tableau.

Para obtener más información sobre la extensión, consulte la guía de *la extensión de Informatica Enterprise Data Catalog 10.2.2 HotFix 1 para Tableau*.

Nuevos recursos

A partir de la versión 10.2.2 HotFix 1, los siguientes recursos nuevos se añaden a Enterprise Data Catalog:

- SAP PowerDesigner. Se puede extraer información de los metadatos, la relación y el linaje de un origen de datos de SAP PowerDesigner.
- SAP HANA. Se pueden extraer metadatos de linaje y objetos de una base de datos de SAP HANA.

Para obtener más información, consulte la *Guía de configuración del escáner de Informatica 10.2.2 HotFix 1*.

Escáner de almacén de creación de perfiles

A partir de la versión 10.2.2 HotFix 1, se puede crear uno o varios recursos del almacén de creación de perfiles. Puede asociar un almacén de creación de perfiles y elegir esquemas en el recurso. Cuando ejecuta el recurso, el escáner extrae y migra los resultados del perfil para los esquemas seleccionados desde el almacén de creación de perfiles asociado al catálogo.

Para obtener más información, consulte el capítulo sobre la configuración de escáneres de Informatica Platform en la *Guía de configuración del escáner de Informatica Enterprise Data Catalog 10.2.2 HotFix 1*.

API de REST

A partir de la versión 10.2.2 HotFix 1, puede usar las siguientes API de REST de Informatica Enterprise Data Catalog:

- API de REST de provisión de datos. Puede devolver, actualizar o eliminar conexiones y recursos.
- API de REST del modelo de catálogo. Además de las API de REST existentes, puede acceder, actualizar o eliminar las facetas de campos, las facetas de consultas y las fichas de búsqueda.
- API de objetos. Además de las API de REST existentes, puede enumerar las sugerencias y la búsqueda de catálogo.

Para obtener más información sobre las API de REST, consulte la *Referencia de la API de REST de Informatica Enterprise Data Catalog 10.2.2 HotFix 1*.

Búsqueda Mejoras

A partir de la versión 10.2.2 HotFix 1, se pueden realizar mejores búsquedas en Enterprise Data Catalog en función de las siguientes mejoras:

Búsqueda de comillas dobles

Se puede realizar una búsqueda de activos mediante las comillas dobles (") para buscar activos que tengan exactamente el nombre del activo que se incluya entre las comillas dobles, pero no las variaciones del nombre del activo en el catálogo.

Operadores de búsqueda

Se pueden utilizar nuevos operadores de búsqueda para que los resultados de la búsqueda sean más precisos. Los operadores de búsqueda son AND, OR, NOT, título y descripción.

Rango de búsqueda

Enterprise Data Catalog utiliza un algoritmo de clasificación para clasificar los activos de datos en la página de resultados de búsqueda. La clasificación de la búsqueda hace referencia a la precedencia de un activo cuando se compara con otros activos que forman parte de los resultados de búsqueda específicos.

Búsqueda relacionada

La opción **Mostrar búsqueda relacionada** se puede habilitar en la página **Resultados de búsqueda** para ver los activos relacionados.

Si desea más información sobre las mejoras en las búsquedas, consulte el capítulo "Buscar activos" en la *Guía del usuario de Informatica Enterprise Data Catalog 10.2.2 Hotfix 1*.

Búsqueda Tabuladores

A partir de la versión 10.2.2 HotFix 1, se pueden usar las fichas de búsqueda para buscar activos sin tener que establecer continuamente los mismos criterios de búsqueda cuando se realice una búsqueda para los activos. Las fichas de búsqueda son filtros predefinidos en el catálogo.

Si desea más información sobre las fichas de búsqueda, consulte el capítulo "Personalizar búsqueda" en la *Guía del usuario de Informatica Enterprise Data Catalog 10.2.2 Hotfix 1*.

Tipos de recursos compatibles en la utilidad de escáner independiente

A partir de la versión 10.2.2 HotFix 1, se pueden extraer metadatos de los siguientes orígenes externos cuando no se pueda acceder a ellos en tiempo de ejecución o cuando estén desconectados:

- Apache Atlas
- Cloudera Navigator
- Sistema de archivos
- HDFS
- Hive
- Informatica Platform
- MicroStrategy
- OneDrive
- Oracle Business Intelligence
- SharePoint
- Sybase
- Tableau

Para obtener más información, consulte el capítulo sobre extracción de metadatos de recursos inaccesibles o sin conexión en la *Guía del administrador de Informatica Enterprise Data Catalog 10.2.2 HotFix 1*.

Versión preliminar técnica

La versión 10.2.2 HotFix 1 de Enterprise Data Catalog incluye la funcionalidad que está disponible para la versión preliminar técnica.

La funcionalidad de versión preliminar técnica es compatible, pero no se garantiza ni está preparada para usarla en producción. Informatica recomienda utilizarla solo en entornos que no sean de producción. Informatica tiene la intención de incluir la funcionalidad de versión preliminar en una próxima versión de disponibilidad general (GA) para poder usarla en producción, pero podría decidir no hacerlo en función de cuáles sean los requisitos del mercado o las circunstancias técnicas en ese momento. Para obtener más información, póngase en contacto con el servicio internacional de atención al cliente de Informatica.

A partir de la versión 10.2.2 HotFix 1, las siguientes funciones están disponibles para la versión preliminar técnica:

- A partir de la versión 10.2.2 HotFix 1, se pueden extraer metadatos para el linaje de datos en el nivel de columna, incluida la lógica de transformación de un origen de datos de Oracle Data Integrator.
- A partir de la versión 10.2.2 HotFix 1, se pueden extraer metadatos para el linaje de datos en el nivel de columna, incluida la lógica de transformación de un origen de datos de IBM InfoSphere DataStage.
- A partir de la versión 10.2.2 HotFix 1, se puede extraer el linaje de datos en el nivel de columna para los procedimientos almacenados en Oracle y SQL Server.
- A partir de la versión 10.2.2 HotFix 1, se puede realizar el aprovisionamiento de datos después de completar la obtención de datos en el catálogo. Gracias al aprovisionamiento de datos, podrá mover los datos a un destino para realizar otros análisis.

Si desea más información sobre la obtención de una vista previa de los datos, consulte la *Guía del administrador de Informatica Catalog 10.2.2 HotFix 1* y la *Guía del usuario de Informatica Enterprise Data Catalog 10.2.2 Hotfix 1*.

- A partir de la versión 10.2.2 HotFix 1, se puede obtener una vista previa de los datos para evaluarlos antes de trasladarlos al destino. Solo se pueden obtener vistas previas de los datos para los activos tabulares de los recursos de Oracle y Microsoft SQL Server.

Si desea más información sobre la obtención de una vista previa de los datos, consulte la *Guía del administrador de Informatica Catalog 10.2.2 HotFix 1* y la *Guía del usuario de Informatica Enterprise Data Catalog 10.2.2 Hotfix 1*.

CAPÍTULO 3

Nuevas características (10.2.2 Service Pack 1)

Este capítulo incluye los siguientes temas:

- [Big Data Management, 44](#)
- [Big Data Transmisión, 45](#)
- [Enterprise Data Catalog, 46](#)
- [Enterprise Data Preparation, 46](#)
- [Adaptadores de PowerExchange para Informatica, 47](#)

Big Data Management

En esta sección se describen las nuevas características de Big Data Management en la versión 10.2.2 Service Pack 1.

Sqoop

A partir de la versión 10.2.2 Service Pack 1, puede utilizar las siguientes nuevas características de Sqoop.

Conexión con bases de datos de Oracle y Microsoft SQL Server habilitadas con SSL

Puede configurar Sqoop para que se conecte con bases de datos relacionales seguras como Oracle y Microsoft SQL Server.

Utilice el argumento de conexión y la cadena de conexión de JDBC adecuados en la conexión JDBC para conectarse a una base de datos de Oracle o Microsoft SQL Server habilitada con SSL.

Para obtener más información, consulte la *Guía del usuario de Informatica Big Data Management 10.2.2 Service Pack 1*.

Compatibilidad con el argumento `connection-param-file` de Sqoop

Puede definir el argumento `connection-param-file` para conectarse con una base de datos de Oracle habilitada con SSL. El argumento `connection-param-file` es un archivo de propiedad que especifica los parámetros JDBC adicionales que Sqoop debe utilizar para conectarse con la base de datos habilitada con SSL.

El contenido de este archivo se analiza como una propiedad Java estándar y pasa al controlador cuando se crea una conexión.

El argumento `connection-param-file` se puede especificar en el campo **Argumentos de Sqoop** de la conexión JDBC.

Utilice la siguiente sintaxis:

```
--connection-param-file <nombre_archivo_parámetro>
```

Para obtener más información, consulte la *Guía del usuario de Informatica Big Data Management 10.2.2 Service Pack 1*.

Big Data Transmisión

En esta sección se describen las nuevas características de Big Data Streaming en la versión 10.2.2 Service Pack 1.

Destino de Amazon S3

A partir de la versión 10.2.2 Service Pack 1, puede crear una asignación de transmisión para escribir datos en Amazon S3.

Cree un objeto de datos de Amazon S3 para escribir datos en Amazon S3. Puede crear una conexión de Amazon S3 para utilizar Amazon S3 como destinos. Una conexión de Amazon S3 se puede crear y administrar en Developer tool o a través de infacmd.

Para obtener más información, consulte la *Guía del usuario de Informatica Big Data Streaming 10.2.2 Service Pack 1*.

Función TIME_RANGE

A partir de la versión 10.2.2 Service Pack 1, puede utilizar la función TIME_RANGE en una transformación de combinación que determina el intervalo de tiempo de los eventos de transmisión que se van a unir.

La función TIME_RANGE es válida únicamente en transformaciones de combinación en una asignación de transmisión.

Sintaxis

```
TIME_RANGE(EventTime1,EventTime2,Format,Interval)
```

Si desea más información sobre la función TIME_RANGE, consulte la *Guía de referencia de lenguaje de transformación de Informatica 10.2.2 Service Pack 1*.

Escritura de datos en varios archivos HDFS

A partir de la versión 10.2.2 Service Pack 1, puede utilizar un objeto de datos de archivo complejo para escribir los datos de origen en varios archivos HDFS. Según los valores de puerto del encabezado FileName del objeto de datos de archivo complejo, el servicio de integración de datos crea varios archivos HDFS en tiempo de ejecución en la ubicación de destino.

Para obtener más información, consulte la *Guía del usuario de Informatica Big Data Streaming 10.2.2 Service Pack 1*.

Enterprise Data Catalog

En esta sección se describe la nueva característica de Enterprise Data Catalog en la versión 10.2.2 Service Pack 1.

