



Informatica®
10.5

Guide de configuration du scanner d'Enterprise Data Catalog

© Copyright Informatica LLC 2015, 2022

Ce logiciel et la documentation associée sont fournis uniquement sous un accord de licence séparé contenant des restrictions d'utilisation et de divulgation. Il est interdit de reproduire ou de transmettre sous quelle que forme et par quel que moyen que ce soit (électronique, photocopie, enregistrement ou autre) tout ou partie de ce document sans le consentement préalable d'Informatica LLC.

Informatica et le logo Informatica sont des marques ou des marques déposées d'Informatica LLC aux États-Unis et dans de nombreux autres pays. La liste actuelle des marques de commerce de Informatica est disponible sur le Web à l'adresse <https://www.informatica.com/trademarks.html>. Les autres noms de société ou de produit peuvent être des marques de commerce ou des marques déposées de leurs détenteurs respectifs.

U.S. GOVERNMENT RIGHTS Les programmes, les logiciels, les bases de données et les documents connexes et les données techniques fournis aux clients du gouvernement américain sont des « logiciels commerciaux » ou des « données techniques commerciales », conformément au règlement fédéral sur les acquisitions et aux règlements supplémentaires propres à l'Agence. En tant que tel, l'utilisation, la duplication, la divulgation, la modification et l'adaptation sont assujetties aux restrictions et aux conditions de licence énoncées dans le contrat gouvernemental applicable et, dans la mesure applicable par les termes du contrat gouvernemental, les droits additionnels énoncés dans la réglementation FAR 52.227-19, licence de logiciel d'ordinateur commercial.

Des portions de ce logiciel et/ou de la documentation sont sujettes au copyright détenu par des tierces parties, dont Copyright DataDirect Technologies. Tous droits réservés. Copyright © Sun Microsystems. Tous droits réservés. Copyright © RSA Security Inc. Tous droits réservés. Copyright © Ordinal Technology Corp. Tous droits réservés. Copyright © Aandacht c.v. Tous droits réservés. Copyright Genivia, Inc. Tous droits réservés. Copyright Isomorphic Software. Tous droits réservés. Copyright © Meta Integration Technology, Inc. Tous droits réservés. Copyright © Intalio. Tous droits réservés. Copyright © Oracle. Tous droits réservés. Copyright © Adobe Systems Incorporated. Tous droits réservés. Copyright © DataArt, Inc. Tous droits réservés. Copyright © ComponentSource. Tous droits réservés. Copyright © Microsoft Corporation. Tous droits réservés. Copyright © Rogue Wave Software, Inc. Tous droits réservés. Copyright © Teradata Corporation. Tous droits réservés. Copyright © Yahoo! Inc. Tous droits réservés. Copyright © Glyph & Cog, LLC. Tous droits réservés. Copyright © Thinkmap, Inc. Tous droits réservés. Copyright © Clearpace Software Limited. Tous droits réservés. Copyright © Information Builders, Inc. Tous droits réservés. Copyright © OSS Nokalva, Inc. Tous droits réservés. Copyright Edifecs, Inc. Tous droits réservés. Copyright Cleo Communications, Inc. Tous droits réservés. Copyright © International Organization for Standardization 1986. Tous droits réservés. Copyright © ej-technologies GmbH. Tous droits réservés. Copyright © Jaspersoft Corporation. Tous droits réservés. Copyright © International Business Machines Corporation. Tous droits réservés. Copyright © yWorks GmbH. Tous droits réservés. Copyright © Lucent Technologies. Tous droits réservés. Copyright © University of Toronto. Tous droits réservés. Copyright © Daniel Veillard. Tous droits réservés. Copyright © Unicode, Inc. Copyright IBM Corp. Tous droits réservés. Copyright © MicroQuill Software Publishing, Inc. Tous droits réservés. Copyright © PassMark Software Pty Ltd. Tous droits réservés. Copyright © LogiXML, Inc. Tous droits réservés. Copyright © 2003-2010 Lorenzi Davide. Tous droits réservés. Copyright © Red Hat, Inc. Tous droits réservés. Copyright © The Board of Trustees of the Leland Stanford Junior University. Tous droits réservés. Copyright © EMC Corporation. Tous droits réservés. Copyright © Flexera Software. Tous droits réservés. Copyright © Jinfonet Software. Tous droits réservés. Copyright © Apple Inc. Tous droits réservés. Copyright © Telerik Inc. Tous droits réservés. Copyright © BEA Systems. Tous droits réservés. Copyright © PDFlib GmbH. Tous droits réservés. Copyright © Orientation in Objects GmbH. Tous droits réservés. Copyright © Tanuki Software, Ltd. Tous droits réservés. Copyright © Ricebridge. Tous droits réservés. Copyright © Sencha, Inc. Tous droits réservés. Copyright © Scalable Systems, Inc. Tous droits réservés. Copyright © jQWidgets. Tous droits réservés. Copyright © Tableau Software, Inc. Tous droits réservés. Copyright © MaxMind, Inc. Tous droits réservés. Copyright © TMate Software s.r.o. Tous droits réservés. Copyright © MapR Technologies Inc. Tous droits réservés. Copyright © Amazon Corporate LLC. Tous droits réservés. Copyright © Highsoft. Tous droits réservés. Copyright © Python Software Foundation. Tous droits réservés. Copyright © BeOpen.com. Tous droits réservés. Copyright © CNRI. Tous droits réservés.

Ce produit inclut des logiciels développés par Apache Software Foundation (<http://www.apache.org/>), et/ou d'autres logiciels sous licence et sous diverses versions Apache License (la « Licence »). Vous pouvez obtenir une copie de ces licences à l'adresse suivante : <http://www.apache.org/licenses/>. Sauf dispositions contraires de la loi en vigueur ou accord écrit, le logiciel distribué sous cette licence est livré « EN LÉTAT », SANS GARANTIE NI CONDITION D'AUCUNE SORTE, expresse ou implicite. Se reporter aux Licences pour la langue spécifique régissant les droits et limitations dans le cadre des Licences.

Ce produit inclut des logiciels développés par Mozilla (<http://www.mozilla.org/>), copyright de logiciel The JBoss Group, LLC, tous droits réservés ; copyright de logiciel © 1999-2006 de Bruno Lowagie et Paulo Soares et d'autres logiciels sous licence et sous diverses versions du GNU Lesser General Public License Agreement, accessible sur <http://www.gnu.org/licenses/lgpl.html>. Les matériaux sont fournis gratuitement par Informatica, « en l'état », sans garantie d'aucune sorte, expresse ou implicite, notamment les garanties implicites de conformité légale et d'usage normal.

Le produit inclut les logiciels ACE(TM) et TAO(TM), copyright Douglas C. Schmidt et son groupe de recherche à Washington University, University of California, Irvine, et Vanderbilt University, Copyright (©) 1993-2006, tous droits réservés.

Ce produit inclut des logiciels développés par OpenSSL Project pour une utilisation dans OpenSSL Toolkit (copyright The OpenSSL Project. Tous droits réservés) et la redistribution de ce logiciel est sujette aux termes publiés sur <http://www.openssl.org> et <http://www.openssl.org/source/license.html>.

Ce produit inclut le logiciel Curl, copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>. Tous droits réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://curl.haxx.se/docs/copyright.html>. L'autorisation d'utiliser, copier, modifier et distribuer ce logiciel à toute fin, avec ou sans rémunération, est accordée par les présentes, à la condition que la notification de copyright ci-dessus et cette notification d'autorisation apparaissent dans toutes les copies.

Le produit inclut des logiciels sous copyright 2001-2005 (©) MetaStuff, Ltd. Tous droits réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://www.dom4j.org/license.html>.

Le produit inclut des logiciels sous copyright © 2004-2007, The Dojo Foundation. Tous droits réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://dojotoolkit.org/license>.

Ce produit inclut le logiciel ICU sous copyright de International Business Machines Corporation et autres. Tous droits réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://source.icu-project.org/repos/icu/icu/trunk/license.html>.

Ce produit inclut des logiciels sous copyright © 1996-2006 Per Bothner. Tous droits réservés. Votre droit à utiliser de tels matériels est défini dans la licence qui peut être consultée sur <http://www.gnu.org/software/kawa/Software-License.html>.

Ce produit inclut le logiciel OSSP UUID sous copyright © 2002 Ralf S. Engelschall, copyright © 2002 The OSSP Project Copyright © 2002 Cable & Wireless Deutschland. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://www.opensource.org/licenses/mit-license.php>.

Ce produit inclut des logiciels développés par Boost (<http://www.boost.org/>) ou sous licence de logiciel Boost. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur http://www.boost.org/LICENSE_1_0.txt.

Ce produit inclut des logiciels sous copyright © 1997-2007 University of Cambridge. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://www.pcre.org/license.txt>.

Ce produit inclut des logiciels sous copyright © 2007 The Eclipse Foundation. Tous droits réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://www.eclipse.org/org/documents/epl-v10.php> et <http://www.eclipse.org/org/documents/edl-v10.php>.

Ce produit comprend des logiciels sous licence dont les conditions se trouvent aux adresses : <http://www.tcl.tk/software/tcltk/license.html>, <http://www.bosrup.com/web/overlib/?License>, <http://www.stlport.org/doc/license.html>, <http://asm.ow2.org/license.html>, <http://www.cryptix.org/LICENSE.TXT>, <http://www.cryptix.org/LICENSE.TXT>, <http://www.cryptix.org/LICENSE.TXT>.

hsqldb.org/web/hsqLicense.html, <http://httpunit.sourceforge.net/doc/license.html>, <http://jung.sourceforge.net/license.txt>, http://www.gzip.org/zlib/zlib_license.html, <http://www.openldap.org/software/release/license.html>, <http://www.libssh2.org>, <http://slf4j.org/license.html>, <http://www.sente.ch/software/OpenSourceLicense.html>, <http://fusesource.com/downloads/license-agreements/fuse-message-broker-v-5-3-1-license-agreement> ; <http://antlr.org/license.html> ; <http://aopalliance.sourceforge.net/> ; <http://www.bouncycastle.org/license.html> ; <http://www.jgraph.com/jgraphdownload.html> ; <http://www.jcraft.com/jsch/LICENSE.txt> ; http://jotm.objectweb.org/bsd_license.html ; <http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231> ; <http://www.slf4j.org/license.html> ; <http://nanoxml.sourceforge.net/orig/copyright.html> ; <http://www.json.org/license.html> ; <http://forge.ow2.org/projects/jaservice/> , <http://www.postgresql.org/about/licence.html>, <http://www.sqlite.org/copyright.html>, <http://www.tcl.tk/software/tcltk/license.html>, <http://www.jaxen.org/faq.html>, <http://www.jdom.org/docs/faq.html>, <http://www.slf4j.org/license.html> ; <http://www.iodbc.org/dataspace/iodbc/wiki/IODBC/License> ; <http://www.keplerproject.org/md5/license.html> ; <http://www.toedter.com/en/jcalendar/license.html> ; <http://www.edankert.com/bounce/index.html> ; <http://www.net-snmp.org/about/license.html> ; <http://www.openmdx.org/#FAQ> ; http://www.php.net/license/3_01.txt ; <http://srp.stanford.edu/license.txt> ; <http://www.schneier.com/blowfish.html> ; <http://www.jmock.org/license.html> ; <http://xsom.java.net> ; <http://benalman.com/about/license/> ; <https://github.com/CreateJS/EaselJS/blob/master/src/easeljs/display/Bitmap.js> ; <http://www.h2database.com/html/license.html#summary> ; <http://jsoncpp.sourceforge.net/LICENSE> ; <http://jdbc.postgresql.org/license.html> ; <http://protobuf.googlecode.com/svn/trunk/src/google/protobuf/descriptor.proto> ; <https://github.com/rantav/hector/blob/master/LICENSE> ; <http://web.mit.edu/Kerberos/krb5-current/doc/mitK5license.html> ; <http://jibx.sourceforge.net/jibx-license.html> ; <https://github.com/lyokato/libgeohash/blob/master/LICENSE> ; <https://github.com/hjiang/jsonxx/blob/master/LICENSE> ; <https://code.google.com/p/lz4/> ; <https://github.com/jedisct1/libsodium/blob/master/LICENSE> ; <http://one-jar.sourceforge.net/index.php?page=documents&file=license> ; <https://github.com/EsotericSoftware/kryo/blob/master/license.txt> ; <http://www.scala-lang.org/license.html> ; <https://github.com/tinkerpop/blueprints/blob/master/LICENSE.txt> ; <http://gee.cs.oswego.edu/dl/classes/EDU/oswego/cs/dl/util/concurrent/intro.html> ; <https://aws.amazon.com/asl/> ; <https://github.com/twbs/bootstrap/blob/master/LICENSE> ; <https://sourceforge.net/p/xmlunit/code/HEAD/tree/trunk/LICENSE.txt> ; <https://github.com/documentcloud/underscore-contrib/blob/master/LICENSE>, and <https://github.com/apache/hbase/blob/master/LICENSE.txt>.

Ce produit inclut un logiciel sous licence Academic Free License (<http://www.opensource.org/licenses/afl-3.0.php>), licence Common Development Distribution License (<http://www.opensource.org/licenses/cddl1.php>) licence Common Public License (<http://www.opensource.org/licenses/cpl1.0.php>), licence Sun Binary Code License Agreement Supplemental License Terms, licence BSD (<http://www.opensource.org/licenses/bsd-license.php>), le nouvelle licence BSD License (<http://opensource.org/licenses/BSD-3-Clause>), la licence MIT (<http://www.opensource.org/licenses/mit-license.php>), la licence Artistic License (<http://www.opensource.org/licenses/artistic-license-1.0>) et la licence publique du développeur initial Version 1.0 (<http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/>).

Ce produit inclut des logiciels sous copyright © 2003-2006 Joe Walnes, 2006-2007 XStream Committers. Tous droits réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://xstream.codehaus.org/license.html>. Ce produit inclut des logiciels développés par Indiana University Extreme! Lab. Pour plus d'informations, veuillez vous rendre sur <http://www.extreme.indiana.edu/>.

Ce produit inclut des logiciels sous copyright © 2013 Frank Balluffi et Markus Moeller. Tous droits réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions de la licence MIT.

Consultez les brevets applicables à l'adresse <https://www.informatica.com/legal/patents.html>.

EXCLUSION DE RESPONSABILITÉ : Informatica LLC fournit cette documentation « en l'état », sans garantie d'aucune sorte, explicite ou implicite, notamment les garanties implicites de non-infraction, de conformité légale ou d'usage normal. Informatica LLC ne garantit pas que ce logiciel et cette documentation sont exempts d'erreurs. Les informations fournies dans ce logiciel ou cette documentation peuvent inclure des inexactitudes techniques ou des erreurs typographiques. Les informations contenues dans ce logiciel et sa documentation sont sujettes à modification à tout moment sans préavis.

AVIS

Ce produit Informatica (le « Logiciel ») inclut certains pilotes (les « Pilotes DataDirect ») de DataDirect Technologies, une société de Progress Software Corporation (« DataDirect ») qui sont sujets aux conditions suivantes :

1. LES PILOTES DATADIRECT SONT FOURNIS « EN L'ÉTAT », SANS GARANTIE D'AUCUNE SORTE, EXPRESSE OU IMPLICITE, NOTAMMENT LES GARANTIES IMPLICITES DE CONFORMITÉ LÉGALE, D'USAGE NORMAL ET DE NON-INFRACTION.
2. DATADIRECT OU SES FOURNISSEURS TIERS NE POURRONT EN AUCUN CAS ÊTRE TENUS RESPONSABLES ENVERS LE CLIENT UTILISATEUR FINAL DE TOUT DOMMAGE DIRECT, ACCESSOIRE, INDIRECT, SPÉCIAL, CONSÉCUTIF OU AUTRE RÉSULTANT DE L'UTILISATION DES PILOTES ODBC, QU'ILS SOIENT INFORMÉS OU NON À L'AVANCE DE LA POSSIBILITÉ DE TELS DOMMAGES. CES LIMITATIONS S'APPLIQUENT À TOUTES LES CAUSES D'ACTION, NOTAMMENT TOUTE INFRACTION AU CONTRAT, INFRACTION À LA GARANTIE, NÉGLIGENCE, RESPONSABILITÉ STRICTE, REPRÉSENTATION INCORRECTE ET AUTRES TORTS.

Les renseignements contenus dans cette documentation sont sujets à modification sans préavis. Si vous constatez des problèmes liés à la documentation, merci de les signaler par courriel à l'adresse info_documentation@Informatica.com.

Les produits Informatica sont garantis conformément aux termes et conditions des accords en vertu desquels ils sont fournis. INFORMATICA FOURNIT LES INFORMATIONS DE CE DOCUMENT « EN L'ÉTAT » SANS GARANTIE D'AUCUNE SORTE, EXPRESSE OU IMPLICITE, NOTAMMENT AUCUNE GARANTIE DE QUALITÉ MARCHANDE, D'ADAPTATION À UN USAGE PARTICULIER ET D'ABSENCE DE CONTREFAÇON

Date de publication: 2022-03-15

Sommaire

Préface.....	7
Ressources Informatica.	7
Informatica Network.	7
Base de connaissances Informatica.	7
Documentation Informatica.	7
Matrices de disponibilité des produits Informatica.	8
Informatica Velocity.	8
Informatica Marketplace.	8
Support client international Informatica.	8
 Chapitre 1: Présentation des ressources.....	 9
Introduction aux scanners.	9
Ressources et scanners.	9
Types de ressources.	10
Ressource.	10
Ressources et planifications.	10
Ressources et attributs.	11
Création d'une ressource.	11
Filtre de profil de données et métadonnées source.	12
 Chapitre 2: Configuration des ressources de nuage.....	 14
Présentation des ressources du nuage.	14
Amazon Redshift.	15
Amazon S3.	18
Microsoft Azure Data Lake Store.	22
Microsoft Azure SQL Data Warehouse.	27
Microsoft Azure SQL Server.	30
Google BigQuery.	33
Microsoft Azure Blob Storage.	35
Salesforce.	39
Workday.	41
AWS Glue.	43
Snowflake.	46
Google Cloud Storage.	49
 Chapitre 3: Configuration des ressources Data Engineering.....	 52
Présentation des ressources Data Engineering.	52
Apache Atlas.	52
Cloudera Navigator.	54
HDFS.	57

Hive.	63
Chapitre 4: Configuration des ressources d'application.....	67
Présentation des ressources d'application.	67
SAP R/3.	67
SAP S/4HANA.	70
Chapitre 5: Configuration des ressources Business Glossary.....	76
Présentation des ressources Business Glossary.	76
Informatica Axon.	76
Business Glossary.	79
Chapitre 6: Configuration des ressources Business Intelligence.....	81
Présentation des ressources Business Intelligence.	81
IBM Cognos.	82
MicroStrategy.	86
Oracle Business Intelligence.	89
QlikView.	94
SAP Business Objects.	95
Tableau Server.	100
SAP Business Warehouse (SAP BW).	103
SAP BW/4HANA.	109
Microsoft Power BI.	114
Qlik Sense.	117
Chapitre 7: Configuration des ressources d'intégration de données.....	120
Présentation des ressources d'intégration de données.	120
Lignage personnalisé.	120
Data Integration Hub.	122
Scripts de base de données.	124
Service Informatica Cloud.	125
Informatica Master Data Management.	127
Informatica Platform.	130
PowerCenter.	133
Service d'intégration de SQL Server.	139
IBM InfoSphere DataStage.	143
Chapitre 8: Configuration des ressources de modélisation de données.....	147
Présentation des ressources de modélisation de données.	147
Erwin.	147
SAP PowerDesigner.	151

Chapitre 9: Configuration des ressources de gestion de la base de données 154

Présentation des ressources de gestion de la base de données.	154
IBM DB2.	155
IBM DB2 for z/OS.	157
IBM Netezza.	160
JDBC.	164
Microsoft SQL Server.	167
Oracle.	170
SAP HANA.	173
Sybase.	175
Teradata.	178

Chapitre 10: Configuration des ressources du gestionnaire de fichiers à parties multiples Hadoop..... 182

Présentation des ressources du gestionnaire de fichiers à parties multiples Hadoop	182
Gestionnaire de fichiers à parties multiples Hadoop.	182

Chapitre 11: Configuration des ressources de gestion de fichiers..... 184

Présentation des ressources de gestion des fichiers.	184
Système de fichiers.	184
OneDrive.	188
Microsoft SharePoint.	192

Chapitre 12: Configuration des ressources de source de diffusion..... 196

.	196
Apache Kafka.	196

Chapitre 13: Configuration des ressources NoSQL..... 200

Présentation des ressources NoSQL.	200
Apache Cassandra.	200
MongoDB.	203

Chapitre 14: Configuration des ressources Informatica Platform..... 206

Présentation des ressources Informatica Platform.	206
Informatica Data Quality.	206
Tâches de ressource Informatica Data Quality.	211
Ressource Informatica Similarity Discovery.	212
Dépannage.	214

Index..... 218

Préface

Pour créer, configurer et gérer plusieurs ressources dans Enterprise Data Catalog, suivez les instructions fournies dans le *Guide de configuration du scanneur d'Enterprise Data Catalog*. Outre les étapes de configuration, le guide contient également les autorisations requises et les prérequis pour les différentes ressources. Ce guide est destiné aux administrateurs de catalogue.

Ressources Informatica

Informatica vous fournit toute une gamme de ressources de produits via Informatica Network et autres portails en ligne. Utilisez ces ressources pour tirer le meilleur parti de vos produits et solutions Informatica, et pour apprendre d'autres utilisateurs et experts en la matière d'Informatica.

Informatica Network

Informatica Network est la passerelle à de nombreuses ressources, y compris la base de connaissances Informatica et le support client international Informatica. Pour accéder à Informatica Network, visitez le site <https://network.informatica.com>.

En tant que membre d'Informatica Network, vous disposez des options suivantes :

- Rechercher les ressources de produits dans la base de connaissances.
- Afficher les informations de disponibilité des produits.
- Créer et vérifier vos dossiers de support.
- Rechercher votre réseau de groupe d'utilisateurs local Informatica et collaborer avec vos pairs.

Base de connaissances Informatica

Utilisez la base de connaissances Informatica pour rechercher des ressources de produits telles que des articles pratiques, des meilleures pratiques, des didacticiels vidéo et des questions fréquemment posées.

Pour rechercher dans la base de connaissances, visitez le site <https://search.informatica.com>. N'hésitez pas à contacter l'équipe Base de connaissances Informatica à l'adresse KB_Feedback@informatica.com pour lui faire part de vos questions, commentaires ou suggestions concernant la base de connaissances.

Documentation Informatica

Utilisez le portail de documentation Informatica pour explorer une vaste bibliothèque de documentation pour les versions de produits actuelles et récentes. Pour explorer le portail de documentation, visitez le site <https://docs.informatica.com>.

N'hésitez pas à contacter l'équipe Documentation Informatica à l'adresse infa_documentation@informatica.com pour lui faire part de vos questions, commentaires ou suggestions concernant la documentation des produits.

Matrices de disponibilité des produits Informatica

Les matrices de disponibilité des produits (PAM) indiquent les versions des systèmes d'exploitation, les bases de données et les types de source et cible de données pris en charge par une version d'un produit. Vous pouvez parcourir les PAM Informatica à l'adresse <https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Velocity

Informatica Velocity est un ensemble de conseils et de meilleures pratiques développés par les services professionnels d'Informatica et basés sur les expériences réelles de centaines de projets de gestion des données. Informatica Velocity représente le savoir collectif de consultants d'Informatica qui collaborent avec des organisations du monde entier pour planifier, développer, déployer et gérer des solutions performantes de gestion des données.

Vous trouverez les ressources d'Informatica Velocity à l'adresse <http://velocity.informatica.com>. Si vous avez des questions, des commentaires ou des suggestions sur Informatica Velocity, contactez les services professionnels d'Informatica à l'adresse ips@informatica.com.

Informatica Marketplace

Informatica Marketplace est un forum dans lequel vous pouvez trouver des solutions qui permettent d'augmenter et d'améliorer vos implémentations Informatica. Exploitez les centaines de solutions de développeurs et de partenaires Informatica sur Marketplace pour améliorer votre productivité et accélérer le délai d'implémentation de vos projets. Vous trouverez Informatica Marketplace à l'adresse <https://marketplace.informatica.com>.

Support client international Informatica

Vous pouvez contacter un centre de support international par téléphone ou via le réseau Informatica.

Pour rechercher le numéro de téléphone du support client international Informatica local, visitez le site Web Informatica à l'adresse <https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

Pour trouver des ressources de support en ligne sur le réseau Informatica, visitez le site <https://network.informatica.com> et sélectionnez l'option eSupport.

CHAPITRE 1

Présentation des ressources

Ce chapitre comprend les rubriques suivantes :

- [Introduction aux scanners, 9](#)
- [Ressources et scanners, 9](#)
- [Types de ressources, 10](#)
- [Ressource, 10](#)
- [Ressources et planifications, 10](#)
- [Ressources et attributs, 11](#)
- [Création d'une ressource, 11](#)
- [Filtre de profil de données et métadonnées source, 12](#)

Introduction aux scanners

Un scanner est un composant enfichable d'Enterprise Data Catalog qui extrait les métadonnées spécifiques, telles que des métadonnées source ou des métadonnées de profil, à partir de sources de données externes et les stocke dans le catalogue. Un scanner correspond généralement à un type de ressource unique. Toutefois, il peut y avoir plusieurs scanners pour un type de ressource.

Un scanner effectue une tâche d'analyse dans les sources de métadonnées afin de récupérer les métadonnées dans le catalogue. Lorsqu'Informatica publie des nouveaux types de ressources pour des sources de données spécifiques, vous pouvez relier ces scanners à l'installation actuelle d'Enterprise Data Catalog sans devoir mettre à niveau l'installation.

Ressources et scanners

Les scanners attachés aux ressources identifient le type de ressource, y compris le nom du type de ressource, le nom d'affichage, la description et les versions prises en charge. Vous pouvez configurer plusieurs scanners pour différents types de métadonnées. Vous pouvez configurer les propriétés du scanner pour une ressource lorsque vous créez la ressource. Chaque scanner possède un ID unique qui correspond au type de ressource.

Un type de ressource peut avoir plusieurs scanners. Par exemple, un scanner peut extraire les métadonnées d'une source de données Oracle directement à partir de la source. Un autre scanner peut dériver les métadonnées relatives à la source de données Oracle à partir d'autres référentiels de métadonnées, tels qu'un référentiel PowerCenter ou un référentiel modèle.

Types de ressources

Le type de ressource est un type de source de données externe ou de référentiel de métadonnées à partir duquel les ressources extraient des métadonnées.

Enterprise Data Catalog prend en charge une liste de ressources, qui sont largement réparties dans les groupes suivants :

- Ressources Business Intelligence
- Ressources d'intégration de données
- Ressources de gestion de fichiers
- Ressources d'application
- Ressources de données Hadoop
- Ressources de données de nuage
- Ressources Data Engineering
- Ressources Business Glossary
- Ressources de modélisation de données
- Ressources de gestion de données
- Ressources personnalisées
- Ressources NoSQL
- Ressources de diffusion de données
- Informatica Platform

Ressource

Une ressource est un objet de catalogue qui représente une source de données externe ou un référentiel de métadonnées à partir duquel les scanners extraient les métadonnées. Une ressource représente une instance d'un type de ressource spécifique. Les opérations de métadonnées de base, telles que l'extraction et le stockage des métadonnées, sont effectuées au niveau des ressources.

Une ressource peut avoir un type de ressource, comme les bases de données relationnelles, la classification Business Glossary et les sources de veille stratégique. Une ressource peut avoir une planification associée. Chaque ressource peut extraire les métadonnées source et les métadonnées de profil des sources de données externes.

Ressources et planifications

Les planifications de ressources déterminent la fréquence à laquelle les scanners extraient les métadonnées des sources de métadonnées. Vous pouvez créer des planifications pour des ressources spécifiques ou créer des planifications réutilisables à attribuer à plusieurs ressources.

Vous pouvez créer des planifications qui ont une date de fin ou des planifications qui s'exécutent indéfiniment. En plus d'une date de début, vous pouvez également configurer l'heure de début en fonction du moment où vous voulez que les scanners commencent à extraire des métadonnées.

Ressources et attributs

Les ressources disposent d'attributs qui représentent des propriétés spécifiques des types d'objet, tels que Colonne Alias, Catégorie, Jeu de données et Table, en fonction des types de données des attributs. Vous pouvez utiliser les attributs de ressource dans Enterprise Data Catalog pour rechercher des informations pertinentes parmi un grand nombre de ressources de données et de relations de données.

La ressource peut avoir des attributs personnalisés associés. Les attributs système incluent des attributs prédéfinis qui représentent des propriétés spécifiques des types d'objet, telles que la procédure stockée, la table, le déclencheur, la requête XML, le schéma et la ressource. Par exemple, Commentaire est un attribut système attribué à une table Hive ou à une vue Hive. Vous pouvez créer des attributs personnalisés en fonction de vos besoins professionnels et les attribuer à différentes ressources. Vous pouvez choisir de créer des attributs personnalisés en fonction des types de scanners utilisés.

Création d'une ressource

Lorsque vous créez une ressource, vous pouvez spécifier le type de ressource, le type de métadonnées qu'Enterprise Data Catalog extrait et une planification facultative de la ressource. Vous pouvez également attribuer des attributs personnalisés à la ressource lorsque vous la créez.

Remarque: Il est recommandé de valider la connexion après avoir configuré une ressource. Cliquez sur **Tester la connexion** pour valider les paramètres de connexion.

1. Lancez Catalog Administrator.
2. Cliquez sur **Nouveau > Ressource**.
L'assistant **Nouvelle ressource** s'affiche.
3. Entrez un nom et une description facultative pour la ressource.
4. Si vous souhaitez affecter des valeurs d'attribut personnalisées à la ressource, entrez des valeurs pour les attributs, tels que le département, le propriétaire des données, le gestionnaire de données et les experts techniques dans la section **Propriétés supplémentaires**.
 - Si vous souhaitez appliquer des valeurs d'attribut personnalisées à toutes les éléments chargés via la ressource, sélectionnez **Propager** sous chaque type d'attribut personnalisé.
5. Cliquez sur **Sélectionner** pour ouvrir la boîte de dialogue **Sélectionner le type de ressource**.
6. Sélectionnez un type de ressource, puis cliquez sur **Sélectionner**.
D'autres champs apparaissent dans l'assistant en fonction du type de ressource que vous avez sélectionné.
7. En fonction du type de ressource, configurez les propriétés de la connexion.
8. Cliquez sur **Suivant** pour passer à la page **Paramètres de chargement des métadonnées**.
Vous pouvez choisir le type de métadonnées que vous voulez voir extraites par Enterprise Data Catalog des systèmes source.
9. Configurez les métadonnées source requises et les paramètres de métadonnées de profilage.
10. Cliquez sur **Suivant** pour passer à la page **Attributs personnalisés**, et configurez les paramètres d'attribut. Vous pouvez sélectionner les attributs personnalisés à associer à la ressource.
11. Cliquez sur **Suivant** pour passer à la page **Provisionnement de données**.
12. Pour créer des tâches de provisionnement dans Enterprise Data Catalog, activez éventuellement le provisionnement de données.

13. Cliquez sur **Suivant** pour accéder à la page **Autorisations** et limitez les autorisations pour la ressource. Vous pouvez également attribuer ou modifier les autorisations de la ressource à partir des paramètres de sécurité.
14. Cliquez sur **Suivant** pour accéder à la page **Planification**.
15. Vous pouvez éventuellement sélectionner les planifications pour le chargement des métadonnées source et les métadonnées de profilage. Vous pouvez créer une planification globale si nécessaire.
16. Cliquez sur **Enregistrer**, ou cliquez sur **Enregistrer et exécuter**.

Filtre de profil de données et métadonnées source

Lorsqu'une ressource contient des tables et des vues multiples, vous pouvez choisir d'extraire les métadonnées source de tables et vues spécifiques, et d'extraire les données et métadonnées de profil de tables et vues spécifiques.

Vous pouvez configurer le champ de filtre si l'une des conditions suivantes est remplie :

- Vous souhaitez extraire les métadonnées uniquement pour quelques tables et vues d'une liste contenant toutes les tables et vues d'une ressource.
- Vous souhaitez inclure quelques tables et vues de l'exécution de ressource suivante.
- Vous souhaitez exclure quelques tables et vues de l'exécution de ressource.
- Vous souhaitez exclure quelques tables et vues de l'exécution de ressource suivante.

Dans une ressource, vous pouvez utiliser les filtres suivants pour inclure et exclure des tables et des vues :

- Filtre de métadonnées source. Dans l'onglet **Paramètres de chargement des métadonnées** de l'assistant Ressource, indiquez les tables et les vues à inclure ou exclure dans une exécution de ressource. Le scanneur extrait les métadonnées source en fonction des tables et des vues indiquées dans le champ Filtre des métadonnées source.
- Filtre de profil de données. Dans l'onglet **Paramètres de chargement des métadonnées** de l'assistant Ressource, indiquez les tables et les vues à inclure ou exclure dans une exécution de profil. Le scanneur de profil extrait les métadonnées de profil en fonction des tables et des vues indiquées dans le champ Filtre de profil de données.

Pour inclure des tables et des vues, vous pouvez indiquer leurs noms complets ou utiliser le caractère générique % dans le nom ou le modèle. Utilisez des points-virgules (;) pour séparer les noms des tables et les noms des vues. Si le modèle contient un caractère spécial, tel qu'un blanc, placez-le entre des guillemets simples. Par exemple, HR_id ; 'HR Emp%' ; HR_job%.

Remarque: Le filtre de profil de données prend en charge les expressions régulières telles que %, ^, \$, ., | et (a | b), tandis que le filtre de métadonnées source prend en charge l'expression %. La contrainte de filtre s'applique à Microsoft Azure SQL Server, Microsoft Azure SQL Data Warehouse, Hive ainsi qu'à toutes les sources de bases de données relationnelles.

Vous pouvez choisir un ou plusieurs des modèles suivants pour identifier les tables et les vues :

- <name>%. Inclut les tables et les vues qui commencent par <name>.
Par exemple, A% inclut toutes les tables et les vues qui commencent par le caractère A.
- %<name>. Inclut les tables et les vues qui se terminent par <name>.
Par exemple, %id inclut toutes les tables et les vues qui se terminent par le texte 'id'.
- %<name>%. Inclut les tables et les vues qui contiennent <name>.
Par exemple, %DB% inclut toutes les tables et les vues qui contiennent le mot clé « DB ».

- <name>. Inclut les tables et les vues qui correspondent à <name>.

Par exemple, le filtre HR_id inclut n'importe quelle table ou vue de la ressource qui correspond à HR_id.

Pour exclure des tables et des vues de l'analyse, faites précéder le nom ou le modèle du terme NOT. Par exemple, NOT %ID ; NOT 'ID%'. Utilisez des points-virgules (;) pour séparer les noms des tables et les noms des vues.

Dépannage

Que se passe-t-il si j'inclus dans le filtre de métadonnées source une table liée à un objet externe ?