Tipos de recursos compatibles en la utilidad de escáner independiente

A partir de la versión 10.2.2 Service Pack 1, puede extraer metadatos de los siguientes orígenes externos cuando no se pueda acceder a ellos en tiempo de ejecución o cuando estén desconectados:

- IBM DB2
- IBM DB2 for z/OS
- IBM Netezza
- JDBC
- PowerCenter
- SQL Server Integration Services

Para obtener más información, consulte el capítulo sobre extracción de metadatos de recursos inaccesibles o sin conexión en la *Guía del administrador de Enterprise Data Catalog en Informática 10.2.2 Service Pack 1*.

Enterprise Data Preparation

En esta sección se describen las nuevas características de Enterprise Data Preparation en la versión 10.2.2 Service Pack 1.

Revertir todos los tipos de datos inferidos

A partir de la versión 10.2.2 Service Pack 1, todos los dominios de datos y tipos inferidos que se han aplicado a las columnas de origen de una hoja de trabajo se pueden revertir a sus tipos originales.

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Enterprise Data Preparation de Informática 10.2.2 Service Pack 1*.

Preparar archivos Avro y Parquet

A partir de la versión 10.2.2 Service Pack 1, puede muestrear los datos jerárquicos de los archivos Avro y Parquet que agregue al proyecto como primer paso en la preparación de datos. Enterprise Data Preparation convierte la estructura del archivo Avro o Parquet en una estructura sin formato, y presenta los datos en una hoja de cálculo que sirve para preparar los datos.

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Enterprise Data Preparation de Informática 10.2.2 Service Pack 1*.

Adaptadores de PowerExchange para Informatica

En esta sección se describen las nuevas características del adaptador de Informatica en la versión 10.2.2 Service Pack 1.

PowerExchange for Hive

A partir de la versión 10.2.2 Service Pack 1, cuando importe una tabla de Hive, puede utilizar la opción **Mostrar solo esquema predeterminado** para borrar los resultados de búsqueda y mostrar solo las tablas que usan el esquema predeterminado, o para buscar las tablas que usan un esquema que no sea el esquema predeterminado.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for Hive 10.2.2 Service Pack 1*.

CAPÍTULO 4

Nuevas características (10.2.2)

Este capítulo incluye los siguientes temas:

- [servicios de aplicación, 48](#)
- [Big Data Management, 49](#)
- [Big Data Transmisión , 53](#)
- [Programas de la línea de comandos, 55](#)
- [Enterprise Data Catalog, 60](#)
- [Enterprise Data Lake, 63](#)
- [Informatica Developer, 68](#)
- [Asignaciones de Informatica, 68](#)
- [Transformaciones de Informatica, 70](#)
- [Adaptadores de PowerExchange para Informatica, 73](#)

servicios de aplicación

En esta sección se describen las nuevas características de los servicios de aplicación de la versión 10.2.2.

Servicio de ingesta masiva

A partir de la versión 10.2.2, puede habilitar la comunicación segura entre un proceso de servicio de ingesta masiva y componentes externos. Puede especificar el número de puerto HTTPS para el protocolo HTTPS y un archivo de almacén de claves para el certificado SSL.

Para obtener más información, consulte el capítulo "Servicio de ingesta masiva" en la *Guía del servicio de aplicación de Informatica 10.2.2*.

Servicio de acceso a metadatos

A partir de la versión 10.2.2, puede configurar el servicio de acceso a metadatos para usar perfiles del sistema operativo. Use perfiles del sistema operativo para aumentar la seguridad y aislar el entorno del usuario en tiempo de diseño cuando se importan metadatos y se obtiene una vista previa de ellos desde un clúster de Hadoop.

Para obtener más información, consulte el capítulo sobre usuarios y grupos en la *Guía de seguridad de Informatica 10.2.2*.

Servicio del concentrador de operaciones REST

A partir de la versión 10.2.2, puede configurar el servicio del concentrador de operaciones REST. Se trata de un servicio de sistema REST en el dominio de Informatica que expone la funcionalidad de los productos de Informatica a clientes externos mediante las API de REST.

Puede configurar el servicio del concentrador de operaciones REST a través de la Herramienta del administrador o infacmd. Puede usar el servicio del concentrador de operaciones REST para ver estadísticas de ejecución de asignaciones para los trabajos de asignación implementados en la aplicación.

Puede usar el servicio del concentrador de operaciones REST para obtener estadísticas de ejecución de asignaciones para asignaciones de Big Data que se ejecutan en el servicio de integración de datos o en el entorno de Hadoop.

Para obtener más información acerca de la API de REST, consulte la *Guía del administrador de Big Data Management 10.2.2*.

Big Data Management

En esta sección se describen las nuevas características de Big Data Management en la versión 10.2.2.

Integración de Azure Databricks

A partir de la versión 10.2.2, puede integrar el dominio de Informatica con el entorno de Azure Databricks.

Azure Databricks es una plataforma de nube de análisis que está optimizada para los servicios de nube de Microsoft Azure. Incorpora las funciones y tecnologías de clúster de Apache Spark de código abierto.

El dominio de Informatica puede instalarse en una máquina virtual de Azure o en las instalaciones. El proceso de integración es similar a la integración con el entorno de Hadoop. Se pueden realizar tareas de integración, incluso la importación de la configuración del clúster desde el entorno de Databricks. El dominio de Informatica usa autenticación de token para acceder al entorno de Databricks. El ID del token de Databricks se almacena en la conexión de Databricks.

Orígenes y destinos

Puede ejecutar asignaciones contra los siguientes orígenes y destinos dentro del entorno de Databricks:

- Microsoft Azure Data Lake Store
- Microsoft Azure Blob Storage
- Almacén de datos de Microsoft Azure SQL
- Microsoft Azure Cosmos DB

Transformaciones

Puede agregar las siguientes transformaciones a una asignación de Databricks:

- Agregación
- Expresión
- Filtro
- Unión
- Búsqueda

Normalizador

Rango

Enrutador

Ordenación

Unión

El motor de Databricks Spark procesa la transformación de una manera muy similar a como lo hace el motor de Spark en el entorno de Hadoop.

Tipos de datos

Se admiten los siguientes tipos de datos:

Array

Bigint

Date/time

Decimal

Doble

Entero

Map

Struct

Text

String

Asignaciones

Cuando se configura una asignación, puede elegir validar y ejecutar la asignación en el entorno de Databricks. Cuando ejecuta la asignación, el servicio de integración de datos genera un código de Scala y lo pasa al motor de Databricks Spark.

Flujos de trabajo

Puede desarrollar flujos de trabajo de clúster para crear clústeres efímeros en el entorno de Databricks.

Para obtener más información, consulte las siguientes guías:

Guía de integración de Big Data Management 10.2.2

Guía del administrador de Informatica Big Data Management 10.2.2

Guía del usuario de Big Data Management 10.2.2

Datos jerárquicos

En esta sección se describen las nuevas características para los datos jerárquicos en la versión 10.2.2.

Puertos complejos dinámicos

A partir de la versión 10.2.2, puede agregar puertos complejos dinámicos a una asignación dinámica que se ejecuta en el motor de Spark. Use puertos complejos dinámicos para administrar cambios de esquema frecuentes a datos jerárquicos en archivos complejos.

Un puerto complejo dinámico recibe elementos nuevos o cambiados de un puerto complejo según los campos de esquema en tiempo de ejecución. Las reglas de entrada determinan los elementos de un puerto complejo dinámico. Según las reglas de entrada, un puerto complejo dinámico recibe uno o varios elementos

de un puerto complejo desde la transformación de nivel superior. Puede usar puertos complejos dinámicos como matriz dinámica, asignación dinámica y estructura dinámica en algunas transformaciones en el motor de Spark.

Para obtener más información, consulte el capítulo "Procesamiento de datos jerárquicos con cambios en el esquema" de la *Guía del usuario de Informatica Big Data Management 10.2.2*.

Vista previa de datos jerárquicos

A partir de la versión 10.2.2, puede obtener una vista previa de los datos jerárquicos dentro de una asignación en Developer tool. Al obtener una vista previa de los datos jerárquicos, podrá diseñar y depurar asignaciones de Big Data.

Puede elegir orígenes y transformaciones como puntos de vista previa en una asignación que contiene los siguientes tipos jerárquicos:

- Matriz
- Struct
- Asignación

Vista previa de datos jerárquicos para la vista previa técnica. La funcionalidad de vista previa técnica es compatible, pero no se garantiza ni está preparada para usarla en producción. Informatica recomienda utilizar estas características solo en entornos que no sean de producción.

Para obtener más información, consulte la *Guía del usuario de Informatica Big Data Management 10.2.2*.

Alta disponibilidad

En esta sección se describen las nuevas características de alta disponibilidad de la versión 10.2.2.

Big Data Recuperación de trabajos

A partir de la versión 10.2.2, el servicio de integración de datos puede recuperar un trabajo de Big Data configurado para ejecutarse en el motor de Spark cuando un nodo del servicio de integración de datos se detiene inesperadamente. Cuando el nodo del servicio de integración de datos falla antes de completarse un trabajo, este servicio envía el trabajo a otro nodo, el cual reanuda las tareas de procesamiento del trabajo desde el punto en el que se produjo el error en el nodo.

Para recuperar asignaciones de Big Data, debe habilitar la recuperación de trabajos de Big Data en las propiedades del servicio de integración de datos y ejecutar el trabajo desde infacmd.

Para obtener más información, consulte el capítulo sobre procesamiento del servicio de integración de datos en la *Guía del administrador de Informatica Big Data Management 10.2.2*.

Colas distribuidas del servicio de integración de datos

A partir de la versión 10.2.2, el servicio de integración de datos usa una cola distribuida para almacenar información de trabajos cuando está habilitada la recuperación de Big Data para trabajos de Big Data implementados. La cola distribuida se almacena en el repositorio de modelos, y cualquier servicio de integración de datos disponible puede ejecutar trabajos desde la cola cuando los recursos están disponibles.

Para obtener más información, consulte el capítulo sobre procesamiento del servicio de integración de datos en la *Guía del administrador de Informatica Big Data Management 10.2.2*.

Modelo de estructura inteligente

En esta sección se describen las nuevas características del modelo de estructura inteligente en la versión 10.2.2.

Alias en archivos XML

A partir de la versión 10.2.2, Intelligent Structure Discovery puede procesar archivos XML que usan diferentes alias para identificar el mismo espacio de nombres, como se utiliza en el archivo XML con el que se creó un modelo de estructura inteligente.

Tipos de datos

A partir de la versión 10.2.2 y desde la versión de invierno de marzo de 2019 de Informatica Intelligent Cloud Services, cuando un lector de archivos complejos usa un modelo de estructura inteligente, Intelligent Structure Discovery pasa los tipos de datos a los puertos de datos de salida.

Por ejemplo, cuando Intelligent Structure Discovery detecta que un campo contiene una fecha, pasa los datos a los puertos de datos de salida como fecha, no como cadena.

Nombres de campo

A partir de la versión 10.2.2 y desde la versión de invierno de marzo 2019 de Informatica Intelligent Cloud Services, los nombres de campo en los objetos de datos de archivo complejo que se importan desde un modelo de estructura inteligente pueden comenzar con números y palabras reservadas, y pueden contener los siguientes caracteres especiales `\. [] {} () * + - ? . ^ $ |`

Cuando un campo comienza con un número o una palabra reservada, la asignación de Big Data Management agrega un guion bajo (`_`) al comienzo del nombre del campo. Por ejemplo, si un campo en un modelo de estructura inteligente comienza con `OR`, la asignación importa el campo como `_OR`. Cuando el nombre del campo contiene un carácter especial, la asignación convierte el carácter en un guion bajo.

Procesar archivos XML grandes

A partir de la versión 10.2.2, Intelligent Structure Discovery puede transmitir archivos XML y procesar datos para elementos repetitivos en fragmentos. De este modo, el procesamiento de archivos XML grandes se realiza de forma más eficaz.

Desfase de datos

A partir de la versión 10.2.2 y desde la versión de invierno de marzo de 2019 de Informatica Intelligent Cloud Services, Intelligent Structure Discovery mejora la gestión de los desfases de datos.

En Intelligent Structure Discovery, los desfases de datos se producen cuando los datos de entrada contienen campos que no contenía el archivo de muestra. En este caso, Intelligent Structure Discovery pasa los datos sin definir a un puerto de datos no asignado en el destino, en lugar de descartarlos.

Ingesta masiva

A partir de la versión 10.2.2, puede ejecutar una carga incremental para ingerir datos incrementales. Cuando ejecuta una carga incremental, el motor de Spark captura datos incrementales según una marca de tiempo o una columna de ID, y los carga en el destino de Hive o HDFS. Si ingiere los datos en un destino de Hive, el motor de Spark también puede propagar los cambios de esquema que se realizaron en las tablas de origen.

Si ingiere datos incrementales, el servicio de ingesta masiva aprovecha el modo de importación incremental de Sqoop.

Para obtener más información, consulte la *Guía de ingesta masiva de Informatica Big Data Management 10.2.2*.

Supervisión

En esta sección se describen las nuevas características relacionadas con la supervisión en Big Data Management en la versión 10.2.2.

Supervisión de Spark

A partir de la versión 10.2.2, puede ver tanto las tareas previas al trabajo como las tareas posteriores al trabajo dentro del panel Estadísticas de resumen para la supervisión de Spark.

Para obtener más información sobre el uso de las tareas previas al trabajo y las tareas posteriores al trabajo, consulte la *Guía del usuario de Informatica Big Data Management 10.2.2*.

Seguridad

En esta sección se describen las nuevas características relacionadas con la seguridad en Big Data Management en la versión 10.2.2.

Paquete de seguridad empresarial

A partir de la versión 10.2.2, Informatica admite un clúster de Azure HDInsight con paquete de seguridad empresarial.

Este paquete usa Kerberos para autenticación y Apache Ranger para autorización.

Para obtener más información sobre el paquete de seguridad empresarial, consulte la *Guía del administrador de Informatica Big Data Management 10.2.2*.

Destinos

En esta sección se describen las nuevas características para los destinos en la versión 10.2.2.

Destinos de archivos sin formato HDFS

A partir de la versión 10.2.2, puede anexar datos de salida a archivos de destino HFS y archivos de rechazo. Para anexar los datos de salida, elija hacerlo si existe el destino HDFS.

Para ayudar a administrar los archivos que contienen datos anexados, el servicio de integración de datos anexa el ID de la ejecución de la asignación a los nombres de los archivos de destino y los archivos de rechazo.

Para obtener más información, consulte el capítulo "Destinos" en la *Guía del usuario de Informatica Big Data Management 10.2.2*.

Big Data Transmisión

En esta sección se describen las nuevas características de Big Data Streaming en la versión 10.2.2.

Objetos de datos de centros de eventos de Azure

A partir de la versión 10.2.2, puede implementar una asignación de transmisión que tenga un centro de eventos como origen en las siguientes distribuciones:

- Amazon EMR
- Azure HDInsight con almacenamiento de ADLS
- Cloudera CDH
- Hortonworks HDP

Roles IAM de cuentas cruzadas en Amazon Kinesis Connection

A partir de la versión 10.2.2, puede usar el rol IAM de cuentas cruzadas para autenticar un origen de Amazon Kinesis.

Use el rol IAM de cuentas cruzadas para compartir recursos en una cuenta de AWS con usuarios de una cuenta de AWS diferente sin crear usuarios en cada cuenta.

Para obtener más información, consulte la *Guía del usuario de Informática Big Data Streaming 10.2.2*.

Modelo de estructura inteligente

A partir de la versión 10.2.2, puede utilizar modelos de estructura inteligente en Big Data Streaming.

Puede incorporar un modelo de estructura inteligente en un objeto de datos de centros de eventos de Azure, Kinesis o Kafka. Cuando se agrega el objeto de datos a una asignación, se puede procesar cualquier tipo de entrada que el modelo pueda analizar.

El objeto de datos puede aceptar y analizar archivos de entrada de formularios PDF, JSON, Microsoft Excel, tablas de Microsoft Word, CSV, texto o XML, según el archivo que se utilizó para crear el modelo.

Para obtener más información, consulte la *Guía del usuario de Informática Big Data Streaming 10.2.2*.

Puertos de encabezado para objetos de datos de Big Data Streaming

A partir de la versión 10.2.2, algunos objetos de datos contienen puertos de encabezado predeterminados que representan metadatos asociados con eventos. Por ejemplo, el puerto de marca de tiempo contiene la hora a la que se generó el evento. Puede usar los puertos de encabezado para agrupar y procesar los datos.

Para obtener más información acerca de los puertos de encabezado, consulte la *Guía del usuario de Informática Big Data Streaming 10.2.2*.

Perfil de credenciales de AWS en la conexión de Amazon Kinesis

A partir de la versión 10.2.2, puede usar la autenticación basada en perfiles de credenciales de AWS en la conexión de Amazon Kinesis.

Cuando crea una conexión de Amazon Kinesis, puede introducir un nombre de perfil de credenciales de AWS. La asignación accede a las credenciales de AWS a través del nombre del perfil que aparece en el archivo de credenciales de AWS durante el tiempo de ejecución.

Para obtener más información, consulte la *Guía del usuario de Informática Big Data Streaming 10.2.2*.

Transmisión estructurada de Spark

A partir de la versión 10.2.2, Big Data Streaming usa la transmisión estructurada de Spark para procesar los datos de transmisión.

La transmisión estructurada de Spark es un motor de procesamiento de transmisiones de código abierto con tolerancia a errores y escalable desarrollado a partir del motor de Spark. Puede administrar la recepción tardía de eventos de transmisión y procesar datos de transmisión según la marca de tiempo de origen.

El motor de Spark ejecuta continuamente la asignación de transmisión. Lee los datos, los divide en microlotes, procesa dichos microlotes, publica los resultados y después escribe en un destino.

Para obtener más información, consulte la *Guía del usuario de Informatica Big Data Streaming 10.2.2*.

Transformación de ventana

A partir de la versión 10.2.2 puede usar las siguientes características cuando crea una transformación de ventana:

Retraso de marca de agua

El retraso de marca de agua define el tiempo de umbral para que un evento retrasado se acumule en un grupo de datos.

El retraso de marca de agua es un umbral donde puede especificar el tiempo necesario para que los datos que llegan con retraso puedan agruparse y procesarse. Si los datos del evento llegan dentro del tiempo de umbral, los datos se procesan y se acumulan en el grupo de datos correspondiente.

Puerto de ventana

El puerto de ventana especifica la columna que contiene los valores de marca de tiempo según los que puede agrupar los eventos. Los datos acumulados contienen el valor de marca de tiempo. Use la columna Puerto de ventana para agrupar los datos de tiempo del evento que se reciben con retraso.

Para obtener más información, consulte la *Guía del usuario de Informatica Big Data Streaming 10.2.2*.

Programas de la línea de comandos

En esta sección se describen los nuevos comandos de la versión 10.2.2.

Comandos infacmd dis

En este tema se describen las nuevas opciones de comando infacmd dis y las opciones de ejecución.

[Actualizar opciones del servicio](#)

El comando Actualizar opciones del servicio actualiza las propiedades de servicio de integración de datos.

En la siguiente tabla se describen las nuevas opciones de comando infacmd dis updateServiceOptions:

Nueva opción	Descripción
-RecycleMode -rm	Opcional. El modo de reciclaje reinicia el servicio y aplica el último servicio y las propiedades de este. Seleccione Anular o Completar. <ul style="list-style-type: none">- Completar. Detiene todas las aplicaciones y cancela todos los trabajos dentro de cada aplicación. Espera a que todos los trabajos se cancelen antes de deshabilitar el servicio.- Anular. Detiene todas las aplicaciones e intenta cancelar todos los trabajos antes de anularlas y deshabilitar el servicio. Predeterminado: Completar

En la siguiente tabla se describen las nuevas opciones de ejecución del comando infacmd dis updateServiceOptions:

Nueva opción	Descripción
ExecutionOptions.BigDataJobRecovery	Se habilita Big Data Job Recovery. Establezca en "true" o "false". Predeterminado: false.

Para obtener más información, consulte el capítulo "Referencia del comando infacmd dis" en la *Referencia de comando de Informatica 10.2.2*.

Comandos infacmd ihs

En la siguiente tabla se describen las nuevas opciones de comando infacmd ihs:

Nueva opción	Descripción
-PrimaryNode -nm	Opcional. Nodo principal en que se ejecuta el servicio.
-BackupNodes -bn	Opcional. Los nodos en los que se puede ejecutar el servicio cuando el nodo principal no está disponible. Puede configurar nodos de copia de seguridad si dispone de alta disponibilidad.

En la siguiente tabla se describen los nuevos comandos infacmd ihs:

Comando	Descripción
cleanCluster	Limpia el servicio de clúster de Informatica.

Para obtener más información, consulte el capítulo "Referencia del comando infacmd ihs" de la *Referencia de comandos de Informatica 10.2.2*.

Comandos infacmd ipc

En la siguiente tabla se describen los comandos infacmd ipc obsoletos:

Comando	Descripción
ExportToPC	Exporta objetos desde el repositorio de modelos o un archivo de exportación y los convierte en objetos de PowerCenter.