Lorsqu'une table ou une vue filtrée dépend d'un objet externe tel qu'une base de données, le scanneur de ressource extrait les métadonnées de l'objet externe même si celui-ci ne remplit pas les conditions de filtre, car le filtre ne s'applique pas à l'objet externe dépendant.

Comment puis-je identifier les tables et les vues qui ont été exclues de l'analyse de profil ?

Par défaut, les fichiers journaux ne répertorient pas toutes les tables et les vues qui ont été exclues d'une analyse. Pour afficher la liste des tables et des vues exclues, activez le niveau de journal *Débogage* puis exécutez l'analyse. Les fichiers journaux répertorient toutes les tables et les vues exclues de l'analyse.

CHAPITRE 2

Configuration des ressources de nuage

Ce chapitre comprend les rubriques suivantes :

- [Présentation des ressources du nuage, 14](#)
- [Amazon Redshift, 15](#)
- [Amazon S3, 18](#)
- [Microsoft Azure Data Lake Store, 22](#)
- [Microsoft Azure SQL Data Warehouse, 27](#)
- [Microsoft Azure SQL Server, 30](#)
- [Google BigQuery, 33](#)
- [Microsoft Azure Blob Storage, 35](#)
- [Salesforce, 39](#)
- [Workday, 41](#)
- [AWS Glue, 43](#)
- [Snowflake, 46](#)
- [Google Cloud Storage, 49](#)

Présentation des ressources du nuage

Le groupe de ressources du nuage inclut des ressources associées au nuage. Lorsque vous créez et exécutez une ressource de nuage, le scanner de nuage extrait et migre les métadonnées de la ressource associée vers le catalogue. Vous pouvez afficher les résultats migrés dans Enterprise Data Catalog.

Amazon Redshift

Amazon Redshift est un service d'hébergement Internet et un produit d'entreposage de données. Amazon Redshift fait partie d'Amazon Web Services, la plate-forme de cloud computing proposée par Amazon.

Objets extraits

- Tables
- Vues

Autorisations de configuration de la ressource

Si vous créez un utilisateur, veillez à configurer l'autorisation `read` dans la source de données Amazon Redshift pour le compte utilisateur.

Prérequis

Téléchargez les fichiers JAR du pilote JDBC pour Amazon Redshift depuis Internet et copiez-les dans le répertoire `RedShiftScanner` du répertoire `<INFA_HOME>/services/CatalogService/ScannerJars/externalDependencies`. Il n'est pas nécessaire de redémarrer le service de catalogue.

Pour respecter les prérequis, vous pouvez également effectuer les étapes suivantes :

1. Téléchargez les pilotes JDBC Amazon Redshift depuis Internet et incluez-les dans le fichier `redshiftJars.zip`.
2. Copiez le fichier `redshiftJars.zip` dans le répertoire `<INFA_HOME>/services/CatalogService/ScannerBinaries`.
3. Ouvrez le fichier `<INFA_HOME>/services/CatalogService/ScannerBinaries/CustomDeployer/scannerDeployer.xml` et ajoutez-y les lignes suivantes :

```
</ExecutionContextProperty>
  <ExecutionContextProperty isLocationProperty="true"
    dependencyToUnpack="redshiftJars.zip">
    <PropertyName>RedShiftScanner_DriverLocation</PropertyName>
    <PropertyValue>scanner_miti/RedShift/Drivers</PropertyValue>
  </ExecutionContextProperty>
```
4. Enregistrez le fichier `scannerDeployer.xml`.
5. Si vous souhaitez activer le profilage pour la ressource Amazon Redshift, copiez le fichier `redshift.jar` dans le répertoire `<INFA_HOME>/connectors/thirdparty/informatica.amazonredshift/common/`.
6. Redémarrez le service de catalogue.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Utilisateur	Nom d'utilisateur utilisé pour accéder à la base de données.
Mot de passe	Mot de passe associé au nom d'utilisateur.
Hôte	Nom d'hôte ou adresse IP du service Amazon Redshift.
Port	Numéro de port du serveur Amazon Redshift. La valeur par défaut est 5439.
Base de données	Nom de l'instance de base de données.

L'image suivante montre l'exemple de propriétés de connexion dans l'onglet **Général** :

Connection Properties

User*:	<input type="text" value="Administrator"/>
Password*:	<input type="password" value="....."/>
Host*:	<input type="text" value="Administrator.Redshift.Amazonws.com"/>
Port*:	<input type="text" value="0002"/>
Database*:	<input type="text" value="Finance"/>
<input type="button" value="Test Connection"/>	

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Importer des objets système	Sélectionnez cette option pour spécifier que les objets système doivent être importés.
Schéma	<p>Cliquez sur Sélectionner... pour spécifier les schémas Amazon Redshift à importer. Vous pouvez utiliser l'une des options suivantes dans la boîte de dialogue Sélectionner un schéma pour importer les schémas :</p> <ul style="list-style-type: none"> - Sélectionner dans la liste : utilisez cette option pour sélectionner les schémas requis à partir d'une liste de schémas disponibles. - Sélectionner à l'aide d'une expression régulière : fournissez une expression régulière SQL pour sélectionner des schémas qui correspondent à l'expression.
Nom du compartiment S3	Indiquez un nom de compartiment Amazon S3 valide pour la source de données Amazon Redshift. Vous devez fournir cette valeur si vous voulez activer le profilage pour Amazon Redshift. Si vous ne voulez pas activer le profilage, conservez la valeur par défaut.
Sensible à la casse	<p>Spécifie que la ressource est configurée pour l'insensibilité à la casse. Sélectionnez l'une des valeurs suivantes :</p> <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. <p>La valeur par défaut est False.</p>
Mémoire	<p>Sélectionnez la mémoire requise pour exécuter le travail du scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le réglage des performances d'Enterprise Data Catalog sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>

Propriété	Description
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none"> -Dscannerloglevel=<DEBUG/INFO/ERROR>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. -Dscanner.container.core=<No. of core>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. -Dscanner.yarn.app.environment=<key=value>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. -Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0>. Augmente la mémoire du conteneur du scanner lorsque pmem est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.
Options d'agent	Spécifiez les options de l'agent Enterprise Data Catalog pour l'exécution de la tâche de scanner.

L'image suivante montre les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

▼ Source Metadata

☒ Enable Source Metadata

Additional properties

Import system objects: ☐

Schema:

S3 Bucket Name*:

Advanced Properties

Memory:

JVM Options:

Vous pouvez activer la découverte de données pour une ressource Amazon Redshift. Pour plus d'informations, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource Amazon Redshift. Pour plus d'informations, consultez la section [Composite Data Domain Discovery](#).

Amazon S3

Amazon S3 est un service de stockage simple proposé par Amazon Web Services (AWS) qui fournit le stockage des objets via une interface de service Web.

Objets extraits

Enterprise Data Catalog extrait des fichiers de la source de données Amazon S3. Vous pouvez également extraire des métadonnées d'un stockage compatible avec Amazon S3, par exemple Scality RING.

Remarque: Dans Enterprise Data Catalog version 10.4.1.2 ou version ultérieure, vous pouvez utiliser la ressource Amazon S3 pour vous connecter à une source de données Amazon S3 à l'aide d'un jeton de session temporaire.

Autorisations de configuration de la ressource

Configurez l'autorisation `read` dans la source de données Amazon S3 pour le compte utilisateur permettant d'accéder à la source de données. Configurez l'autorisation d'accès du compte utilisateur si celui utilisé est différent du compte utilisateur permettant de créer la source de données Amazon S3.

Types de fichiers pris en charge

La ressource Amazon S3 permet d'extraire les métadonnées de fichiers structurés, non structurés et non structurés étendus.

Fichiers structurés pris en charge :

- Fichiers AVRO
- Fichiers délimités
- Fichiers texte
- Fichiers JSON
- Fichiers Parquet
- Fichiers XML

Fichiers non structurés pris en charge :

- Fichiers Apple
- Fichiers compressés
- E-mail

Fichiers non structurés étendus :

- Fichiers VB
- Fichiers ASP
- Fichiers TIF
- Fichiers LOG
- Fichiers CSS
- Fichiers ASPX
- Fichiers DLL
- Fichiers GIF
- Fichiers SQL

Attribuez l'autorisation `read` aux fichiers pour l'extraction des métadonnées.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Type compatible S3	Indique si la source de données est un stockage compatible Amazon S3 tel que Scalify RING.
URL du point de terminaison REST	URL du point de terminaison REST du stockage compatible Amazon S3 permettant d'extraire des métadonnées.
URL de compartiment Amazon Web Services	URL Amazon Web Services permettant d'accéder à un compartiment.
Informations d'identification temporaires ?	Active la ressource pour la connexion et l'extraction de métadonnées d'une source de données Amazon S3 en utilisant des informations d'identification temporaires.
ID de clé d'accès AWS	ID de clé d'accès Amazon Web Services permettant de signer les demandes que vous envoyez à Amazon Web Services. Remarque: Indiquez l'ID de clé d'accès au stockage compatible avec Amazon S3 pour signer les demandes que vous envoyez à ce stockage.
Clé d'accès secrète AWS	Clé d'accès secrète Amazon Web Services permettant de signer les demandes que vous envoyez à Amazon Web Services. Remarque: Indiquez la clé d'accès secrète au stockage compatible avec Amazon S3 pour signer les demandes que vous envoyez à ce stockage.
Jeton de session AWS	Jeton de session temporaire pour la connexion à la source de données Amazon S3.
Nom du compartiment Amazon Web Services	Nom de compartiment Amazon Web services qu'Enterprise Data Catalog doit analyser.
Répertoire source	Répertoire source à partir duquel les métadonnées doivent être extraites.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Types de fichiers	<p>Sélectionnez un ou tous les types de fichiers suivants à partir desquels vous voulez extraire les métadonnées :</p> <ul style="list-style-type: none"> - Tous. Utilisez cette option pour spécifier si vous voulez extraire des métadonnées de tous les types de fichiers. - Sélectionner. Utilisez cette option pour spécifier que vous voulez extraire des métadonnées à partir de types de fichiers spécifiques. Effectuez les étapes suivantes pour spécifier les types de fichiers : <ol style="list-style-type: none"> 1. Cliquez sur Sélectionner. La boîte de dialogue Sélectionner des types de fichiers spécifiques s'affiche. 2. Sélectionnez les fichiers requis parmi les options suivantes : <ul style="list-style-type: none"> - Formats non structurés étendus Utilisez cette option pour extraire les métadonnées des types de fichiers tels que les fichiers audio, vidéo, image et les ebooks. - Types de fichiers structurés. Utilisez cette option pour extraire des métadonnées à partir de types de fichiers tels que Avro, Parquet, JSON, XML, fichiers texte et fichiers délimités. - Types de fichiers non structurés. Utilisez cette option pour extraire des métadonnées à partir de types de fichiers tels que Microsoft Excel, Microsoft PowerPoint, Microsoft Word, pages Web, fichiers compressés, e-mails et PDF. 3. Cliquez sur Sélectionner. <p>Remarque: Vous pouvez sélectionner l'option Types de fichiers spécifiques dans la boîte de dialogue pour sélectionner les fichiers de toutes les catégories.</p>
Autres types de fichiers	Extrait les métadonnées de base des fichiers (taille de fichier, chemin et horodatage) pour les types de fichiers qui ne figurent pas dans la propriété Types de fichiers .
Traiter les fichiers sans extension As	<p>Sélectionnez l'une des options suivantes pour identifier les fichiers sans extension :</p> <ul style="list-style-type: none"> - Aucun - Avro - Parquet
Entrer le délimiteur de fichier	<p>Spécifiez le délimiteur de fichier si le fichier à partir duquel vous extrayez des métadonnées utilise un délimiteur autre qu'un délimiteur de la liste suivante :</p> <ul style="list-style-type: none"> - Virgule (,) - Tabulation horizontale (\t) - Point-virgule (;) - Deux points (:) - Barre verticale () <p>Assurez-vous de placer le délimiteur entre des guillemets simples. Par exemple : '\$ '. Utilisez une virgule pour séparer plusieurs délimiteurs. Par exemple : '\$ ', '% ', '& '</p>
Répertoire de premier niveau	<p>Spécifiez un répertoire ou une liste de répertoires sous le répertoire source. Si vous ne spécifiez pas cette option, Enterprise Data Catalog importe tous les fichiers du répertoire source spécifié. Pour spécifier un répertoire ou une liste de répertoires, vous pouvez effectuer les étapes suivantes :</p> <ol style="list-style-type: none"> 1. Cliquez sur Sélectionner.... La boîte de dialogue Sélectionner répertoire de premier niveau s'affiche. 2. Utilisez l'une des options suivantes pour sélectionner les répertoires requis : <ul style="list-style-type: none"> - Sélectionner dans la liste : sélectionnez les répertoires requis à partir d'une liste de répertoires. - Sélectionner à l'aide d'une expression régulière : fournissez une expression régulière SQL pour sélectionner des schémas qui correspondent à l'expression. <p>Remarque: Si vous voulez sélectionner plusieurs répertoires, vous devez les séparer par un point-virgule (;).</p>

Propriété	Description
Analyse récursive	Analyse les sous-répertoires de manière récursive sous les répertoires de premier niveau sélectionné. L'analyse récursive est requise pour la découverte des fichiers partitionnés.
Activer la découverte des fichiers partitionnés	Identifie et publie les fichiers partitionnés horizontalement sous le même répertoire et les fichiers organisés dans des structures de répertoire hiérarchique de type Hive dans un fichier partitionné unique.
Sensible à la casse	Spécifie que la ressource est configurée pour la sensibilité à la casse. Sélectionnez l'une des valeurs suivantes : <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. La valeur par défaut est True.
Mémoire	Sélectionnez la mémoire requise pour exécuter le travail du scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé : <ul style="list-style-type: none"> - Faible - Moyenne - Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1. - <code>-DmaxPartFilesToValidatePerTable=<number></code>. Valide le nombre spécifié de fichiers de partition dans la table partitionnée. La valeur par défaut est 10. - <code>-DmaxPartFilesToValidatePerPartition=<number></code>. Valide le nombre spécifié de fichiers de partition pour chaque partition dans la table partitionnée. La valeur par défaut est 5. - <code>-DexcludePatterns=<comma separated regex patterns></code>. Exclut les fichiers lors de l'analyse des table de partition selon le modèle d'expression régulière. Par défaut, les noms de fichier qui commencent par un point et un trait de soulignement sont exclus.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource Amazon S3. Pour plus d'informations, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource Amazon S3. Pour plus d'informations, consultez la section [Composite Data Domain Discovery](#).

Microsoft Azure Data Lake Store

Microsoft Azure Data Lake est un service d'analyse et de stockage de données évolutif. Ce service est hébergé sur Azure. Configurez un type de ressource Azure Data Lake Store (ADLS) pour extraire les métadonnées de Data Lake Store.

Lorsque vous créez une ressource Azure Data Lake Store, vous pouvez accéder aux fichiers et aux dossiers dans les produits de stockage Azure suivants :

Azure Data Lake Store ou Data Lake Storage Gen1

Pour accéder à ce référentiel, Enterprise Data Catalog utilise l'authentification de service à service ou OAuth 2.0. Pour utiliser l'authentification OAuth 2.0, vous devez créer une application Azure Active Directory (AD) et utiliser l'ID de client et la clé du client depuis l'application pour l'authentification. Enterprise Data Catalog utilise SDK pour accéder au contenu du référentiel.

Azure Data Lake Storage Gen2

Le stockage Blob Azure prend en charge Azure Data Lake Storage Gen2. Il s'agit d'un système de fichiers hiérarchique. Lorsque vous créez une ressource Azure Data Lake Store et que vous choisissez l'option **Azure Data Lake Storage Gen2**, vous devez entrer l'ID de compte utilisateur et l'une des clés fournies dans la section **Clés d'accès**. Dans le portail Azure, vous pouvez afficher les deux clés générées pour chaque compte de stockage Azure Data Lake Storage Gen2 dans la section **Paramètres > Clés d'accès**. Pour accéder aux fichiers et aux dossiers de ce système de fichiers hiérarchique, Enterprise Data Catalog utilise des API REST. Azure utilise l'autorisation de clé partagée afin d'authentifier les demandes. Dans Enterprise Data Catalog, l'accès et l'exécution sont deux fois plus rapides pour *Azure Data Lake Storage Gen2* par rapport au stockage de *Data Lake Storage Gen1*.

Objets extraits

- Fichier
- Dossier
- Conteneur

Autorisations de configuration de la ressource

Si vous créez un utilisateur, veillez à configurer l'autorisation `read` dans la source de données pour le nouveau compte utilisateur.

Types de fichiers pris en charge

La ressource Microsoft Azure Data Lake Store permet d'extraire les métadonnées de fichiers structurés, non structurés et non structurés étendus.

Fichiers structurés pris en charge :

- Fichiers AVRO
- Fichiers délimités
- Fichiers texte
- Fichiers JSON
- Fichiers Parquet
- fichiers XML

Fichiers non structurés pris en charge :

- Fichiers Apple
- Fichiers compressés

- E-mail

Fichiers non structurés étendus :

- Fichiers VB
- Fichiers ASP
- Fichiers TIF
- Fichiers LOG
- Fichiers CSS
- Fichiers ASPX
- Fichiers DLL
- Fichiers GIF
- Fichiers SQL

Attribuez les autorisations `read` et `write` aux fichiers pour l'extraction des métadonnées.

Prérequis

Assurez-vous que vous avez effectué les prérequis suivants avant de créer la ressource :

1. Fusionnez les certificats dans `<INFA_HOME>/java/jre/lib/security/cacerts` avec le fichier `<INFA_HOME>/services/shared/security/ infa_truststore.jks`.
2. Déplacez le fichier `infa_truststore.jks` vers un emplacement commun accessible à tous les nœuds du cluster.
3. Dans les propriétés de configuration HDFS de l'interface Ambari, mettez à jour le chemin d'accès au fichier `infa_truststore.jks` dans la propriété **ssl.client.truststore.location** et mettez à jour le mot de passe `infa_truststore.jks` dans la propriété **ssl.client.truststore.password**.
4. Redémarrez le service de cluster Informatica.

Remarque: Veillez à configurer les autorisations requises pour le stockage ADLS dans Azure Active Directory.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Nom du compte	Entrez le nom du compte de stockage que vous avez créé dans le portail Azure.
Type de source ADLS	Choisissez l'option Data Lake Store Gen 1 ou Data Lake Store Gen 2 .
ID du client	Entrez l'ID de client pour vous connecter à Microsoft Azure Data Lake Store. Utilisez la valeur indiquée pour l'ID d'application dans le portail Azure. Cette option s'affiche lorsque vous choisissez l'option Data Lake Store Gen 1 comme type de source ADLS.
Clé du client	Entrez la clé du client pour vous connecter à Microsoft Azure Data Lake Store. Utilisez la valeur de la clé d'application Azure Active Directory dans le portail Azure comme clé du client. Cette option s'affiche lorsque vous choisissez l'option Data Lake Store Gen 1 comme type de source ADLS.

Propriété	Description
Nom de répertoire	Nom de répertoire de Azure Data Lake Store.
URL du point de terminaison d'autorisation	URL du point de terminaison de jeton OAuth 2.0 dans le portail Azure. Cette option s'affiche lorsque vous choisissez l'option Data Lake Store Gen 1 comme type de source ADLS.
Clé de compte de stockage	Entrez key1 ou key2 comme clé de compte de stockage. Accédez à la section Paramètres > Clés d'accès dans le portail Azure pour afficher les clés de compte de stockage. Cette option s'affiche lorsque vous choisissez l'option Data Lake Store Gen 2 comme type de source ADLS.

L'image suivante montre un exemple de propriétés de connexion dans l'onglet **Général** :

Connection Properties

Account name:	<input type="text" value="azureqadatalake"/>
Client Id*:	<input type="text" value="....."/>
Client Key*:	<input type="text" value="....."/>
Directory Name*:	<input type="text" value="/BDM"/>
Auth Endpoint URL*:	<input type="text" value="https://login.microsoftonline.com/2638f43e-f77d-4fc7-ab92-7b753b7876fd/oauth2/token"/>
<input type="button" value="Test Connection"/>	

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Types de fichiers	<p>Sélectionnez un ou tous les types de fichiers suivants à partir desquels vous voulez extraire les métadonnées :</p> <ul style="list-style-type: none"> - Tous. Utilisez cette option pour spécifier si vous voulez extraire des métadonnées de tous les types de fichiers. - Sélectionner. Utilisez cette option pour spécifier que vous voulez extraire des métadonnées à partir de types de fichiers spécifiques. Effectuez les étapes suivantes pour spécifier les types de fichiers : <ol style="list-style-type: none"> 1. Cliquez sur Sélectionner. La boîte de dialogue Sélectionner des types de fichiers spécifiques s'affiche. 2. Sélectionnez les fichiers requis parmi les options suivantes : <ul style="list-style-type: none"> - Formats non structurés étendus Utilisez cette option pour extraire les métadonnées des types de fichiers tels que les fichiers audio, vidéo, image et les ebooks. - Types de fichiers structurés. Utilisez cette option pour extraire des métadonnées à partir de types de fichiers tels que Avro, Parquet, JSON, XML, fichiers texte et fichiers délimités. - Types de fichiers non structurés. Utilisez cette option pour extraire des métadonnées à partir de types de fichiers tels que Microsoft Excel, Microsoft PowerPoint, Microsoft Word, pages Web, fichiers compressés, e-mails et PDF. 3. Cliquez sur Sélectionner. <p>Remarque: Vous pouvez sélectionner l'option Types de fichiers spécifiques dans la boîte de dialogue pour sélectionner les fichiers de toutes les catégories.</p>
Autres types de fichiers	Extrayez les métadonnées de base des fichiers (taille de fichier, chemin et horodatage) pour les types de fichiers qui ne figurent pas dans la propriété Types de fichiers .
Traiter les fichiers sans extension As	<p>Sélectionnez l'une des options suivantes pour identifier les fichiers sans extension :</p> <ul style="list-style-type: none"> - Aucun - Avro - Parquet
Entrer le délimiteur de fichier	<p>Spécifiez le délimiteur de fichier si le fichier à partir duquel vous extrayez des métadonnées utilise un délimiteur autre qu'un délimiteur de la liste suivante :</p> <ul style="list-style-type: none"> - Virgule (,) - Tabulation horizontale (\t) - Point-virgule (;) - Deux points (:) - Barre verticale () <p>Assurez-vous de placer le délimiteur entre des guillemets simples. Par exemple : '\$ '. Utilisez une virgule pour séparer plusieurs délimiteurs. Par exemple : '\$ ', '% ', '& '</p>
Répertoire de premier niveau	<p>Spécifiez un répertoire ou une liste de répertoires sous le répertoire source. Si vous ne spécifiez pas cette option, Enterprise Data Catalog importe tous les fichiers du répertoire source spécifié. Pour spécifier un répertoire ou une liste de répertoires, vous pouvez effectuer les étapes suivantes :</p> <ol style="list-style-type: none"> 1. Cliquez sur Sélectionner.... La boîte de dialogue Sélectionner répertoire de premier niveau s'affiche. 2. Utilisez l'une des options suivantes pour sélectionner les répertoires requis : <ul style="list-style-type: none"> - Sélectionner dans la liste : sélectionnez les répertoires requis à partir d'une liste de répertoires. - Sélectionner à l'aide d'une expression régulière : fournissez une expression régulière SQL pour sélectionner des schémas qui correspondent à l'expression. <p>Remarque: Si vous voulez sélectionner plusieurs répertoires, vous devez les séparer par un point-virgule (;).</p>

Propriété	Description
Analyse récursive	Analyse les sous-répertoires de manière récursive sous les répertoires de premier niveau sélectionné. L'analyse récursive est requise pour la découverte des fichiers partitionnés.
Activer la découverte des fichiers partitionnés	Identifie et publie les fichiers partitionnés horizontalement sous le même répertoire et les fichiers organisés dans des structures de répertoire hiérarchique de type Hive dans un fichier partitionné unique.
Sensible à la casse	Spécifie que la ressource est configurée pour la sensibilité à la casse. Sélectionnez l'une des valeurs suivantes : <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. La valeur par défaut est True.
Mémoire	Sélectionnez la mémoire requise pour exécuter le travail du scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé : <ul style="list-style-type: none"> - Faible - Moyenne - Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1. - <code>-DmaxPartFilesToValidatePerTable=<number></code>. Valide le nombre spécifié de fichiers de partition dans la table partitionnée. La valeur par défaut est 10. - <code>-DmaxPartFilesToValidatePerPartition=<number></code>. Valide le nombre spécifié de fichiers de partition pour chaque partition dans la table partitionnée. La valeur par défaut est 5. - <code>-DexcludePatterns=<comma separated regex patterns></code>. Exclut les fichiers lors de l'analyse des table de partition selon le modèle d'expression régulière. Par défaut, les noms de fichier qui commencent par un point et un trait de soulignement sont exclus.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour Azure Data Lake Store. Pour plus d'informations, consultez la section [Enable Data Discovery](#). Vous pouvez activer la découverte de domaines de données composites pour Azure Data Lake Store. Pour plus d'informations, consultez la section [Composite Data Domain Discovery](#).

Microsoft Azure SQL Data Warehouse

Microsoft Azure SQL Data Warehouse est un entrepôt de données pour les entreprises, entièrement géré dans le nuage.

Objets extraits

- Tables
- Vues
- Procédure stockée étendue
- Table (définie par l'utilisateur)
- Synonyme
- Fonctions
- Contrainte UNIQUE
- Déclencheurs
- Procédures
- Index
- Fonction table SQL
- Utilisateurs
- Schéma de partition
- Fonction de partition
- Collections de schémas XML
- Type de données défini par l'utilisateur
- Règles
- Valeurs par défaut
- Contraintes
- SQL_INLINE_TABLE_VALUED_FUNCTION
- SQL_SCALAR_FUNCTION
- SEQUENCE
- Base de données
- Schéma
- Procédures stockées

Autorisations de configuration de la ressource

Avant de créer une ressource Microsoft Azure SQL, configurez les autorisations du compte utilisateur de la base de données Microsoft Azure SQL Server que vous utilisez pour vous connecter à Microsoft Azure SQL Server. Enterprise Data Catalog utilise l'authentification SQL Server pour se connecter à la base de données Microsoft Azure SQL Server. Le compte utilisateur que vous utilisez pour vous connecter à Microsoft Azure SQL Server doit être un compte de connexion SQL Server. Configurez l'autorisation VIEW DEFINITION pour le compte utilisateur. La configuration de cette autorisation répertorie tous les schémas à partir desquels l'utilisateur peut charger des métadonnées.

Prérequis

Vérifiez que vous avez configuré les autorisations VIEW DEFINITION et CONNECT pour l'utilisateur de la base de données Microsoft SQL Server. Vous devez vérifier que vous configurez `sqljdbc_auth.dll` dans la

variable d'environnement PATH. La version du fichier .dll doit correspondre à la version du fichier sqljdbc4.jar que vous utilisez.

Propriétés de connexion à la ressource

Les tableaux suivants répertorient les propriétés à configurer pour ajouter une ressource Microsoft Azure SQL Data Warehouse :

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Utilisateur	Nom d'utilisateur que vous devez spécifier pour vous connecter à la base de données Microsoft SQL Server. Si vous utilisez un pont Microsoft SQL pour vous connecter à la base de données et que vous laissez le nom d'utilisateur vide, par défaut, la signature de sécurité intégrée est utilisée pour la connexion à la base de données. Une sécurité intégrée utilise la signature suivante pour la connexion : jdbc:sqlserver://;integratedSecurity=true instead of the jdbc:sqlserver://;user=userid;password=userpassword.
Mot de passe	Mot de passe de l'utilisateur de la base de données.
Hôte	Nom d'hôte ou adresse IP de la machine sur laquelle Microsoft SQL Server est en cours d'exécution.
Port	Numéro de port de service du moteur de base de données Microsoft SQL Server. La valeur par défaut est 1433. Il est recommandé de spécifier le numéro de port lorsque vous vous connectez à l'aide de la propriété Instance . Si vous spécifiez à la fois les propriétés Port et Instance , Enterprise Data Catalog utilise la propriété Port .
Base de données	Nom de la base de données depuis laquelle vous voulez importer des métadonnées. Enterprise Data Catalog importe les tables et les schémas de la base de données.
Instance	Facultatif. Nom de l'instance de Microsoft SQL Server. Vous pouvez également spécifier le numéro de port de l'instance. Il est recommandé de spécifier le numéro de port de l'instance.

L'image suivante montre l'exemple de propriétés de connexion dans l'onglet **Général** :

Connection Properties

User*:

Administrator

Password*:

Host*:

AzureMicrosoftSQLDW.informatica.com

Port*:

1433

Database*:

HR_AMSDW

Instance:

0001

Test Connection

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Activer les métadonnées source	Sélectionnez cette option pour activer et configurer la ressource pour extraire des métadonnées.
Schéma	Facultatif. Spécifiez le schéma à importer. Utilisez des points-virgules (;) pour séparer plusieurs schémas. Si vous ne spécifiez pas de schéma, Enterprise Data Catalog importe tous les schémas.
Filtre de métadonnées source	Vous pouvez inclure ou exclure des tables et des vues depuis l'exécution de la ressource. Utilisez des points-virgules (;) pour séparer les noms des tables et les noms des vues. Pour plus d'informations sur le champ de filtre, voir "Filtre de profil de données et métadonnées source" à la page 12 .
Nom de conteneur BLOB Azure	Nom du conteneur de stockage BLOB Azure.
Sensible à la casse	Spécifie que la ressource est configurée pour l'insensibilité à la casse. Sélectionnez l'une des valeurs suivantes : <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. La valeur par défaut est False.
Mémoire	Sélectionnez la mémoire requise pour exécuter le travail du scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé : <ul style="list-style-type: none"> - Faible - Moyenne - Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource Microsoft Azure SQL Data Warehouse. Pour plus d'informations, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource Microsoft Azure SQL Data Warehouse. Pour plus d'informations, consultez la section [Composite Data Domain Discovery](#).

Microsoft Azure SQL Server

Microsoft Azure SQL Database est une base de données gérée dans le nuage.

Objets extraits

- Tables
- Vues
- Procédure stockée étendue
- Table (définie par l'utilisateur)
- Synonyme
- Contrainte UNIQUE
- Déclencheurs
- Procédures
- Index
- Fonction table SQL
- Utilisateurs
- Schéma de partition
- Fonction de partition
- Collections de schémas XML
- Type de données défini par l'utilisateur
- Règles
- Valeurs par défaut
- Contraintes
- SQL_INLINE_TABLE_VALUED_FUNCTION
- SQL_SCALAR_FUNCTION
- SEQUENCE
- Base de données
- Schéma
- Procédures stockées
- Fonctions

Autorisations de configuration de la ressource

Avant de créer une ressource Microsoft Azure SQL, configurez les autorisations du compte utilisateur de la base de données Microsoft Azure SQL Server que vous utilisez pour vous connecter à Microsoft Azure SQL Server. Enterprise Data Catalog utilise l'authentification SQL Server pour se connecter à la base de données Microsoft Azure SQL Server. Le compte utilisateur que vous utilisez pour vous connecter à Microsoft Azure

SQL Server doit être un compte de connexion SQL Server. Configurez l'autorisation VIEW DEFINITION pour le compte utilisateur. La configuration de cette autorisation répertorie tous les schémas à partir desquels l'utilisateur peut charger des métadonnées.

Prérequis

Veillez à configurer les autorisations VIEW DEFINITION et CONNECT pour l'utilisateur de la base de données Microsoft SQL Server. Vous devez également vous assurer que vous configurez `sqljdbc_auth.dll` dans la variable d'environnement PATH. La version du fichier .dll doit correspondre à la version du fichier sqljdbc4.jar que vous utilisez.

Propriétés de connexion à la ressource

Les tableaux suivants répertorient les propriétés à configurer pour ajouter une ressource Microsoft Azure SQL Server.

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Utilisateur	Nom d'utilisateur que vous devez spécifier pour vous connecter à la base de données Microsoft SQL Server. Si vous utilisez un pont Microsoft SQL pour vous connecter à la base de données et que vous laissez le nom d'utilisateur vide, par défaut, Enterprise Data Catalog utilise la signature de sécurité intégrée pour la connexion à la base de données. Une sécurité intégrée utilise la signature suivante pour la connexion : <code>jdbc:sqlserver://;integratedSecurity=true</code> instead of the <code>jdbc:sqlserver://;user=userid;password=userpassword</code> .
Mot de passe	Mot de passe que vous utilisez pour l'utilisateur de la base de données.
Hôte	Nom d'hôte ou adresse IP de la machine sur laquelle Microsoft SQL Server est en cours d'exécution.
Port	Numéro de port de service du moteur de base de données Microsoft SQL Server. La valeur par défaut est 1433. Il est recommandé de spécifier le numéro de port lorsque vous vous connectez à l'aide de la propriété Instance . Si vous spécifiez à la fois les propriétés Port et Instance , Enterprise Data Catalog utilise la propriété Port .
Base de données	Nom de la base de données depuis laquelle vous voulez importer des métadonnées. Enterprise Data Catalog importe les tables et les schémas de la base de données.
Instance	Facultatif. Nom de l'instance de Microsoft SQL Server. Vous pouvez également spécifier le numéro de port de l'instance. Il est recommandé de spécifier le numéro de port de l'instance.

L'image suivante montre l'exemple de propriétés de connexion dans l'onglet **Général** :

Connection Properties

User*:	<input type="text" value="Administrator"/>
Password*:	<input type="password" value="*****"/>
Host*:	<input type="text" value="INLDMDB01.informatica.com"/>
Port*:	<input type="text" value="1433"/>
Database*:	<input type="text" value="Finance_AMSQLServer"/>
Instance:	<input type="text" value="0006"/>
<input type="button" value="Test Connection"/>	

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Importer des objets système	Facultatif. Sélectionnez cette option pour importer des objets système. Par défaut, Enterprise Data Catalog n'importe pas les objets système.
Schéma	(Facultatif) Cliquez sur Sélectionner pour spécifier les schémas que vous souhaitez importer. Vous pouvez utiliser l'une des options suivantes dans la boîte de dialogue Sélectionner un schéma pour importer les schémas : <ul style="list-style-type: none">- Sélectionner dans la liste. Sélectionnez les schémas requis dans une liste de schémas disponibles.- Sélectionner à l'aide d'une expression régulière. Indiquez une expression régulière SQL pour sélectionner des schémas qui correspondent à l'expression. Si vous ne spécifiez pas de schéma, Enterprise Data Catalog importe tous les schémas.
Filtre de métadonnées source	Vous pouvez inclure ou exclure des tables et des vues depuis l'exécution de la ressource. Utilisez des points-virgules (;) pour séparer les noms des tables et les noms des vues.
Importer des procédures stockées	Sélectionnez cette propriété pour importer des procédures stockées. Enterprise Data Catalog n'importe pas les procédures stockées par défaut.
Sensible à la casse	Spécifie que la ressource est configurée pour l'insensibilité à la casse. Sélectionnez l'une des valeurs suivantes : <ul style="list-style-type: none">- True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse.- False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. La valeur par défaut est False.

Propriété	Description
Mémoire	<p>Sélectionnez la mémoire requise pour exécuter le travail du scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource Microsoft Azure SQL Server. Pour plus d'informations, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource Microsoft Azure SQL Server. Pour plus d'informations, consultez la section [Composite Data Domain Discovery](#).

Google BigQuery

Vous pouvez utiliser la ressource Google BigQuery pour collecter les métadonnées des actifs contenus dans Google BigQuery.

Objets extraits

La ressource Google BigQuery extrait les métadonnées des actifs suivants dans une source de données Google BigQuery :

- Projet
- Jeu de données
- Table
- Afficher

Autorisations de configuration de la ressource

Assurez-vous d'effectuer les étapes suivantes avant de configurer la ressource Google BigQuery :

- Attribuez le rôle `dataViewer` au compte de service que vous utilisez pour accéder au projet Google Cloud Platform.
- Configurez les autorisations `bigquery.tables.list` et `bigquery.jobs.get` pour le compte de service que vous utilisez pour accéder au projet Google Cloud Platform.

Connexion à une source de données Google BigQuery activée pour SSL

Pour vous connecter à une source de données Google BigQuery activée pour SSL, procédez comme suit :

1. Téléchargez les certificats SSL Google BigQuery à l'aide d'un navigateur Web.

Remarque: Veillez à importer le certificat de Google Trust Services dans le chemin de certification.

2. Copiez les certificats dans le répertoire `<INFA_HOME>/services/shared/security/`.
3. Accédez au répertoire `<INFA_HOME>/source/java/jre/bin/`, puis exécutez la commande `keytool` pour importer chaque certificat copié en tant que certificat approuvé dans le fichier keystore du domaine Informatica :

```
keytool -import -file <INFA_HOME>/services/shared/security/<certificate>.cer -alias  
<alias name> -keystore <INFA_HOME>/services/shared/security/infa_truststore.jks -  
storepass <Informatica domain keystore password>
```

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
ID projet	Nom du projet Google Cloud Platform auquel vous voulez accéder.
Clé privée	Clé privée associée au compte de service.
E-mail du client	Adresse e-mail du client associée au compte de service.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Analyser les jeux de données masqués	Extrait les métadonnées des jeux de données masqués et anonymes.
Jeu de données	Sélectionnez les jeux de données que vous souhaitez utiliser pour importer les métadonnées des tables Google BigQuery du projet. La valeur par défaut est tous les jeux de données.
Filtre de métadonnées source	Vous pouvez inclure ou exclure des tables et des vues depuis l'exécution de la ressource. Utilisez des points-virgules (;) pour séparer les noms des tables et les noms des vues. Pour plus d'informations sur le champ de filtre, voir "Filtre de profil de données et métadonnées source" à la page 12.

Propriété	Description
Sensible à la casse	<p>Spécifie que la ressource est configurée pour la sensibilité à la casse. Sélectionnez l'une des valeurs suivantes :</p> <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. <p>La valeur par défaut est True.</p>
Mémoire	<p>Spécifie la mémoire requise pour exécuter la tâche de scanneur. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données que vous prévoyez d'importer dans le catalogue :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. La paire clé-valeur que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Remarque: La liste suivante répertorie les fonctionnalités que la ressource Google BigQuery ne prend pas en charge :

- Lors du profilage, les colonnes de type de données Géographie et Octet sont ignorées.
- La ressource Google BigQuery ne lit pas les enregistrements multiples lorsque vous utilisez la requête SQL héritée.
- Google BigQuery ne prend pas en charge une table partitionnée par un champ.

Microsoft Azure Blob Storage

Configurez un type de ressource Microsoft Azure Blob Storage pour extraire des métadonnées de Microsoft Azure Blob Storage.

Objets extraits

Enterprise Data Catalog n'extrait que les fichiers de la source Microsoft Azure Blob Storage.

Autorisations de configuration de la ressource

Configurez l'autorisation `read` dans la source de données Microsoft Azure Blob Storage pour le compte utilisateur permettant d'accéder à la source de données.

Types de fichiers pris en charge

La ressource Microsoft Azure Blob Storage permet d'extraire les métadonnées de fichiers structurés, non structurés et non structurés étendus.

Fichiers structurés pris en charge :

- Fichiers AVRO
- Fichiers délimités
- Fichiers texte
- Fichiers JSON
- Fichiers Parquet
- Fichiers XML

Fichiers non structurés pris en charge :

- Fichiers Apple
- Fichiers compressés
- E-mail

Fichiers non structurés étendus :

- Fichiers VB
- Fichiers ASP
- Fichiers TIF
- Fichiers LOG
- Fichiers CSS
- Fichiers ASPX
- Fichiers DLL
- Fichiers GIF
- Fichiers SQL

Attribuez les autorisations `read` et `write` aux fichiers pour l'extraction des métadonnées.

Propriétés de connexion à la ressource

Indiquez la **Clé de compte** de Microsoft Azure Blob Storage dans Informatica Administrator. Utilisez la valeur indiquée pour la **clé 1** sous **Clés d'accès** dans le portail Azure.

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
URL du point de terminaison d'objet blob	URL de Microsoft Azure Blob Storage pour accéder à un conteneur. Utilisez la valeur indiquée pour l'URL du point de terminaison de service BLOB principal dans le portail Azure.
Nom du compte	Nom du compte Microsoft Azure Blob Storage.

Propriété	Description
Nom du conteneur	Nom du conteneur d'objets blob.
Répertoire source	Répertoire source à partir duquel les métadonnées doivent être extraites.
Jeton de signature d'accès partagé	Jeton permettant d'accéder à la ressource Microsoft Azure Blob Storage. Utilisez la valeur indiquée pour la signature d'accès partagé dans le portail Azure.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Préfixe de l'objet blob	Utilisez cette option pour trier les objets blob en fonction du préfixe du nom Microsoft Azure Blob. Cette valeur est sensible à la casse.
Types de fichiers	<p>Sélectionnez un ou tous les types de fichiers suivants à partir desquels vous voulez extraire les métadonnées :</p> <ul style="list-style-type: none"> - Tous. Utilisez cette option pour spécifier si vous voulez extraire des métadonnées de tous les types de fichiers. - Sélectionner. Utilisez cette option pour spécifier que vous voulez extraire des métadonnées à partir de types de fichiers spécifiques. Effectuez les étapes suivantes pour spécifier les types de fichiers : <ol style="list-style-type: none"> 1. Cliquez sur Sélectionner. La boîte de dialogue Sélectionner des types de fichiers spécifiques s'affiche. 2. Sélectionnez les fichiers requis parmi les options suivantes : <ul style="list-style-type: none"> - Formats non structurés étendus Utilisez cette option pour extraire les métadonnées des types de fichiers tels que les fichiers audio, vidéo, image et les ebooks. - Types de fichiers structurés. Utilisez cette option pour extraire des métadonnées à partir de types de fichiers tels que Avro, Parquet, JSON, XML, fichiers texte et fichiers délimités. - Types de fichiers non structurés. Utilisez cette option pour extraire des métadonnées à partir de types de fichiers tels que Microsoft Excel, Microsoft PowerPoint, Microsoft Word, pages Web, fichiers compressés, e-mails et PDF. 3. Cliquez sur Sélectionner. <p>Remarque: Vous pouvez sélectionner l'option Types de fichiers spécifiques dans la boîte de dialogue pour sélectionner les fichiers de toutes les catégories.</p>
Autres types de fichiers	Extrait les métadonnées de base des fichiers (taille de fichier, chemin et horodatage) pour les types de fichiers qui ne figurent pas dans la propriété Types de fichiers .
Traiter les fichiers sans extension As	<p>Sélectionnez l'une des options suivantes pour identifier les fichiers sans extension :</p> <ul style="list-style-type: none"> - Aucun - Avro - Parquet

Propriété	Description
Entrer le délimiteur de fichier	<p>Spécifiez le délimiteur de fichier si le fichier à partir duquel vous extrayez des métadonnées utilise un délimiteur autre qu'un délimiteur de la liste suivante :</p> <ul style="list-style-type: none"> - Virgule (,) - Tabulation horizontale (\t) - Point-virgule (;) - Deux points (:) - Barre verticale () <p>Assurez-vous de placer le délimiteur entre des guillemets simples. Par exemple : '\$ '. Utilisez une virgule pour séparer plusieurs délimiteurs. Par exemple : '\$ ', '% ', '& '</p>
Répertoire de premier niveau	<p>Spécifiez un répertoire ou une liste de répertoires sous le répertoire source. Si vous ne spécifiez pas cette option, Enterprise Data Catalog importe tous les fichiers du répertoire source spécifié. Pour spécifier un répertoire ou une liste de répertoires, vous pouvez effectuer les étapes suivantes :</p> <ol style="list-style-type: none"> 1. Cliquez sur Sélectionner.... La boîte de dialogue Sélectionner répertoire de premier niveau s'affiche. 2. Utiliser l'une des options suivantes pour sélectionner les répertoires requis : <ul style="list-style-type: none"> - Sélectionner dans la liste : sélectionnez les répertoires requis à partir d'une liste de répertoires. - Sélectionner à l'aide d'une expression régulière : fournissez une expression régulière SQL pour sélectionner des schémas qui correspondent à l'expression. <p>Remarque: Si vous voulez sélectionner plusieurs répertoires, vous devez les séparer par un point-virgule (;).</p>
Inclure le sous-répertoire	Sélectionnez cette option pour importer tous les fichiers des sous-répertoires du répertoire source.
Sensible à la casse	<p>Spécifie que la ressource est configurée pour la sensibilité à la casse. Sélectionnez l'une des valeurs suivantes :</p> <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. <p>La valeur par défaut est True.</p>
Mémoire	<p>Sélectionnez la mémoire requise pour exécuter le travail du scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>

Propriété	Description
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> -Dscannerloglevel=<DEBUG/INFO/ERROR>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. -Dscanner.container.core=<No. of core>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. -Dscanner.yarn.app.environment=<key=value>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. -Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0>. Augmente la mémoire du conteneur du scanneur lorsque pmem est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource Microsoft Azure Blob Storage. Pour plus d'informations, consultez la section [Enable Data Discovery](#). Vous pouvez activer la découverte de domaines de données composites pour une ressource Microsoft Azure Blob Storage. Pour plus d'informations, consultez la section [Composite Data Domain Discovery](#).

Salesforce

Utilisez une connexion Salesforce pour vous connecter à un objet Salesforce. La connexion Salesforce est un type de connexion d'application.

Objets extraits

Enterprise Data Catalog n'extrait que les tables de la source Salesforce.

Prérequis

Pour respecter les prérequis, effectuez les étapes suivantes :

1. Exportez les certificats Salesforce au format DER à partir des pages Web suivantes :
 - <https://login.salesforce.com>
 - Page utilisateur de Salesforce
2. Importez les certificats dans le fichier <INFA_HOME>/services/shared/security/infa_truststore.jks.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de connexion du type de ressource Salesforce :

Propriété	Description
Type d'authentification	Le type d'authentification pour se connecter à Salesforce.
Jeton d'actualisation	Entrez le jeton d'actualisation de Salesforce.
Clé du consommateur	Clé du consommateur obtenue à partir de Salesforce, requise pour générer le jeton d'actualisation. Pour plus d'informations sur la génération de la clé du consommateur, consultez la documentation de Salesforce.
Secret du consommateur	Secret du consommateur obtenu à partir de Salesforce, requis pour générer le jeton d'actualisation. Pour plus d'informations sur la génération du secret du consommateur, consultez la documentation de Salesforce.
Nom d'utilisateur	Nom d'utilisateur Salesforce.
Mot de passe	Mot de passe et jeton de sécurité Salesforce pour le nom d'utilisateur Salesforce. Remarque: Assurez-vous de fournir le mot de passe et le jeton Salesforce en tant que chaîne.
URL de service	URL du service Salesforce auquel vous voulez accéder.

Le tableau suivant décrit la propriété **avancée** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Sensible à la casse	Spécifie que la ressource est configurée pour l'insensibilité à la casse. - Cochez cette case pour indiquer que la ressource est configurée comme étant sensible à la casse. - Décochez cette case pour indiquer que la ressource est configurée comme étant insensible à la casse. Par défaut, la ressource est configurée comme étant insensible à la casse.
Mémoire	Valeur de mémoire requise pour exécuter une tâche de scanneur. Spécifiez l'une des valeurs de mémoire suivantes : - Faible - Moyenne - Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .

Propriété	Description
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> -Dscannerloglevel=<DEBUG/INFO/ERROR>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. -Dscanner.container.core=<No. of core>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. -Dscanner.yarn.app.environment=<key=value>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. -Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0>. Augmente la mémoire du conteneur du scanneur lorsque pmem est activé. La valeur par défaut est 1. -DenableSalesforceRelationships=true. Permet d'afficher toutes les relations pour un objet.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource Salesforce. Pour plus d'informations, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource Salesforce. Pour plus d'informations, consultez la section [Composite Data Domain Discovery](#).

Configuration d'une connexion par proxy pour accéder à la ressource Salesforce

Vous pouvez configurer une connexion par proxy pour que la ressource accède à la source de données Salesforce. Une connexion proxy sécurise l'accès à la source de données.

Pour plus d'informations sur la configuration d'une connexion par proxy pour accéder à la source de données Salesforce, consultez l'article *Configuration d'une connexion par proxy pour accéder à la source de données Salesforce* de la base de connaissances.

Workday

Vous pouvez utiliser la ressource Workday pour collecter les métadonnées depuis des actifs contenus dans Workday.

Objets extraits

La ressource Workday extrait les métadonnées des actifs suivants dans une source de données Workday :

- Service
- Entité*
- Rapport
- Opération
- Source de données
- Propriété*
- Objets d'entreprise

Remarque: *Enterprise Data Catalog extrait les métadonnées du service Workday en tant que schéma XSD. Dans le schéma, une entité représente les éléments XML complexes, tandis que la propriété représente les éléments XML simples.

Autorisations de configuration de la ressource

Lorsque vous configurez une ressource Workday, veillez à utiliser un compte utilisateur disposant des privilèges d'administrateur Workday.

Propriétés de connexion à la ressource

Le tableau suivant répertorie les propriétés de connexion que vous devez configurer dans l'onglet **Général** pour la ressource :

Propriété	Description
Nom d'utilisateur	Nom d'utilisateur pour accéder au locataire Workday. Un locataire Workday représente une instance sécurisée du logiciel Workday.
Mot de passe	Mot de passe pour accéder au locataire Workday.
Locataire	URL de l'instance Workday à laquelle vous voulez accéder.
Domaine Workday	Nom du domaine Workday qui contient les actifs Workday auxquels vous voulez accéder. Dans Workday, un domaine représente une collection d'actifs associés qui sont sécurisés à l'aide d'une stratégie de sécurité de domaine. Par exemple, un domaine peut inclure des actifs tels que des rapports, des sources de données pour le rapport et des propriétés des sources de données.

Le tableau suivant répertorie les paramètres d'extraction des métadonnées que vous pouvez configurer dans l'onglet **Paramètres de chargement des métadonnées** pour la ressource :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Version	Version de l'API Workday Web Services. La liste des services Workday disponibles varie en fonction de la version que vous sélectionnez.
Services	Services Workday à partir desquels vous voulez extraire les métadonnées. Dans Workday, un service représente un module qui répond aux besoins d'une fonction. Par exemple, la fonction de gestion des services dans le domaine de la paie peut être fournie par le module Paie.
Importer des objets d'entreprise, des sources de données et des rapports	Importe les métadonnées associées aux objets d'entreprise, aux sources de données et aux rapports depuis la source de données Workday.

Propriété	Description
Sensible à la casse	<p>Spécifie que la ressource est configurée pour la sensibilité à la casse. Sélectionnez l'une des valeurs suivantes :</p> <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. <p>La valeur par défaut est True.</p>
Mémoire	<p>Spécifie la mémoire requise pour exécuter la tâche de scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données que vous prévoyez d'importer dans le catalogue :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. La paire clé-valeur que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	<p>Sélectionnez cette option pour afficher les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.</p>

AWS Glue

Amazon Web Services Glue est un service ETL de l'écosystème Amazon Web Services qui utilise les données déplacées sur différents magasins. Glue capture les métadonnées des différents magasins de données qui font partie de l'écosystème Amazon Web Services.

Objets extraits

La ressource Glue extrait les métadonnées des actifs suivants dans une source de données Glue :

- Bases de données
- Tables
- Colonnes
- Métadonnées dérivées du catalogue Glue. Vous pouvez également afficher le lignage entre les actifs Glue et les actifs dérivés.

Remarque: Si l'actif dérivé est un type de données de structure, le catalogue affiche cet actif sous forme de champ aux côtés de la hiérarchie structurelle existante. Par exemple, un fichier `person.json` contient des informations relatives au prénom, au nom de famille, au numéro de téléphone et à l'adresse. L'adresse comprend des champs imbriqués tels que ville, État et code postal. Enterprise Data Catalog extrait l'actif de la personne sous forme de champ, aux côtés des champs imbriqués.

- Tâches

Remarque: Athena, un service basé sur les requêtes fourni par Amazon S3, a été créé en tant que ressource JDBC de référence qui peut interagir avec les sources de données, telles que Tableau et autres.

Autorisations de configuration de la ressource

Pour accéder à Glue, assurez-vous d'effectuer les étapes suivantes avant de configurer la ressource :

- Attribuez le rôle `AWSGlueServiceRole` à l'utilisateur IAM Amazon Web Services.
- Créez une stratégie personnalisée avec les autorisations suivantes sur le service Glue, puis attribuez-la à un utilisateur IAM Amazon Web Services :
 - `GetConnection`
 - `GetConnections`
 - `GetDatabase`
 - `GetDatabases`
 - `GetJob`
 - `GetJobs`
 - `GetPartition`
 - `GetPartitions`
 - `GetTable`
 - `GetTables`

Connexion à une source de données Glue activée pour SSL

Pour vous connecter à une source de données Glue activée pour SSL, procédez comme suit :

1. Téléchargez les certificats SSL Glue à l'aide d'un navigateur Web.

Remarque: Veillez à importer le certificat Glue Trust Services dans le répertoire Certificats.
2. Copiez les certificats dans le répertoire `<INFA_HOME>/services/shared/security/`.
3. Accédez au répertoire `<INFA_HOME>/source/java/jre/bin`, puis exécutez la commande `keytool` pour importer chaque certificat copié en tant que certificat approuvé dans le fichier keystore du domaine Informatica :

```
keytool -import -file <INFA_HOME>/services/shared/security/<certificate>.cer -alias  
<alias name> -keystore <INFA_HOME>/services/shared/security/infa_truststore.jks -  
storepass <Informatica domain keystore password>
```

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Authentification basée sur un rôle	Option permettant d'utiliser les informations d'identification du profil de l'instance d'Amazon Elastic Compute Cloud (Amazon EC2) lorsqu'Enterprise Data Catalog est installé sur une instance d'Amazon EC2.
Clé d'accès AWS	Clé d'accès du compte Amazon Web Services.
Clé secrète AWS	Clé secrète du compte Amazon Web Services.
Région AWS	Région Amazon Web Services à partir de laquelle vous souhaitez scanner le catalogue Glue.

Le tableau suivant décrit les propriétés que vous pouvez configurer dans la section **Métadonnées source** de l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Active l'extraction des métadonnées.
Filtre de base de données	Filtre qui vous permet d'inclure ou d'exclure des bases de données de l'exécution de la ressource. Vous pouvez également spécifier une expression régulière qui représente les bases de données à inclure ou exclure.
Filtre de table	Filtre qui vous permet d'entrer une combinaison appropriée d'expression régulière et de caractères spéciaux pour inclure ou exclure des actifs spécifiques de l'exécution de la ressource qui correspond au format de l'expression régulière. Vous pouvez également entrer des noms de table pour les inclure dans l'exécution de la ressource. Utilisez un point-virgule pour séparer les modèles de caractères spéciaux et les noms de table.
Activer les ressources de référence	Option permettant d'extraire les métadonnées des actifs qui ne sont pas inclus dans la ressource, mais qui y sont référés. Par exemple, des tables source et cible dans des mappages PowerCenter, et des tables et fichiers source des rapports Tableau.
Créer des ressources Athena	Spécifie si une source de données Athena doit être créée.
Conserver les actifs de référence non résolus	Option permettant de conserver les actifs de référence non résolus dans le catalogue après l'attribution des connexions. La conservation des actifs de référence non résolus vous permet de visualiser l'intégralité du lignage. Les actifs non résolus incluent les fichiers supprimés, les tables temporaires et d'autres actifs qui ne figurent pas dans la ressource principale.
Attribution automatique des connexions	Spécifie si les connexions doivent être attribuées automatiquement.

Propriété	Description
Mémoire	<p>Valeur de mémoire requise pour exécuter une tâche de scanneur.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, INFO ou ERROR. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Paire clé-valeur que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur pour les paires clé-valeur multiples. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Snowflake

Snowflake est un entrepôt de données analytique fourni sous la forme d'une application SaaS (Software-as-a-Service). Snowflake est plus rapide, plus facile à utiliser et plus flexible que les autres offres d'entrepôt de données standard. L'entrepôt de données Snowflake utilise un moteur de base de données SQL avec une architecture unique conçue pour le nuage.

Objets extraits

La ressource Snowflake extrait les métadonnées des actifs suivants dans une source de données Snowflake :

- Lignage récapitulatif de la vue à la table
- Base de données permanente
- Schéma permanent
- Table permanente
- Table temporaire
- Table externe
- Colonnes
- Vues avec informations de lignage

- Vues sécurisées sans informations de lignage
- Vue matérialisée avec informations de lignage
- Afficher le lignage détaillé des objets Snowflake à l'aide de Tableau, PowerCenter ou Informatica Intelligent Cloud Services
- Les commentaires des tables, des vues et des colonnes sont affichés sous la forme de descriptions de source dans Enterprise Data Catalog
- Les relations entre la clé primaire et la clé étrangère des tables utilisant les contraintes en ligne et hors ligne sont affichées dans Enterprise Data Catalog

Les objets Snowflake suivants ne sont pas extraits par la ressource Snowflake :

- Base de données temporaire
- Procédure stockée
- Fonctions
- Schéma temporaire
- Table temporaire
- Séquences

Connexion à une source de données Snowflake activée pour SSL

Pour vous connecter à une source de données Snowflake activée pour SSL, procédez comme suit :

1. Téléchargez les certificats SSL Snowflake suivants à l'aide d'un navigateur Web :
 - Starfield Class 2 Certification Authority
 - Starfield Services Root Certificate Authority - G2
 - *.snowflakecomputing.com

Remarque: Veillez à importer le certificat Snowflake Trust Services dans le répertoire Certificats.
2. Téléchargez les certificats à partir de la plate-forme cloud sur laquelle le compte Snowflake est hébergé. Le compte Snowflake est hébergé sur l'une des plates-formes cloud suivantes :
 - Amazon Web Services (AWS).
 - Google Cloud Platform (GCP)
 - Microsoft Azure (Azure)
3. Copiez les certificats dans le répertoire <INFA_HOME>/services/shared/security/.
4. Accédez au répertoire <INFA_HOME>/source/java/jre/bin, puis exécutez la commande `keytool` pour importer chaque certificat copié en tant que certificat approuvé dans le fichier keystore du domaine Informatica :

```
keytool -import -file <INFA_HOME>/services/shared/security/<certificate>.cer -alias
<alias name> -keystore <INFA_HOME>/services/shared/security/infa_truststore.jks -
storepass <Informatica domain keystore password>
```

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Compte	Nom du compte Snowflake.
Paramètres supplémentaires	Spécifiez au moins un paramètre de connexion JDBC au format suivant : <param1>=<value>&<param2>=<value>&<param3>=<value> Par exemple : user=jon&warehouse=mywh&db=mydb&schema=public
Mot de passe	Mot de passe associé au compte utilisateur Snowflake.
Rôle	Rôle Snowflake attribué à l'utilisateur.
Utilisateur	Nom d'utilisateur pour la connexion au compte Snowflake.

Le tableau suivant décrit les propriétés que vous pouvez configurer dans la section **Métadonnées source** de l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Active l'extraction des métadonnées..
Sensible à la casse	Spécifie que la ressource est configurée pour l'insensibilité à la casse. Sélectionnez l'une des valeurs suivantes : True. Cochez cette case pour indiquer que la ressource est configurée comme étant sensible à la casse. False. Décochez cette case pour indiquer que la ressource est configurée comme étant insensible à la casse. La valeur par défaut est True.
Schéma	Sélectionnez les schémas de base de données sur lesquels vous voulez appliquer la contrainte.
Mémoire	Valeur de mémoire requise pour exécuter une tâche de scanneur. Spécifiez l'une des valeurs de mémoire suivantes : <ul style="list-style-type: none">- Faible- Moyenne- Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .

Propriété	Description
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> -Dscanner.loglevel=<DEBUG/INFO/ERROR>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, INFO ou ERROR. La valeur par défaut est INFO. -Dscanner.container.core=<No. of core>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. -Dscanner.yarn.app.environment=<key=value>. Paire clé-valeur que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur pour les paires clé-valeur multiples. -Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0>. Augmente la mémoire du conteneur du scanner lorsque pmem est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource Snowflake. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource Snowflake. Pour plus d'informations sur l'activation de la découverte de données composites, consultez la section [Composite Data Domain Discovery](#).

Google Cloud Storage

Google Cloud Storage est un service Web de stockage de fichier en ligne qui stocke les données et y accède à partir de Google Cloud Platform.

Objets extraits

La ressource Google Cloud Storage extrait les métadonnées des actifs suivants dans une source de données Google Cloud Storage :

- File
- Dossier

Connexion à une source de données Google Cloud Storage activée pour SSL

Pour vous connecter à une source de données Google Cloud Storage activée pour SSL, procédez comme suit :

1. Téléchargez les certificats SSL Google Cloud Storage à l'aide d'un navigateur Web.

Remarque: Veillez à importer le certificat de Google Cloud Storage Trust Services dans le répertoire Certificats.

2. Copiez les certificats dans le répertoire <INFA_HOME>/services/shared/security/.

3. Accédez au répertoire <INFA_HOME>/source/java/jre/bin, puis exécutez la commande `keytool` pour importer chaque certificat copié en tant que certificat approuvé dans le fichier keystore du domaine Informatica :

```
keytool -import -file <INFA_HOME>/services/shared/security/<certificate>.cer -alias  
<alias name> -keystore <INFA_HOME>/services/shared/security/infa_truststore.jks -  
storepass <Informatica domain keystore password>
```

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
ID de projet	ID du projet Google Cloud Platform
Clé privée	Clé privée du compte de service Google Cloud Platform
E-mail du client	ID de l'e-mail du client du compte de service Google Cloud Platform
Nom du compartiment	Nom du compartiment des objets stockés qui doivent être extraits.
Répertoire source	Chemin du répertoire source de l'extraction de métadonnées. Utilisez une barre oblique en tant que suffixe vers le chemin du répertoire source.

Le tableau suivant décrit les propriétés que vous pouvez configurer dans la section **Métadonnées source** de l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Active l'extraction des métadonnées.
Types de fichiers	Permet d'extraire les métadonnées de tous les fichiers ou de fichiers spécifiques.
Types de fichiers spécifiques	Sélectionnez le type de fichier pour l'extraction de métadonnées.
Entrer le délimiteur de fichier	Spécifiez le délimiteur de fichier. Enterprise Data Catalog prend en charge les délimiteurs suivants : Virgule (,), Deux points (:), Point-virgule (;), Tabulation (\t) et Barre verticale (). Utilisez des guillemets simples pour spécifier un autre délimiteur de fichier.
Autres types de fichiers	Extrait les métadonnées de fichier, telles que la taille du fichier, le chemin d'accès au fichier et l'horodatage d'autres types de fichiers.
Répertoire de premier niveau	Option permettant d'ajouter une contrainte lors de l'importation d'un répertoire de premier niveau à partir du répertoire source. Une chaîne vide indique que les fichiers du répertoire source ont été importés. Un nom de fichier indique que tous les fichiers du répertoire source ainsi que les fichiers du répertoire spécifié ont été importés.
Inclure le sous-répertoire	Option permettant d'inclure tous les répertoires dans le répertoire de premier niveau sélectionné lors de l'extraction des métadonnées. Si le répertoire de premier niveau est vide, tous les répertoires dans le répertoire source sont inclus dans l'extraction de métadonnées.

Propriété	Description
Sensible à la casse	<p>Spécifie que la ressource est configurée pour l'insensibilité à la casse.</p> <p>Cochez cette case pour configurer la ressource comme étant sensible à la casse. Décochez cette case pour configurer la ressource comme étant insensible à la casse.</p> <p>Par défaut, la ressource est configurée comme étant sensible à la casse.</p>
Mémoire	<p>Valeur de mémoire requise pour exécuter une tâche de scanneur.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, INFO ou ERROR. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Paire clé-valeur que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur pour les paires clé-valeur multiples. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	<p>Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.</p>

CHAPITRE 3

Configuration des ressources Data Engineering

Ce chapitre comprend les rubriques suivantes :

- [Présentation des ressources Data Engineering, 52](#)
- [Apache Atlas, 52](#)
- [Cloudera Navigator, 54](#)
- [HDFS, 57](#)
- [Hive, 63](#)

Présentation des ressources Data Engineering

Le groupe de ressources Data Engineering inclut les ressources liées à l'ingénierie des données. Lorsque vous créez et exécutez une ressource Data Engineering, le scanneur Data Engineering extrait et migre les résultats de métadonnées de la ressource associée vers le catalogue. Vous pouvez afficher les résultats migrés dans Enterprise Data Catalog.

Apache Atlas

Apache Atlas est l'infrastructure de gouvernance et de métadonnées pour Hadoop. Apache Atlas possède une architecture évolutive et extensible qui peut être reliée à de nombreux composants Hadoop pour gérer leurs métadonnées dans un référentiel central.

Objets extraits

La ressource Apache Atlas extrait les liens d'objets de données des sources de données suivantes :

- Bases de données relationnelles
- Hive
- HDFS

Remarque:

- Veuillez à configurer les ressources dans Enterprise Data Catalog pour les sources de données à partir desquelles vous voulez extraire les liens d'objets de données.
- Le cluster doit avoir accès à Internet, car l'extraction des métadonnées des ressources Apache Atlas nécessite quelques fichiers binaires.

Prérequis

Si le cluster est activé pour Kerberos, effectuez les étapes suivantes :

1. Copiez les fichiers `krb5.conf` et `keytab` du cluster vers les emplacements suivants sur le domaine Informatica :
 - `<INFA_HOME>/services/shared/security`
 - `<INFA_HOME>/java/jre/lib/security`
2. Spécifiez le chemin d'accès au fichier `krb5.conf` dans le domaine Informatica et les nœuds du service de cluster Informatica.
3. Ajoutez le Nom de principal du service (SPN) et les propriétés du keytab aux propriétés du service d'intégration de données.
4. Exécutez la commande `kinit` sur le domaine Informatica avec le SPN et le fichier `keytab` requis.
5. Exécutez la commande `kinit` sur le cluster avec le SPN et le fichier `keytab` requis.
6. Redémarrez le service d'intégration de données et le service de catalogue.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de connexion du type de ressource Atlas :

Propriété	Description
URL	URL permettant d'accéder à Atlas Apache.
Authentification	Sélectionnez l'une des options suivantes pour spécifier le type d'authentification configuré pour Apache Atlas : <ul style="list-style-type: none">- Simple. Spécifiez les paramètres suivants :<ul style="list-style-type: none">- Connexion. Spécifiez le nom d'utilisateur configuré pour accéder à Apache Atlas.- Mot de passe. Spécifiez le mot de passe configuré pour accéder à Apache Atlas.- Kerberos. Spécifiez les paramètres suivants :<ul style="list-style-type: none">- Fichier de configuration Kerberos. Cliquez sur Sélectionner pour sélectionner et charger le fichier de configuration Kerberos utilisé pour l'authentification.- Fichier Keytab Kerberos. Cliquez sur Sélectionner pour sélectionner et charger le fichier keytab Kerberos utilisé pour l'authentification.- Principal. Spécifiez le principal Kerberos utilisé pour l'authentification.

L'image suivante montre un exemple de propriétés de connexion dans l'onglet **Général** :

Connection Properties

URL*:

Authentication*: ☐ Simple ☒ Kerberos

Kerberos configuration file*:

Kerberos Keytab file*:

Principal*:

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Attribution automatique des connexions	Sélectionnez cette option pour spécifier que la connexion doit être attribuée automatiquement.
Activer les ressources de référence	Option permettant d'extraire les métadonnées des actifs qui ne sont pas inclus dans la ressource, mais qui y sont référés. Par exemple, des tables source et cible dans des mappages PowerCenter, et des tables et fichiers source des rapports Tableau.
Conserver les actifs de référence non résolus	Option permettant de conserver les actifs de référence non résolus dans le catalogue après l'attribution des connexions. La conservation des actifs de référence non résolus vous permet de visualiser l'intégralité du lignage. Les actifs non résolus incluent les fichiers supprimés, les tables temporaires et d'autres actifs qui ne figurent pas dans la ressource principale.
Mémoire	<p>Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanneur.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que <code>DEBUG</code>, <code>ERROR</code> ou <code>INFO</code>. La valeur par défaut est <code>INFO</code>. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Options d'agent	Spécifiez les options de l'agent Enterprise Data Catalog pour l'exécution de la tâche de scanneur.

Cloudera Navigator

Cloudera Navigator est un système de sécurité et de gestion de données entièrement intégré pour la plateforme Hadoop.

Objets extraits

La ressource Cloudera Navigator extrait les opérations Hive, Impala et Spark.

Autorisations de configuration de la ressource

Avant de créer une ressource Cloudera Navigator, vous devez configurer la taille des segments de mémoire Java pour le serveur Cloudera Navigator et la taille maximale des segments de mémoire pour le service de catalogue. Si vous ne configurez pas correctement la taille des segments de mémoire, le chargement des métadonnées peut échouer.

Configurez les tailles des segments de mémoire suivantes :

- Taille des segments de mémoire Java pour le serveur Navigator. Avant de créer une ressource Cloudera Navigator, définissez la taille des segments de mémoire Java pour le serveur Cloudera Navigator sur au moins 2 Go. Si la taille des segments de mémoire n'est pas suffisante, le chargement des ressources échoue en générant une erreur de connexion refusée.
- Taille maximale des segments de mémoire pour le service de catalogue. Avant de créer une ressource Cloudera Navigator, ouvrez l'outil Administrator tool et vérifiez la valeur de la propriété Taille maximale des segments de mémoire du service de catalogue. Définissez la taille maximale des segments de mémoire sur au moins 4 096 Mo (4 Go).

Si vous effectuez plusieurs chargements de ressources simultanément, augmentez la taille maximale des segments de mémoire d'au moins 1 024 Mo (1 Go) pour chaque chargement de ressources. Pour charger simultanément, par exemple, deux ressources Cloudera Navigator, augmentez la taille maximale des segments de mémoire de 2 048 Mo (2 Go). Par conséquent, vous définissez l'option Taille maximale des segments de mémoire sur 6 144 Mo.

Remarque: Il est possible que certaines distributions Cloudera nécessitent une taille maximale des segments de mémoire supérieure à 4 Go. Si le chargement des métadonnées échoue avec une erreur de mémoire insuffisante, augmentez la taille maximale des segments de mémoire.