Comandos infacmd Idm

En la siguiente tabla se describen las nuevas opciones de comando infacmd Idm:

Nueva opción	Descripción
-PrimaryNode -nm	Opcional. Si quiere configurar la alta disponibilidad para Enterprise Data Catalog, especifique el nodo principal.
-BackupNodes -bn	Opcional. Si quiere configurar la alta disponibilidad para Enterprise Data Catalog, especifique una lista de nombres de nodo de copia de seguridad separada por comas.
-isNotifyChangeEmailEnabled -cne	Opcional. Especifique True si quiere habilitar las notificaciones de cambio de activo. El valor predeterminado es false.
-ExtraJarsPath -ejp	Opcional. Ruta al directorio en el equipo donde instaló el dominio de Informatica. El directorio debe incluir los archivos JAR necesarios para implementar Enterprise Data Catalog en un clúster existente con WANDisco Fusion.
-ExtraJarsPath -ejp	Opcional. Ruta al directorio en el equipo donde instaló el dominio de Informatica. El directorio debe incluir los archivos JAR necesarios para implementar Enterprise Data Catalog en un clúster existente con WANDisco Fusion.

En la siguiente tabla se describen los nuevos comandos infacmd Idm:

Comando	Descripción
collectAppLogs	Recopila archivos de registro para aplicaciones YARN que se ejecutan para habilitar el servicio de catálogo.
publishArchive	Crea un recurso en modo sin conexión y ejecuta la exploración.

Para obtener más información, consulte el capítulo "Referencia del comando infacmd Idm" de la *Referencia de comandos de Informatica 10.2.2*.

Comandos infacmd mi

En la tabla siguiente se describen los cambios introducidos en los comandos infacmd mi:

Comando	Cambiar descripción
createService	A partir de la versión 10.2.2, puede usar las opciones -HttpsPort, -KeystoreFile y -KeystorePassword para especificar si los procesos del servicio de ingesta masiva usan una conexión segura para comunicarse con componentes externos.
extendedRunStats	<p>A partir de la versión 10.2.2, debe usar la opción -RunID para especificar el RunID de la especificación de ingesta masiva y la opción -SourceName para especificar el nombre de una tabla de origen para ver las estadísticas de ejecución extendida para la tabla de origen. Si la tabla de origen se ingestó usando una carga incremental, la estadística de ejecución mostrará la clave incremental y el valor de inicio.</p> <p>Anteriormente, especificaba el JobID para el trabajo de asignación de ingesta que ingestaba la tabla de origen.</p> <p>Si actualiza a 10.2.2, debe actualizar cualquier script que ejecute infacmd mi extendedRunStats para usar las nuevas opciones.</p>
listSpecRuns	A partir de la versión 10.2.2, el comando devuelve adicionalmente el tipo de carga que el motor de Spark usa para ejecutar una especificación de ingesta masiva.
runSpec	A partir de la versión 10.2.2, puede usar la opción -LoadType para especificar el tipo de carga para ejecutar una especificación de ingesta masiva. El tipo de carga puede ser una carga completa o una carga incremental.

Para obtener más información, consulte el capítulo "Referencia del comando infacmd mi" de la *Referencia de comandos de Informatica 10.2.2*.

Comandos infacmd ms

En la siguiente tabla se describen los nuevos comandos infacmd ms:

Comando	Descripción
abortAllJobs	Anula todos los trabajos de asignación que están configurados para ejecutarse en el motor de Spark. Puede elegir anular trabajos en cola, trabajos en ejecución o ambos.
createConfigurationWithParams	Crea una configuración de clúster mediante parámetros de clúster que se especifican en la línea de comandos.
listMappingOptions	Enumera las opciones de asignación de una aplicación.
purgeDatabaseWorkTables	Purga la información de todos los trabajos desde la cola cuando habilita la recuperación de Big Data para el servicio de integración de datos.
updateMappingOptions	Actualiza las opciones de asignación de una aplicación.
updateOptimizationLevel	Actualiza el nivel de optimización para varias asignaciones de una aplicación.

Para obtener más información, consulte el capítulo "Referencia del comando infacmd ms" de la *Referencia de comandos de Informatica*.

Comandos infacmd oie

A partir de la versión 10.2.2, ha dejado de usarse el complemento oie y la compatibilidad con este complemento se eliminará en una versión futura. Los comandos infacmd oie se han migrado al complemento tools. Para obtener detalles, consulte ["Comandos infacmd tools" en la página 59](#).

Comandos infacmd tools

A partir de la versión 10.2.2, el complemento de infacmd tools reemplaza al complemento oie. El complemento tools realiza las operaciones de importación y exportación de objetos que lleva a cabo el complemento oie, y puede usarse para realizar algunas operaciones adicionales.

La siguiente tabla enumera los comandos infacmd oie que migraron al complemento tools:

Comando anterior	Comando actual
infacmd oie deployApplication	infacmd tools deployApplication
infacmd oie ExportObjects	infacmd tools exportObjects
infacmd oie exportResources	infacmd tools exportResources
infacmd oie ImportObjects	infacmd tools importObjects

En la siguiente tabla se describen los nuevos comandos infacmd tools:

Comando	Descripción
patchApplication	Implementa una revisión de aplicación con un archivo .piar en un servicio de integración de datos.

Para obtener más información, consulte el capítulo "Referencia del comando infacmd tools" en la *Referencia de comandos de Informatica 10.2.2*.

Comandos de infasetup

La siguiente tabla describe los comandos infasetup cambiados:

Comando	Descripción
DefineDomain	A partir de la versión 10.2.2, la opción -spid se ha agregado al comando DefineDomain.
updateDomainSamlConfig	A partir de la versión 10.2.2, la opción -spid se ha agregado al comando updateDomainSamlConfig.

Para obtener más información, consulte el capítulo "Referencia de comando infasetup" de la *Referencia de comandos de Informatica 10.2.2*.

Enterprise Data Catalog

En esta sección se describen las nuevas características de Enterprise Data Catalog en la versión 10.2.2.

Asignar automáticamente un título empresarial a una columna

A partir de la versión 10.2.2, Enterprise Data Catalog infiere un término empresarial como el título empresarial para los activos de columna. Un término empresarial con puntuación de confianza del 80 por ciento o más se infiere como un título empresarial de una columna. La puntuación de confianza se calcula mediante un algoritmo interno basado en términos empresariales aceptados en dominios de datos, la similitud del nombre de columna y la similitud del nombre entre la columna y el término empresarial.

Para obtener más información, consulte el capítulo sobre la realización de tareas de activo en la *Guía del usuario de Informatica Enterprise Data Catalog 10.2.2*.

Colaboración del usuario en activos

A partir de la versión 10.2.2, puede colaborar con otros usuarios de Enterprise Data Catalog en los activos. La colaboración en activos le permite interactuar con otros usuarios, compartir comentarios acerca de los activos, hacer consultas relacionadas con estos, realizar un seguimiento de todos los cambios en ellos y certificarlos.

Puede colaborar en activos de las siguientes maneras:

Seguir activos

Puede seguir activos para supervisar los cambios en activos en el catálogo. Siga un activo para estar informado acerca de los cambios que otros usuarios realizan a fin de poder supervisar el activo y realizar las acciones que sean necesarias.

Calificar y revisar un activo

Puede calificar y revisar activos según una escala de cinco estrellas en el catálogo. Califique y revise un activo para proporcionar comentarios acerca de este según diferentes aspectos, como la calidad, la aplicabilidad, la posibilidad de uso y la disponibilidad del activo.

Consultas de activos

Puede realizar preguntas acerca de un activo si quiere obtener un mejor conocimiento del activo en el catálogo. Realice preguntas que sean descriptivas, exploratorias, predictivas o causales por naturaleza.

Certificar activo

Puede certificar un activo para aprobarlo a fin de que otros usuarios puedan usarlo como confiable en contraposición a los que no están certificados.

Para obtener más información, consulte el capítulo sobre la colaboración del usuario en activos en la *Guía de Informatica Enterprise Data Catalog 10.2.2*.

Crear servicios de aplicación de Enterprise Data Catalog usando el programa de instalación

A partir de la versión 10.2.2, puede usar el programa de instalación para crear los servicios de aplicación de Enterprise Data Catalog después de instalar Enterprise Data Catalog. Puede usar el programa de instalación si había instalado Enterprise Data Catalog sin crear los servicios de aplicación.

Para obtener más información acerca de cómo usar el programa de instalación para crear los servicios de aplicación, consulte la *Guía de instalación y configuración de Informatica Enterprise Data Catalog 10.2.2*.

Utilidad de validación de metadatos personalizada

A partir de la versión 10.2.2, puede usar una utilidad independiente de línea de comandos de validación basada en Java para validar la sintaxis y la semántica de metadatos personalizados que quiera ingerir en el catálogo. Los metadatos personalizados representan los metadatos que quiere ingerir desde orígenes de datos personalizados para los que Enterprise Data Catalog no proporciona un recurso.

Para obtener más información acerca de cómo usar la utilidad, consulte el artículo de la base de conocimientos *Procedimientos para validar metadatos personalizados antes de ingerirlos en el catálogo*. Póngase en contacto con el servicio internacional de atención al cliente de Informatica para obtener instrucciones para descargar la utilidad.

Notificaciones de cambio

A partir de la versión 10.2.2, Enterprise Data Catalog muestra notificaciones cuando se realizan cambios en los activos seguidos. Los tipos de notificaciones incluyen notificaciones de aplicación, notificación de cambio por correo electrónico y notificación de resumen por correo electrónico.

Para obtener más información, consulte el capítulo sobre la colaboración del usuario en activos en la *Guía de Informatica Enterprise Data Catalog 10.2.2*.

Business Glossary Informe de asignación

A partir de la versión 10.2.2, puede aceptar o rechazar varios términos empresariales recomendados para un recurso.

Para obtener más información, consulte el capítulo sobre la realización de tareas de activo en la *Guía de Informatica Enterprise Data Catalog 10.2.2*.

Perfiles de sistema operativo

A partir de la versión 10.2.2, puede elegir un perfil del sistema operativo si no tiene un perfil del sistema operativo predeterminado. El servicio de integración de datos usa las credenciales de usuario del perfil del sistema operativo para realizar la detección de datos. La detección de datos incluye perfiles de columna y perfiles de detección del dominio de datos.

Para obtener más información acerca de cómo usar los perfiles del sistema operativo en Enterprise Data Catalog, consulte el capítulo sobre los conceptos de Enterprise Data Catalog en la *Guía del administrador de Informatica Catalog 10.2.2*.

API de REST

A partir de la versión 10.2.2, puede usar las siguientes API de REST de Informatica Enterprise Data Catalog:

- API de REST de términos empresariales. Puede devolver, actualizar o eliminar un término empresarial aceptado, inferido o rechazado.
- API de REST de eventos de catálogo. Puede acceder, actualizar o eliminar la configuración del usuario, la configuración del correo electrónico y las suscripciones de usuario.
- API de certificación de objetos. Puede enumerar, actualizar y eliminar las propiedades de certificación para un objeto.
- API de comentarios de objetos. Puede enumerar, crear, actualizar y eliminar comentarios, respuestas y votos para un objeto de datos.

- API de revisiones de objetos. Puede enumerar, crear, actualizar y eliminar revisiones, calificaciones y votos para una revisión.