Sur la source de données Cloudera Navigator, attribuez l'autorisation `Navigator Administrator` pour le compte utilisateur que vous utilisez pour accéder à la source de données. Il s'agit de l'autorisation minimale requise pour extraire les métadonnées à partir de la source de données Cloudera Navigator.

Propriétés de connexion à la ressource

Configurez les propriétés de connexion lorsque vous créez ou modifiez une ressource Cloudera Navigator.

Le tableau suivant décrit les propriétés de connexion :

Propriété	Description
URL Navigator	<p>URL du serveur Navigator Cloudera.</p> <p>Vous pouvez fournir les options suivantes dans l'URL pour ajuster la ressource :</p> <ul style="list-style-type: none">- <code>searchLimit</code>. Spécifie la limite de recherche Solr. Utilisez ce paramètre si vous souhaitez restreindre le nombre d'objets récupérés dans Cloudera Navigator. La valeur par défaut est 10 000.- <code>maximumPoolSize</code>. Spécifie le nombre maximal de connexions à Cloudera Navigator. La valeur par défaut est 20. La valeur minimale est 10.- <code>queueSize</code>. Spécifie le nombre de tâches conservées dans la file d'attente avant l'exécution des tâches. La valeur par défaut est cinq fois la valeur de <code>maximumPoolSize</code>. Une valeur plus élevée pour l'option <code>queueSize</code> nécessite plus de mémoire.- <code>disableIncremental</code>. Définissez cette option sur <code>true</code> pour désactiver les analyses incrémentielles par la ressource. <p>Fournissez l'URL avec les options de syntonisation comme indiqué dans l'exemple de format suivant :</p> <pre>http://<host name:<port>>/? searchLimit=10000&maximumPoolSize=15&queueSize=1000&disableIncremental=true</pre>
Utilisateur	Nom du compte utilisateur permettant de se connecter à Cloudera Navigator.
Mot de passe	Mot de passe du compte d'utilisateur permettant de se connecter à Cloudera Navigator.

L'image suivante montre un exemple de propriétés de connexion dans l'onglet **Général** :

Connection Properties

Navigateur URL*:

User*:

Password*:

Disable SSL Validation: ☐

Test Connection

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Attribution automatique des connexions	Sélectionnez cette option pour spécifier que la connexion doit être attribuée automatiquement.
Activer les ressources de référence	Option permettant d'extraire les métadonnées des actifs qui ne sont pas inclus dans la ressource, mais qui y sont référés. Par exemple, des tables source et cible dans des mappages PowerCenter, et des tables et fichiers source des rapports Tableau.
Conserver les actifs de référence non résolus	Option permettant de conserver les actifs de référence non résolus dans le catalogue après l'attribution des connexions. La conservation des actifs de référence non résolus vous permet de visualiser l'intégralité du lignage. Les actifs non résolus incluent les fichiers supprimés, les tables temporaires et d'autres actifs qui ne figurent pas dans la ressource principale.
Base de données Hive	Nom de la base de données Hive ou d'un schéma à partir duquel vous voulez importer une table.
Lignage détaillé	<p>Sélectionnez cette option pour extraire et ingérer les métadonnées liées à la logique de transformation des actifs qui incluent des transformations.</p> <p>Transformation</p> <p>Une transformation indique la génération, la modification ou le passage de données entre les connexions source et cible.</p> <p>Logique de transformation</p> <p>Une logique de transformation affiche les mappages ou les types de relations de flux de données entre les ressources source et cible liées à la ressource sélectionnée dans Enterprise Data Catalog.</p>

Propriété	Description
Mémoire	<p>Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanner.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal..</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

HDFS

Utilisez une ressource HDFS pour importer des métadonnées à partir de fichiers CSV, XML et JSON.

Objets extraits

La ressource HDFS extrait les métadonnées de fichiers dans une source de données HDFS.

Autorisations de configuration de la ressource

Configurez l'autorisation `read` dans la source de données HDFS pour le compte utilisateur permettant d'accéder à la source de données.

Types de fichiers pris en charge

La ressource HDFS permet d'extraire les métadonnées de fichiers structurés, non structurés et non structurés étendus.

Fichiers structurés pris en charge :

- Fichiers AVRO
- Fichiers délimités
- Fichiers texte
- Fichiers JSON
- Fichiers Parquet
- fichiers XML

Fichiers non structurés pris en charge :

- Fichiers Apple

- Fichiers compressés
- E-mail

Fichiers non structurés étendus :

- Fichiers VB
- Fichiers ASP
- Fichiers TIF
- Fichiers LOG
- Fichiers CSS
- Fichiers ASPX
- Fichiers DLL
- Fichiers GIF
- Fichiers SQL

Attribuez les autorisations `read` et `write` aux fichiers pour l'extraction des métadonnées.

Prérequis

Si le domaine et le cluster sont activés respectivement pour SSL et Kerberos, effectuez les étapes suivantes :

1. Copiez les fichiers `krb5.conf`, `infa_truststore.jks` et `keytab` du cluster vers les emplacements suivants sur le domaine Informatica :
 - `<INFA_HOME>/services/shared/security`
 - `<INFA_HOME>/java/jre/lib/security`
2. Spécifiez le chemin d'accès au fichier `krb5.conf` dans le domaine Informatica et les nœuds du service de cluster Informatica.
3. Ajoutez le Nom de principal du service (SPN) et les propriétés du `keytab` aux propriétés du service d'intégration de données.
4. Exécutez la commande `kinit` sur le domaine Informatica avec le SPN et le fichier `keytab` requis.
5. Exécutez la commande `kinit` sur le cluster avec le SPN et le fichier `keytab` requis.
6. Redémarrez le service d'intégration de données et le service de catalogue.

Choisir la connexion Hadoop pour les ressources HDFS existantes

Si vous disposez d'une ressource HDFS qui utilise le moteur Hive comme environnement d'exécution, modifiez la ressource, choisissez Blaze, Spark ou Databricks comme option **Exécuter sur**, sélectionnez une connexion Hadoop, puis exécutez la ressource pour afficher les résultats de profil.

Importer la ressource HDFS

Effectuez les étapes suivantes si vous utilisez une ressource HDFS :

1. Si vous souhaitez activer le profilage pour la ressource HDFS, créez une configuration de cluster basée sur le type de cluster que vous utilisez.
2. Modifiez la ressource HDFS pour attribuer la nouvelle connexion HDFS dans le champ Connexion source.
3. Sélectionnez **Blaze**, **Spark** ou **Databricks** comme option **Exécuter sur**.
4. Sélectionnez le nom de la connexion Hadoop.

5. Si la source de données Hive se trouve sur un cluster compatible avec Kerberos, assurez-vous d'exécuter les étapes suivantes :
 - a. Utilisez Informatica Administrator pour configurer le nom principal du service Kerberos Hadoop et le fichier Keytab Kerberos Hadoop dans les propriétés du service d'intégration de données.
 - b. Définissez la valeur du répertoire de distribution Hadoop du service d'intégration de données sur /data/
 - c. Créez la configuration de cluster si vous ne l'avez pas fait avant la mise à niveau. Si vous l'avez créée avant la mise à niveau, actualisez-la à l'aide d'Informatica Administrator ou de la CLI.
 - d. Utilisez Informatica Administrator pour ajouter la ligne suivante aux options de ligne de commande JVM pour le service d'intégration de données : -Djava.security.krb5.conf=/data. Les options JVM se trouvent dans la section Processus du service d'intégration de données.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Configuration de grappe Détails	Sélectionnez l'une des options suivantes pour spécifier les détails de configuration de la ressource : <ul style="list-style-type: none"> - Fournir les détails de configuration du cluster - Charger à partir du fichier d'archivage de configuration
Type de stockage	Sélectionnez l'une des options suivantes pour spécifier le type de stockage à partir duquel vous voulez extraire les métadonnées : <ul style="list-style-type: none"> - DFS. Distributed File System - WASB. Stockage blob Windows Azure. Configurez les options suivantes si vous sélectionnez WASB : <ul style="list-style-type: none"> - URI du compte de stockage Azure. URI entièrement qualifié pour accéder aux données stockées dans WASB. - Nom du compte de stockage Azure. Nom du compte de stockage. - Clé du compte de stockage Azure. Clé permettant d'accéder au compte de stockage. - ABFS. Système de fichiers Azure Blob File System. Configurez les options suivantes si vous sélectionnez ABFS : <ul style="list-style-type: none"> - URI du compte de stockage Azure. URI complet pour accéder aux données stockées dans ABFS. - Nom du compte de stockage Azure. Nom du compte de stockage. - Clé du compte de stockage Azure. Clé permettant d'accéder au compte de stockage.
URI 1 du nœud de nom.	URI du NameNode HDFS actif. Le NameNode HDFS actif gère toutes les opérations clientes dans le cluster.
Cluster HA	Sélectionnez Oui si le cluster est configuré pour la haute disponibilité et configurez les propriétés suivantes : <ul style="list-style-type: none"> - URI 2 du nœud de nom. URI du NameNode HDFS secondaire. Le NameNode HDFS secondaire stocke des modifications apportées à HDFS en tant que fichier journal ajouté à un fichier système de fichier natif. - Nom du service HDFS. Nom du service configuré pour HDFS.

Propriété	Description
Type de distribution	<p>Sélectionnez l'un des types de distribution Hadoop suivants pour la ressource HDFS :</p> <ul style="list-style-type: none"> - Cloudera - Hortonworks - IBM BigInsights - Azure HDInsight - Amazon EMR - MapR FS. Remplissez les prérequis avant de sélectionner MapR comme type de distribution. Pour plus d'informations, consultez la section <i>Prérequis pour le type de distribution MapR FS</i>. <p>Remarque: Cette propriété est utilisée si vous sélectionnez l'option Charger à partir du fichier d'archivage de configuration pour configurer la ressource.</p>
Accueil MAPR	<p>Spécifiez le chemin d'installation du client MapR.</p> <p>Remarque: Cette propriété est utilisée si vous avez sélectionné MapR FS comme type de distribution Hadoop.</p>
Nom d'utilisateur/ principal d'utilisateur	Nom d'utilisateur permettant de se connecter à HDFS. Spécifiez le principal Kerberos si le cluster est compatible avec Kerberos.
Répertoire source	Emplacement source à partir duquel les métadonnées doivent être extraites.
Fichier d'archivage de configuration	<p>Un fichier ZIP qui contient les propriétés de configuration des ressources dans les fichiers XML suivants :</p> <ul style="list-style-type: none"> - core-site.xml - hdfs-site.xml - mapred-site.xml - yarn-site.xml <p>Remarque: Cette propriété est utilisée si vous sélectionnez l'option Charger à partir du fichier d'archivage de configuration pour configurer la ressource.</p>
Cryptage transparent HDFS	<p>Sélectionnez Oui si le chiffrement transparent est activé pour HDFS. Fournissez l'URI complet au fournisseur de clés du serveur gestionnaire de clés dans la zone URI du fournisseur de serveur de gestion de clés.</p> <p>Si vous sélectionnez Oui, vous devez importer le certificat SSL dans le fichier truststore <code>infa_truststore.jks</code> du domaine Informatica qui se trouve à l'emplacement <code><INFA_HOME>/services/shared/security</code>.</p>
Cluster Kerberos	<p>Sélectionnez Oui si le cluster est compatible avec Kerberos. Si le cluster est compatible avec Kerberos, indiquez les informations suivantes :</p> <ul style="list-style-type: none"> - Protection Hadoop RPC. Sélectionnez l'une des options suivantes en fonction de la valeur de protection des appels de procédure distante (RPC) configurée pour le cluster : <ul style="list-style-type: none"> - authentification - intégrité - confidentialité <p>La valeur par défaut est authentification.</p> <ul style="list-style-type: none"> - Principal du service HDFS. Nom du principal du service HDFS. - Fichier Keytab. Chemin d'accès au fichier Keytab du principal Kerberos. Assurez-vous que le fichier Keytab se trouve à l'emplacement spécifié sur l'hôte du domaine Informatica et les hôtes du cluster du service de catalogue.

Pour plus d'informations sur la configuration du type de distribution d'accès au stockage Azure pour Azure HDInsight, reportez-vous au *Guide de Data Engineering Integration*.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Types de fichiers	<p>Sélectionnez un ou tous les types de fichiers suivants à partir desquels vous voulez extraire les métadonnées :</p> <ul style="list-style-type: none"> - Tous. Utilisez cette option pour spécifier si vous voulez extraire des métadonnées de tous les types de fichiers. - Sélectionner. Utilisez cette option pour spécifier que vous voulez extraire des métadonnées à partir de types de fichiers spécifiques. Effectuez les étapes suivantes pour spécifier les types de fichiers : <ol style="list-style-type: none"> 1. Cliquez sur Sélectionner. La boîte de dialogue Sélectionner des types de fichiers spécifiques s'affiche. 2. Sélectionnez les fichiers requis parmi les options suivantes : <ul style="list-style-type: none"> - Formats non structurés étendus Utilisez cette option pour extraire les métadonnées des types de fichiers tels que les fichiers audio, vidéo, image et les ebooks. - Types de fichiers structurés. Utilisez cette option pour extraire les métadonnées des types de fichiers tels que JSON, Avro, Parquet, XML, texte et les fichiers délimités. - Types de fichiers non structurés. Utilisez cette option pour extraire des métadonnées à partir de types de fichiers tels que Microsoft Excel, Microsoft PowerPoint, Microsoft Word, pages Web, fichiers compressés, e-mails et PDF. 3. Cliquez sur Sélectionner. <p>Remarque: Vous pouvez sélectionner l'option Types de fichiers spécifiques dans la boîte de dialogue pour sélectionner les fichiers de toutes les catégories.</p>
Autres types de fichiers	Extrait les métadonnées de base des fichiers (taille de fichier, chemin et horodatage) pour les types de fichiers qui ne sont pas répertoriés dans la propriété Types de fichiers .
Traiter les fichiers sans extension As	<p>Sélectionnez l'une des options suivantes pour identifier les fichiers sans extension :</p> <ul style="list-style-type: none"> - Aucun - Avro - Parquet
Entrer le délimiteur de fichier	<p>Spécifiez le délimiteur de fichier si le fichier à partir duquel vous extrayez des métadonnées utilise un délimiteur autre qu'un délimiteur de la liste suivante :</p> <ul style="list-style-type: none"> - Virgule (,) - Tabulation horizontale (\t) - Point-virgule (;) - Deux points (:) - Barre verticale () <p>Assurez-vous de placer le délimiteur entre des guillemets simples. Par exemple : '\$ '. Utilisez une virgule pour séparer plusieurs délimiteurs. Par exemple : '\$ ', '% ', '& '</p>
Répertoire de premier niveau	Spécifie que tous les répertoires doivent être sélectionnés. Si vous voulez sélectionner des répertoires spécifiques, utilisez l'option Sélectionner un répertoire . Cette option est désactivée si vous avez sélectionné l'option Inclure les sous-répertoires dans l'onglet Général .
Inclure le sous-répertoire	Entrez les répertoires requis dans la zone de texte ou cliquez sur Sélectionner... pour choisir les répertoires requis. Cette option est désactivée si vous avez sélectionné l'option Inclure les sous-répertoires dans l'onglet Général ou l'option Sélectionner tous les répertoires répertoriée ci-dessus.

Propriété	Description
Sensible à la casse	Indique si la ressource est configurée pour la sensibilité à la casse. Sélectionnez l'une des valeurs suivantes : <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. La valeur par défaut est True.
Mémoire	Sélectionnez la mémoire requise pour exécuter le travail du scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé : <ul style="list-style-type: none"> - Faible - Moyenne - Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1. - <code>Djava.security.auth.login.config={MAPR_HOME}/conf/mapr.login.conf</code>. Extrait les métadonnées d'une distribution MapR Hadoop qui accède aux données d'une source de données HDFS. Le client MapR installé sur le domaine Informatica et le cluster doit être à la version 6.1.0, et la version distribuée de MapR Hadoop doit être MEP 6.x.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource HDFS. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource HDFS. Pour plus d'informations sur l'activation de la découverte de domaines de données, consultez la section [Composite Data Domain Discovery](#).

Prérequis pour le type de distribution MapR FS

Remplissez les prérequis suivants avant de sélectionner MapR FS comme type de distribution :

- Vérifiez que le client MapR est installé sur les nœuds du service de cluster Informatica.
- Vérifiez que le client MapR est installé dans le même emplacement sur les nœuds du service de cluster Informatica.
- Configurez le client MapR pour générer le fichier `mapr-clusters.conf` et les autres fichiers requis pour se connecter au cluster MapR. Exécutez la commande `/opt/mapr/server/configure.sh -N my.cluster.com -c -C mynode01:7222 -HS mynode02` pour configurer le client MapR.

Hive

Apache Hive est un logiciel d'entreposage de données reposant sur Apache Hadoop afin de fournir la requête et l'analyse de données. Apache Hive fournit une interface semblable à SQL afin d'interroger les données stockées dans des bases de données et des systèmes de fichiers différents qui s'intègrent à Hadoop.

Objets extraits

La ressource Hive extrait les métadonnées des actifs suivants dans une source de données Hive :

- Tables
- Vues
- Base de données
- Schéma
- Détails de connexion pour les vues de différents schémas

Autorisations de configuration de la ressource

Configurez l'autorisation `read` dans la source de données Hive pour le compte utilisateur permettant d'accéder à la source de données.

Prérequis

Si le domaine et le cluster sont activés respectivement pour SSL et Kerberos, effectuez les étapes suivantes :

1. Copiez les fichiers `krb5.conf` et `infa_truststore.jks` à partir du cluster, et le fichier `hive.service.keytab` à partir de `/etc/security/keytabs` dans les emplacements suivants sur le domaine Informatica :
 - `<INFA_HOME>/services/shared/security`
 - `<INFA_HOME>/java/jre/lib/security`
2. Spécifiez le chemin d'accès au fichier `krb5.conf` dans le domaine Informatica et les nœuds du service de cluster Informatica.
3. Ajoutez le Nom de principal du service (SPN) et les propriétés du keytab aux propriétés du service d'intégration de données.
4. Exécutez la commande `kinit` sur le domaine Informatica avec le SPN et le fichier keytab requis.
5. Exécutez la commande `kinit` sur le cluster avec le SPN et le fichier keytab requis.
6. Redémarrez le service d'intégration de données et le service de catalogue.

Choisir la connexion Hadoop pour les ressources Hive existantes

Si vous disposez d'une ressource Hive qui utilise le moteur Hive comme environnement d'exécution, modifiez la ressource, choisissez Blaze, Spark ou Databricks comme option **Exécuter sur**, sélectionnez une connexion Hadoop, puis exécutez la ressource pour afficher les résultats de profil.

Importer la ressource Hive

Effectuez les étapes suivantes si vous utilisez une ressource Hive :

1. Si vous souhaitez activer le profilage pour la ressource Hive, créez une configuration de cluster basée sur le type de cluster que vous utilisez.
2. Modifiez la ressource Hive pour attribuer la nouvelle connexion Hive dans le champ Connexion source.
3. Sélectionnez **Blaze**, **Spark** ou **Databricks** comme option **Exécuter sur**.
4. Sélectionnez le nom de la connexion Hadoop.

5. Si la source de données Hive se trouve sur un cluster compatible avec Kerberos, assurez-vous d'exécuter les étapes suivantes :
 - a. Utilisez Informatica Administrator pour configurer le nom principal du service Kerberos Hadoop et le fichier Keytab Kerberos Hadoop dans les propriétés du service d'intégration de données.
 - b. Définissez la valeur du répertoire de distribution Hadoop du service d'intégration de données sur /data/
 - c. Créez la configuration de cluster si vous ne l'avez pas fait avant la mise à niveau. Si vous l'avez créée avant la mise à niveau, actualisez-la à l'aide d'Informatica Administrator ou de la CLI.
 - d. Utilisez Informatica Administrator pour ajouter la ligne suivante aux options de ligne de commande JVM pour le service d'intégration de données : -Djava.security.krb5.conf=/data. Les options JVM se trouvent dans la section Processus du service d'intégration de données.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de la connexion :

Propriété	Description
Distribution Hadoop	Sélectionnez l'un des types de distribution Hadoop suivants pour la ressource Hive : <ul style="list-style-type: none">- Cloudera- Hortonworks- MapR- Amazon EMR- Azure HDInsight- IBM BigInsights
URL	URL de connexion JDBC utilisée pour accéder au serveur Hive.
Utilisateur	Nom d'utilisateur Hive.
Mot de passe	Mot de passe du nom d'utilisateur Hive.
Fichier Keytab	Chemin d'accès au fichier Keytab si Hive utilise Kerberos pour l'authentification.
Proxy de l'utilisateur	Nom d'utilisateur proxy à utiliser si Hive utilise Kerberos pour l'authentification.
Fichier de configuration Kerberos	Spécifiez le chemin d'accès au fichier de configuration Kerberos si vous utilisez l'authentification basée sur Kerberos pour Hive.
Activer le débogage pour Kerberos	Sélectionnez cette option pour activer les options de débogage pour l'authentification basée sur Kerberos.

Pour plus d'informations sur la configuration du type de distribution d'accès au stockage Azure pour Azure HDInsight, reportez-vous au *Guide de Data Engineering Integration*.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Schéma	<p>Cliquez sur Sélectionner... pour spécifier les schémas Hive que vous voulez importer. Vous pouvez utiliser l'une des options suivantes dans la boîte de dialogue Sélectionner un schéma pour importer les schémas :</p> <ul style="list-style-type: none"> - Sélectionner dans la liste : utilisez cette option pour sélectionner les schémas requis à partir d'une liste de schémas disponibles. - Sélectionner à l'aide d'une expression régulière : fournissez une expression régulière SQL pour sélectionner des schémas qui correspondent à l'expression.
Filtre de métadonnées source	<p>Vous pouvez inclure ou exclure des tables et des vues depuis l'exécution de la ressource. Utilisez des points-virgules (;) pour séparer les noms des tables et les noms des vues.</p> <p>Pour plus d'informations sur le champ de filtre, voir "Filtre de profil de données et métadonnées source" à la page 12.</p>
Table	Spécifiez le nom de la table Hive à importer. Si vous laissez cette propriété vide, Enterprise Data Catalog importe toutes les tables Hive.
Liste des fichiers JAR SerDe	Spécifiez le chemin d'accès à la liste de fichiers JAR du sérialiseur/désérialiseur (SerDe). Vous pouvez spécifier plusieurs fichiers JAR en séparant les chemins de fichiers à l'aide d'un point-virgule (;).
Threads de travailleur	<p>Spécifiez le nombre de threads de travailleur permettant de traiter les métadonnées de façon asynchrone. Ne spécifiez pas la valeur si vous souhaitez qu'Enterprise Data Catalog la calcule. Enterprise Data Catalog attribue une valeur comprise entre un et six en fonction de l'architecture JVM et du nombre de cœurs de processeur disponibles.</p> <p>Vous pouvez utiliser les points suivants pour décider de la valeur à utiliser :</p> <ul style="list-style-type: none"> - Vous pouvez fournir une valeur supérieure ou égale à un et inférieure à six pour spécifier le nombre de threads de travailleur requis. - Si vous spécifiez une valeur non valide, Enterprise Data Catalog affiche un avertissement et utilise la valeur un. - Si votre machine a plus de mémoire, vous pouvez spécifier une valeur plus élevée pour traiter plus de métadonnées de façon asynchrone. <p>Remarque: La spécification d'une valeur plus élevée peut affecter les performances du système.</p>
Sensible à la casse	<p>Spécifie que la ressource est configurée pour l'insensibilité à la casse. Sélectionnez l'une des valeurs suivantes :</p> <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. <p>La valeur par défaut est False.</p>

Propriété	Description
Mémoire	<p>Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanner.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.
Attribution automatique des connexions	Spécifie si les connexions doivent être attribuées automatiquement.
Activer les ressources de référence	Extrait les métadonnées des actifs non inclus dans cette ressource, mais qui y sont référencés. Par exemple, des tables source et cible dans des mappages PowerCenter, et des tables et fichiers source des rapports Tableau.
Conserver les actifs de référence non résolus	Conserve les actifs de référence non résolus dans le catalogue après l'attribution des connexions. La conservation des actifs de référence non résolus vous permet de visualiser l'intégralité du lignage. Les actifs non résolus incluent les fichiers supprimés, les tables temporaires et d'autres actifs qui ne figurent pas dans la ressource principale.

Vous pouvez activer la découverte de données pour une ressource Hive. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource Hive. Pour plus d'informations sur l'activation de la découverte de données composites, consultez la section [Composite Data Domain Discovery](#).

Configurer la ressource Hive avec la passerelle Apache Knox

Enterprise Data Catalog prend en charge Knox si vous configurez Hive pour Knox. Assurez-vous d'installer Informatica et le service d'hébergement Hive sur le même cluster.

Remarque: Vous ne pouvez pas déployer Enterprise Data Catalog sur un cluster si vous configurez tous les services sur les nœuds pour Knox. Vérifiez que vous configurez Knox pour le service Hive et non pour les autres services qui s'exécutent sur les nœuds.

CHAPITRE 4

Configuration des ressources d'application

Ce chapitre comprend les rubriques suivantes :

- [Présentation des ressources d'application, 67](#)
- [SAP R/3, 67](#)
- [SAP S/4HANA, 70](#)

Présentation des ressources d'application

Le groupe de ressources de l'application Informatica inclut la ressource SAP. Lorsque vous créez et exécutez la ressource SAP, le scanner SAP extrait et migre les résultats de métadonnées dans l'instance associée de SAP vers le catalogue. Vous pouvez afficher les résultats de métadonnées migrés dans Enterprise Data Catalog.

SAP R/3

Vous devez remplir les conditions préalables répertoriées avant de configurer les propriétés d'une ressource SAP R/3.

Objets extraits

La ressource SAP R/3 extrait les métadonnées des actifs suivants dans une source de données SAP R/3 :

- Composants de l'application
- Packages
- Tables
- Colonnes
- Clés primaires et étrangères

Autorisations de configuration de la ressource

Pour configurer SAP R/3, effectuez les tâches suivantes :

1. Installez les transports de scanner.

2. Configurez un profil d'autorisation de l'utilisateur.
3. Installez les bibliothèques de SDK SAP.

Installation des transports de scanneur

Vous devez installer les transports de scanneur suivants, qui sont utilisés par Enterprise Data Catalog pour extraire des métadonnées à partir de SAP R/3 :

- ZINFABC_RUN
- TBL_READ_RUN
- TRANS_VER_RUN
- TBL_READ_RUN_CMP
- TBL_DESIGN_PROGINFO
- TBL_DESIGN
- TBL_READ_RUN_V2

Les transports sont situés dans le dossier suivant à l'emplacement où vous avez téléchargé PowerCenter :

<download location>\saptrans\mySAP.

Pour plus d'informations sur l'installation des transports, consultez la *Notice d'installation des versions de transports de PowerExchange for SAP NetWeaver*.

Configuration d'un profil d'autorisation de l'utilisateur

L'administrateur SAP R/3 doit créer le profil d'autorisation de l'utilisateur de produit et de développement. Le tableau suivant décrit le profil d'autorisation de l'utilisateur :

Objet d'autorisation	Description	Classe	Valeurs de champ
S_RFC	Vérification de l'autorisation pour l'accès à RFC.	Projets d'autorisation inter-applications	Activité : 16 (Exécuter) Nom de RFC à protéger : *. Type d'objet RFC à protéger : FUGR

Prérequis

1. Créez une connexion Event Details Record (EDR) pour SAP R/3 à l'aide d'Informatica Administrator. Pour en savoir plus sur la création de connexions, consultez le *Guide d'Informatica Administrator*.
2. Désactivez le service de catalogue.
3. Supprimez tous les fichiers binaires de ressources dans HDFS depuis le répertoire /Informatica/LDM/<service cluster name>/scanner/*. Vous pouvez utiliser la commande `hdfs dfs -rm -R /Informatica/LDM/<service cluster name>/scanner/*` pour supprimer tous les fichiers binaires de ressources.
4. Téléchargez le fichier `sapjco3.jar` et copiez-le dans l'emplacement suivant : <Install_directory>/services/CatalogService/access/WEB-INF/lib.
5. Si vous souhaitez activer le profilage, copiez les fichiers suivants aux emplacements indiqués :
 - Pour Windows, copiez le fichier `sapjco3.jar` dans l'emplacement suivant : <Install_directory>/services/shared/jars/thirdparty.
 - Pour Linux, copiez le fichier `libsapjco3.so` dans l'emplacement suivant : <Install_directory>/server/bin/.

6. Pour Linux, téléchargez le fichier `libsapjco3.so` et copiez-le dans l'emplacement suivant :
<Install_directory>/services/shared/bin.
7. Pour Windows, téléchargez le fichier `sapjco3.dll` et copiez-le dans l'emplacement suivant :
<Install_directory>/services/shared/bin.
8. Téléchargez le pilote JDBC pour SAP R/3 à partir du site Web de SAP et copiez-le dans le répertoire suivant : <Install_directory>/services/CatalogService/ScannerBinaries.
9. Intégrez les fichiers `libsapjco3.so` et `sapjco3.jar` au fichier `SAPJCO.zip` et copiez ce dernier dans l'emplacement suivant : <Install_directory>/services/CatalogService/ScannerBinaries.
10. Activez le service de catalogue.
11. Redémarrez le domaine Informatica.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Nom d'utilisateur	Nom d'utilisateur permettant d'accéder au système SAP R/3.
Mot de passe	Mot de passe associé au nom d'utilisateur.
Hôte du serveur d'applications	Nom d'hôte ou adresse IP du système qui héberge SAP R/3.
Numéro du système	Numéro de système du système SAP R/3 auquel vous voulez vous connecter.
Client	Client SAP R/3 permettant d'accéder aux données du système SAP R/3.
Langue	Spécifiez la langue à utiliser lors de l'importation des métadonnées à l'aide de la ressource.
Activer la diffusion en continu pour l'accès aux données	Choisissez cette option pour extraire des données à l'aide du protocole HTTP. Si vous ne choisissez pas cette option, le scanner extrait les données SAP dans un fichier intermédiaire.
Codage	Par défaut, le codage de caractères est UTF-8 pour les métadonnées importées depuis la ressource. Vous ne pouvez pas modifier le paramètre par défaut pour cette propriété.

Le tableau suivant décrit les propriétés que vous pouvez configurer dans la section **Métadonnées source** de l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Objets de référentiel	Importe les objets de référentiel tels que les ressources, les informations et les activités du système SAP R/3.

Propriété	Description
Mémoire	<p>Spécifie la mémoire requise pour exécuter la tâche de scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

Vous pouvez activer la découverte de données pour une ressource SAP R/3. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource SAP R/3. Pour plus d'informations sur l'activation de la découverte de données composites, consultez la section [Composite Data Domain Discovery](#).

SAP S/4HANA

SAP Business Suite 4 SAP HANA (SAP S/4HANA) est une suite ERP professionnelle que vous pouvez utiliser pour effectuer des transactions et analyser les données d'entreprise en temps réel.

Objets extraits

Vous pouvez utiliser la ressource SAP S/4HANA pour extraire les métadonnées des actifs suivants dans une source de données SAP S/4HANA :

- Composants de l'application
- Packages
- Classes
- Éléments de données
- Sources de données
- Domaines
- Modules de fonction
- Programmes
- Connexions RFC

- Tables
- Champs
- Codes de transaction
- Vues telles que des vues de base de données, vues de CDS, vues d'ajout et vues externes

Remarque: Si le type d'actif n'inclut pas d'informations de lignage, de hiérarchie, de catégories et de relation, vous ne pouvez pas afficher les onglets suivants dans le catalogue :

- Lignage et impact
- Relation
- Contient

Vous pouvez également afficher les détails du lignage entre le module de fonction et les actifs de source de données dans le catalogue.

Autorisations de configuration de la ressource

Pour configurer SAP S/4HANA, procédez comme suit :

1. Installez les transports de scanner.
2. Configurez un profil d'autorisation de l'utilisateur.

Installation des transports de scanner

Pour extraire les métadonnées de SAP S/4HANA, vous devez importer les transports de scanner qu'Enterprise Data Catalog utilise pour extraire les métadonnées de SAP S/4HANA. Les transports se trouvent dans le fichier `SAP_Scanner_Binaries.zip` téléchargé à partir de l'emplacement du programme d'installation Informatica.

Remarque: Pour extraire les données de profilage de SAP S/4HANA, vous devez importer les transports de scanner situés dans le dossier suivant dans lequel vous avez téléchargé PowerCenter : `<download location>\saptrans\mySAP`.

Configuration de profils d'autorisation de l'utilisateur

Pour accéder aux métadonnées à partir de SAP S/4HANA, l'administrateur SAP S/4HANA doit créer les profils d'autorisation d'utilisateur suivants :

Objet d'autorisation	Description	Valeurs de champ
S_RFC	Vérification de l'autorisation pour l'accès à RFC	ACTVT: 16 (Exécuter) RFC_NAME : DDIF_FIELDINFO_GET RFCPING RFC_GET_FUNCTION_INTERFACE SEU_COMPONENT ZINFA_ERP RFC_TYPE: FUGR, FUNC
S_RFC_ADM	Destination d'administration pour RFC	ACTVT: 03

Prérequis

1. Désactivez le service de catalogue.

2. Téléchargez le fichier `sapjco3.jar` et copiez-le dans les emplacements suivants :

- `<Install_directory>/services/CatalogService/access/WEB-INF/lib`
- `<Install_directory>/services/shared/jars/thirdparty`

Remarque: La version de SAP Java Connector prise en charge est la version 3.0.19.

3. Téléchargez le fichier `libsapjco3.so` et copiez-le dans les emplacements suivants :

- `<Install_directory>/services/shared/bin`
- `<Install_directory>/server/bin`

4. Intégrez le fichier `libsapjco3.so` au fichier `SAPJCO.zip` et copiez ce dernier dans l'emplacement suivant : `<Install_directory>/services/CatalogService/ScannerBinaries`.

5. Activez le service de catalogue.

Remarque: Si vous utilisez la ressource SAP S/4HANA pour extraire les données de profilage, vous devez redémarrer le domaine Informatica.

Prérequis à l'utilisation de la connexion SNC (Secure Network Communications)

1. Désactivez le service de catalogue.

2. Téléchargez le fichier `sapjco3.jar` et copiez-le dans les emplacements suivants :

- `<Install_directory>/services/CatalogService/access/WEB-INF/lib`
- `<Install_directory>/services/shared/jars/thirdparty`

Remarque: La version de SAP Java Connector prise en charge est la version 3.0.19.

3. Téléchargez le fichier `libsapjco3.so` et copiez-le dans les emplacements suivants :

- `<Install_directory>/services/shared/bin`
- `<Install_directory>/server/bin`

4. Configurez le type de connexion SNC avec le certificat X.509. Pour plus d'informations sur les étapes de configuration du type de connexion SNC, consultez la documentation de SAP.

5. Créez et définissez la variable d'environnement `SECUDIR` pour le service de catalogue comme emplacement du fichier `libsapcrypto.so`.