Para obtener más información acerca de las API de REST, consulte la *Referencia de la API de REST de Informatica 10.2.2 Enterprise Data Catalog*.

Filtro de perfiles de datos y metadatos de origen

A partir de la versión 10.2.2, puede usar el filtro de metadatos de origen y filtro de perfiles de datos para especificar vistas y tablas de origen en la ejecución de un recurso. Cuando use estos filtros, Enterprise Data Catalog extraerá los metadatos de origen y del perfil de vistas y tablas específicas.

Para obtener más información acerca del filtro de perfiles de datos y metadatos de origen, consulte el capítulo sobre la administración de recursos en la *Guía del administrador de Informatica Catalog 10.2.2*.

Utilidad de analizador

A partir de la versión 10.2.2, Informatica brinda una utilidad de analizador independiente que se puede usar para extraer metadatos de recursos inaccesibles o sin conexión. La utilidad contiene un script que debe ejecutarse junto con los comandos asociados en una secuencia.

Para obtener más información acerca de la utilidad de analizador independiente, consulte el apéndice "Extracción de metadatos de recursos inaccesibles o sin conexión" de la *Guía del administrador de Informatica Catalog 10.2.2*.

Tipos de recursos

A partir de la versión 10.2.2, puede crear recursos para los siguientes tipos de origen de datos:

Google BigQuery

Puede extraer información de los metadatos, la relación y el linaje desde los siguientes activos en un origen de datos de Google BigQuery:

- Proyecto
- Conjunto de datos
- Tabla
- Ver

Para obtener más información acerca de cómo configurar un origen de datos de Google BigQuery, consulte la *Guía del administrador de Informatica Catalog 10.2.2*.

Workday

Puede extraer información de los metadatos, la relación y el linaje desde los siguientes activos en un origen de datos de Workday:

- Servicio
- Entidad
- Informe
- Operación
- Origen de datos
- Propiedad

- Objetos de negocio

Para obtener más información acerca de cómo configurar un origen de datos de Workday, consulte la *Guía del administrador de Informatica Catalog 10.2.2*.

Enterprise Data Lake

En esta sección se describen las nuevas características de Enterprise Data Lake en la versión 10.2.2.

Aplicar reglas activas

A partir de la versión 10.2.2, puede usar reglas activas en proyectos.

Las reglas activas son mapplets desarrollados mediante el uso de Developer tool. Puede usar reglas activas para aplicar transformaciones complejas como transformaciones de agregador y calidad de datos a hojas de cálculo para coincidencia y consolidación.

Una regla activa usa todas las filas dentro de un conjunto de datos como entrada. Puede seleccionar varias hojas de cálculo para usar como entradas en la regla. La aplicación agrega una hoja de cálculo que contiene la salida de la regla para el proyecto.

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Informatica Enterprise Data Lake 10.2.2*.

Eliminar filas duplicadas

A partir de la versión 10.2.2, puede eliminar las filas que contengan valores duplicados de una hoja de cálculo.

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Informatica Enterprise Data Lake 10.2.2*.

Agrupar y categorizar datos de columna

A partir de la versión 10.2.2, puede agrupar los valores similares en una columna y categorizarlos después según las recomendaciones de Enterprise Data Lake. La aplicación usa un algoritmo fonético para agrupar los valores similares y sugiere después que reemplace los valores menos frecuentes con el valor más frecuente.

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Informatica Enterprise Data Lake 10.2.2*.

Recomendaciones basadas en CLAIRE

A partir de la versión 10.2.2 la aplicación usa el motor de detección de aprendizaje del equipo CLAIRE incrustado para proporcionar recomendaciones al preparar los datos.

Cuando ve la página del proyecto, la aplicación muestra recomendaciones alternativas y adicionales derivadas de los orígenes de datos de nivel superior según un linaje de datos, así como relaciones de clave externa y principal.

Al seleccionar una columna en una hoja de cálculo durante la preparación de los datos, la aplicación muestra sugerencias para mejorar los datos según el tipo de datos de columna en el panel Resumen de columna.

Al realizar una operación de combinación en dos hojas de cálculo, la aplicación utiliza relaciones de clave externa y primaria para indicar un muestreo incompatible cuando se produce una superposición baja para los pares de claves deseados.

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Informatica Enterprise Data Lake 10.2.2*.

Agregación condicional

A partir de la versión de 10.2.2, puede usar la lógica de AND y OR para aplicar varias condiciones sobre los cálculos de IF que se usan cuando se crea una hoja de cálculo agregada en un proyecto.

- Use AND con todos los operadores para incluir más de una columna en una condición.
- Use OR con los operadores IS, IS NOT e IS BETWEEN para incluir más de un valor dentro de una columna en una condición.

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Informatica Enterprise Data Lake 10.2.2*.

Enmascaramiento de datos

A partir de la versión 10.2.2, Enterprise Data Lake se integra con Informatica Dynamic Data Masking, un producto de seguridad de datos, para habilitar el enmascaramiento de datos confidenciales en los activos de datos.

Para habilitar el enmascaramiento de datos en Enterprise Data Lake, configure el servidor de Dynamic Data Masking para aplicar reglas de enmascaramiento a activos de datos en Data Lake. También debe configurar el dominio de Informatica para habilitar Enterprise Data Lake para que se conecte con dicho servidor.

Dynamic Data Masking intercepta solicitudes enviadas a Data Lake desde Enterprise Data Lake y aplica las reglas de enmascaramiento a columnas en el activo solicitado. Cuando los usuarios de Enterprise Data Lake ven o realizan operaciones en columnas que contienen datos enmascarados, los datos reales se confunden total o parcialmente según las reglas de enmascaramiento aplicadas.

Para obtener más información, consulte el capítulo sobre datos de enmascaramiento confidenciales en la *Guía del administrador de Informatica Enterprise Data Lake 10.2.2*.

Localización

A partir de la versión 10.2.2, la interfaz de usuario admite el japonés. También puede usar caracteres no latinos en nombres y descripciones de proyecto.

Orígenes y destinos de la partición

A partir de la versión 10.2.2, Enterprise Data Lake puede leer datos desde orígenes particionados durante las operaciones de importación, publicación o copia. La aplicación también puede anexar datos a destinos particionados en Data Lake durante las operaciones de carga, copia, publicación e importación.

Agregar comentarios a pasos de la receta

A partir de la versión 10.2.2, puede agregar un comentario a un paso de la receta. Use comentarios para mejorar la colaboración y proporcionar detalles para cumplir con los requisitos de auditoría.

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Informatica Enterprise Data Lake 10.2.2*.

Guardar una receta como asignación

A partir de la versión 10.2.2, puede guardar una receta como asignación en lugar de publicar la receta y crear una nueva tabla de salida.

Puede guardar la asignación en el repositorio de modelos asociado con el servicio de Enterprise Data Lake o puede guardar la asignación en un archivo .xml. Los desarrolladores pueden usar Developer tool para revisar y modificar la asignación, y ejecutarla después cuando sea conveniente según la disponibilidad de recursos del sistema.

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Informatica Enterprise Data Lake 10.2.2*.

Amazon S3, ADLS, WASB, MapR-FS como orígenes de datos

A partir de la versión 10.2.2, puede preparar los datos en archivos almacenados en los siguientes orígenes de datos:

- Amazon S3
- MapR-FS
- Almacenamiento de Microsoft Azure Data Lake
- Windows Azure Storage Blob

Debe crear un recurso en Enterprise Data Catalog para cada origen de datos que contenga los datos que quiera preparar. Un recurso es un objeto de repositorio que representa un repositorio de metadatos u origen de datos externo. Los analizadores adjuntados a un recurso extraen los metadatos del recurso y los almacenan en Enterprise Data Catalog.

Para obtener información acerca de cómo crear recursos en Enterprise Data Catalog, consulte el capítulo sobre la administración de recursos en la *Guía del administrador de Informatica Catalog 10.2.2*.

Funciones estadísticas

A partir de la versión 10.2.2, puede aplicar las siguientes funciones estadísticas a columnas de una hoja de cálculo cuando prepara datos:

- AVG
- AVGIF
- COUNT
- COUNTIF
- COUNTDISTINCT
- COUNTDISTINCTIF
- MAX
- MAXIF
- MIN
- MINIF
- STDDEV
- STDDEVIF
- SUM
- SUMIF

- VARIANCE
- VARIANCEIF

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Informatica Enterprise Data Lake 10.2.2*.

Funciones de fecha y hora

A partir de la versión 10.2.2, puede aplicar las siguientes funciones de fecha y hora a las columnas de la hoja de cálculo cuando prepara los datos:

- ADD_TO_DATE
- CURRENT_DATETIME
- DATETIME
- DATE_DIFF
- DATE_TO_UNIXTIME
- EXTRACT_MONTH_NAME
- UNIXTIME_TO_DATE
- Convertir fecha en texto
- Convertir texto en fecha

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Informatica Enterprise Data Lake 10.2.2*.

Funciones matemáticas

A partir de la versión 10.2.2, puede aplicar las siguientes funciones matemáticas a las columnas cuando prepara los datos:

- EXP
- LN
- LOG
- PI
- POWER
- SQRT

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Informatica Enterprise Data Lake 10.2.2*.

Funciones de texto

A partir de la versión 10.2.2, puede aplicar las siguientes funciones de texto a las columnas cuando prepara los datos:

- ENDSWITH
- ENDSWITH_IGNORE_CASE
- FIND_IGNORE_CASE
- FIND_REGEX
- FIRST_CHARACTER_TO_NUMBER

- NUMBER_TO_CHARACTER
- PROPER_CASE
- REMOVE_NON_ALPHANUMERIC_CHARACTERS
- STARTSWITH
- STARTSWITH_IGNORE_CASE
- SUBSTITUTE_REGEX
- TRIM_ALL
- Convertir fecha en texto
- Convertir número en texto
- Convertir texto en fecha
- Convertir texto en número

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Informatica Enterprise Data Lake 10.2.2*.

Funciones de ventana

A partir de la versión 10.2.2 puede usar las funciones de ventana para realizar operaciones en grupos de fila dentro de una hoja de cálculo. El grupo de filas sobre el cuál actúa una función se llama "ventana", y se define con una clave de partición, un orden por clave y desplazamientos opcionales. Una función de ventana calcula un valor de devolución para cada fila de entrada dentro del contexto de la ventana.

Puede utilizar funciones de ventana para realizar las siguientes tareas:

- Recuperar datos de filas anteriores o posteriores.
- Calcular una suma acumulada o un porcentaje acumulado según un grupo de filas.
- Asignar un número de filas secuencial a cada fila de un grupo de filas.
- Reemplazar valores nulos en las filas con el valor no nulo precedente dentro de un grupo de filas.
- Generar identificadores de sesión que pueda usar para agrupar filas según un período de tiempo específico, como visitas a un sitio web registradas en un archivo de registro.