6. Incluez les fichiers suivants dans le fichier `SAPJCO.zip` :

- `cred_v2`
- `libsapcrypto.so`
- `sapjco3.jar`
- `libsapjco3.so`
- Certificat X.509

7. Copiez le fichier `SAPJCO.zip` à l'emplacement suivant : `<Install_directory>/services/CatalogService/ScannerBinaries`

8. Activez le service de catalogue.

9. Dans un cluster à plusieurs nœuds, vérifiez que vous avez effectué les tâches suivantes :

- Copier le fichier PSE (Personal Security Environment) au même chemin sur tous les nœuds du cluster.
- Configurer l'autorisation `read` dans la source de données SAP S/4HANA pour le compte utilisateur permettant d'accéder à la source de données.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Type de serveur SAP	Type d'installation de serveur SAP : <ul style="list-style-type: none">- Instance centrale SAP- Équilibrage de charge SAP (Serveur de message)
Hôte du serveur	Nom d'hôte ou adresse IP du système qui héberge SAP S/4HANA.
Groupe de connexion	Le groupe de connexion est un groupe de serveurs qui appartient à un système SAP. Les groupes de connexion distribuent les connexions des utilisateurs dans des instances individuelles ou des groupes d'instances SAP ou de serveurs d'application. Ils sont gérés dans les groupes SMLG (SAP Logon Groups) de transaction. Cette option s'affiche lorsque vous sélectionnez SAP Load Balancer (Message Server) comme type de serveur SAP.
Port du serveur de message	Port du serveur de message par défaut. Cette option s'affiche lorsque vous sélectionnez SAP Load Balancer (Message Server) comme type de serveur SAP.
Numéro du système	Numéro du système SAP S/4HANA. Cette option s'affiche lorsque vous sélectionnez SAP Central Instance comme type de serveur SAP.
SID SAP	Identifiant système SAP de l'instance de SAP.
Client	Client SAP S/4HANA permettant d'accéder aux données depuis SAP S/4HANA.
Langue	Langue pour l'importation des métadonnées à partir de la ressource. La langue de connexion sélectionnée est utilisée pour extraire les descriptions SAP.
Type de connexion	Type de connexion au serveur SAP : <ul style="list-style-type: none">- De base- SNC (Secure Network Communications). Vous pouvez utiliser le type de connexion pour extraire les métadonnées de la source de données SAP S/4HANA.
Nom d'utilisateur	Nom d'utilisateur du système SAP S/4HANA.
Mot de passe	Mot de passe du système SAP S/4HANA.
Nom du partenaire SNC	Le nom SNC du partenaire de communication.
Nom du fichier de certificat X.509	Le nom du fichier du certificat X.509 comme ticket de connexion.
Mon nom SNC	Le nom SNC de l'utilisateur qui envoie l'appel RFC (Remote Function Call) à partir d'Enterprise Data Catalog.

Propriété	Description
QoP SNC	<p>Spécifie le niveau de protection à utiliser pour la connexion.</p> <p>Voici les niveaux de protection disponibles :</p> <ul style="list-style-type: none"> - 1. Authentification uniquement - 2. Protection de l'intégrité - 3. Protection de la vie privée - 8. Utiliser la valeur de l'emplacement <code>snc/data_protection/use</code> sur le serveur d'applications - 9. Utiliser la valeur de l'emplacement <code>snc/data_protection/max</code> sur le serveur d'applications <p>Le niveau de protection par défaut est de 3.</p>
Attribution automatique des connexions	Spécifie si les connexions doivent être attribuées automatiquement.

Le tableau suivant décrit les propriétés que vous pouvez configurer dans la section **Métadonnées source** de l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Active l'extraction des métadonnées.
Objets de référentiel	Appliquez un filtre pour extraire les métadonnées basées sur les composants d'application et les packages SAP S/4HANA. Si vous ne sélectionnez pas un composant d'application ou un package SAP S/4HANA, la ressource effectue une analyse complète des métadonnées.
Types d'actifs	Sélectionnez les types d'actifs à charger dans le catalogue. Utilisez les options permettant d'affiner les objets de référentiel.
Activer la ressource de référence	Extrait les métadonnées relatives aux actifs qui ne sont pas incluses dans le composant ou le package sélectionné, mais qui y sont référencées.
Conserver les actifs de référence non résolus	Conserve les actifs de référence non résolus dans le catalogue après l'attribution des connexions. La conservation des actifs de référence non résolus vous permet de visualiser l'intégralité du lignage. Les actifs non résolus incluent des actifs non présents dans la ressource principale.
Mémoire	<p>Valeur de mémoire requise pour exécuter une tâche de scanneur.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>

Propriété	Description
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, INFO ou ERROR. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur pour les paires clé-valeur multiples. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1. <p>Remarque: Contactez le support client international Informatica avant de modifier les paramètres de la JVM par défaut.</p>
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez utiliser la ressource SAP S/4HANA pour extraire les informations de profilage suivantes pour un actif de table lors de la gestion intermédiaire :

- Découverte de domaine
- Profilage de colonne
- Découverte de relation

Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

CHAPITRE 5

Configuration des ressources Business Glossary

Ce chapitre comprend les rubriques suivantes :

- [Présentation des ressources Business Glossary, 76](#)
- [Informatica Axon, 76](#)
- [Business Glossary, 79](#)

Présentation des ressources Business Glossary

Le groupe de ressources Business Glossary inclut des ressources associées au glossaire d'entreprise. Lorsque vous créez et exécutez une ressource de glossaire d'entreprise, le scanner de glossaire d'entreprise extrait et migre les résultats de métadonnées de la ressource associée vers le catalogue. Vous pouvez afficher les résultats migrés dans Enterprise Data Catalog.

Informatica Axon

Informatica Axon est un référentiel de connaissances et un outil de gouvernance qui stocke les éléments de données de base et le contexte professionnel de votre organisation. Les propriétaires, gestionnaires, experts et autres intervenants responsables collaborent pour tracer progressivement la réalité commerciale des données, leur lignage et leur utilisation à travers des processus, des stratégies, des projets et des réglementation.

Objets extraits

La ressource Informatica Axon extrait les métadonnées des actifs suivants :

- Glossaire
- Stratégie
- Jeu de données
- Système
- Attribut. Enterprise Data Catalog extrait les attributs créés manuellement et organisés.

Vous pouvez consulter les détails des utilisateurs natifs, des utilisateurs LDAP et des utilisateurs SSO associés aux actifs dans le catalogue.

Vous pouvez afficher les relations suivantes dans le catalogue :

- Les relations parent-enfant entre des artefacts Axon et des objets extraits
- Les relations sémantiques entre les actifs du glossaire

Autorisations

Enterprise Data Catalog vous recommande de configurer l'autorisation `SuperAdmin` du compte utilisateur que vous utilisez pour accéder à la source de données Axon.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de connexion de la ressource Informatica Axon :

Propriété	Description
Nom d'utilisateur	Nom d'utilisateur Axon ou LDAP pour accéder à l'application Axon.
Mot de passe	Mot de passe Axon ou LDAP pour accéder à l'application Axon.
Hôte	Nom d'hôte complet pour accéder à l'application Axon.
Port	Numéro de port pour accéder à l'application Axon.
Activer la communication sécurisée	Sélectionnez la propriété si l'application Axon est activée pour SSL.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Statut d'Axon	Sélectionnez les objets comportant les valeurs de statut d'Axon à charger dans le catalogue.
Cycle de vie du glossaire	Sélectionnez les objets du glossaire comportant les valeurs de cycle de vie à charger dans le catalogue.
Cycle de vie du système	Sélectionnez les objets du système comportant les valeurs de cycle de vie à charger dans le catalogue.
Cycle de vie du jeu de données	Sélectionnez les objets du jeu de données comportant les valeurs de cycle de vie à charger dans le catalogue.
Cycle de vie de la stratégie	Sélectionnez les objets de la stratégie comportant les valeurs de cycle de vie à charger dans le catalogue.

Propriété	Description
Mémoire	<p>Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanner.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

Fonctionnalités non prises en charge

- Vous ne pouvez pas afficher la relation logique-physique entre les artefacts Axon suivants et les objets extraits dans le catalogue :
 - Système et ressource
 - Attribut et champ
- Vous ne pouvez pas afficher la relation de flux de données logique entre les actifs suivants dans le catalogue :
 - Attribut et attribut
 - Système et système
- Vous ne pouvez pas afficher la relation sémantique entre les actifs suivants dans le catalogue :
 - Stratégie et glossaire
 - Stratégie et système
 - Système et glossaire
- Vous ne pouvez pas afficher la relation glossaire-artefact Axon entre les actifs suivants dans le catalogue :
 - Glossaire et jeu de données
 - Glossaire et attribut

Business Glossary

Business Glossary comprend des glossaires en ligne de termes d'entreprise et de stratégies qui définissent les concepts importants d'une organisation. Configurez un type de ressource de classification Business Glossary pour extraire les métadonnées de Business Glossary.

Objets extraits

La ressource Business Glossary extrait des métadonnées des actifs suivants :

- Catégorie
- Terme
- Glossaire
- Stratégie

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de connexion du de ressource de classification Business Glossary :

Propriété	Description
Nom d'utilisateur	Nom du compte utilisateur permettant de se connecter à l'outil Analyst tool.
Mot de passe	Mot de passe du compte utilisateur permettant de se connecter à l'outil Analyst tool.
Hôte	Nom du glossaire d'entreprise de l'outil Analyst tool à partir duquel extraire les métadonnées. Chaque ressource peut extraire les métadonnées d'un glossaire d'entreprise.
Port	Numéro de port sur lequel l'outil Analyst tool s'exécute.
Espace de noms	Nom du domaine de sécurité auquel l'utilisateur de l'outil Analyst tool appartient. Si le domaine utilise l'authentification LDAP ou l'authentification Kerberos, entrez le nom du domaine de sécurité. Dans le cas contraire, entrez <code>Natif</code> .
Activer la communication sécurisée	Activez la communication sécurisée de l'outil Analyst tool au service Analyst.
Importer le contenu publié uniquement	Sélectionnez cette option pour spécifier que vous voulez importer uniquement le contenu publié. Si vous ne sélectionnez pas cette option, Enterprise Data Catalog importe tout le contenu.

L'image suivante montre un exemple de propriétés de connexion dans l'onglet **Général** :

Connection Properties

Username*:	Administrator
Password*:	*****
Host*:	BG.informatica.com
Port*:	8085
Namespace*:	Native
Enable Secure Communication:	<input checked="" type="checkbox"/>
Import Published content only:	<input checked="" type="checkbox"/>
Test Connection	

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Glossaire	Nom de la ressource du glossaire d'entreprise à importer.
Mémoire	<p>Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanner.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

Remarque:

- Assurez-vous de vous connecter à l'outil Analyst tool avant d'exécuter la ressource Business Glossary.
- Si vous modifiez les propriétés d'une ressource Business Glossary après la création de la ressource, assurez-vous de réexécuter la ressource.

CHAPITRE 6

Configuration des ressources Business Intelligence

Ce chapitre comprend les rubriques suivantes :

- [Présentation des ressources Business Intelligence, 81](#)
- [IBM Cognos, 82](#)
- [MicroStrategy, 86](#)
- [Oracle Business Intelligence, 89](#)
- [QlikView, 94](#)
- [SAP Business Objects, 95](#)
- [Tableau Server, 100](#)
- [SAP Business Warehouse \(SAP BW\), 103](#)
- [SAP BW/4HANA, 109](#)
- [Microsoft Power BI, 114](#)
- [Qlik Sense, 117](#)

Présentation des ressources Business Intelligence

Le groupe de ressources Business Intelligence inclut des ressources associées à la veille stratégique. Lorsque vous créez et exécutez une ressource de veille stratégique, le scanner de veille stratégique extrait et migre les résultats de métadonnées de la ressource associée vers le catalogue. Vous pouvez afficher les résultats migrés dans Enterprise Data Catalog.

IBM Cognos

IBM Cognos est une suite de veille stratégique intégrée Web par IBM. Elle fournit un jeu d'outils pour la création de rapports, l'analyse, la création de fiches d'évaluation et la surveillance des événements et des scores.

Objets extraits

Enterprise Data Catalog extrait les métadonnées des objets de schéma suivants d'une ressource IBM Cognos Business Intelligence Suite :

- Axe
- Corps
- Analyse croisée
- Colonnes d'analyse croisée
- Nœud d'analyse croisée
- Membre de nœud d'analyse croisée
- Sources de données
- DataAttribute
- Requête de base de données
- Dimension
- Champ
- Filtre
- Dossier
- Pied de page
- En-tête
- Hiérarchie
- Espace de noms
- Package
- Page
- PromptAnswer
- Requêtes
- Requête
- Élément de requête
- Package de requête
- Rectangle
- Rapport
- Table
- Texte
- Variables
- Colonne Cognos
- Espace de noms OlapSchema
- Modèle de connexion

- Coin d'analyse croisée
- Modèle FrameworkManager
- Modèle QueryStudio
- ReportStudioModel
- Espaces de noms
- Source de données ConnectionPackage
- Référentiel
- Sources de données StoreConnection
- Colonne Alias
- Table d'alias
- Vue Couche d'entreprise
- Cellule
- Classe
- Businessobjectscolumn
- Dossier de connexion
- Contextes
- Crystal Reports
- Modèle Crystal Reports
- Hiérarchies personnalisées
- Fournisseur de données
- Connexion à la base de données
- Hiérarchies par défaut
- Cible
- Résultat de dimension
- Variable de dimension
- Dossier d'entreprise
- Formules
- Champs de nom de groupe
- Jointure
- Clé
- Colonne logique
- Résultat de mesure
- Filtre de requête
- Graphique de rapport
- Champ Rapport
- Dossier de rapport
- Page de rapport
- Table de rapports
- Référentiel

- Schéma
- Section
- Champs spéciaux
- Champs récapitulatifs
- Champs de nom de groupe
- Dossier d'univers
- Modèle d'univers
- Requête d'univers
- Vue
- Document Web Intelligence
- Modèle Web Intelligence
- Connexion
- Développer le mappage
- Expression
- Groupe
- Jointure
- Recherche
- Port de recherche
- Mappage
- Qualificateur source
- Table dérivée

Autorisations de configuration de la ressource

Avant de créer une ressource IBM Cognos, vous devez configurer les autorisations du compte utilisateur que vous utilisez pour vous connecter à IBM Cognos.

Vous devez configurer les autorisations `read`, `execute` et `traverse` de tous les objets pour lesquels Enterprise Data Catalog extrait les métadonnées. Par exemple, des rapports, des requêtes, des analyses, des packages et des connexions.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
URL d'agent	URL de l'agent Enterprise Data Catalog qui s'exécute sur un serveur Microsoft Windows. Remarque: Assurez-vous de spécifier l'URL au format suivant : <code>http://<hostname>:<connector_port>/MIMBWebServices</code> .
Version	Indique la version du serveur Cognos.
URL du répartiteur	URL utilisée par le gestionnaire de framework pour envoyer des demandes à Cognos.
Espace de noms	Définit une collection de comptes utilisateurs à partir d'un fournisseur d'authentification.
Utilisateur	Nom d'utilisateur permettant de se connecter au serveur Cognos.

Propriété	Description
Mot de passe	Mot de passe du compte utilisateur permettant de se connecter au serveur Cognos.
Lignage détaillé	Sélectionnez cette option pour extraire et ingérer les métadonnées liées à la logique de transformation des ressources qui incluent des transformations. Une transformation indique la génération, la modification ou le passage de données entre les connexions source et cible. Une logique de transformation affiche les mappages ou les types de relations de flux de données entre les ressources source et cible liées à la ressource sélectionnée dans Enterprise Data Catalog.
Ajouter des objets dépendants	Utilisez cette option pour importer des objets dépendants dans la sélection. La sélection de cette option nécessite une analyse complète des dépendances de rapport sur le serveur Cognos. Vous pouvez sélectionner l'une des options suivantes pour cette propriété : <ul style="list-style-type: none"> - Aucun : importe uniquement les objets Cognos sélectionnés. - Packages référencés par des rapports sélectionnés : importe les rapports et les packages source associés. - Tous : importe les packages source lorsqu'un rapport est sélectionné et importe les rapports dépendants lorsqu'un package source est sélectionné.
Représentation de dossier	Indique comment les dossiers du gestionnaire de framework Cognos doivent être représentés. Vous pouvez sélectionner l'une des options suivantes : <ul style="list-style-type: none"> - Ignorer : ignore les dossiers. - Plat : représente les dossiers en tant que diagrammes, mais ne conserve pas la hiérarchie. - Hiérarchique : représente les dossiers en tant que diagrammes et conserve la hiérarchie.
Configuration d'importation de transformateur	Fichier XML qui décrit les mappages entre les sources de données du gestionnaire de contenu Cognos et les modèles de transformateurs PowerPlay.
Threads de travailleur	Nombre de threads de travailleur requis pour récupérer les métadonnées de façon asynchrone.
Attribution automatique des connexions	Permet d'attribuer automatiquement la connexion.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Mode de navigation de contenu	Spécifie le contenu à récupérer lors de la recherche dans le référentiel Cognos. Vous pouvez sélectionner l'une des options suivantes : <ul style="list-style-type: none"> - Packages uniquement : récupère les packages et les dossiers, mais pas les rapports. - Connexions uniquement : récupère la liste des connexions. - Tous : récupère les packages, les dossiers, les requêtes et les rapports. - Contenu : vous permet de réduire l'étendue de l'importation à un plus petit ensemble d'objets que l'ensemble des objets sur le serveur.
Contenu	Spécifie la hiérarchie des objets de contenu.

Propriété	Description
Mémoire	<p>Spécifie la mémoire requise pour exécuter la tâche de scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

MicroStrategy

MicroStrategy est un logiciel d'application de veille stratégique d'entreprise. La plate-forme MicroStrategy prend en charge les tableaux de bord interactifs, les fiches d'évaluation, les rapports hautement formatés, les requêtes ad hoc, les seuils et les alertes, et la distribution automatique des rapports.

Objets extraits

Enterprise Data Catalog extrait les métadonnées de rapports dans une ressource MicroStrategy Business Intelligence.

Vous pouvez afficher les actifs suivants dans le catalogue pour une ressource MicroStrategy Business Intelligence :

- (Projet) Score
- (Projet) Sous-total
- (Rapport) Score
- (Rapport) Sous-total
- Attribut
- Formulaire d'attribut
- Champ Rapport calculé
- Consolidation
- Élément de consolidation
- Groupe personnalisé
- Instance de base de données
- Document

- Rapport de document
- Fait
- Mesure de fait
- Filtre
- Dossier
- Dossier caché
- Hiérarchie
- Document HTML
- Rapport de document HTML
- Table logique
- Recherche
- Mesure de score
- Colonne MicroStrategy
- Projet
- Dossier du projet
- Invite
- Réponse à l'invite
- Rapport
- Champ Rapport
- Graphique de rapport
- Grille de rapport
- Page de rapport
- Corps de page de rapport
- Texte de rapport
- Ressource
- Schéma
- Filtre de sécurité
- Sous-totaux
- Mesure de sous-total
- Source du projet
- Table d'entrepôt

Autorisations de configuration de la ressource

Configurez les privilèges d'administration suivants pour le compte utilisateur de projet MicroStrategy pour MicroStrategy version 7.x, et `SecurityRole` pour MicroStrategy version 9.x ou ultérieure :

- Utiliser le cache de serveur
- Utiliser Developer tool
- Afficher le SQL
- Administrer les caches
- Charger et télécharger le projet

- Contourner toutes les vérifications d'accès de sécurité aux objets

Remarque: Configurez ce privilège pour attribuer des privilèges d'administration pour tous les objets dans le projet MicroStrategy que vous souhaitez importer.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
URL d'agent	URL de l'agent Enterprise Data Catalog qui s'exécute sur un serveur Microsoft Windows. Remarque: Assurez-vous de spécifier l'URL au format suivant : <code>http://<hostname>:<connector_port>/MIMBWebServices</code> .
Version	Sélectionnez la version de MicroStrategy dans la liste déroulante. Vous pouvez sélectionner l'option Détection automatique si vous souhaitez qu'Enterprise Data Catalog détecte automatiquement la version de la ressource MicroStrategy.
Source du projet	Nom de la source de projet MicroStrategy à laquelle vous voulez vous connecter.
Utilisateur de connexion	Nom d'utilisateur permettant de se connecter à la source du projet.
Mot de passe de connexion	Mot de passe associé au nom d'utilisateur.
Langue par défaut	Spécifiez la langue à utiliser lors de l'importation des métadonnées de la ressource.
Importer le schéma uniquement	Sélectionnez cette option pour importer le schéma du projet sans les rapports ni les documents.
Import incrémentiel	Permet de charger les modifications apportées depuis le dernier chargement des ressources ou de charger l'intégralité des métadonnées. - Sélectionnez cette option pour charger uniquement les modifications récentes. - Désélectionnez-la pour effectuer un chargement complet des métadonnées.
Niveau de conception des tableaux des modèles de données	Sélectionnez l'une des options suivantes pour spécifier la conception des tables importées : - Physique : les tables importées apparaissent dans l'affichage physique du modèle. - Logique et physique : les tables importées apparaissent dans l'affichage logique et physique du modèle.
Lignage détaillé	Sélectionnez cette option pour extraire et ingérer les métadonnées liées à la logique de transformation des ressources qui incluent des transformations. Une transformation indique la génération, la modification ou le passage de données entre les connexions source et cible. Une logique de transformation affiche les mappages ou les types de relations de flux de données entre les ressources source et cible liées à la ressource sélectionnée dans Enterprise Data Catalog.
Projet(s)	Sélectionnez les noms des projets auxquels vous voulez vous connecter à partir de la source du projet.
Attribution automatique des connexions	Permet d'attribuer automatiquement la connexion.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Mémoire	<p>Spécifie la mémoire requise pour exécuter la tâche de scanneur. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Oracle Business Intelligence

Oracle Business Intelligence est un outil de veille stratégique qui inclut des capacités d'analyse et de création de rapports. Vous pouvez créer et configurer une ressource afin d'extraire des métadonnées d'Oracle Business Intelligence Enterprise Edition. Enterprise Data Catalog extrait des objets de rapport qui incluent des présentations, des modèles d'entreprise et des modèles physiques.

Objets extraits

Enterprise Data Catalog extrait les métadonnées des objets de schéma suivants d'une ressource Oracle Business Intelligence :

- Alias
- Colonne Alias
- Colonne de clé d'alias
- AliasPhysicalTable
- Table d'alias
- AnswersPage

- AnswersReport
- Axe
- BusinessModel
- Modèle et mappage d'entreprise
- Catalogue
- Modèle de connexion
- Condition
- Vue composée
- Critères
- Colonne de critères
- Colonne hiérarchique de critères
- Colonne Attribut de données
- Tableau de bord
- Page Tableau de bord
- Base de données
- Colonne logique dérivée
- Colonne Présentation dérivée
- Domaine de package de conception
- Colonne logique de dimension
- DynamicVariable
- Colonne externe
- Connexion externe
- Schéma externe
- Table externe
- FactLogicalTable
- Champ
- Vue de filtres
- Dossier
- Entonnoir
- Indicateur
- Graphique
- Hiérarchie
- Vue HTML
- Key
- Colonne logique de clé
- LogicalColumn
- LogicalDimension
- Table de faits logiques
- LogicalJoin

- LogicalKeyColumn
- LogicalLevel
- LogicalTable
- Mesure
- Contenu de modèle
- Physique
- PhysicalColumn
- Colonne de clé physique
- PhysicalForeignKey
- PhysicalJoin
- PhysicalSelectTable
- PhysicalTable
- PivotTable
- Présentation
- Table de présentation
- Table de faits de présentation
- Niveau de présentation
- Variable de présentation
- PresentationColumn
- Hiérarchie de présentation
- PresentationTable
- Colonne de clé primaire
- Rapport
- Champ Rapport
- Référentiel
- RepositoryInitializationBlock
- RepositoryInitializationBlocks
- RepositoryVariables
- Schéma
- SessionInitBlock
- Blocs d'initialisation de session
- SessionVariable
- SessionVariables
- StaticVariable
- StaticVariables
- Domaine
- TableView
- Code
- TitleView

- Vue supérieure
- Variables
- Afficher

Autorisations de configuration de la ressource

Enterprise Data Catalog se connecte au serveur de présentation OBIEE pour extraire les métadonnées du rapport. Assurez-vous que le compte utilisateur OBIEE pour lequel vous fournissez le nom d'utilisateur et le mot de passe dispose de l'autorisation de lecture dans les métadonnées qu'Enterprise Data Catalog extrait d'OBIEE.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Version	Sélectionnez l'une des options suivantes pour spécifier la version d'OBIEE : <ul style="list-style-type: none"> - Détection automatique. Sélectionnez cette option pour qu'Enterprise Data Catalog détecte la version d'OBIEE. - OBIEE 11.x. Sélectionnez cette option pour spécifier la version d'OBIEE comme étant 11.x.
URL du serveur	URL du serveur de présentation OBIEE. Si vous utilisez SSL, assurez-vous qu'Enterprise Data Catalog approuve le certificat du serveur de présentation OBIEE.
Utilisateur de connexion	Nom d'utilisateur employé pour ouvrir une session sur le serveur de présentation OBIEE. Assurez-vous que le nom d'utilisateur que vous utilisez possède les autorisations nécessaires pour importer des métadonnées.
Mot de passe de connexion	Mot de passe associé au nom d'utilisateur.
Optimiser pour les grands modèles	Sélectionnez cette option pour optimiser l'importation des métadonnées pour les grands modèles de référentiel OBIEE. Si vous sélectionnez cette option, Enterprise Data Catalog n'importe pas de métadonnées pour les éléments suivants : <ul style="list-style-type: none"> - Clés étrangères - Jointures - Relations - Clés étrangères logiques De plus, Enterprise Data Catalog ne stocke pas d'objets d'arborescence d'expression avec des liens de lignage. Si vous ne sélectionnez pas cette option, Enterprise Data Catalog importe l'intégralité du modèle de référentiel, ce qui entraîne une forte consommation de mémoire.
Threads de travailleur	Spécifiez le nombre de threads de travailleur permettant de traiter les métadonnées de façon asynchrone. Ne spécifiez pas la valeur si vous souhaitez qu'Enterprise Data Catalog la calcule. Enterprise Data Catalog attribue une valeur comprise entre un et six en fonction de l'architecture JVM et du nombre de cœurs de processeur disponibles. Vous pouvez utiliser les points suivants pour décider de la valeur à utiliser : <ul style="list-style-type: none"> - Vous pouvez fournir une valeur supérieure ou égale à un et inférieure à six pour spécifier le nombre de threads de travailleur requis. - Si vous spécifiez une valeur non valide, Enterprise Data Catalog affiche un avertissement et utilise la valeur un. - Si votre machine a plus de mémoire, vous pouvez spécifier une valeur plus élevée pour traiter plus de métadonnées de façon asynchrone. Remarque: La spécification d'une valeur plus élevée peut affecter les performances du système.

Propriété	Description
Fichier	<p>Fichier RPD du référentiel Oracle Business Intelligence où les métadonnées sont stockées.</p> <p>Remarque:</p> <ul style="list-style-type: none"> - Pour OBIEE version 11.x, vous devez convertir le fichier RPD en format XML. Enterprise Data Catalog ne prend pas en charge le format de fichier UDML pour OBIEE version 11.x. - Pour OBIEE version 10.x, vous devez convertir le fichier RPD en fichier UDML. <p>Cliquez sur Sélectionner pour sélectionner le fichier RPD. Vous pouvez également faire glisser le fichier RPD dans la zone de texte Fichier de valeurs de variables à partir de votre explorateur de fichiers.</p>
Fichier de valeurs de variables	<p>(Facultatif) Le fichier qui définit la liste des valeurs variables du RPD.</p> <p>Cliquez sur Sélectionner pour sélectionner le fichier contenant les valeurs de variables. Vous pouvez également faire glisser le fichier dans la zone de texte Fichier de valeurs de variables à partir de votre explorateur de fichiers.</p>
Attribution automatique des connexions	Sélectionnez cette option pour spécifier que la connexion doit être attribuée automatiquement.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Sous-ensemble du référentiel	Cliquez sur Sélectionner . La boîte de dialogue Sélectionner le sous-ensemble du référentiel s'affiche. Sélectionnez les dossiers à partir desquels vous voulez importer des métadonnées pour les rapports du serveur de présentation Oracle Business Intelligence.
Mémoire	<p>Spécifie la mémoire requise pour exécuter la tâche de scanneur. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

QlikView

QlikView est une plate-forme de veille stratégique flexible qui transforme les données en connaissances. Configurez les propriétés de la connexion pour créer une ressource QlikView.

Objets extraits

La ressource QlikView extrait les objets suivants :

- Connexion
- Expressions
- Champs
- Groupes
- Champs système
- Variables système
- Tables
- Variables
- Champs de texte

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
AgentURL	URL de l'agent Enterprise Data Catalog qui s'exécute sur un serveur Microsoft Windows.
Répertoire	Chemin d'accès au répertoire à partir duquel Enterprise Data Catalog doit importer les fichiers QlikView. Le répertoire doit être présent sur la machine sur laquelle l'agent Microsoft Windows s'exécute.
Divers Options	Spécifiez l'une des options suivantes : <ul style="list-style-type: none">- Type de base de données de connexion ODBC- Instruction Replace DIRECTORY- Demander le journal d'exécution QVW Utilisez une virgule pour séparer les valeurs de chaîne. Par exemple, DatabaseType=Oracle, RequireLog=no, DIRECTORY []=[E:\\Qlik_View].

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Attribution automatique des connexions	Sélectionnez cette option pour spécifier que la connexion doit être attribuée automatiquement.

Propriété	Description
Mémoire	<p>Spécifie la mémoire requise pour exécuter la tâche de scanneur. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

SAP Business Objects

Vous pouvez configurer un type de ressource SAP BusinessObjects pour extraire des métadonnées de SAP BusinessObjects. SAP BusinessObjects est un outil de veille stratégique qui comprend des composants pour la gestion des performances, la planification, la création de rapports, l'analyse et la gestion de données d'entreprise.

Objets extraits

Enterprise Data Catalog extrait les métadonnées des objets de schémas suivants d'une source de données SAP Business Objects :

- Corps
- Cellule
- Classe
- Contexte
- Crystal Reports
- Modèle Crystal Reports
- Dossier de hiérarchies personnalisées
- Fournisseurs de données
- Table dérivée
- EnterpriseFolder
- Champ
- Filtre
- Dossier

- Pied de page
- En-tête
- Hiérarchie
- Jointure
- Clé
- Mesure
- Page
- Requête
- Rectangle
- Graphique de rapport
- Schéma
- Requête d'univers
- UniverseFolder
- Document Web Intelligence
- Modèle Web Intelligence
- Dimension
- Résultat de dimension
- Rectangle
- Variable de dimension
- colonne Alias
- section
- champ Rapport
- Texte
- Résultat de mesure
- Table Business Object
- Page de rapport
- Axe
- Table de rapports
- Formules
- Champ spécial
- Champs de nom de groupe
- Table d'alias
- Dossier de rapport
- Champs récapitulatifs
- Colonne logique
- Filtre de requête
- Connectionfolder
- Modèle de connexion
- Connexion à la base de données

- Vue Couche d'entreprise
- Contexte
- Contextes
- Hiérarchies par défaut
- Détail
- Référentiel
- Vue
- Requêtes
- Modèle de connexion
- Source de données ConnectionPackage
- Analyse croisée
- Package
- Page
- Colonnes d'analyse croisée
- Coin d'analyse croisée
- Nœud d'analyse croisée
- Membre de nœud d'analyse croisée
- DataAttribute
- Sources de données
- Requête de base de données
- Modèle FrameworkManager
- Espace de noms
- Élément de requête
- Package de requête
- Rapport
- Modèle ReportStudio
- Référentiel
- Sources de données StoreConnection
- Texte
- Connexion
- Mappage développé
- Expression
- Groupe
- Jointure
- Recherche
- Port de recherche
- Mappage
- Qualificateur source
- Cible

- Espace de noms OlapSchema
- UniverseModel

Autorisations de configuration de la ressource

Vous devez supprimer les autorisations `write-protection` de tout univers SAP Business Object à exporter. Pour plus d'informations sur la suppression des autorisations de Business Object, consultez la documentation de SAP Business Objects.

Assurez-vous que l'utilisateur qui se connecte au référentiel SAP Business Objects appartient au groupe d'utilisateurs du concepteur d'univers et qu'il dispose d'un accès en lecture à toutes les métadonnées de Business Objects. Si le référentiel Business Objects contient des rapports Web Intelligence dans le dossier **Favoris** ou **Personnel**, veillez à inclure l'utilisateur dans le groupe **Administrateurs**. Vous devez créer un groupe de sécurité personnalisé et fournir à l'utilisateur l'autorisation d'affichage des rapports Web Intelligence.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés du type de ressource BusinessObjects :

Propriété	Description
URL d'agent	Nom d'hôte et numéro de port de l'agent Enterprise Data Catalog qui s'exécute sur un serveur Microsoft Windows. Remarque: Assurez-vous de spécifier l'URL au format suivant : <code>http://<hostname>:<connector_port>/MIMBWebServices</code> .
Version	Version du référentiel SAP Business Objects.
Système	Nom du référentiel BusinessObjects. Pour BusinessObjects 11.x et 12.x, spécifiez le nom du serveur de gestion centralisée BusinessObjects. Spécifiez le nom du serveur dans le format suivant : <code><server name>:<port number></code> . Si le serveur de gestion centralisée est configuré sur un cluster, spécifiez le nom du cluster au format suivant : <code><host name>:<port>@<cluster name></code> . Le port par défaut est 6400. Remarque: Si la version du référentiel BusinessObjects est 14.0.6, ne spécifiez pas de numéro de port dans le nom du référentiel. Si vous spécifiez le numéro de port, Enterprise Data Catalog ne peut pas extraire les rapports Web Intelligence.
Mode d'authentification	Mode d'authentification du compte utilisateur qui se connecte au référentiel BusinessObjects. Spécifiez l'une des valeurs suivantes : - Entreprise. Connectez-vous à l'aide du mode d'authentification BusinessObjects Entreprise. - LDAP. Connectez-vous à l'aide de l'authentification LDAP configurée sur BusinessObjects. La valeur par défaut est Entreprise.
Nom d'utilisateur	Nom d'utilisateur permettant de se connecter au référentiel BusinessObjects.
Mot de passe	Mot de passe du compte utilisateur pour le référentiel BusinessObjects.

Propriété	Description
Ajouter des objets dépendants	<p>Choisissez les documents qui dépendent de l'univers sélectionné. Spécifiez l'une des valeurs suivantes :</p> <ul style="list-style-type: none"> - True. Importe les documents qui dépendent de l'univers spécifié. - False. Ignore les documents qui dépendent de l'univers spécifié. <p>Remarque: Les informations de dépendance sont extraites du cache de métadonnées de référentiel des objets d'entreprise. Si le chargement Enterprise Data Catalog ne reflète pas les rapports modifiés ou déplacés, actualisez le cache en chargeant ces rapports et en actualisant les requêtes.</p>
Ajouter des objets spécifiques	<p>Spécifie des objets supplémentaires dans l'univers. Spécifiez l'une des valeurs suivantes :</p> <ul style="list-style-type: none"> - Aucune. Ignore tous les objets. - Documents d'univers indépendants. Importe des documents qui ne dépendent d'aucun univers. <p>La valeur par défaut est Aucun.</p>
Port CORBA Crystal	<p>Spécifie le numéro de port du client sur lequel le SDK Crystal communique avec le serveur d'applications de rapports (RAS).</p> <p>Le serveur RAS utilise le port pour envoyer des métadonnées à la machine cliente locale. Si vous ne spécifiez pas de port, le serveur en sélectionne au hasard un pour chaque exécution.</p>
Représentation de classe	<p>Contrôle la façon dont l'importation de la structure de l'arborescence des classes et des sous-classes se produit.</p> <p>L'agent Enterprise Data Catalog importe chaque classe contenant des objets sous la forme d'une dimension ou d'une arborescence de packages. Spécifiez l'une des valeurs suivantes :</p> <ul style="list-style-type: none"> - Comme une structure plate. Ne crée pas de package. - Comme une structure plate simplifiée. Crée un package pour chaque classe avec une sous-classe. - Comme une structure arborescente complète. Crée un package pour chaque classe. <p>La valeur par défaut est Comme une structure plate.</p>
Lignage détaillé	<p>Sélectionnez cette option pour extraire et ingérer les métadonnées liées à la logique de transformation des ressources qui incluent des transformations. Une transformation indique la génération, la modification ou le passage de données entre les connexions source et cible. Une logique de transformation affiche les mappages ou les types de relations de flux de données entre les ressources source et cible liées à la ressource sélectionnée dans Enterprise Data Catalog.</p>
Threads de travailleur	<p>Nombre de threads de travail que l'agent Enterprise Data Catalog utilise pour l'extraction asynchrone de métadonnées. Laissez le champ vide ou entrez une valeur entière positive.</p> <p>Si vous laissez le champ vide, l'agent Enterprise Data Catalog calcule le nombre de threads de travail. L'agent Enterprise Data Catalog utilise l'architecture JVM et le nombre de cœurs de processeur disponibles sur la machine de l'agent Enterprise Data Catalog pour calculer le nombre de threads. Si vous spécifiez une valeur non valide, l'agent Enterprise Data Catalog utilise un thread de travail.</p> <p>Réduisez le nombre de threads de travail si l'agent Enterprise Data Catalog génère des erreurs d'erreur de mémoire insuffisante lors de l'extraction des métadonnées. Augmentez le nombre de threads de travail si la machine de l'agent Enterprise Data Catalog possède une grande quantité de mémoire disponible, par exemple, 10 Go ou plus. Si vous spécifiez trop de threads de travailleur, les performances peuvent diminuer.</p> <p>Par défaut, la valeur est vide.</p>
Attribution automatique des connexions	<p>Choisissez d'attribuer automatiquement les schémas de base de données à la ressource que vous créez pour la source SAP BusinessObjects.</p>

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Mode de navigation du référentiel	Spécifie les objets disponibles dans le référentiel SAP BusinessObjects. Sélectionnez l'une des options suivantes : <ul style="list-style-type: none"> - Univers uniquement. Importe les métadonnées des univers SAP BusinessObjects. - Connexion uniquement. Importe les métadonnées des connexions aux bases de données qui sont publiées dans le référentiel SAP BusinessObjects. - Tous. Importe à la fois les univers et les connexions du référentiel SAP BusinessObjects.
Sous-ensemble du référentiel	Spécifie les objets stockés dans un référentiel SAP BusinessObjects distant.
Mémoire	Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanner. Spécifiez l'une des valeurs de mémoire suivantes : <ul style="list-style-type: none"> - Faible - Moyenne - Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

Tableau Server

Tableau est un outil de veille stratégique qui permet de vous connecter aux données et de créer des tableaux de bord qui peuvent être partagés. Tableau comprend un composant poste de travail et un composant serveur.

Objets extraits

Enterprise Data Catalog extrait les métadonnées du composant serveur de Tableau :

- Bin
- Dimension de calcul
- Mesure de calcul
- Colonne Tableau

- Colonnes
- Condition
- Connexion
- Tableau de bord
- Source de données
- Filtre de source de données
- Jointures de source de données
- Sources de données
- Dimension
- Table de dimensions
- Répertoire
- Table de faits
- Champ
- File
- Filtre
- Group
- Hiérarchie
- Champ de hiérarchie
- Jointure
- Repère
- Repères
- Mesure
- Espace de noms
- Page
- Dimension de paramètre
- Mesure de paramètre
- Paramètres
- Projet
- Champ de proxy
- Table de proxy
- Source de données publiées
- Requête
- Relation
- Ligne
- Lignes
- Session
- Ensemble
- Filtre partagé
- Scénario

- Table Tableau
- Source de données Tableau
- Tableau Server
- Classeur Tableau
- Champ temporaire
- Table temporaire
- Feuille de calcul non publiée
- Classeur
- Source de données de classeur
- Feuille de calcul

Autorisations de configuration de la ressource

Pour se connecter à Tableau Server, Enterprise Data Catalog utilise les informations d'identification d'un utilisateur créé dans Tableau Server. Veillez à configurer le compte utilisateur avec les autorisations et les niveaux de licence suivants :

- Niveau de licence d'interacteur.
- Affichez et téléchargez les autorisations de l'ensemble des projets, classeurs et sources de données pour lesquels vous voulez extraire des métadonnées.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Serveur	Nom d'hôte ou adresse IP où le serveur Tableau s'exécute.
Site	Spécifiez le site si le serveur Tableau a plusieurs sites installés. Cette valeur est sensible à la casse.
Nom d'utilisateur	Nom d'utilisateur pour la connexion au serveur Tableau.
Mot de passe	Mot de passe associé au nom d'utilisateur.
Threads de travailleur	Nombre de threads de travailleur requis pour récupérer les métadonnées de façon asynchrone.
Attribution automatique des connexions	Permet d'attribuer automatiquement la connexion.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Activer les ressources de référence	Option permettant d'extraire les métadonnées des actifs qui ne sont pas inclus dans la ressource, mais qui y sont référés. Par exemple, des tables source et cible dans des mappages PowerCenter, et des tables et fichiers source des rapports Tableau.
Conserver les actifs de référence non résolus	Option permettant de conserver les actifs de référence non résolus dans le catalogue après l'attribution des connexions. La conservation des actifs de référence non résolus vous permet de visualiser l'intégralité du lignage. Les actifs non résolus incluent les fichiers supprimés, les tables temporaires et d'autres actifs qui ne figurent pas dans la ressource principale.
Grouper par	Utilisez cette option pour regrouper les classeurs dans les catégories suivantes : <ul style="list-style-type: none"> - Projet - Mes classeurs - Tous les classeurs
Objets de référentiel	Importe les objets de référentiel tels que les classeurs et les sources de données. Pour tous les classeurs, les sources de données dépendantes sont également importées.
Mémoire	Spécifie la mémoire requise pour exécuter la tâche de scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé : <ul style="list-style-type: none"> - Faible - Moyenne - Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none"> -Dscannerloglevel=<DEBUG/INFO/ERROR>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. -Dscanner.container.core=<No. of core>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. -Dscanner.yarn.app.environment=<key=value>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. -Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0>. Augmente la mémoire du conteneur du scanner lorsque pmem est activé. La valeur par défaut est 1. -DhandleIdCollisionForTableau=true. Traite les conflits d'objets en double avec les objets Tableau extraits.

SAP Business Warehouse (SAP BW)

SAP Business Warehouse est un entrepôt de données orienté modèle basé sur l'environnement SAP NetWeaver ABAP. SAP Business Warehouse collecte, transforme et stocke les données générées par les applications SAP et non-SAP. Ces données sont accessibles via les rapports intégrés, les outils analytiques

et de veille stratégique, et les logiciels tiers. La ressource SAP Business Warehouse est également appelée Classic SAP Business Warehouse.

Objets extraits

La ressource SAP Business Warehouse extrait les métadonnées des actifs suivants dans une source de données SAP Business Warehouse :

- Zone d'informations
- Source de données
- Objets d'informations
- Catalogue d'objets d'informations
- Système source
- Caractéristiques
- Chiffres clés
- Dimension
- Attributs de navigation
- InfoSet
- DSO
- InfoCube
- Fournisseurs multiples
- Transformation
- Transformation Règle
- Requête
- Élément ou champ de requête
- Classeur
- DSO avancée
- Fournisseur composite
- DTP
- InfoObject as InfoProvider

Remarque: Si le type d'actif n'inclut pas d'informations de lignage, de hiérarchie, de catégories et de relation, vous ne pouvez pas afficher les onglets suivants dans le catalogue :

- Lignage et impact
- Relation
- Contient

Autorisations de configuration de la ressource

Pour configurer SAP Business Warehouse, effectuez les tâches suivantes :

1. Installez les transports de scanneur.
2. Configurez un profil d'autorisation de l'utilisateur.

Installation des transports de scanneur

Pour extraire les métadonnées de SAP Business Warehouse, vous devez importer les transports de scanneur qu'Enterprise Data Catalog utilise pour extraire les métadonnées de SAP Business Warehouse.

Les transports se trouvent dans le fichier `SAP_Scanner_Binaries.zip` téléchargé à partir de l'emplacement du programme d'installation Informatica.

Configuration de profils d'autorisation de l'utilisateur

L'administrateur SAP Business Warehouse doit créer les profils d'autorisation de l'utilisateur de produit et de développement. Le tableau suivant décrit les profils d'autorisation de l'utilisateur :

Objet d'autorisation	Description	Classe	Valeurs de champ
S_RFC	Vérification de l'autorisation pour l'accès à RFC	Objets d'autorisation interapplications	ACTVT: 16 (Exécuter) RFC_NAME : DDIF_FIELDINFO_GET RFCPING RFC_GET_FUNCTION_INTERFACE ZINFA_BW_FG RFC_TYPE: FUGR, FUNC
S_RS_COMP	Composants de Business Explorer	Business Warehouse	ACTVT: 03,16 RSINFOAREA: * (Autorisation complète) RSINFOCUBE: * (Autorisation complète) RSZCOMPID: * (Autorisation complète) RSZCOMPTP: Autorisation complète
S_RS_COMP1	Composants de Business Explorer : améliorations apportées pour le propriétaire	Business Warehouse	ACTVT: 03,16 RSZCOMPID: * (Autorisation complète) RSZOWNER: * (Autorisation complète) RSZCOMPTP: * (Autorisation complète)
S_RS_ISET	InfoSet de Data Warehouse Workbench	Business Warehouse	ACTVT: 03 RSINFOAREA - * (Autorisation complète) RSINFOSET - * (Autorisation complète) RSISETOBJ - * (Autorisation complète)
S_RS_ADMWB	Objets de Data Warehouse Workbench	Business Warehouse	ACTVT: 03 RSADMWBBOBJ - * (Autorisation complète)
S_RFC_ADM	Destination d'administration pour RFC	Business Warehouse	ACTVT: 03 ICF_VALUE: * RFCDEST: * RFCTYPE: * (Autorisation complète)

Prérequis

1. Désactivez le service de catalogue.
2. Téléchargez le fichier `sapjco3.jar` et copiez-le dans l'emplacement suivant : `<Install_directory>/services/CatalogService/access/WEB-INF/lib`.

Remarque: La version de SAP Java Connector prise en charge est la version 3.0.19.

3. Téléchargez le fichier `libsapjco3.so` et copiez-le dans l'emplacement suivant : `<Install_directory>/services/shared/bin`.
4. Intégrez le fichier `libsapjco3.so` au fichier `SAPJCO.zip` et copiez ce dernier dans l'emplacement suivant : `<Install_directory>/services/CatalogService/ScannerBinaries`.
5. Activez le service de catalogue.

Prérequis à l'utilisation de la connexion SNC (Secure Network Communications)

1. Désactivez le service de catalogue.
 2. Téléchargez le fichier `sapjco3.jar` et copiez-le à l'emplacement suivant : `<Install_directory>/services/CatalogService/access/WEB-INF/lib`.
- Remarque:** La version de SAP Java Connector prise en charge est la version 3.0.19.
3. Téléchargez le fichier `libsapjco3.so` et copiez-le à l'emplacement suivant : `<Install_directory>/services/shared/bin`.
 4. Configurez le type de connexion SNC avec le certificat X.509. Pour plus d'informations sur les étapes de configuration du type de connexion SNC, consultez la documentation de SAP.
 5. Créez et définissez la variable d'environnement `SECUDIR` pour le service de catalogue comme emplacement du fichier `libsapcrypto.so`.
 6. Incluez les fichiers suivants dans le fichier `SAPJCO.zip` :
 - `cred_v2`
 - `libsapcrypto.so`
 - `sapjco3.jar`
 - `libsapjco3.so`
 - Certificat X.509
 7. Copiez le fichier `SAPJCO.zip` à l'emplacement suivant : `<Install_directory>/services/CatalogService/ScannerBinaries`
 8. Activez le service de catalogue.
 9. Dans un cluster à plusieurs nœuds, vérifiez que vous avez effectué les tâches suivantes :
 - Copier le fichier PSE (Personal Security Environment) au même chemin sur tous les nœuds du cluster.
 - Configurer l'autorisation `read` dans la source de données SAP BW pour le compte utilisateur permettant d'accéder à la source de données.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Type de serveur SAP	Type d'installation de serveur SAP : <ul style="list-style-type: none">- Instance centrale SAP- Équilibrage de charge SAP (Serveur de message)
Hôte du serveur	Nom d'hôte ou adresse IP du système qui héberge SAP Business Warehouse.
Groupe de connexion	Le groupe de connexion est un groupe de serveurs qui appartient à un système SAP. Les groupes de connexion distribuent les connexions des utilisateurs dans des instances individuelles ou des groupes d'instances SAP ou de serveurs d'application. Ils sont gérés dans les groupes SMLG (SAP Logon Groups) de transaction. Cette option s'affiche lorsque vous sélectionnez SAP Load Balancer (Message Server) comme type de serveur SAP.
Port du serveur de message	Port du serveur de message par défaut. Cette option s'affiche lorsque vous sélectionnez SAP Load Balancer (Message Server) comme type de serveur SAP.
Numéro du système	Numéro du système de SAP Business Warehouse. Cette option s'affiche lorsque vous sélectionnez SAP Central Instance comme type de serveur SAP.
SID SAP	Identifiant système SAP de l'instance de SAP.
Client	Client SAP Business Warehouse pour l'accès aux données à partir de SAP Business Warehouse.
Langue	Langue pour l'importation des métadonnées à partir de la ressource. La langue de connexion sélectionnée est utilisée pour extraire les descriptions SAP.
Type de connexion	Type de connexion au serveur SAP : <ul style="list-style-type: none">- De base- SNC (Secure Network Communications). Vous pouvez utiliser ce type de connexion pour extraire les métadonnées de la source de données SAP BW.
Nom d'utilisateur	Nom d'utilisateur du système SAP Business Warehouse.
Mot de passe	Mot de passe du système SAP Business Warehouse.
Nom du partenaire SNC	Le nom SNC du partenaire de communication.
Nom du fichier de certificat X.509	Le nom du fichier du certificat X.509 comme ticket de connexion.
Mon nom SNC	Le nom SNC de l'utilisateur qui envoie l'appel RFC (Remote Function Call) à partir d'Enterprise Data Catalog.

Propriété	Description
QoP SNC	<p>Spécifie le niveau de protection à utiliser pour la connexion.</p> <p>Voici les niveaux de protection disponibles :</p> <ul style="list-style-type: none"> - 1. Authentification uniquement - 2. Protection de l'intégrité - 3. Protection de la vie privée - 8. Utiliser la valeur de l'emplacement <code>snc/data_protection/use</code> sur le serveur d'applications - 9. Utiliser la valeur de l'emplacement <code>snc/data_protection/max</code> sur le serveur d'applications <p>Le niveau de protection par défaut est de 3.</p>
Attribution automatique des connexions	Spécifie si les connexions doivent être attribuées automatiquement.

Le tableau suivant décrit les propriétés que vous pouvez configurer dans la section **Métadonnées source** de l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Active l'extraction des métadonnées.
Extraction source	<p>Type de l'extraction des métadonnées source :</p> <ul style="list-style-type: none"> - Complet. Extrait l'intégralité des métadonnées à partir de la source. - Basé sur la date de modification de l'objet. Extrait les métadonnées modifiées ou créées après la date spécifiée. Vous devez effectuer au moins une extraction complète avant d'extraire les métadonnées en fonction de la date de modification de l'objet. - Basé sur la date d'exécution de la requête. Extrait les métadonnées en fonction de la date d'exécution de la requête. Les objets de métadonnées dépendants correspondants sont également extraits.
Mémoire	<p>Valeur de mémoire requise pour exécuter une tâche de scanneur.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>

Propriété	Description
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> -Dscannerloglevel=<DEBUG/INFO/ERROR>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, INFO ou ERROR. La valeur par défaut est INFO. -Dscanner.container.core=<No. of core>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. -Dscanner.yarn.app.environment=<key=value>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur pour les paires clé-valeur multiples. -Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0>. Augmente la mémoire du conteneur du scanner lorsque pmem est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

SAP BW/4HANA

SAP BW/4HANA est une plate-forme d'entrepôt de données d'entreprise qui capture, stocke et consolide les données d'entreprise.

Objets extraits

La ressource SAP BW/4HANA extrait les métadonnées des actifs suivants dans une source de données SAP BW/4HANA :

- Zone d'informations
- Source de données
- Objets d'informations
- Système source
- Source d'informations
- Caractéristiques
- Chiffres clés
- Attributs de navigation
- Transformation
- Transformation Règle
- Requête
- Élément ou champ de requête
- DSO avancée
- Fournisseur composite
- DTP
- Vue ODS ouverte

- Niveau d'agrégation
- Destination de hub ouvert
- InfoObject as InfoProvider

Remarque: Si le type d'actif n'inclut pas d'informations de lignage, de hiérarchie, de catégories et de relation, vous ne pouvez pas afficher les onglets suivants dans le catalogue :

- Lignage et impact
- Relation
- Contient

Autorisations de configuration de la ressource

Pour configurer SAP BW/4HANA, procédez comme suit :

1. Installez les transports de scanneur.
2. Configurez un profil d'autorisation de l'utilisateur.

Installation des transports de scanneur

Pour extraire les métadonnées de SAP BW/4HANA, vous devez importer les transports de scanneur qu'Enterprise Data Catalog utilise pour extraire les métadonnées de SAP BW/4HANA. Les transports se trouvent dans le fichier `SAP_Scanner_Binaries.zip` téléchargé à partir de l'emplacement du programme d'installation Informatica.

Configuration de profils d'autorisation de l'utilisateur

L'administrateur SAP Business Warehouse doit créer les profils d'autorisation de l'utilisateur de produit et de développement. Le tableau suivant décrit les profils d'autorisation de l'utilisateur :

Objet d'autorisation	Description	Classe	Valeurs de champ
S_RFC	Vérification de l'autorisation pour l'accès à RFC	Objets d'autorisation interapplications	ACTVT: 16 (Exécuter) RFC_NAME : DDIF_FIELDINFO_GET RFCPING RFC_GET_FUNCTION_INTERFACE ZINFA_BW4H_FG RFC_TYPE: FUGR, FUNC
S_RS_COMP	Composants de Business Explorer	Business Warehouse	ACTVT: 03,16 RSINFOAREA: * (Autorisation complète) RSINFOCUBE: * (Autorisation complète) RSZCOMPID: * (Autorisation complète) RSZCOMPTP: Autorisation complète

Objet d'autorisation	Description	Classe	Valeurs de champ
S_RS_COMP1	Composants de Business Explorer : améliorations apportées pour le propriétaire	Business Warehouse	ACTVT: 03,16 RSZCOMPID: * (Autorisation complète) RSZOWNER: * (Autorisation complète) RSZCOMPTP: * (Autorisation complète)
S_RFC_ADM	Destination d'administration pour RFC	Business Warehouse	ACTVT: 03 ICF_VALUE: * RFCDEST: * RFCTYPE: * (Autorisation complète)

Prérequis

1. Désactivez le service de catalogue.
2. Téléchargez le fichier `sapjco3.jar` et copiez-le dans l'emplacement suivant : `<Install_directory>/services/CatalogService/access/WEB-INF/lib`.
Remarque: La version de SAP Java Connector prise en charge est la version 3.0.19.
3. Téléchargez le fichier `libsapjco3.so` et copiez-le dans l'emplacement suivant : `<Install_directory>/services/shared/bin`.
4. Intégrez le fichier `libsapjco3.so` au fichier `SAPJCO.zip` et copiez ce dernier dans l'emplacement suivant : `<Install_directory>/services/CatalogService/ScannerBinaries`.
5. Activez le service de catalogue.

Prérequis à l'utilisation de la connexion SNC (Secure Network Communications)

1. Désactivez le service de catalogue.
2. Téléchargez le fichier `sapjco3.jar` et copiez-le à l'emplacement suivant : `<Install_directory>/services/CatalogService/access/WEB-INF/lib`.
Remarque: La version de SAP Java Connector prise en charge est la version 3.0.19.
3. Téléchargez le fichier `libsapjco3.so` et copiez-le à l'emplacement suivant : `<Install_directory>/services/shared/bin`.
4. Configurez le type de connexion SNC avec le certificat X.509. Pour plus d'informations sur les étapes de configuration du type de connexion SNC, consultez la documentation de SAP.
5. Créez et définissez la variable d'environnement `SECUDIR` pour le service de catalogue comme emplacement du fichier `libsapcrypto.so`.
6. Incluez les fichiers suivants dans le fichier `SAPJCO.zip` :
 - `cred_v2`
 - `libsapcrypto.so`
 - `sapjco3.jar`
 - `libsapjco3.so`

- Certificat X.509
7. Copiez le fichier `SAPJCO.zip` à l'emplacement suivant : `<Install_directory>/services/CatalogService/ScannerBinaries`
 8. Activez le service de catalogue.
 9. Dans un cluster à plusieurs nœuds, vérifiez que vous avez effectué les tâches suivantes :
 - Copier le fichier PSE (Personal Security Environment) au même chemin sur tous les nœuds du cluster.
 - Configurez l'autorisation `read` dans la source de données SAP BW/4HANA pour le compte utilisateur permettant d'accéder à la source de données.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Type de serveur SAP	Type d'installation de serveur SAP : <ul style="list-style-type: none"> - Instance centrale SAP - Équilibrage de charge SAP (Serveur de message)
Hôte du serveur	Nom d'hôte ou adresse IP du système qui héberge SAP Business Warehouse.
Groupe de connexion	Le groupe de connexion est un groupe de serveurs qui appartient à un système SAP. Les groupes de connexion distribuent les connexions des utilisateurs dans des instances individuelles ou des groupes d'instances SAP ou de serveurs d'application. Ils sont gérés dans les groupes SMLG (SAP Logon Groups) de transaction. Cette option s'affiche lorsque vous sélectionnez SAP Load Balancer (Message Server) comme type de serveur SAP.
Port du serveur de message	Port du serveur de message par défaut. Cette option s'affiche lorsque vous sélectionnez SAP Load Balancer (Message Server) comme type de serveur SAP.
Numéro du système	Numéro du système de SAP Business Warehouse. Cette option s'affiche lorsque vous sélectionnez SAP Central Instance comme type de serveur SAP.
SID SAP	Identifiant système SAP de l'instance de SAP.
Client	Client SAP Business Warehouse pour l'accès aux données à partir de SAP Business Warehouse.
Langue	Langue pour l'importation des métadonnées à partir de la ressource. La langue de connexion sélectionnée est utilisée pour extraire les descriptions SAP.
Type de connexion	Type de connexion au serveur SAP : <ul style="list-style-type: none"> - De base - SNC (Secure Network Communications). Vous pouvez utiliser ce type de connexion pour extraire les métadonnées de la source de données SAP BW/4HANA.
Nom d'utilisateur	Nom d'utilisateur du système SAP Business Warehouse.
Mot de passe	Mot de passe du système SAP Business Warehouse.

Propriété	Description
Nom du partenaire SNC	Le nom SNC du partenaire de communication.
Nom du fichier de certificat X.509	Le nom du fichier du certificat X.509 comme ticket de connexion.
Mon nom SNC	Le nom SNC de l'utilisateur qui envoie l'appel RFC (Remote Function Call) à partir d'Enterprise Data Catalog.
QoP SNC	<p>Spécifie le niveau de protection à utiliser pour la connexion.</p> <p>Voici les niveaux de protection disponibles :</p> <ul style="list-style-type: none"> - 1. Authentification uniquement - 2. Protection de l'intégrité - 3. Protection de la vie privée - 8. Utiliser la valeur de l'emplacement <code>snc/data_protection/use</code> sur le serveur d'applications - 9. Utiliser la valeur de l'emplacement <code>snc/data_protection/max</code> sur le serveur d'applications <p>Le niveau de protection par défaut est de 3.</p>
Attribution automatique des connexions	Spécifie si les connexions doivent être attribuées automatiquement.

Le tableau suivant décrit les propriétés que vous pouvez configurer dans la section **Métadonnées source** de l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Active l'extraction des métadonnées.
Extraction source	<p>Type de l'extraction des métadonnées source :</p> <ul style="list-style-type: none"> - Complet. Extrait l'intégralité des métadonnées à partir de la source. - Basé sur la date de modification de l'objet. Extrait les métadonnées modifiées ou créées après la date spécifiée. Vous devez effectuer au moins une extraction complète avant d'extraire les métadonnées en fonction de la date de modification de l'objet. - Basé sur la date d'exécution de la requête. Extrait les métadonnées en fonction de la date d'exécution de la requête. Les objets de métadonnées dépendants correspondants sont également extraits.
Mémoire	<p>Valeur de mémoire requise pour exécuter une tâche de scanneur.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>

Propriété	Description
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> -Dscannerloglevel=<DEBUG/INFO/ERROR>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, INFO ou ERROR. La valeur par défaut est INFO. -Dscanner.container.core=<No. of core>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. -Dscanner.yarn.app.environment=<key=value>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur pour les paires clé-valeur multiples. -Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0>. Augmente la mémoire du conteneur du scanner lorsque pmem est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Microsoft Power BI

Microsoft Power BI est un outil de veille stratégique qui permet de vous connecter aux données et de créer des tableaux de bord qui peuvent être partagés.

Objets extraits

La ressource Microsoft Power BI extrait les métadonnées des actifs suivants dans une source de données Microsoft Power BI :

- Espace de travail
- Rapport
- Jeu de données
- Flux de données
- Mosaïque
- Tableau de bord

Autorisations de configuration de la ressource

Vous avez besoin d'un compte utilisateur disposant de privilèges d'administrateur Power BI.

Remarque: En raison de la contrainte imposée par l'API REST Microsoft Power BI, vous avez besoin d'un compte utilisateur disposant de privilèges d'administrateur Power BI.

Dans Azure Active Directory, configurez l'autorisation `read` et les autorisations API suivantes sur l'application enregistrée pour le service Power BI :

- Capacity.Read.All
- Dashboard.Read.All
- Dataflow.Read.All
- Dataset.Read.All

- Group.Read
- Group.Read.All
- Gateway.Read.All
- Metadata.View_Any
- Report.Read.All
- Tenant.Read.All
- Workspace.Read.All

Prérequis

Avant de créer une ressource Power BI, effectuez les étapes suivantes :

1. Créez un locataire Azure Active Directory.
Remarque: Vous pouvez utiliser un locataire existant s'il est configuré dans Azure Active Directory.
Pour plus d'informations sur la création d'un locataire dans Azure Active Directory, reportez-vous à la documentation de Microsoft.
2. Créez un utilisateur dans Azure Active Directory ou utilisez-en un existant.
3. Attribuez les privilèges d'administrateur Power BI à l'utilisateur Azure Active Directory, qu'il soit nouveau ou existant :
Pour plus d'informations sur l'attribution de privilèges d'administrateur Power BI, consultez la documentation Microsoft.
4. Enregistrez l'application dans Azure Active Directory pour accéder au contenu de Microsoft Power BI. Utilisez l'ID client reçu pour créer la ressource Power BI.
Remarque: Pour enregistrer une application dans Azure Active Directory, le compte utilisateur doit disposer des privilèges `Application administrator`.
Pour plus d'informations sur l'enregistrement d'une application dans Azure Active Directory, reportez-vous à la documentation de Microsoft.
5. Fournissez les autorisations de l'API Service Microsoft Power BI à l'application, puis accordez l'accès à toutes les API Service Microsoft Power BI.
6. Sélectionnez le nouvel utilisateur Azure Active Directory ou l'actuel comme propriétaire de l'application.
7. Définissez la propriété **Traiter l'application comme client public** sur Oui dans l'application enregistrée.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Nom d'utilisateur	Nom d'utilisateur de l'administrateur pour la connexion au serveur Microsoft Power BI.
Mot de passe	Mot de passe du compte d'administrateur pour la connexion au serveur Microsoft Power BI.
ID du client	ID du client pour la connexion au serveur Microsoft Power BI.

Le tableau suivant décrit les propriétés que vous pouvez configurer dans la section **Métadonnées source** de l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Active l'extraction des métadonnées.
Espace de travail	Type d'espace de travail. L'espace de travail est un environnement partagé dans lequel un groupe d'utilisateurs peut créer des collections de tableaux de bord et de rapports.
Mémoire	Valeur de mémoire requise pour exécuter une tâche de scanner. Spécifiez l'une des valeurs de mémoire suivantes : <ul style="list-style-type: none"> - Faible - Moyenne - Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, INFO ou ERROR. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Paire clé-valeur que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur pour les paires clé-valeur multiples. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Fonctionnalités non prises en charge

- Vous ne pouvez pas afficher le lignage externe et le lignage au niveau de l'élément de données pour une ressource Power BI.
- La ressource Power BI ne prend pas en charge l'authentification multifacteur.
- La ressource Power BI extrait des métadonnées uniquement du nuage Power BI.

Qlik Sense

Qlik Sense est une plate-forme de veille stratégique et d'analyse visuelle qui prend en charge des cas d'utilisation incluant des applications d'analyse déployées de manière centralisée, des tableaux de bord ainsi que des analyses personnalisées et intégrées dans une structure évolutive.

Objets extraits

La ressource Qlik Sense extrait les métadonnées des actifs suivants dans une source de données Qlik Sense :

- Applications
- Feuilles
- Scénarios
- Mesures
- Dimensions
- Graphes
- Champs
- Tables (table de données temporaire)

Autorisations de configuration de la ressource

Attribuez le rôle `RootAdmin` à l'utilisateur Qlik Sense pour accéder à tous les actifs dans la ressource Qlik Sense.

Prérequis

Ajoutez les détails de la machine de l'hôte Qlik Sense au fichier `/etc/hosts` dans la machine du domaine et la machine du cluster sur laquelle le service de catalogue est en cours d'exécution.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
URL d'agent	URL de l'agent MITI externe. Remarque: Pour extraire les métadonnées de la source de données Qlik Sense, vous devez installer Enterprise Data Catalog Agent sur le serveur Qlik Sense.
URL de l'API de moteur	URL du moteur de l'API Qlik.
Utilisateur	Nom d'utilisateur de l'administrateur pour la connexion au serveur Qlik Sense.
Annuaire d'utilisateurs	Annuaire d'utilisateurs du serveur Qlik Sense.
Certificat	Emplacement du certificat du client Qlik Sense.
Mot de passe du certificat	Mot de passe du certificat du client Qlik Sense. Spécifiez le mot de passe si le certificat du client Qlik Sense est généré avec un mot de passe.

Le tableau suivant décrit les propriétés que vous pouvez configurer dans la section **Métadonnées source** de l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Active l'extraction des métadonnées..
Applications	Sélectionnez les applications requises créées dans le système source Qlik Sense.
Dossier du journal pour les applications sélectionnées	Chemin du dossier du journal Qlik Sense commun aux applications sélectionnées.
Mémoire	<p>Valeur de mémoire requise pour exécuter une tâche de scanner.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, INFO ou ERROR. La valeur par défaut est INFO. - <code>-DagentLogLevel=<NONE/FATAL/ERROR/WARNING/STATUS/INFO/DEBUG></code>. Définit le niveau de journal de l'agent Enterprise Data Catalog sur des valeurs telles que NONE, FATAL, ERROR, WARNING, STATUS, INFO et DEBUG. Le niveau de journal par défaut est DEBUG. Utilisez le niveau de journal ERROR lorsque le jeu de données Qlik Sense est volumineux. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur pour les paires clé-valeur multiples. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1. <p>Remarque: Vous pouvez activer diverses options telles que <code>DatabaseType</code>, <code>FILEPATH</code>, <code>RequireLog</code> et <code>m</code> pour une ressource Qlik Sense utilisant les options de JVM.</p>
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.
Attribution automatique des connexions	Spécifie si les connexions doivent être attribuées automatiquement.

Propriété	Description
Activer les ressources de référence	Extrait les métadonnées des actifs non inclus dans cette ressource, mais qui y sont référencés. Par exemple, des tables source et cible dans des mappages PowerCenter, et des tables et fichiers source des rapports Tableau.
Conserver les actifs de référence non résolus	Conserve les actifs de référence non résolus dans le catalogue après l'attribution des connexions. La conservation des actifs de référence non résolus vous permet de visualiser l'intégralité du lignage. Les actifs non résolus incluent les fichiers supprimés, les tables temporaires et d'autres actifs qui ne figurent pas dans la ressource principale.

Fonctionnalités non supportées

Vous ne pouvez pas afficher le lignage externe pour les applications qui génèrent un fichier QVD pour une ressource Qlik Sense.

CHAPITRE 7

Configuration des ressources d'intégration de données

Ce chapitre comprend les rubriques suivantes :

- [Présentation des ressources d'intégration de données, 120](#)
- [Lignage personnalisé, 120](#)
- [Data Integration Hub, 122](#)
- [Scripts de base de données, 124](#)
- [Service Informatica Cloud, 125](#)
- [Informatica Master Data Management, 127](#)
- [Informatica Platform, 130](#)
- [PowerCenter, 133](#)
- [Service d'intégration de SQL Server, 139](#)
- [IBM InfoSphere DataStage, 143](#)

Présentation des ressources d'intégration de données

Le groupe de ressources d'intégration de données inclut les ressources d'intégration de données. Lorsque vous créez et exécutez une ressource d'intégration de données, le scanner d'intégration de données extrait et migre les métadonnées de la ressource associée vers le catalogue. Vous pouvez afficher les résultats migrés dans Enterprise Data Catalog.

Lignage personnalisé

Vous pouvez créer une ressource de lignage personnalisée pour afficher les informations du lignage de données pour les ressources de votre organisation. Une ressource de lignage personnalisée utilise les

fichiers CSV que vous lui fournissez et qui incluent des données de lignage de votre entreprise. Vous pouvez utiliser cette option si vous n'avez pas d'outil ETL pris en charge par Enterprise Data Catalog.

Propriétés de connexion à la ressource

Les tableaux suivants répertorient les propriétés que vous devez configurer pour ajouter une ressource de lignage personnalisé :

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Fichier	<p>Fichier CSV ou .zip qui inclut les fichiers CSV avec les données de lignage. Cliquez sur Sélectionner pour sélectionner le fichier CSV requis ou le fichier .zip à charger. Assurez-vous que les fichiers CSV du fichier .zip ne sont pas stockés dans un répertoire du fichier .zip.</p> <p>Si vous voulez sélectionner plusieurs fichiers CSV, vous devez inclure les fichiers CSV requis dans un fichier .zip, puis sélectionner le fichier .zip à télécharger.</p> <p>Remarque: Assurez-vous que l'en-tête du fichier CSV inclut les paramètres suivants :</p> <ul style="list-style-type: none">- Association- Connexion De- Connexion À- Objet De- Objet À

L'image suivante montre l'exemple de propriétés de connexion dans l'onglet **Général** :

Connection Properties

File*:  [Choose...](#)

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Attribution automatique des connexions	Permet d'attribuer automatiquement la connexion.