Puede aplicar varias funciones de ventana a una hoja de cálculo. Por ejemplo, es posible aplicar una función para calcular la suma de valores de cada fila a continuación de la fila actual dentro de una ventana y aplicar después otra función para calcular el promedio de los mismos valores.

Enterprise Data Lake agrega una columna que contiene los resultados de cada función que se aplica a la hoja de cálculo.

Para obtener más información, consulte el capítulo sobre preparación de datos en la *Guía del usuario de Informatica Enterprise Data Lake 10.2.2*.

Purgar eventos de auditoría

A partir de la versión 10.2.2, puede ejecutar el comando `infacmd edl purgeevents` para eliminar eventos de actividad del usuario de la base de datos de historial de auditorías. También puede ejecutar el comando para eliminar eventos del historial de proyectos de la base de datos.

Motor de ejecución de Spark

A partir de la versión 10.2.2, Enterprise Data Lake usa el motor de Spark para actividades de consumo alto de recursos como la publicación de un activo o para ejecutar mapplets de reglas activos que usen la

transformación de Python. El uso del motor de Spark para actividades de consumo alto de recursos brinda un mejor rendimiento y habilita una implementación de Enterprise Data Lake en Amazon Elastic MapReduce (EMR) para aprovechar el ajuste de escala automático.

Informatica Developer

En esta sección se describen las nuevas características de Developer tool en la versión 10.2.2.

Aplicaciones

A partir de la versión 10.2.2, puede crear aplicaciones incrementales. Una aplicación incremental es una aplicación que puede actualizarse mediante la implementación de una revisión de aplicación que actualiza un subconjunto de objetos. El servicio de integración de datos actualiza los objetos de la revisión mientras siguen ejecutándose los objetos de aplicación.

Si actualiza a la versión 10.2.2, las aplicaciones existentes se marcan como "aplicaciones completas". Puede seguir creando aplicaciones completas en la versión 10.2.2, pero no puede convertir una aplicación completa en una aplicación incremental.

Para obtener más información, consulte los capítulos sobre implementación de aplicaciones e implementación de revisiones de aplicación en la *Guía de Informatica Developer Tool 10.2.2*.

Asignaciones de Informatica

En esta sección se describen las nuevas características de las asignaciones de Informatica en la versión 10.2.2.

Tipos de datos

A partir de la versión 10.2.2, puede habilitar el modo de alta precisión en asignaciones por lote que se ejecutan en el motor de Spark. Con el motor de Spark, puede procesar los valores decimales con una precisión de hasta 38 dígitos.

Para obtener más información, consulte la *Guía del usuario de Informatica Big Data Management 10.2.2*.

Salidas de asignación

A partir de la versión 10.2.2, puede usar salidas de asignación en asignaciones por lote que se ejecuten como tareas de asignación en flujos de trabajo en el motor de Spark. Puede conservar las salidas de asignación en el repositorio de modelos o vincular las salidas de asignación a variables de flujo de trabajo.

Para obtener más información, consulte el capítulo "Salidas de asignación" en la *Guía de asignaciones de Informatica Developer 10.2.2* y el capítulo "Tarea de asignación" en la *Guía de flujos de trabajo de Informatica Developer 10.2.2*.

Parámetros de asignación

A partir de la versión 10.2.2, puede asignar parámetros de expresión a las expresiones de puerto en las transformaciones de agregación, expresión y clasificación que se ejecutan en los entornos nativos y no nativos.

Para obtener más información, consulte el capítulo sobre dónde se asignan los parámetros y "Asignaciones dinámicas" en la *Guía de asignaciones de Informatica Developer 10.2.2*.

Niveles del optimizador

A partir de la versión 10.2.2, puede configurar el nivel automático del optimizador para las asignaciones y las tareas de asignación. Con el nivel de optimizador automático, el servicio de integración de datos aplica optimizaciones según el modo de ejecución y los contenidos de la asignación.

El valor predeterminado del nivel del optimizador para las aplicaciones nuevas es Automático.

Cuando actualiza a la versión 10.2.2, los niveles del optimizador configurados en las asignaciones siguen siendo los mismos. Para usar el nivel automático del optimizador con asignaciones de actualización, debe cambiar manualmente el nivel del optimizador.

Para obtener más información, consulte el capítulo "Niveles del optimizador" en la *Guía de asignaciones de Informatica Developer 10.2.2*.

Sqoop

A partir de la versión 10.2.2, puede utilizar las siguientes nuevas características de Sqoop:

Compatibilidad con la extracción de datos incrementales

Puede configurar una asignación de Sqoop para realizar una extracción de datos incrementales según un ID o una marca de tiempo. Con la extracción de datos incrementales, Sqoop extrae solamente los datos que hayan cambiado desde la última extracción de datos. La extracción de datos incrementales aumenta el rendimiento de la asignación.

Soporte de conectividad de Vertica

Puede configurar Sqoop para que lea datos desde un origen de Vertica o escriba datos en un destino de Vertica.

Optimización del motor de Spark para asignaciones de transferencia de Sqoop

Cuando ejecuta una asignación de transferencia con un origen de Sqoop en el motor de Spark, el servicio de integración de datos optimiza el rendimiento de la asignación en los siguientes escenarios:

- Escribe datos en un destino de Hive que se creó con una consulta de DDL personalizada.
- Escribe datos en un destino de Hive existente que está particionado con una consulta de DLL personalizada o particionada, y con depósitos con una consulta de DDL personalizada.
- Escribe datos en un destino de Hive existente que está particionado y tiene depósitos.

Compatibilidad con el argumento --infaownername

Puede configurar el argumento --infaownername para indicar si Sqoop debe ajustarse al nombre del propietario para un objeto de datos.

Para obtener más información, consulte la *Guía del usuario de Informatica Big Data Management 10.2.2*.

Transformaciones de Informatica

En esta sección se describen las nuevas características de transformación de Informatica en la versión 10.2.2.

Transformación del validador de direcciones

En esta sección se describen las nuevas características de la transformación del validador de direcciones.

La transformación del validador de direcciones contiene funciones de dirección adicionales para los siguientes países:

Todos los países

A partir de la versión 10.2.2, la transformación del validador de direcciones admite la verificación de direcciones de una sola línea en todos los países para los que Informatica proporciona datos de referencia de direcciones.

En las versiones anteriores, la transformación admitía la verificación de direcciones de una sola línea de 26 países.

Para verificar una dirección de una sola línea, introduzca la dirección en el puerto Dirección completa. Si la dirección identifica un país para el que el alfabeto predeterminado preferido no es un alfabeto latino u occidental, use la propiedad predeterminada Alfabeto preferido en la transformación con la dirección.

Australia

A partir de la versión 10.2.2, puede configurar la transformación del validador de direcciones para que agregue enriquecimientos de direcciones a direcciones de Australia. Los enriquecimientos sirven para detectar los sectores geográficos y las regiones a las que el Australia Bureau of Statistics asigna las direcciones. Los sectores y las regiones incluyen distritos de realización de censo, bloques de malla y áreas estadísticas.

La transformación usa los siguientes puertos para entregar los enriquecimientos:

- Código de distrito de colección de censo 2006
- Identificador de archivo de direcciones nacionales geocodificadas
- Área de estadística de área metropolitana de capital de 5 dígitos
- Nombre de área de estadísticas de área metropolitana de capital
- Área de estadísticas de nivel 1 de 11 dígitos
- Área de estadísticas de nivel 1 de 7 dígitos
- Área de estadísticas de nivel 2 de 9 dígitos
- Área de estadísticas de nivel 2 de 5 dígitos
- Nombre de área de estadísticas de nivel 2
- Área de estadísticas de nivel 3 de 5 dígitos
- Nombre de área de estadísticas de nivel 3
- Área de estadísticas de nivel 4 de 3 dígitos
- Nombre de área de estadísticas de nivel 4
- Bloque de malla de 11 dígitos 2011
- Bloque de malla de 11 dígitos 2016

- Código de estado o territorio
- Nombre de estado o territorio
- Estado complementario de AU

Busque los puertos en el grupo de puertos complementarios de Australia.

Bolivia

A partir de la versión 10.2.2, la transformación del validador de direcciones mejora el análisis y la validación de las direcciones de Bolivia. Además, Informatica actualiza los datos de referencia de Bolivia.

La transformación también incluye las siguientes mejoras para Bolivia:

- Validación de direcciones a nivel de calle.
- Geocoordenadas a nivel intermedio de calle para las direcciones en las grandes ciudades.

Canadá

Informatica introduce las siguientes características y mejoras para Canadá:

Compatibilidad de la propiedad Descriptor global preferido en direcciones de Canadá

A partir de la versión 10.2.2, puede configurar la transformación del validador de direcciones para que devuelva la forma corta o la forma larga de un descriptor de elementos.

La transformación puede devolver la forma corta o larga de los siguientes descriptores:

- Descriptores de calle
- Valores direccionales
- Descriptores de edificios
- Descriptores de elementos por debajo del nivel de edificio

Para especificar el formato de salida de los descriptores, configure la propiedad Descriptor global preferido en la transformación. La propiedad se aplica a los descriptores de los idiomas inglés y francés. De forma predeterminada, la transformación devuelve el descriptor en el formato que especifiquen los datos de referencia. Si selecciona la opción PRESERVE INPUT en la propiedad, la propiedad Alfabeto preferido tendrá prioridad sobre la propiedad Descriptor global preferido.

Compatibilidad de CH y CHAMBER como descriptores de elementos por debajo del nivel de edificio

A partir de la versión 10.2.2, la transformación del validador de direcciones reconoce CH y CHAMBER como descriptores de elementos por debajo del nivel de edificio en las direcciones de Canadá.

Colombia

A partir de la versión 10.2.2, la transformación del validador de direcciones mejora el procesamiento de los datos de calles en las direcciones de Colombia. Además, Informatica actualiza los datos de referencia de Colombia.

Francia

A partir de la versión 10.2.2, Informatica introduce las siguientes mejoras para las direcciones de Francia:

- Informatica mejora los datos de referencia complementarios de Francia.
- La transformación del validador de direcciones asigna direcciones a las unidades IRIS en Francia con una mayor precisión. La transformación usa el número de casa en la dirección para verificar la unidad IRIS a la que pertenece la dirección. El uso de los números de casa puede mejorar la precisión de la asignación cuando la dirección está cerca del borde entre unidades diferentes.

Israel

A partir de la versión 10.2.2, Informatica presenta las siguientes características y mejoras para Israel:

Soporte multilingüe para direcciones de Israel

Puede configurar la transformación del validador de direcciones para devolver una dirección de Israel en idioma inglés o hebreo.

Utilice la propiedad Idioma preferido para seleccionar el idioma preferido para las direcciones que devuelve la transformación.

El idioma predeterminado para las direcciones de Israel es el hebreo. Para devolver la información de dirección en hebreo, establezca la propiedad Idioma preferido en DATABASE o ALTERNATIVE_1. Para devolver la información de dirección en inglés, establezca la propiedad en ENGLISH o ALTERNATIVE_2.