Propriété	Description
Mémoire	<p>Spécifie la mémoire requise pour exécuter la tâche de scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Enterprise Data Catalog restreint la taille du fichier <code>lineage.csv</code> en fonction de la taille de mémoire suivante que vous spécifiez :</p> <ul style="list-style-type: none"> - Faible. La taille maximale de fichier est 500 Mo. - Mémoire moyenne. La taille maximale de fichier est 1 024 Mo. - Mémoire élevée. La taille maximale de fichier est 2 048 Mo. <p>Pour remplacer la restriction imposée sur la taille du fichier, utilisez l'option de JVM – <code>DcustomScannerMaxFileSize</code>.</p> <p>Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que <code>DEBUG</code>, <code>ERROR</code> ou <code>INFO</code>. La valeur par défaut est <code>INFO</code>. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1. - <code>-DcustomScannerMaxFileSize=<file size></code>. Remplace la restriction de taille imposée au fichier <code>lineage.csv</code>. <p>Remarque: Si vous indiquez une taille de fichier très élevée, une erreur de manque de mémoire Java risque de se produire.</p>

Data Integration Hub

Data Integration Hub est une solution d'intégration d'application que vous pouvez utiliser pour partager et synchroniser les données entre plusieurs applications.

Objets extraits

La ressource Data Integration Hub extrait les métadonnées des actifs suivants dans une source de données Data Integration Hub :

- Publication
- Thème
- Abonnement
- Application

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
URL de la console DIH	URL de la console Data Integration Hub.
Nom d'utilisateur	Nom d'utilisateur permettant de se connecter au serveur Data Integration Hub.
Mot de passe	Mot de passe du compte utilisateur qui se connecte au serveur Data Integration Hub.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Mémoire	Valeur de mémoire requise pour exécuter une tâche de scanneur. Spécifiez l'une des valeurs de mémoire suivantes : <ul style="list-style-type: none">- Faible- Moyenne- Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Scripts de base de données

Créez une ressource de scripts de base de données pour l'extraction des informations de lignage des scripts de base de données.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Type de script	Spécifiez le type de script de base de données pour l'extraction des informations de lignage. Sélectionnez l'un des types de scripts de base de données suivants dans la liste déroulante : <ul style="list-style-type: none">- Oracle PL/SQL- Teradata- Apache HiveQL- IBM DB2
Fichier	Fichier ZIP contenant les scripts de base de données. Le fichier ZIP peut contenir des scripts de base de données dans des formats de fichier, tels que .txt, .bteq, .sql, et .plsql. Assurez-vous que les fichiers du fichier .zip ne sont pas stockés dans un répertoire du fichier .zip. Si le fichier ZIP est sensible à la casse, vous devez mettre le nom du schéma, de la table et de la colonne entre guillemets doubles.
Divers Options	Spécifiez les options diverses identifiées par un trait d'union (-) suivi d'une lettre et de sa valeur. Par exemple, <code>-d schema</code> . <ul style="list-style-type: none">- <code>-d</code> fournit un schéma par défaut. La lettre <code>-d</code> permet de spécifier un nom de schéma pour les objets qui ne l'ont pas défini explicitement.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Attribution automatique des connexions	Indique que les connexions sont attribuées automatiquement.

Propriété	Description
Mémoire	<p>Spécifie la mémoire requise pour exécuter la tâche de scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

Remarque: Le type de ressource scripts de base de données est disponible pour l'évaluation technique. La fonctionnalité d'évaluation technique est prise en charge mais elle n'est pas justifiée, ni prête pour la production. Informatica vous recommande d'utiliser cette fonctionnalité uniquement dans des environnements autres que de production.

Service Informatica Cloud

Informatica Cloud est un service d'abonnement à la demande qui permet d'accéder aux applications, bases de données, plates-formes et fichiers plats hébergés sur site ou sur un nuage. Informatica Cloud s'exécute dans un centre d'hébergement.

Autorisations de configuration de la ressource

Pour accéder aux objets de tâche de synchronisation, de tâche de mappage et de tâche PowerCenter dans les services Informatica Cloud, l'utilisateur du nuage doit disposer de la licence de l'API REST dans le nuage et de l'autorisation `read` sur les tâches.

Prérequis

Avant d'ajouter une ressource du service Informatica Cloud, effectuez les étapes suivantes :

1. Créez une organisation pour votre entreprise sur le site Web d'Informatica Cloud, définissez la hiérarchie de l'organisation et configurez les propriétés de l'organisation. Vous devez effectuer cette étape avant de pouvoir utiliser Informatica Cloud.

Remarque: Pour créer une organisation, vous devez disposer d'une licence API REST. Si vous n'avez pas de licence API REST, contactez le support client international Informatica.

2. Créez un compte d'abonnement sur Informatica Cloud.

3. Vérifiez que la machine sur laquelle vous installez l'agent sécurisé d'Informatica Cloud répond à la configuration minimale requise. L'agent sécurisé d'Informatica Cloud est un programme léger qui exécute toutes les tâches et permet la communication sécurisée sur le pare-feu entre votre organisation et Informatica Cloud.
4. Téléchargez, installez et enregistrez l'agent sécurisé d'Informatica Cloud à l'aide du nom d'utilisateur et du mot de passe Informatica Cloud.
5. Créez les tâches suivantes dans Informatica Cloud :
 - a. Tâches de mappage
Un mappage définit la logique de flux de données réutilisable que vous pouvez utiliser dans les tâches de configuration de mappage. Utilisez un mappage pour définir la logique de flux de données qui n'est pas disponible dans les tâches de synchronisation de données, telles que le tri spécifique de la logique ou la jointure de sources à partir de différents systèmes. Lorsque vous configurez un mappage, vous décrivez le flux de données depuis la source et la cible. Vous pouvez ajouter des transformations pour transformer des données, comme une transformation Expression pour les calculs au niveau de la ligne ou une transformation Filtre pour supprimer des données du flux de données.
 - b. Tâches PowerCenter
La tâche PowerCenter vous permet d'importer des flux de travail PowerCenter dans Informatica Cloud et de les exécuter en tant que tâches Informatica Cloud.
 - c. Tâches de synchronisation de données
La tâche de synchronisation de données vous permet de synchroniser des données entre une source et une cible.

Remarque: Une ressource de service Informatica Cloud importe toutes les tâches dans Enterprise Data Catalog la première fois que les métadonnées sont extraites de la ressource. Pendant les opérations d'extraction suivantes, la ressource importe uniquement les tâches mises à jour dans Enterprise Data Catalog.

Pour plus d'informations sur les prérequis, consultez le *Guide de l'utilisateur d'Informatica Cloud* et le *Guide de l'administrateur d'Informatica Cloud*.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
URL du cloud	URL permettant d'accéder au service Informatica Cloud.
Nom d'utilisateur	Nom d'utilisateur permettant de se connecter au service Informatica Cloud.
Mot de passe	Mot de passe associé au nom d'utilisateur.
Attribution automatique des connexions	Sélectionnez cette option pour spécifier que la connexion doit être attribuée automatiquement.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Lignage détaillé	Sélectionnez cette option pour extraire et ingérer les métadonnées liées à la logique de transformation des ressources qui incluent des transformations. Une transformation indique la génération, la modification ou le passage de données entre les connexions source et cible. Une logique de transformation affiche les mappages ou les types de relations de flux de données entre les ressources source et cible liées à la ressource sélectionnée dans Enterprise Data Catalog.
Mémoire	Spécifie la mémoire requise pour exécuter la tâche de scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé : <ul style="list-style-type: none">- Faible- Moyenne- Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none">- <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO.- <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre.- <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur.- <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

Informatica Master Data Management

Multidomain MDM crée et gère les données principales, c'est-à-dire une collection d'entités principales et leurs attributs et valeurs critiques pour l'entreprise.

Objets extraits

Vous pouvez utiliser la ressource Informatica MDM pour extraire les métadonnées des actifs suivants dans une source de données Informatica MDM :

- Entités d'entreprise
- Vues d'entité d'entreprise
- Tables temporaires
- Tables d'arrivée
- Relations
- Objets de base
- Tables de références croisées

- Informations d'attribut de toutes les entités commerciales

Vous pouvez afficher le lignage des actifs suivants du catalogue :

- Tables d'arrivée et tables temporaires
- Tables temporaires et tables de références croisées
- Tables de références croisées et objets de base
- Objets de base et entités commerciales

Remarque: Enterprise Data Catalog version 10.4.1 prend en charge Multidomain MDM version 10.4.0 ou version ultérieure.

Prérequis

- Enregistrez le stockage de référence opérationnelle (Operational Reference Store - ORS) dans la console Informatica MDM Hub.
- Configurez l'application Informatica Data Director (IDD) sur l'ORS à l'aide du gestionnaire de configuration Informatica Data Director.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
SSL activé	Indique si l'outil de provisionnement MDM utilise SSL.
Hôte de provisionnement	Entrez le nom d'hôte pour la connexion à l'outil de provisionnement MDM.
Port de provisionnement	Entrez le numéro de port de l'outil de provisionnement MDM.
Nom d'utilisateur de connexion du provisionnement	Entrez le nom d'utilisateur pour la connexion à l'outil de provisionnement MDM.
Mot de passe de connexion du provisionnement	Entrez le mot de passe pour la connexion à l'outil de provisionnement MDM.
Identifiant de la base de données	Entrez l'ID de base de données du schéma enregistré dans la console Informatica MDM Hub. Remarque: L'ID de base de données du schéma doit être en minuscules.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Mémoire	<p>Spécifie la mémoire requise pour exécuter la tâche de scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none">- Faible- Moyenne- Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none">- <code>-Dscanner.loglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO.- <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre.- <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur.- <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Affichage du lignage détaillé entre les ressources Informatica MDM et Informatica Platform

Pour afficher le lignage détaillé entre les ressources Informatica MDM et Informatica Platform pour la qualité des données et les mappages d'intégration de données, procédez comme suit :

1. Créez et exécutez une ressource Informatica MDM.
2. Créez une ressource Informatica Platform et activez l'option **Lignage détaillé** pour la ressource.
3. Exécutez la ressource Informatica Platform.

Remarque: Vérifiez que les mappages Informatica Platform et MDM sont intégrés pour la gestion intermédiaire. Pour plus d'informations, consultez l'article [How to Integrate the MDM Hub with Informatica Platform for Staging](#) dans Informatica Doc Portal.

4. Utilisez l'option **Attribution de connexion** pour configurer les ressources Informatica Platform et MDM en tant que sources de données cible ou de destination. Pour plus d'informations sur l'attribution de connexions, consultez la section [Managing Connections](#).

Remarque: Vérifiez que le statut de l'analyse des deux ressources est terminé avant d'effectuer l'attribution de connexions.

Informatica Platform

Créez une ressource basée sur le type de ressource Informatica Platform pour extraire des métadonnées du référentiel modèle. Vous devez spécifier les informations de connexion au service d'intégration de données lorsque vous configurez la ressource.

Objets extraits

Enterprise Data Catalog n'extrait que les métadonnées de mappage d'Informatica Platform.

Prérequis

Si le domaine et le cluster sont compatibles respectivement avec SSL et Kerberos, effectuez les étapes suivantes :

- Copiez le certificat SSL sur toutes les machines du cluster. Assurez-vous que le chemin d'accès au fichier `infa_truststore.jks` est commun à toutes les machines du cluster.
- Pour obtenir le mot de passe chiffré, accédez au chemin d'accès au fichier `<INFAHOME>/server/bin`, puis tapez `pmpasswd <Truststore Password>` dans la ligne de commande.
- Configurez les paramètres JVM suivants pour la ressource :
 - `--DINFA_TRUSTSTORE=<trust store path>`
 - `--DINFA_TRUSTSTORE_PASSWORD=<trust store encrypted key>`

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de connexion au service d'intégration de données :

Propriété	Description
Version cible	Numéro de version d'Informatica Platform. Vous pouvez choisir l'une des versions Informatica suivantes : <ul style="list-style-type: none">- 9.6.1 HotFix 3- 9.6.1 HotFix 4- 10.0- 10.1- 10.1.1- 10.1.1. HF1- 10.2- 10.2 HF2- 10.2.1- 10.4.1
Nom du domaine	Nom du domaine Informatica.
Nom du service d'intégration de données	Nom du service d'intégration de données.
Nom d'utilisateur	Nom d'utilisateur de la connexion au service d'intégration de données.
Mot de passe	Mot de passe de la connexion du service d'intégration de données.
Domaine de sécurité	Nom du domaine de sécurité LDAP si le domaine Informatica contient un domaine de sécurité LDAP.
Hôte	Nom d'hôte du domaine Informatica.

Propriété	Description
Port	Numéro de port du domaine Informatica.
Nom de l'application	Nom de l'application du service d'intégration de données. Par défaut, tous les noms d'application sont sélectionnés. Pour sélectionner un seul nom d'application, cliquez sur Sélectionner... et sélectionnez le nom de l'application dans la boîte de dialogue Sélectionner le nom de l'application . Remarque: Cette propriété s'applique aux versions cible 10.0 ou ultérieures.
Param défini pour les mappages dans l'application	Jeu de paramètres des mappages configurés dans l'application du service d'intégration de données. Cliquez sur Sélectionner... pour sélectionner le jeu de paramètres dans la boîte de dialogue Sélectionner les jeux de paramètres pour les mappages dans l'application . Remarque: Cette propriété s'applique aux versions cible 10.0, 10.1, 10.1.1 ou 10.2.
Jeu de paramètres ou fichier de paramètres	Indiquez si vous souhaitez utiliser un jeu de paramètres ou un fichier de paramètres dans l'application du service d'intégration de données. Remarque: Cette propriété s'applique aux versions cible 10.2.1 ou ultérieures.
Jeux de paramètres	Jeu de paramètres pour les mappages ou les flux de travail configurés pour l'application du service d'intégration de données. L'application du service d'intégration de données peut contenir un jeu de paramètres unique ou plusieurs jeux de paramètres basés sur les jeux configurés pour l'application du service d'intégration de données. Vous pouvez utiliser cette propriété si vous sélectionnez une seule application de service d'intégration de données. Cliquez sur Sélectionner... pour sélectionner le jeu de paramètres dans la boîte de dialogue Sélectionner les jeux de paramètres . Remarque: Cette propriété s'applique aux versions cible 10.2.1 ou ultérieures.
Fichier de jeux de paramètres	Fichier ZIP contenant tous les jeux de paramètres pour les mappages ou les flux de travail configurés pour les applications de service d'intégration de données dans un fichier. config. Créez le fichier de jeux de paramètre dans un éditeur de texte, nommez-le <code>ApplicationParameterSet.config</code> , puis compressez-le. Dans le fichier <code><Parameter Sets></code> , définissez les jeux de paramètres à configurer pour les applications de service d'intégration de données sélectionnées. Séparez les jeux de paramètres par une virgule dans le fichier. Remarque: Cette propriété s'applique aux versions cible 10.2.1 ou ultérieures.
Fichier de paramètres	Fichier XML qui répertorie les paramètres définis par l'utilisateur et les valeurs qui leur sont attribuées pour les mappages ou les flux de travail configurés pour les applications de service d'intégration de données. Pour créer un fichier de paramètres, créez un dossier portant le nom de l'application du service d'intégration de données dans lequel vous déplacerez les fichiers XML contenant les paramètres de mappage ou de flux de travail. puis compressez le dossier. Remarque: Cette propriété s'applique aux versions cible 10.2.1 ou ultérieures.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Attribution automatique des connexions	Indique que les connexions sont attribuées automatiquement.

Propriété	Description
Activer les ressources de référence	Extrait les métadonnées des actifs non inclus dans cette ressource, mais qui y sont référencés. Par exemple, des tables source et cible dans des mappages de streaming, des sources de streaming telles qu'Apache Kafka et des tables et fichiers source de rapports Tableau.
Conserver les actifs de référence non résolus	Conserve les actifs de référence non résolus dans le catalogue après l'attribution des connexions. La conservation des actifs de référence non résolus vous permet de visualiser l'intégralité du lignage. Les actifs non résolus incluent les fichiers supprimés, les tables temporaires et d'autres actifs qui ne figurent pas dans la ressource principale.
Lignage détaillé	Active le lignage détaillé pour la ressource. Sélectionnez cette option pour extraire et ingérer les métadonnées liées à la logique de transformation des ressources qui incluent des transformations. Une transformation indique la génération, la modification ou le passage de données entre les connexions source et cible. Une logique de transformation affiche les mappages ou les types de relations de flux de données entre les ressources source et cible liées à la ressource sélectionnée dans Enterprise Data Catalog.
Mémoire	Valeur de mémoire requise pour exécuter une tâche de scanneur. Spécifiez l'une des valeurs de mémoire suivantes : <ul style="list-style-type: none"> - Faible - Moyenne - Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none"> - Dscanner.loglevel=<DEBUG/INFO/ERROR>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - Dscanner.container.core=<No. of core>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - Dscanner.yarn.app.environment=<key=value>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0>. Augmente la mémoire du conteneur du scanneur lorsque pmem est activé. La valeur par défaut est 1. - DINFA_TRUSTSTORE=<trust store path>. Chemin d'accès au fichier Informatica truststore. - DINFA_TRUSTSTORE_PASSWORD=<trust store encrypted key>. Mot de passe chiffré du truststore.

Exemple de fichier de jeux de paramètres

L'exemple suivant montre un fichier de jeux de paramètres :

```
Application_Map_rel_conn_param,Parameter_Set_Rel_Conn,Parameter_Set_Rel_Conn1
Application_Map_rel_schema_param,Parameter_Set_Rel_Schema
Application_Map_rel_srctable_param,Parameter_Set_Rel_ScrTblname
Application_Map_rel_conn_param
```

Configuration requise du fichier de paramètres

Un fichier de paramètres est un fichier XML qui répertorie les paramètres définis par l'utilisateur et les valeurs qui leur sont attribuées pour les mappages ou les flux de travail configurés pour les applications de service d'intégration de données. Utilisez un éditeur de texte ou le client Informatica Developer pour créer un fichier de paramètres. Pour permettre à la ressource Informatica Platform de lire les valeurs de paramètres à partir d'un fichier de paramètres, une extension .xml doit être affectée à ce fichier. Vous pouvez regrouper les paramètres et les variables dans différentes sections du fichier de paramètres.

Remarque: Un fichier de paramètres ne doit inclure qu'un flux de travail ou un mappage. Enterprise Data Catalog ne prend pas en charge les fichiers de paramètres qui comprennent plusieurs flux de travail ou mappages.

Création d'un fichier de paramètres

Pour créer un fichier de paramètres, effectuez les étapes suivantes :

1. Créez un dossier avec le même nom que l'application Service d'intégration de données.
2. Créez des fichiers de paramètres au format suivant : `<file name>.xml`.
Remarque: Chaque fichier de paramètres doit inclure un flux de travail ou un mappage spécifique.
3. Placez toutes les définitions de fichier de paramètres du flux de travail ou du mappage dans le dossier.
4. Compressez le dossier dans un fichier ZIP. Le fichier ZIP peut contenir plusieurs dossiers contenant plusieurs fichiers de paramètres.

Exemple de fichier de paramètres

Vous trouverez ci-après un exemple de fichier de paramètres qui contient un flux de travail :

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<root>
  <project name="ABC">
    <workflow name="Name1">
      <parameter name="Param1">Parameter1</parameter>
    </workflow>
  </project>
</root>
```

PowerCenter

Vous pouvez configurer un type de ressource PowerCenter pour extraire des métadonnées des objets de référentiel PowerCenter. Utilisez PowerCenter pour extraire des données de plusieurs sources, transformer les données selon la logique d'entreprise que vous élaborez dans l'application cliente et charger les données transformées dans un fichier et des cibles relationnelles.

Objets extraits

Enterprise Data Catalog extrait les métadonnées de mappage à partir de la source PowerCenter.

Autorisations de configuration de la ressource

- Veillez à configurer le privilège `Access Repository Manager` de l'utilisateur qui accède au référentiel PowerCenter.
- Configurez l'autorisation de lecture dans la source de données PowerCenter pour le compte utilisateur permettant d'accéder à la source de données PowerCenter.
- Veillez à exécuter la commande `pmrep ObjectExport` pour exporter les mappages dans PowerCenter.

Prérequis

Si le domaine et le cluster sont activés respectivement pour SSL et Kerberos, effectuez les étapes suivantes :

- Copiez le certificat SSL sur toutes les machines du cluster. Assurez-vous que le chemin d'accès au fichier `infa_truststore.jks` est commun à toutes les machines du cluster.

- Pour obtenir le mot de passe chiffré, accédez au chemin d'accès du fichier <INFAHOME>/server/bin, puis tapez `pmpasswd <Truststore Password>` dans la ligne de commande.
- Configurez les paramètres JVM suivants pour la ressource :
 - `-DINFA_TRUSTSTORE=<trust store path>`
 - `-DINFA_TRUSTSTORE_PASSWORD=<trust store encrypted key>`

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de la ressource PowerCenter :

Propriété	Description
Nom ou adresse de l'hôte de passerelle	Nom d'hôte ou adresse de la passerelle du domaine PowerCenter.
Numéro du port de passerelle	Numéro de port de la passerelle du domaine PowerCenter.
Domaine de sécurité Informatica	Nom de domaine de sécurité LDAP, s'il en existe un. Sinon, entrez "Native."
Nom du référentiel	Nom du référentiel PowerCenter.
Nom d'utilisateur du référentiel	Nom d'utilisateur pour le référentiel PowerCenter.
Mot de passe utilisateur du référentiel	Mot de passe pour le référentiel PowerCenter.
Version de PowerCenter	Version du référentiel PowerCenter. Remarque: Informatica ne prend pas en charge les versions de PowerCenter antérieures à 9.6.0.
Page de code PowerCenter	Page de code du référentiel PowerCenter.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données. Remarque: Vous pouvez exécuter l'analyse de flux de données sur la ressource même si vous n'avez pas activé le paramètre de métadonnées source.
Fichier de paramètres	Spécifiez le fichier ZIP du fichier de paramètres à attacher à partir d'un système local.
Attribution automatique des connexions	Spécifie si Enterprise Data Catalog attribue automatiquement la connexion.

Propriété	Description
Activer les ressources de référence	Option permettant d'extraire les métadonnées des actifs qui ne sont pas inclus dans la ressource, mais qui y sont référés. Par exemple, des tables source et cible dans des mappages PowerCenter, et des tables et fichiers source des rapports Tableau.
Conserver les actifs de référence non résolus	Option permettant de conserver les actifs de référence non résolus dans le catalogue après l'attribution des connexions. La conservation des actifs de référence non résolus vous permet de visualiser l'intégralité du lignage. Les actifs non résolus incluent les fichiers supprimés, les tables temporaires et d'autres actifs qui ne figurent pas dans la ressource principale.
Sous-ensemble du référentiel	Entrez la liste des chemins d'accès de fichiers séparés par des points-virgules pour l'objet de référentiel Informatica PowerCenter. Remarque: Si vous souhaitez exécuter l'analyse de flux de données sur la ressource, vous devez entrer le chemin d'accès au fichier de l'objet de référentiel Informatica PowerCenter. Si vous n'entrez pas le chemin d'accès au fichier pour exécuter l'analyse de flux de données, le catalogue exécute par défaut l'analyse de flux de données sur tous les objets du référentiel PowerCenter.
Lignage détaillé	Sélectionnez cette option pour extraire et ingérer les métadonnées liées à la logique de transformation des ressources qui incluent des transformations. Une transformation indique la génération, la modification ou le passage de données entre les connexions source et cible. Une logique de transformation affiche les mappages ou les types de relations de flux de données entre les ressources source et cible liées à la ressource sélectionnée dans Enterprise Data Catalog. Remarque: Pour afficher le lignage détaillé des actifs qui comprennent des transformations de service Web dans le catalogue, vous devez sélectionner l'option Activer les ressources de référence .
Mémoire	Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanner. Spécifiez l'une des valeurs de mémoire suivantes : <ul style="list-style-type: none"> - Faible - Moyenne - Élevée Remarque: Informatica vous recommande d'utiliser la valeur de mémoire Élevée si vous activez l'analyse de flux de données. Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

Le tableau suivant décrit les détails du **référéntiel** et les propriétés **avancées** pour les paramètres de l'analyse de flux de données dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer l'analyse de flux de données	Sélectionnez cette option pour activer l'analyse de flux de données et afficher les données analytiques sur les mappages PowerCenter dans le catalogue. Remarque: Vous devez configurer l'analyse d'actifs de données pour que le service de catalogue exécute l'analyse de flux de données sur la ressource.
Sélectionner le type de référentiel PowerCenter	Type de base de données pour le référentiel PowerCenter. Sélectionnez parmi les options suivantes : - Oracle - Microsoft SQL Server - IBM DB2 - Sybase - PostgreSQL
Nom d'utilisateur	Nom du compte utilisateur de la base de données du référentiel PowerCenter.
Mot de passe	Mot de passe du compte utilisateur de la base de données du référentiel PowerCenter.
Chaîne de connexion de la base de données	Chaîne de connexion JDBC pour la connexion à la base de données sécurisée qui contient le nom d'hôte, le numéro de port et les paramètres de sécurité de la base de données.
Options personnalisées	Spécifiez les options de personnalisées pour exécuter le travail du scanner.

Remarque : à partir de la version 10.5, l'analyse de flux de données est disponible pour l'évaluation technique. La fonctionnalité d'évaluation technique est prise en charge mais n'est pas garantie, ni prête pour la production. Informatica vous recommande d'utiliser ces fonctionnalités uniquement dans des environnements autres que de production. Informatica a l'intention d'inclure la fonctionnalité d'aperçu dans une prochaine version GA pour une utilisation en production, mais pourrait décider de ne pas le faire en fonction de l'évolution du marché ou des circonstances techniques. Pour plus d'informations, contactez le support client international Informatica.

Connexion à une ressource PowerCenter compatible avec SSL

Effectuez les étapes suivantes avant de créer la ressource PowerCenter si vous souhaitez vous connecter à une source de données PowerCenter compatible avec SSL :

- Désactivez le service de catalogue.
- Exécutez la commande suivante pour exporter les certificats à partir du truststore de domaine PowerCenter :
 - \$INFAINSTALL/java/bin/keytool -export -keystore
 - \$INFAINSTALL/services/shared/security/infa_truststore.jks -alias <alias_name> -file
 - \$INFAINSTALL/services/shared/security/certExportFromSSLDomainTruststore.cert
- Sur le nœud où Enterprise Data Catalog s'exécute, exécutez la commande suivante pour importer le certificat dans le truststore de domaine Informatica :
 - \$LDMINSTALL/java/bin/keytool -import -file certExportFromSSLDomainTruststore.cert -alias <new_alias_name>-keystore
 - \$LDMINSTALL/services/shared/security/infa_truststore.jks -storepass pass2038@infaSSL

Remarque: Veillez à utiliser un nouveau nom d'alias lorsque vous importez le certificat.

4. Copiez le truststore et les fichiers keystore depuis le domaine PowerCenter vers le truststore de domaine Informatica.
5. Exécutez les commandes suivantes pour configurer les variables d'environnement :
 - a. Modifiez le script `infaservice.sh` dans le domaine Informatica à l'aide de la commande suivante : `$INFAINSTALL/tomcat/bin/infaservice.sh`
 - b. Dans le script, ajoutez la ligne suivante pour définir la variable d'environnement : `INFA_TRUSTSTORE=<emplacement où vous avez copié le truststore et les fichiers keystore.>` Veillez à ajouter la ligne après la commande suivante dans le script : `unset INFA_TRUSTSTORE`
`INFA_TRUSTSTORE_PASSWORD INFA_KEYSTORE INFA_KEYSTORE_PASSWORD`
 - c. Pour un cluster Hortonworks, ouvrez l'application Yarn et définissez la variable `INFA_TRUSTSTORE=<emplacement où vous avez copié le truststore et les fichiers keystore.>` dans la section Variables d'environnement Yarn avancées, puis redémarrez Yarn.
 - d. Pour un cluster Cloudera, modifiez le script `yarn-env.sh` et définissez la variable `INFA_TRUSTSTORE=<emplacement où vous avez copié le truststore et les fichiers keystore.>` et redémarrez Yarn.
6. Redémarrez le service de catalogue.

Création d'un fichier de paramètres

Fichier ZIP contenant l'ensemble des paramètres et variables, et leurs valeurs associées, configurés pour les flux de travail, worklets ou sessions dans le référentiel PowerCenter. Si un référentiel source PowerCenter utilise des fichiers de paramètres dans des sessions et flux de travail, vous pouvez configurer Catalog Administrator pour la lecture des fichiers de paramètres lorsque vous créez la ressource PowerCenter. Les paramètres de PowerCenter peuvent représenter des sources de fichiers plats, des recherches de fichiers plats, des cibles de fichiers plats, des connexions relationnelles, des expressions de niveau transformation ou des objets dans les remplacements SQL. La ressource PowerCenter analyse les fichiers de paramètres et remplace les valeurs de paramètres pour extraire les métadonnées pour les sources de fichiers plats, les recherches de fichiers plats, les cibles de fichiers plats, les connexions relationnelles et les objets dans les remplacements SQL.

Vous pouvez créer un fichier de paramètres dans un répertoire semblable à un dossier PowerCenter à l'aide de l'une des méthodes suivantes :

Créer des fichiers de paramètres dans un répertoire

1. Créez un répertoire portant le même nom que le dossier dans le référentiel PowerCenter.
2. Créez des fichiers de paramètres portant le nom `<workflow name>.prm`.
3. Placez toutes les définitions de paramètre requis pour le flux de travail et les sessions correspondantes dans le répertoire que vous avez créé.
4. Compressez le dossier dans un fichier ZIP. Le fichier ZIP peut contenir plusieurs dossiers contenant plusieurs fichiers de paramètres.

Créer plusieurs fichiers de paramètres

Créez des fichiers de paramètres portant le nom `<PowerCenter folder name>.<workflow name>.prm`. Chaque fichier de paramètres s'applique à un flux de travail avec lequel le fichier est nommé.

Créer un fichier de paramètres unique

Créez un fichier de paramètres unique nommé `parameters.prm`, puis compressez-le. Le fichier de paramètres doit contenir les paramètres qui s'appliquent à tous les flux de travail dans les dossiers de PowerCenter.

Configuration requise du fichier de paramètres

Utilisez un éditeur de texte pour créer le fichier de paramètres. Pour permettre à la ressource PowerCenter de lire les valeurs de paramètres à partir d'un fichier de paramètres, une extension .prm doit être affectée à ce fichier. Regroupez les paramètres et variables dans différentes sections du fichier de paramètres. Chaque section est précédée d'un titre qui identifie le dossier, le flux de travail, le worklet et la session auxquels vous souhaitez transmettre des valeurs de paramètres ou de variables. Définissez les paramètres et les variables directement au-dessous du titre, en entrant chaque paramètre ou variable sur une nouvelle ligne.

Le tableau suivant décrit les titres qui définissent chaque section dans le fichier de paramètres et la portée des paramètres et variables que vous définissez dans chaque section :

Titre	Étendue
[Global]	Tous les dossiers, flux de travail, worklets et sessions.
[folder name.WF:workflow name]	Le flux de travail nommé et toutes les sessions dans le flux de travail.
[folder name.WF:workflow name.WT:worklet name]	Le worklet nommé et toutes les sessions dans le worklet.
[folder name.WF:workflow name.WT:worklet name.WT:worklet name...]	Le worklet imbriqué et toutes les sessions dans le worklet imbriqué.
[folder name.WF:workflow name.ST:session name] -ou- [folder name.WF:workflow name.WT:worklet name.ST:session name] -ou- [folder name.WF:workflow name.WT:worklet name.WT:worklet name.ST:session name] -ou- [folder name.session name] -ou- [session name]	La session nommée.

Exemple de fichier de paramètres

L'exemple suivant présente un exemple de fichier parameters.prm :