Compatibilidad con varios conjuntos de caracteres para direcciones de Israel

La transformación del validador de direcciones puede leer y escribir direcciones de Israel en conjuntos de caracteres en hebreo y latín.

Utilice la propiedad Codificación de caracteres preferida para seleccionar el conjunto de caracteres preferido para los datos de dirección.

El conjunto de caracteres predeterminado para las direcciones de Israel es Hebreo. Cuando la propiedad del script preferido se establece en Latin o Latin-1, la transformación transcribe datos de direcciones en hebreo en caracteres en latín.

Perú

A partir de la versión 10.2.2, la transformación del validador de direcciones valida las direcciones de Perú a nivel de número de casa. Además, Informatica actualiza los datos de referencia de Perú.

Suecia

A partir de la versión 10.2.2, la transformación del validador de direcciones mejora la verificación de los nombres de calles en las direcciones de Suecia.

La transformación mejora la verificación de los nombres de calles de las siguientes maneras:

- La transformación puede reconocer un nombre de calle que termine con el carácter G como alias del mismo nombre con los caracteres finales GATAN.
- La transformación puede reconocer un nombre de calle que termine con el carácter V como alias del mismo nombre con los caracteres finales VÄGEN.
- La transformación del validador de direcciones puede reconocer y corregir el nombre de una calle con un descriptor incorrecto cuando se usan la forma larga o la forma corta del descriptor.

Por ejemplo, la transformación puede corregir RUNIUSV o RUNIUSVÄGEN como RUNIUSGATAN en la siguiente dirección:

RUNIUSGATAN 7

SE-112 55 STOCKHOLM

Estados Unidos

A partir de la versión 10.2 HotFix 2, puede configurar la transformación del validador de direcciones para identificar direcciones de Estados Unidos que no reciben correo en uno o varios días de la semana.

Para identificar las direcciones, use el puerto *Días de no entrega*. El puerto contiene una cadena de siete dígitos que representa los días de la semana de domingo a sábado. Cada posición de la cadena representa un día diferente.

La transformación del validador de direcciones devuelve la primera letra del día de la semana en la posición correspondiente del puerto si la dirección no recibe correo ese día. La transformación devuelve un símbolo de guion en la posición correspondiente para otros días de la semana.

Por ejemplo, un valor de S---FS en el puerto Días de no entrega indica que una dirección no recibe correo en domingo, viernes y sábado.

Encuentre el puerto Días de no entrega en el grupo de puertos Específico de EE. UU. en el modelo básico. Para recibir datos en el puerto Días de no entrega, ejecute una transformación del validador de direcciones en modo certificado. La transformación lee los valores del puerto desde los archivos de base de datos USA5C129.MD y USA5C130.MD.

Para obtener información completa acerca de las características y las operaciones de la versión del motor de software de verificación de direcciones de la versión 10.2.2, consulte la *Guía de desarrollador de Informatica Address Verification 5.14.0*.

Transformación de estrategia de actualización

A partir de la versión 10.2.2, puede usar una transformación de estrategia de actualización en una asignación que se ejecuta en el motor de Spark para actualizar destinos relacionales.

Anteriormente, podía usar una transformación de estrategia de actualización en una asignación que se ejecutaba en el motor de Spark solo para actualizar destinos de Hive.

Para obtener más información, consulte el capítulo "Transformación de estrategia de actualización" en la *Guía de transformaciones de Developer*.

Adaptadores de PowerExchange para Informatica

En esta sección se describen las nuevas características del adaptador de Informatica de la versión 10.2.2.

PowerExchange for Amazon Redshift

A partir de la versión 10.2.2, PowerExchange for Amazon Redshift incluye las siguientes características:

- Puede leer datos de las siguientes regiones y escribirlos en ellas:
 - China (Ningxia)
 - UE (París)
- Puede utilizar objetos de Amazon Redshift como orígenes y destinos dinámicos en una asignación.
- Puede usar valores de octal de caracteres ASCII imprimibles y no imprimibles como DELIMITER o QUOTE.
- Puede introducir comandos de pre-SQL y post-SQL para ejecutar consultas para los objetos de origen y destino de una asignación.
- Puede definir una consulta SQL para los objetos de datos de lectura de una asignación para reemplazar la consulta predeterminada. Puede especificar una instrucción SQL que sea compatible con la base de datos de Amazon Redshift.
- Puede especificar el tamaño máximo de un objeto de Amazon S3 en bytes cuando descarga objetos grandes de Amazon S3 en varias partes.
- Puede leer valores únicos al leer datos desde un origen de Amazon Redshift.

- Cuando carga un objeto en Amazon S3, puede especificar el tamaño mínimo del objeto y el número de subprocesos para cargar los objetos en paralelo como un conjunto de piezas independientes.
- Puede elegir conservar la tabla de destino existente, reemplazar una tabla de destino en tiempo de ejecución o crear una nueva tabla de destino si no existe en el destino.
- Puede configurar las transformaciones de estrategia de actualización para un destino de Amazon Redshift en el entorno nativo.
- Cuando escribe datos en Amazon Redshift, puede reemplazar el esquema de tabla de destino de Amazon Redshift y el nombre de la tabla durante el tiempo de ejecución.
- Cuando el tipo de conexión es ODBC, el servicio de integración de datos puede insertar la lógica de transformación en orígenes y destinos de Amazon Redshift usando la optimización de delegación de tareas completa y del lado de origen.
- Puede usar la encriptación del lado del servidor con el servicio de administración de claves de AWS (AWS KMS) en Amazon EMR versión 5.16 y Cloudera CDH versión 5.15 y 5.16.
- PowerExchange for Amazon Redshift admite AWS SDK for Java versión 1.11.354.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for Amazon Redshift 10.2.2*.

PowerExchange for Amazon S3

A partir de la versión 10.2.2, PowerExchange for Amazon S3 incluye las siguientes características:

- Puede leer datos de las siguientes regiones y escribirlos en ellas:
 - China (Ningxia)
 - UE (París)
 - AWS GovCloud (EE. UU.)
- Puede utilizar objetos de Amazon S3 como orígenes y destinos dinámicos en una asignación.
- Cuando se ejecuta una asignación en el entorno nativo o en el motor de Spark para leer datos desde un archivo Avro, sin formato, JSON, ORC o Parquet, puede usar caracteres comodines para especificar el nombre del directorio de origen o el nombre del archivo de origen.
- Puede agregar una sola etiqueta o varias etiquetas a los objetos almacenados en el depósito de Amazon S3 para categorizar los objetos. Cada etiqueta contiene un par clave-valor. Puede especificar los pares clave-valor o especificar la ruta de archivo absoluta que contiene los pares clave-valor.
- Puede especificar el tamaño de umbral máximo para descargar un objeto de Amazon S3 en varias partes.
- Cuando carga un objeto en Amazon S3, puede especificar el tamaño mínimo del objeto y el número de subprocesos para cargar los objetos en paralelo como un conjunto de piezas independientes.
- Cuando crea una operación de lectura o escritura de un objeto de datos, puede leer los datos presentes en el puerto FileName que contiene el nombre del extremo y la ruta de origen del archivo.
- Puede agregar nuevas columnas o modificar las existentes directamente en la pestaña Puerto cuando crea una operación de lectura o escritura de objeto de datos.
- Puede copiar las columnas de las transformaciones de origen, las transformaciones de destino o cualquier otra transformación desde la pestaña Puerto y pegarlas directamente en la operación de lectura o escritura del objeto de datos cuando crea una asignación para leer o escribir en un archivo Avro, JSON, ORC o Parquet.
- Puede actualizar el formato de archivo Amazon S3 sin perder los metadatos de columna en el campo Esquema de las propiedades de proyección de columna incluso después de configurar las propiedades de proyección de columna para otro formato de archivo de Amazon S3.

- Puede usar la encriptación del lado del servidor con el servicio de administración de claves de AWS (AWS KMS) en Amazon EMR versión 5.16 y Cloudera CDH versión 5.15 y 5.16.
- PowerExchange for Amazon S3 admite AWS SDK for Java versión 1.11.354.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for Amazon S3 10.2.2*.

PowerExchange for Google BigQuery

A partir de la versión 10.2.2, puede crear un destino de Google BigQuery usando la opción de botón derecho **Crear destino**.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for Google BigQuery 10.2.2*.

PowerExchange for HBase

A partir de la versión 10.2.2, PowerExchange for HBase incluye estas nuevas características:

- Cuando crea un objeto de datos de HBase, puede seleccionar un perfil del sistema operativo para aumentar la seguridad y aislar el entorno del usuario en tiempo de diseño cuando se importan metadatos y se obtiene una vista previa de ellos desde un clúster de Hadoop.

Nota: Puede elegir un perfil del sistema operativo si el servicio de acceso a metadatos está configurado para usar perfiles del sistema operativo. El servicio de acceso a metadatos ejecuta los metadatos con el perfil del sistema operativo predeterminado asignado al usuario. Puede cambiar el perfil del sistema operativo en la lista de perfiles del sistema operativo disponibles.

- Puede utilizar objetos de HBase como orígenes y destinos dinámicos en una asignación.
- Puede ejecutar una asignación en el motor de Spark para buscar datos en un recurso de HBase.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for HBase 10.2.2*.

PowerExchange for HDFS

A partir de la versión 10.2.2, PowerExchange for HDFS incluye estas nuevas características:

- Cuando crea un objeto de datos de archivo complejo, puede seleccionar un perfil del sistema operativo para aumentar la seguridad y aislar el entorno del usuario en tiempo de diseño cuando se importan metadatos y se obtiene una vista previa de ellos desde un clúster de Hadoop.

Nota: Puede elegir un perfil del sistema operativo si el servicio de acceso a metadatos está configurado para usar perfiles del sistema operativo. El servicio de acceso a metadatos ejecuta los metadatos con el perfil del sistema operativo predeterminado asignado al usuario. Puede cambiar el perfil del sistema operativo en la lista de perfiles del sistema operativo disponibles.

- Cuando se ejecuta una asignación en el entorno nativo o en el motor de Spark para leer datos desde un objeto de datos de archivo complejo, puede usar caracteres comodines para especificar el nombre del directorio de origen o el nombre del archivo de origen.
Puede usar los siguientes caracteres comodines:

? (Signo de interrogación)

El carácter de signo de interrogación (?) permite una ocurrencia de cualquier carácter.

* (Asterisco)

El carácter de asterisco (*) permite una ocurrencia nula o más de una ocurrencia de cualquier carácter.

- Puede utilizar objetos de archivos complejos como orígenes y destinos dinámicos en una asignación.
- Puede utilizar objetos de archivos complejos para leer datos desde un sistema de archivos complejo o escribirlos en él.
- Cuando ejecuta una asignación en el entorno nativo o en el motor de Spark para escribir datos en un objeto de datos de archivo complejo, puede sobrescribir los datos de destino. El servicio de integración de datos elimina los datos de destino antes de escribir nuevos datos.
- Cuando crea una operación de lectura o escritura de un objeto de datos, puede leer los datos presentes en el puerto FileName que contiene el nombre del extremo y la ruta de origen del archivo.
- Ahora puede ver las operaciones de objeto de datos inmediatamente después de crear la operación de lectura o escritura de objeto de datos.
- Puede agregar nuevas columnas o modificar las existentes cuando crea una operación de lectura o escritura de objeto de datos.
- Puede copiar las columnas de las transformaciones de origen, las transformaciones de destino o cualquier otra transformación, y pegarlas en la operación de lectura o escritura del objeto de datos directamente cuando lee o escribe en un archivo Avro, JSON, ORC o Parquet.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for HDFS 10.2.2*.

PowerExchange for Hive

A partir de la versión 10.2.2, PowerExchange for Hive incluye estas nuevas características:

- Puede configurar las siguientes opciones de la estrategia de esquema de destino para un destino de Hive:
 - RETAIN. Mantener esquema de destino existente
 - CREATE. Crear o reemplazar una tabla en tiempo de ejecución
 - APPLYNEWCOLUMNS. Modificar tabla y aplicar nuevas columnas solo
 - APPLYNEWSHEMA. Modificar tabla y aplicar nuevo esquema
 - FAIL. Error de asignación si el esquema de destino es diferente
 - Asignar parámetro
- Puede truncar un destino de Hive particionado interno o externo antes de cargar datos. Esta opción se aplica cuando ejecuta la asignación en el entorno de Hadoop.
- Puede crear una transformación de escritura o lectura para Hive en modo nativo para leer datos desde el origen de Hive o escribirlos en el destino de Hive.
- Cuando escribe datos en un destino de Hive, puede configurar las siguientes propiedades en una conexión de Hive:
 - Directorio de transferencia provisional de Hive en HDFS. Representa el directorio de HDFS para las tablas de transferencia provisional de Hive. Esta opción se aplica y se requiere cuando se escriben datos en un destino de Hive en el entorno nativo.
 - Nombre de la base de datos de transferencia provisional de Hive. Representa el espacio de nombres para las tablas de transferencia provisional de Hive. Esta opción se aplica cuando ejecuta una asignación en el entorno nativo para escribir datos en un destino de Hive. Si ejecuta la asignación en el motor de Blaze o Spark, no tendrá que configurar el nombre de la base de datos de transferencia provisional de Hive en la conexión de Hive. El servicio de integración de datos utiliza el valor que configure en la conexión de Hadoop.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for Hive 10.2.2*.

PowerExchange for MapR-DB

A partir de la versión 10.2.2, cuando crea un objeto de base de datos de HBase para MapR-DB, puede seleccionar un perfil del sistema operativo para aumentar la seguridad y aislar el entorno del usuario en tiempo de diseño cuando se importan metadatos y se obtiene una vista previa de ellos desde un clúster de Hadoop.

Nota: Puede elegir un perfil del sistema operativo si el servicio de acceso a metadatos está configurado para usar perfiles del sistema operativo. El servicio de acceso a metadatos ejecuta los metadatos con el perfil del sistema operativo predeterminado asignado al usuario. Puede cambiar el perfil del sistema operativo en la lista de perfiles del sistema operativo disponibles.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for MapR-DB 10.2.2*.

PowerExchange for Microsoft Azure Blob Storage

A partir de la versión 10.2.2, PowerExchange for Microsoft Azure Blob Storage incluye las siguientes funcionalidades:

- Puede ejecutar asignaciones en el entorno de Azure Databricks.
- Puede configurar los puntos de conexión de Microsoft Azure del gobierno de Estados Unidos.
- Puede comprimir datos en los siguientes formatos cuando lee datos desde Microsoft Azure Blob Storage o los escribe allí:
 - Ninguno
 - Deflate
 - Gzip
 - Bzip2
 - Lzo
 - Snappy
- Puede usar objetos de Microsoft Azure Blob Storage como orígenes y destinos dinámicos en una asignación.
- Puede leer el nombre del archivo desde el que el servicio de integración de datos lee los datos en tiempo de ejecución en el entorno nativo.
- Puede configurar la ruta de acceso relativa en **Reemplazo de contenedor de Blob** en las propiedades avanzadas de origen y destino.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for Microsoft Azure Blob Storage 10.2.2*.

PowerExchange for Microsoft Azure Cosmos DB SQL API

A partir de la versión 10.2.2, PowerExchange for Microsoft Azure Cosmos DB SQL API incluye las siguientes funcionalidades:

- Puede ejecutar asignaciones en el entorno de Azure Databricks. Existe compatibilidad de Databricks con PowerExchange for Microsoft Azure Cosmos DB SQL API para vista previa técnica. La funcionalidad de vista previa técnica es compatible, pero no se garantiza ni está preparada para usarla en producción. Informatica recomienda utilizar estas características solo en entornos que no sean de producción.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for Microsoft Azure Cosmos DB SQL API 10.2.2*.

PowerExchange for Microsoft Azure Data Lake Store

A partir de la versión 10.2.2, PowerExchange for Microsoft Azure Data Store incluye las siguientes características:

- Puede ejecutar asignaciones en el entorno de Azure Databricks.
- Puede utilizar tipos de datos complejos, como array, struct y map, en las asignaciones que se ejecutan en el entorno de Hadoop. Con tipos de datos complejos, el correspondiente motor lee, procesa y escribe directamente los datos jerárquicos en archivos complejos Avro, JSON y Parquet. Para un origen de estructura inteligente, puede configurar solo la operación de lectura.
- Puede crear asignaciones para leer y escribir archivos Avro y Parquet que contienen solo tipos de datos primitivos en el entorno nativo.
- Puede seleccionar un directorio como origen en una asignación para leer varios archivos desde el directorio.
- Puede usar objetos de Microsoft Azure Data Lake Store como orígenes y destinos dinámicos en una asignación.
- Puede crear un destino de Microsoft Azure Data Lake Store usando la opción Crear destino.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for Microsoft Azure Data Lake Store 10.2.2*.

PowerExchange for Microsoft Azure SQL Data Warehouse

A partir de la versión 10.2.2, PowerExchange for Microsoft Azure SQL Data Warehouse incluye las siguientes funcionalidades:

- Puede ejecutar asignaciones en el entorno de Azure Databricks.
- Puede configurar los puntos de conexión de Microsoft Azure del gobierno de Estados Unidos en asignaciones que se ejecutan en el entorno nativo y en el motor de Spark.
- Puede generar archivos de error en el contenedor de Microsoft Azure Blob Storage. Los archivos de error contienen filas rechazadas y la causa relacionada con esas filas.
- Puede definir el tamaño de lote en las propiedades avanzadas del destino en el entorno nativo.
- Puede configurar la optimización de delegación de tareas completa para la lógica de transformación de delegación de tareas en las bases de dato de origen y de destino. Use la optimización de delegación de tareas para mejorar el rendimiento de la tarea usando los recursos de base de datos.
- Puede usar objetos de Microsoft Azure SQL Data Warehouse como orígenes y destinos dinámicos en una asignación.

La optimización de delegación de tareas completa y la funcionalidad de asignaciones dinámicas para PowerExchange for Microsoft Azure SQL Data Warehouse está disponible para vista previa técnica. La funcionalidad de vista previa técnica es compatible, pero no se garantiza ni está preparada para usarla en producción. Informatica recomienda utilizar estas características solo en entornos que no sean de producción.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for Microsoft Azure SQL Data Warehouse 10.2.2*.

PowerExchange for Salesforce

A partir de la versión 10.2.2, PowerExchange for Salesforce incluye estas nuevas características:

- Puede utilizar la versión 43.0 y 44.0 de la API de Salesforce para crear una conexión de Salesforce y acceder a los objetos de Salesforce.
- Puede configurar OAuth para las conexiones de Salesforce.
- Puede configurar el filtro de expresiones nativas para la operación del objeto de datos de origen.
- Puede parametrizar las siguientes propiedades de la operación de lectura para un objeto de datos Salesforce:
 - Condición de filtro de SOQL
 - Marca de hora de inicio de CDC
 - Marca de hora de finalización de CDC
 - Tamaño de fragmentación de clave principal
 - ID de startRow de fragmentación de clave principal

Puede parametrizar las siguientes propiedades de la operación de escritura para un objeto de datos de Salesforce:

- Establezca el prefijo para los archivos de errores y de éxito de BULK
- Directorio de archivos de éxito de SFDC
- Establezca la ubicación de los archivos de error de BULK

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for Salesforce 10.2.2*.

PowerExchange for Snowflake

A partir de la versión 10.2.2, PowerExchange for Snowflake incluye estas nuevas características:

- Puede configurar la autenticación SSO de Okta especificando los detalles de la autenticación en los parámetros de JDBC URL de la conexión de Snowflake.
- También puede configurar un reemplazo de SQL para reemplazar la consulta SQL predeterminada utilizada para extraer datos del origen de Snowflake. Especifique el reemplazo de SQL en las propiedades de operación de lectura del objeto de datos de Snowflake.
- Puede elegir comprimir los archivos antes de escribir en las tablas de Snowflake y optimizar el rendimiento de escritura. En las propiedades avanzadas. Puede establecer el parámetro de compresión en On u Off en el campo **Parámetros adicionales de tiempo de ejecución de escritura** en las propiedades avanzadas de la operación de escritura del objeto de datos de Snowflake.
- El servicio de integración de datos usa las API del conector de Snowflake Spark para ejecutar asignaciones de Snowflake en el motor de Spark.
- Puede leer datos y escribir datos en Snowflake habilitados para los datos de almacenamiento provisional en Azure o Amazon.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for Snowflake 10.2.2*.

PowerExchange para la API del transportador paralelo de Teradata

A partir de la versión 10.2.2, PowerExchange for Teradata Parallel Transporter API incluye las siguientes funciones en las propiedades del destino avanzadas:

- Puede especificar un carácter de reemplazo para usar en lugar de un carácter unicode de Teradata no compatible en la base de datos de Teradata mientras se cargan datos en los destinos.
- Si especificó un carácter usado en lugar de uno no compatible al cargar datos en los destinos de Teradata, puede especificar las versiones 8.x - 13.x o 14.x y posterior para la base de datos de Teradata de destino. Use este atributo junto con el atributo de carácter de reemplazo. El servicio de integración de datos omite este atributo si no especificó un carácter de reemplazo mientras cargaba los datos en los destinos de Teradata.
- Cuando escribe datos en Teradata, puede reemplazar el esquema de tabla de destino de Teradata y el nombre de tabla durante el tiempo de ejecución.

Para obtener más información, consulte la *Guía del usuario de Informatica PowerExchange for Teradata Parallel Transporter API 10.2.2*.