```
[Map_Param.WF:WF_Src_Tgt_map_param_case.ST:s_src_tgt_tbl_override_default_map_param]
$$Src_OwnName=MM_PERF6
$$Src_TblName=TBL_SAME_COL
$$Tgt_Tbl_Prefix=MM_PERF6
$$Tgt_TblName=INVENTORY_Q4_2005
[Param_lookup.WF:wf_M_LKP_schema_tble
$$LKP_SCHEMA=TEST_DATA
$$LKP_TBL=LKP_TBL_PARAM
[Param_lookup.WF:wf_M_LKP_schema_tbl_sess_param]
$Param_Lkp_Schema=TEST_DATA
$Param_Lkp_Tbl=LKP_TBL_PARAM[
Param_session.WF:wf_session_param.ST:s_session_param]
$Param_Schema_Name=CROSS_RESOURCE_LINKING_DUP
$Param_SrcTbl_Name=SRC_TBL_NAME_OVERRIDE_PARAM
$Param_TgtTbl_Name=TGT_TBL_NAME_OVERRIDE_PARAM
[Param_Sql_override.WF:wf_M_schema_table_map_parm_sql.ST:s_M_schema_table_map_parm_sql]
$$Map_Schema_Name=CROSS_RESOURCE_LINKING
$$Map_Tbl_Name=SRC_TBL_NAME_OVERRIDE_DUP
```

Service d'intégration de SQL Server

Le service SSIS (SQL Server Integration Service) est un composant de la base de données Microsoft SQL Server que vous pouvez utiliser pour effectuer des tâches de migration des données.

Objets extraits

La ressource SSIS extrait les tâches de flux de données sous la forme de mappages à partir de la source de données SSIS.

Vous pouvez afficher le résumé et le lignage détaillé, et contrôler le résumé des actifs SSIS dans le catalogue.

Prérequis

Vous devez disposer de la licence de ressource SSIS avancée. Contactez le support client international Informatica pour obtenir cette licence.

Autorisations de configuration de la ressource

Configurez les autorisations `ssisadmin` et `dbdatareader` dans la source de données SSIS pour le compte utilisateur permettant d'accéder à la source de données.

Sources de données prises en charge

Enterprise Data Catalog prend en charge l'extraction des métadonnées à partir des sources de données suivantes :

- Microsoft SQL Server
- Système de fichiers
- Teradata
- IBM DB2
- Oracle

Transformations prises en charge

Enterprise Data Catalog affiche le lignage détaillé des transformations suivantes :

- Agréger
- Audit
- Table des caractères
- Fractionnement conditionnel
- Copier la colonne
- Colonne dérivée
- Exporter la colonne
- Regroupement d'éléments approximatifs
- Recherche d'éléments approximatifs
- Importer la colonne
- Recherche
- Fusionner
- Fusionner/Joindre
- Multidiffusion

- Échantillonnage de pourcentage
- Pivot
- Nombre de lignes
- Échantillonnage de ligne
- Changement lent de dimension
- Trier
- Extraction de termes
- Recherche de termes
- Tout unir
- Supprimer le tableau croisé dynamique

Variables prises en charge

Vous pouvez charger un fichier de valeurs de variable pour remplacer les variables définies dans le mappage SSIS. Vous pouvez remplacer les variables telles que le nom de fichier source, le chemin de fichier, les tables source, les tables cibles et les requêtes SQL.

Exemple de fichier de valeurs de variables :

[Packagename.DataFlowTaskName]

User::Variablename1=values

User::Variablename2=values

Paramètres pris en charge

Enterprise Data Catalog prend en charge les paramètres de connexion.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Type de scanner SSIS	<p>Sélectionnez le type de ressource SSIS :</p> <ul style="list-style-type: none"> - Serveur de référentiel. Sélectionnez cette option si vous avez configuré un référentiel dans la base de données MSDB pour stocker tous les packages dans SSIS. - Fichier. Sélectionnez cette option si vous avez configuré un système de fichiers pour stocker tous les packages dans SSIS. - ServerDB de référentiel. Sélectionnez cette option si les packages SSIS sont déployés dans le catalogue SSIS. <p>Vous devez configurer les propriétés requises pour le type de ressource SSIS sélectionné. Pour plus d'informations, consultez la section "Types de ressource SSIS" à la page 142.</p>
Mot de passe	Mot de passe pour la connexion à la base de données ou au package protégé SQL Server Integration Service.
Fichier de valeurs de variables	Cliquez sur Choisir pour sélectionner le fichier contenant des valeurs pour les variables SSIS. Vous pouvez également faire glisser le fichier dans la zone de texte Fichier de valeurs de variables à partir de votre explorateur de fichiers.
Chemin du répertoire de travail	Spécifiez le chemin du fichier source chargé vers le package SSIS.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Attribution automatique des connexions	Indique que les connexions sont attribuées automatiquement.
Activer les ressources de référence	Par exemple, des tables source et cible dans des mappages PowerCenter, et des tables et fichiers source des rapports Tableau.
Conserver les actifs de référence non résolus	La conservation des actifs de référence non résolus vous permet de visualiser l'intégralité du lignage. Les actifs non résolus incluent les fichiers supprimés, les tables temporaires et d'autres actifs qui ne figurent pas dans la ressource principale.
Mémoire	<p>Spécifie la mémoire requise pour exécuter la tâche de scanneur. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que <code>DEBUG</code>, <code>ERROR</code> ou <code>INFO</code>. La valeur par défaut est <code>INFO</code>. Lorsque le jeu de données SSIS est volumineux, utilisez le niveau de journal de scanneur <code>ERROR</code>. - Utilisez le niveau de journal <code>ERROR</code> lorsque le jeu de données SSIS est volumineux. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

Types de ressource SSIS

Configurez les propriétés suivantes pour le type de ressource SSIS sélectionné :

Type de ressource SSIS	Propriétés
Repository Server	<ul style="list-style-type: none"> - URL d'agent. URL de l'agent Enterprise Data Catalog qui s'exécute sur un serveur Microsoft Windows Server. Remarque: Assurez-vous de spécifier l'URL au format suivant : <code>http://<hostname>:<connector_port>/MIMBWebServices</code>. - Version de SQL Server. Sélectionnez l'une des options suivantes dans la liste déroulante de la version de SQL Server : <ul style="list-style-type: none"> - SQL Server 2012 - SQL Server 2014 - SQL Server 2016 La valeur par défaut est SQL Server 2012. - Nom du package/référentiel. Spécifiez le référentiel ou le package à partir duquel vous voulez importer des métadonnées. Cliquez sur Sélectionner. La boîte de dialogue Sélectionner le nom du package/référentiel s'affiche. Sélectionnez le package requis à l'aide de l'une des options suivantes : <ul style="list-style-type: none"> - Sélectionner dans la liste. Sélectionnez le package ou le référentiel requis à partir d'une liste de packages ou de référentiels. - Sélectionner à l'aide d'une expression régulière. Fournissez une expression régulière SQL pour sélectionner le package ou le nom du référentiel qui correspond à l'expression. - Nom d'hôte. Nom d'hôte ou adresse IP de la machine sur laquelle SSIS est en cours d'exécution.
Fichier	<ul style="list-style-type: none"> - Fichier. Cliquez sur Sélectionner pour sélectionner le fichier contenant les packages que vous souhaitez charger vers Enterprise Data Catalog. Vous pouvez également faire glisser le fichier dans la zone de texte Fichier à partir de votre explorateur de fichiers. Assurez-vous que les fichiers ont l'extension .dtsx. Si vous voulez charger plusieurs fichiers et fichiers de paramètres, ajoutez les dans un fichier .zip et chargez celui-ci. - Codage. Sélectionnez le type de codage souhaité pour les métadonnées extraites dans la liste déroulante.
ServerDB de référentiel	<ul style="list-style-type: none"> - Mode d'authentification. Sélectionnez le type d'authentification pour la connexion à la base de données SQL Server Integration Service. Spécifiez l'une des valeurs suivantes : <ul style="list-style-type: none"> - Authentification SQL Server. Spécifiez le nom d'utilisateur et le mot de passe de la base de données pour la connexion au serveur Microsoft SQL Server. - Authentification Windows. Authentification Windows permettant d'authentifier le nom d'utilisateur et le mot de passe de la base de données pour la connexion au serveur SSIS. La valeur par défaut est Authentification SQL Server. - Nom d'hôte. Nom d'hôte ou adresse IP de l'ordinateur sur lequel SQL Server Integration Service est en cours d'exécution. - Utilisateur. Nom d'utilisateur pour la connexion à la base de données SQL Server Integration Service. - Base de données. Nom de la base de données SQL Server Integration Service. - Parcourir le filtre de dossier. Spécifiez le nom du dossier de package dans le catalogue SSIS. - Packages. Sélectionnez le package SQL Server Integration Service à partir duquel vous souhaitez importer les métadonnées à l'aide de l'une des options suivantes : <ul style="list-style-type: none"> - Sélectionner dans la liste. Sélectionnez le package requis dans une liste de packages. - Sélectionner à l'aide d'une expression régulière. Fournissez une expression régulière SQL pour sélectionner le nom du package qui correspond à l'expression. - URL d'agent. URL de l'agent Enterprise Data Catalog qui s'exécute sur un serveur Microsoft Windows. Remarque: Assurez-vous de spécifier l'URL au format suivant : <code>http://<hostname>:<connector_port>/MIMBWebServices</code>.

Configuration de l'agent Enterprise Data Catalog pour la ressource sur un serveur authentifié par Windows

Pour configurer l'agent Enterprise Data Catalog pour SSIS, procédez comme suit :

1. Téléchargez le pilote JDBC SQL Server (version 4.2) dans un emplacement temporaire.
2. Copiez le fichier `sqljdbc<version>.jar` dans l'emplacement `<Enterprise_Data_Catalog_Agent_location>\java\Jdbc\mssql`.
3. Copiez le fichier `\sqljdbc_4.2\enu\auth\x64\sqljdbc_auth.dll` extrait du fichier `sqljdbc_4.2.jar` dans l'emplacement `<Enterprise_Data_Catalog_Agent_location>\jre\bin`.
4. Exécutez le fichier `RestartServerService.bat`.
5. Dans Catalog Administrator, créez une ressource SSIS avec l'URL de l'agent Enterprise Data Catalog suivante : `http://<HostName>:<Port>/MIMBWebServices/`
6. Exécutez la ressource.

Les limitations de la ressource SSIS sont les suivantes :

- La ressource extrait des connexions incorrectes si vous utilisez des scripts en tant que source ou cible de données.
- La ressource n'utilise pas les valeurs du fichier de valeurs de variable si le package SSIS sélectionné contient des paramètres.
- La ressource ne parvient pas à extraire des mappages avec les valeurs de paramètres si vous utilisez des paramètres pour les tables source ou cible.
- La ressource n'affiche pas les connexions, les mappages de sources de données et le lignage si vous utilisez * dans la requête de recherche ou une commande SQL.

IBM InfoSphere DataStage

IBM® InfoSphere® DataStage® est une plate-forme ETL utilisée pour intégrer des données sur plusieurs systèmes.

Objets extraits

La ressource IBM InfoSphere DataStage prend en charge les sources de données source et cible suivantes incluses dans les mappages ETL dans IBM InfoSphere DataStage :

- Oracle
- Teradata
- Fichiers plats

La ressource IBM InfoSphere DataStage extrait les mappages pour les tâches de serveur et les tâches parallèles à partir du fichier DSX ou XML. Si un fichier DSX ou XML inclut des tâches de séquence ou autonomes, la ressource extrait uniquement les tâches de serveur et les tâches autonomes liées aux tâches de séquence.

La ressource extrait les métadonnées des types de transformation suivants dans les mappages :

- Agrégation*
- Filtre*

- Jointure*
- Entonnoir*
- Tri*
- Recherche*
- Fusion*
- Suppression des doublons*
- Différence*
- Copier
- Aperçu
- Transformateur SQL
- Transformateur de base
- Capture des modifications*
- Tableau croisé dynamique
- Switch
- Résultat XML

Remarque: * Indique les transformations pour lesquelles la ressource affiche les propriétés dans la logique de transformation sous l'onglet **Lignage et impact**.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de connexion à la ressource :

Propriété	Description
Fichier	Fichier au format XML ou DSX qui inclut les métadonnées de tâche opérationnelle. Vérifiez que la taille de fichier ne dépasse pas 250 Mo.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Fichier de valeurs de variables	Sélectionnez le fichier qui inclut toutes les valeurs de paramètres requises par les mappages ou les tâches présentes dans le fichier DSX ou XML d'IBM InfoSphere DataStage.
Attribution automatique des connexions	Indique que les connexions à la ressource doivent être attribuées automatiquement.
Activer les ressources de référence	Par exemple, des tables source et cible dans des mappages PowerCenter, et des tables et fichiers source des rapports Tableau.

Propriété	Description
Conserver les actifs de référence non résolus	Si vous conservez des actifs de référence non résolus, vous pouvez afficher l'intégralité du lignage. Les actifs non résolus incluent les fichiers supprimés, les tables temporaires et d'autres actifs qui ne figurent pas dans la ressource principale.
Lignage détaillé	Extrait et ingère les métadonnées liées à la logique de transformation des actifs qui incluent des transformations ou des étapes. Une transformation ou une étape indique la génération, la modification ou le passage de données entre les connexions source et cible. Une logique de transformation affiche les mappages ou les types de relations de flux de données entre les actifs source et cible liés à l'actif sélectionné dans Enterprise Data Catalog.
Mémoire	Valeur de mémoire requise pour exécuter une tâche de scanner. Spécifiez l'une des valeurs de mémoire suivantes : <ul style="list-style-type: none"> - Faible - Moyenne - Élevée Remarque: Pour obtenir plus d'informations sur les valeurs de mémoire, consultez l'article <i>Tuning Enterprise Data Catalog Performance</i> de la bibliothèque de guides pratiques.
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none"> - <code>Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journal du scanner sur des valeurs spécifiques, telles que <code>DEBUG</code>, <code>ERROR</code> ou <code>INFO</code>. La valeur par défaut est <code>INFO</code>. - <code>Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement YARN. Utilisez une virgule pour séparer plusieurs valeurs de paire de clés. - <code>Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

Directives pour la configuration des paramètres et des variables

Les conditions suivantes s'appliquent lorsque vous utilisez la ressource IBM InfoSphere DataStage pour extraire les métadonnées :

- La ressource prend en charge la substitution de valeur de paramètre lorsque vous utilisez des paramètres au niveau de la connexion.
- Le fichier de paramètres ne prend pas en charge la syntaxe permettant d'isoler les paramètres de substitution nommés de manière identique par les mappages.
- Les entrées dans la propriété **Fichier de valeurs de variables** doivent être au format suivant :
variable1_name=variable1_value
variable2_name=variable2_value
...
variableN_name=variableN_value
- IBM InfoSphere DataStage utilise les variables de substitution dans les paramètres utilisés dans les connexions. Si vous ne fournissez aucune valeur de variable, la ressource enregistre le message d'avertissement impossible de déterminer la valeur d'une variable dans le fichier journal lorsque vous exécutez la ressource. La ressource n'utilise pas de substitution pour le nom de variable dans le modèle.
- Vous devez définir l'ensemble de paramètres au format suivant pour la ressource :

```
parameterset_name.$parameter1_name=parameter1_value
```

- Les noms de variable ne sont pas sensible à la casse et la ressource supprime les espaces de début et de fin.

Limitations de la ressource

Les limitations de la ressource IBM InfoSphere DataStage sont les suivantes :

- Elle extrait uniquement les métadonnées des fichiers XML ou DSX.
- Elle ne prend pas en charge les transformations de conteneur.

CHAPITRE 8

Configuration des ressources de modélisation de données

Ce chapitre comprend les rubriques suivantes :

- [Présentation des ressources de modélisation de données, 147](#)
- [Erwin, 147](#)
- [SAP PowerDesigner, 151](#)

Présentation des ressources de modélisation de données

Le groupe de ressources de modélisation de données inclut les ressources de modélisation de données. Lorsque vous créez et exécutez une ressource de modélisation de données, le scanner de modélisation de données extrait et migre les métadonnées de la ressource associée vers le catalogue. Vous pouvez afficher les résultats migrés dans Enterprise Data Catalog.

Erwin

Erwin est un logiciel informatique pour la modélisation de données.

Objets extraits

Enterprise Data Catalog extrait les métadonnées des objets suivants dans une source de données Erwin :

- IErwin
- DataElement
- LogicalModel
- PhysicalModel
- Colonne ou attribut
- Membre de l'index
- Objet de données

- Index
- RelationshipEnd non nommé
- Groupe de clés de candidat
- Colonne de vue
- Colonne de clé primaire
- Table ou entité
- Groupe de clés étrangères
- Relation
- Colonne
- Identité
- Séquence
- Déclencheur
- Relation de vue
- Vue
- Attribut de clé primaire
- DataSource
- Entité
- Schéma de base de données
- Domaine
- Synonyme
- Diagramme ER
- Domaine
- Défini par l'utilisateur
- Table
- Argument
- Valeur par défaut
- Procédure stockée
- Rôle de sous-type de relation
- Min/Max
- Modèle
- Relation de sous-type
- UDP
- UDP
- Ressource

Propriétés de connexion à la ressource

Les tableaux suivants répertorient les propriétés que vous devez configurer pour ajouter une ressource de modélisation de données Erwin : l'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Type de scanner ERwin	<p>Sélectionnez l'une des options suivantes pour spécifier le type de scanner ERwin :</p> <ul style="list-style-type: none"> - Fichier ERwin 8.x. Cliquez sur Sélectionner pour charger un fichier .xml à partir duquel vous voulez extraire des métadonnées. - Fichier ERwin 9.x. Cliquez sur Sélectionner pour charger un fichier .xml à partir duquel vous voulez extraire des métadonnées. - Modeleur de données ERwin 8.x (modèle unique de Mart). Spécifiez les paramètres suivants si vous sélectionnez cette option : <ul style="list-style-type: none"> - URL d'agent. URL de l'agent Enterprise Data Catalog qui s'exécute sur un serveur Microsoft Windows. <p>Remarque: Assurez-vous de spécifier l'URL au format suivant : <code>http://<hostname>:<connector_port>/MIMBWebServices</code>.</p> <ul style="list-style-type: none"> - Type de base de données. Sélectionnez le type de serveur de base de données configuré pour le serveur Mart dans la liste déroulante. - Serveur de base de données. Spécifiez le serveur Mart auquel vous voulez vous connecter. - Nom de la base de données. Spécifiez le nom du serveur de base de données auquel vous voulez vous connecter. - Authentification. Sélectionnez l'une des options suivantes en fonction du mode d'authentification configuré pour se connecter au serveur Mart : <ul style="list-style-type: none"> - Authentification du serveur. Spécifiez le nom d'utilisateur et le mot de passe permettant de vous connecter au serveur Mart. - Authentification de base de données. Spécifiez le nom d'utilisateur et le mot de passe de la base de données permettant de vous connecter au serveur Mart. - Authentification Windows. Sélectionnez cette option pour spécifier que l'authentification Windows doit être utilisée pour authentifier le nom d'utilisateur et le mot de passe de la base de données permettant de vous connecter au serveur Mart. - Modèle (facultatif). Spécifiez la chaîne de repère de modèle ERwin. Utilisez le format suivant pour spécifier la chaîne de repère : <code>mmart://<database name>/<path>/<model name></code>. - Modeleur de données ERwin 9.x (modèle unique de Mart). Spécifiez les paramètres suivants si vous sélectionnez cette option : <ul style="list-style-type: none"> - URL d'agent. URL de l'agent Enterprise Data Catalog qui s'exécute sur un serveur Microsoft Windows. <p>Remarque: Assurez-vous de spécifier l'URL au format suivant : <code>http://<hostname>:<connector_port>/MIMBWebServices</code>.</p> <ul style="list-style-type: none"> - Nom du serveur. Spécifiez le serveur Mart auquel vous voulez vous connecter. - Port du serveur. Spécifiez le numéro de port sur le serveur Mart à utiliser pour vous connecter. - Utilisez IIS (facultatif). Sélectionnez cette option si vous avez configuré la connexion au serveur Mart avec le serveur Web Microsoft IIS. Le numéro de port est attribué dynamiquement pour les connexions à l'aide d'un serveur IIS. Si un serveur IIS n'est pas utilisé pour la connexion, vous devez spécifier le numéro de port du serveur. - Utilisez SSL (facultatif). Sélectionnez cette option si vous avez configuré l'authentification SSL (Secure Socket Layer) pour vous connecter au serveur Mart. - Nom de l'application. Indiquez le nom de l'application à laquelle vous voulez vous connecter sur le serveur Mart. - Authentification. Sélectionnez l'une des options suivantes pour spécifier le mode d'authentification à utiliser pour la connexion au serveur Mart : <ul style="list-style-type: none"> - Authentification du serveur. Spécifiez le nom d'utilisateur et le mot de passe permettant de vous connecter au serveur Mart. - Authentification de base de données. Spécifiez le nom d'utilisateur et le mot de passe de la base de données permettant de vous connecter au serveur Mart. - Authentification Windows. Sélectionnez cette option pour spécifier que l'authentification Windows doit être utilisée pour authentifier le nom d'utilisateur et le mot de passe de la base de données permettant de vous connecter au serveur Mart.

Propriété	Description
	<ul style="list-style-type: none"> - Modèle (facultatif). Spécifiez la chaîne de repère de modèle ERwin. Utilisez le format suivant pour spécifier la chaîne de repère : <code>mmart://<database name>/<path>/<model name></code>.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Importer des propriétés définies par l'utilisateur (UDP)	<p>Spécifiez la façon dont vous voulez importer les définitions et les valeurs des propriétés définies par l'utilisateur (UDP). Dans ERwin, une propriété définie par l'utilisateur est un objet de définition de propriété qui a une valeur par défaut. L'objet auquel l'UDP s'applique peut avoir une valeur explicite ou aucune valeur attribuée. Si un objet n'a pas de valeur attribuée, ERwin attribue à l'UDP la valeur par défaut de la définition de propriété.</p> <p>Vous pouvez sélectionner l'une des options suivantes :</p> <ul style="list-style-type: none"> - En tant que métadonnées. Importe des valeurs explicites en tant qu'objets Valeur de propriété. Cette option conserve la valeur par défaut du type de propriété et n'importe pas les valeurs implicites. - En tant que métadonnées, migrer les valeurs par défaut. Importe des valeurs explicites et implicites en tant qu'objets Valeur de propriété. - Dans la description, migrer les valeurs par défaut. Ajoute le nom et la valeur à la propriété Description de l'objet pour les valeurs explicites et implicites. - Les deux méthodes, migrer les valeurs par défaut. Importe la valeur de l'UDP en tant que métadonnées et dans la description de l'objet. <p>La valeur par défaut est En tant que métadonnées.</p>
Importer le nom d'une relation	<p>Spécifiez la façon dont vous voulez importer des noms de relations à partir d'ERwin :</p> <ul style="list-style-type: none"> - À partir du nom de la relation. Utilisez la propriété Nom de la relation. - À partir du nom dérivé. Utilisez la propriété Nom dérivé. <p>La valeur par défaut est Du nom de la relation.</p>
Importer l'ordre de colonne de	<p>Spécifiez la façon dont vous voulez importer la position des colonnes dans les tables :</p> <ul style="list-style-type: none"> - Ordre des colonnes. Importe la position des colonnes à partir de l'ordre des colonnes affichées dans la vue physique ERwin. - Ordre physique. Importe la position des colonnes à partir de l'ordre dans lequel les colonnes sont stockées dans la base de données. <p>La valeur par défaut est Ordre physique.</p>
Importer des schémas de propriétaire	Sélectionnez cette option si vous voulez importer des schémas de propriétaire.
Déplacer les entités vers les domaines	Sélectionnez cette option pour déplacer des entités vers des zones d'objet définies par l'utilisateur comme des packages de conception UML conventionnels ayant leur propre espace de noms.
Attribution automatique des connexions	Sélectionnez cette option pour spécifier que la connexion doit être attribuée automatiquement.

Propriété	Description
Mémoire	<p>Spécifie la mémoire requise pour exécuter la tâche de scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

SAP PowerDesigner

Créez une ressource selon le type de ressource SAP PowerDesigner afin d'extraire les métadonnées d'une source de données SAP PowerDesigner. Vous devez spécifier les informations de connexion lorsque vous configurez la ressource.

Objets extraits

Enterprise Data Catalog extrait les métadonnées des objets suivants dans une source de données SAP PowerDesigner :

- *Table
- *Vue
- Procédure
- Séquence
- Utilisateur
- Association
- *Synonyme
- Règle d'entreprise
- Dimension
- Fait
- Modèle physique de données
- Diagramme physique
- Diagramme multidimensionnel

Remarque: * indique les objets pour lesquels la ressource extrait les informations de lignage.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de connexion à la ressource :

Propriété	Description
URL d'agent	URL de l'agent Enterprise Data Catalog qui s'exécute sur un serveur Microsoft Windows.
Type de modèle SAP PowerDesigner	Version de PowerDesigner utilisée pour créer les fichiers PDM (Physical Data Modeling).
Supprimer les caractères binaires	Supprime les caractères binaires du fichier PDM avant l'analyse de celui-ci.
Importer des propriétés définies par l'utilisateur (UDP)	Cette ressource ne prend pas en charge cette option dans la version actuelle d'Enterprise Data Catalog.
Ajouter des informations volumétriques	Importe la valeur du nombre d'occurrences d'un objet et ajoute les informations à la propriété description d'un objet.
Supprimer RTF	Supprime le formatage RTF (Rich Text Formatting) des annotations.
Inverser les noms de rôles	Inverse les noms de rôles de chaque relation du modèle.
Attribution automatique des connexions	Attribue automatiquement les connexions à la ressource.
Fichier	Sélectionnez le fichier PDM qui inclut le modèle de données à partir duquel importer les métadonnées.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Mémoire	Mémoire requise pour exécuter une tâche de scanneur. Spécifiez l'une des valeurs de mémoire suivantes : <ul style="list-style-type: none">- Faible- Moyenne- Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .

Propriété	Description
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journal du scanner sur des valeurs spécifiques, telles que <code>DEBUG</code>, <code>ERROR</code> ou <code>INFO</code>. La valeur par défaut est <code>INFO</code>. - <code>Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule pour séparer plusieurs valeurs de paire de clés. - <code>Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1. - <code>-DINFA_TRUSTSTORE=<trust store path></code>. Chemin d'accès au fichier <code>Informatica truststore</code>. - <code>-DINFA_TRUSTSTORE_PASSWORD=<trust store encrypted key></code>. Mot de passe chiffré du <code>truststore</code>.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

CHAPITRE 9

Configuration des ressources de gestion de la base de données

Ce chapitre comprend les rubriques suivantes :

- [Présentation des ressources de gestion de la base de données, 154](#)
- [IBM DB2, 155](#)
- [IBM DB2 for z/OS, 157](#)
- [IBM Netezza, 160](#)
- [JDBC, 164](#)
- [Microsoft SQL Server, 167](#)
- [Oracle, 170](#)
- [SAP HANA, 173](#)
- [Sybase, 175](#)
- [Teradata, 178](#)

Présentation des ressources de gestion de la base de données

Le groupe de ressources de gestion de la base de données inclut les ressources de gestion de la base de données. Lorsque vous créez et exécutez une ressource de gestion de la base de données, le scanner de gestion de la base de données extrait et migre les métadonnées de la ressource associée vers le catalogue. Vous pouvez afficher les résultats migrés dans Enterprise Data Catalog.

IBM DB2

Vous pouvez configurer un type de ressource IBM DB2 pour extraire les métadonnées des bases de données IBM DB2.

Objets extraits

Enterprise Data Catalog extrait les métadonnées des objets de schéma suivants d'une ressource IBM DB2 :

- Tables
- Table partitionnée
- Table en cluster de plage
- Table de mise en cluster multidimensionnelle
- Contraintes (NOT NULL , UNIQUE , CHECK , PRIMARY KEY , FOREIGN KEY , DEFAULT)
- Alias
- Séquences
- Vues
- Déclencheurs
- Fonction
- Procédure
- Index
- Table de requêtes matérialisée

Autorisations de configuration de la ressource

Le tableau suivant répertorie les objets de schéma et les tables système dans lesquels vous devez configurer les autorisations SELECT pour le compte utilisateur de la base de données IBM DB2 :

Objets de schéma	Tables système
<ul style="list-style-type: none">- Tables- Vues- Index- Packages- Procédures- Fonctions- Séquences- Déclencheurs- Synonymes	<ul style="list-style-type: none">- SYSCAT.SCHEMATA- SYSCAT.TRIGGERS- SYSCAT.INDEXES- SYSCAT.INDEXCOLUSE- SYSCAT.REFERENCES- SYSCAT.TABCONST- SYSCAT.KEYCOLUSE- SYSCAT.COLCHECKS- SYSCAT.DATATYPES- SYSCAT.COLUMNS- SYSCAT.TABLES- SYSCAT.FUNCTIONS- SYSCAT.PROCEDURES- SYSCAT.PACKAGES- SYSCAT.TRIGGERS- SYSCAT.SEQUENCES- SYSCAT.INDEXES- SYSCAT.CHECKS- SYSCAT.VIEWS- SYSCAT.PACKAGES- SYSCAT.FUNCDEP- SYSCAT.TRIGDEP- SYSCAT.VIEWDEP

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de connexion du type de ressource IBM DB2 :

Propriété	Description
Utilisateur	Nom du compte utilisateur qui se connecte à la base de données IBM DB2.
Mot de passe	Mot de passe du compte d'utilisateur qui se connecte à la base de données IBM DB2.
Hôte	Nom d'hôte complet de la machine sur laquelle la base de données IBM DB2 est hébergée.
Port	Numéro de port de la base de données IBM DB2.
Base de données	URL de connexion DB2 utilisée pour accéder aux métadonnées de la base de données.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Importer des objets système	Spécifie les objets système à importer.
Schéma	Spécifie une liste de schémas de base de données.
Filtre de métadonnées source	<p>Vous pouvez inclure ou exclure des tables et des vues depuis l'exécution de la ressource. Utilisez des points-virgules (;) pour séparer les noms des tables et les noms des vues.</p> <p>Pour plus d'informations sur le champ de filtre, voir "Filtre de profil de données et métadonnées source" à la page 12.</p>
Importer des procédures stockées	Spécifie les procédures stockées à importer.
Sensible à la casse	<p>Spécifie que la ressource est configurée pour l'insensibilité à la casse. Sélectionnez l'une des valeurs suivantes :</p> <ul style="list-style-type: none">- True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse.- False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. <p>La valeur par défaut est False.</p>
Mémoire	<p>Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanneur.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none">- Faible- Moyenne- Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le réglage des performances d'Enterprise Data Catalog sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>

Propriété	Description
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> -Dscannerloglevel=<DEBUG/INFO/ERROR>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. -Dscanner.container.core=<No. of core>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. -Dscanner.yarn.app.environment=<key=value>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. -Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0>. Augmente la mémoire du conteneur du scanner lorsque pmem est activé. La valeur par défaut est 1. -DskipExternalDatabaseObjects=true. Exclut tous les objets externes tels que le schéma, la table, la base de données et la vue.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource IBM DB2. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource IBM DB2. Pour plus d'informations sur l'activation de la découverte de domaines de données, consultez la section [Composite Data Domain Discovery](#).

IBM DB2 for z/OS

Configurez un type de ressource IBM DB2 for z/OS afin d'extraire des métadonnées des bases de données IBM DB2 for z/OS.

Objets extraits

Enterprise Data Catalog extrait les métadonnées des objets de schéma suivants d'une ressource IBM DB2 for z/OS :

- Procédure stockée
- Fonction
- Synonyme
- Déclencheur
- Table
- Vues
- Schéma

Autorisations de configuration de la ressource

Lorsque vous chargez une ressource IBM DB2 for z/OS, Enterprise Data Catalog extrait les métadonnées du sous-système IBM DB2 for z/OS. Enterprise Data Catalog sélectionne les métadonnées des tables du catalogue DB2.

Configurez les privilèges SELECT dans les tables système suivantes du compte utilisateur IBM DB2 for z/OS que vous utilisez pour vous connecter à la source de données IBM DB2 for z/OS :

Tables système	Métadonnées extraites
<ul style="list-style-type: none"> - SYSIBM.SYSSYNONYMS - SYSIBM.SYSTABLES - SYSIBM.SYSROUTINES - SYSIBM.SYSSEQUENCES - SYSIBM.SYSDATATYPES 	Schémas
<ul style="list-style-type: none"> - SYSIBM.SYSTABLES - SYSIBM.SYSTABLESPACE 	<ul style="list-style-type: none"> - Tables - Vues
<ul style="list-style-type: none"> - SYSIBM.SYSVIEWS 	<ul style="list-style-type: none"> - Alias
<ul style="list-style-type: none"> - SYSIBM.SYSTABLES - SYSIBM.SYSCOLUMNS - SYSIBM.SYSTABLESPACE 	Colonnes
<ul style="list-style-type: none"> - SYSIBM.SYSKEYS - SYSIBM.SYSINDEXES 	Clés primaires
<ul style="list-style-type: none"> - SYSIBM.SYSRELS - SYSIBM.SYSFOREIGNKEYS - SYSIBM.SYSINDEXES 	Clés étrangères
<ul style="list-style-type: none"> - SYSIBM.SYSCHECKS - SYSIBM.SYSCHECKDEP 	Vérifier les contraintes
<ul style="list-style-type: none"> - SYSIBM.SYSINDEXES - SYSIBM.SYSKEYS 	Index
<ul style="list-style-type: none"> - SYSIBM.SYSSYNONYMS - SYSIBM.SYSTABLES 	Synonymes
SYSIBM.SYSTRIGGERS	Déclencheurs
SYSIBM.SYSTRIGGERS	Procédures stockées
SYSIBM.SYSROUTINES	Fonctions
SYSIBM.SYSSEQUENCES	Séquences
<ul style="list-style-type: none"> - SYSIBM.SYSSEQUENCESDEP - SYSIBM.SYSSEQUENCES 	Identité
SYSIBM.SYSDATATYPES	Types définis par l'utilisateur

Prérequis

Pour accéder aux métadonnées à partir des sous-systèmes z/OS à l'aide de PowerExchange, procédez comme suit :

- Vérifiez que le client PowerExchange est installé dans la machine du domaine.
- Incluez le chemin PWX_HOME dans le fichier <INFA_HOME>/tomcat/temp/envvars.txt.

Pour plus d'informations sur l'installation et la configuration de PowerExchange, consultez les rubriques *Installer PowerExchange* et *Configurer PowerExchange* de l'article de la bibliothèque de guides pratiques [Configuring an IBM DB2 for z/OS Resource in Metadata Manager 9.0](#).

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de la ressource IBM DB2 for z/OS :

Propriété	Description
Emplacement	Nom du nœud dans le fichier dbmover.cfg de la machine sur laquelle le service de catalogue s'exécute et qui pointe vers l'écouteur PowerExchange sur le système z/OS. Remarque: Enterprise Data Catalog utilise PowerExchange for DB2 for z/OS pour accéder aux métadonnées des sous-systèmes z/OS.
Utilisateur	Nom du compte utilisateur permettant de se connecter à la base de données IBM DB2 for z/OS.
Mot de passe	Mot de passe du compte utilisateur permettant de se connecter à la base de données IBM DB2 for z/OS.
Codage	Page de code du sous-système IBM DB2 for z/OS.
ID du sous-système	Nom du sous-système DB2.

Le tableau suivant décrit la propriété **supplémentaire** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Active l'extraction des métadonnées.
Base de données	Cliquez sur Sélectionner pour spécifier une liste de bases de données à importer. Vous pouvez utiliser l'une des options suivantes dans la boîte de dialogue Sélectionner une base de données pour importer les bases de données : <ul style="list-style-type: none"> - Sélectionner dans la liste. Sélectionnez la base de données requise dans une liste de bases de données disponibles. - Sélectionner à l'aide d'une expression régulière. Fournissez des expressions régulières SQL telles que * ou ? pour inclure des bases de données correspondant à l'expression.
Schéma	Cliquez sur Sélectionner pour spécifier les schémas à importer. Vous pouvez utiliser l'une des options suivantes dans la boîte de dialogue Sélectionner un schéma pour importer les schémas : <ul style="list-style-type: none"> - Sélectionner dans la liste. Sélectionnez les schémas requis dans une liste de schémas disponibles. - Sélectionner à l'aide d'une expression régulière. Fournissez des expressions régulières SQL telles que * ou ? pour inclure des schémas correspondant à l'expression. Remarque: Si vous utilisez une expression régulière pour sélectionner les bases de données à importer, vous ne devez pas utiliser l'expression régulière pour sélectionner les schémas.

Propriété	Description
Filtre de métadonnées source	<p>Vous pouvez inclure des tables et des vues à partir de l'exécution de la ressource. Utilisez des points-virgules (;) pour séparer les noms des tables et les noms des vues. Lorsque la liste est vide, l'exécution des ressources inclut toutes les tables et vues disponibles. Vous pouvez utiliser des expressions régulières telles que * et ? pour spécifier la liste des tables et des vues.</p> <p>Pour plus d'informations sur le champ de filtre, voir "Filtre de profil de données et métadonnées source" à la page 12.</p>
Sensible à la casse	<p>Indique que la ressource est configurée pour la sensibilité à la casse.</p> <ul style="list-style-type: none"> - Cochez cette case pour indiquer que la ressource est configurée comme étant sensible à la casse. - Décochez cette case pour indiquer que la ressource est configurée comme étant insensible à la casse. <p>Par défaut, la ressource est configurée comme étant sensible à la casse.</p>

Pour consulter les informations des journaux de la ressource IBM DB2 for z/OS, reportez-vous au fichier `LDM.log`.

Vous pouvez activer la découverte de données pour une ressource IBM DB2 for z/OS. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource IBM DB2 for z/OS. Pour plus d'informations sur l'activation de la découverte de domaines de données, consultez la section [Composite Data Domain Discovery](#).

IBM Netezza

Vous devez configurer plusieurs propriétés de configuration lorsque vous créez une ressource afin d'extraire des métadonnées des bases de données IBM Netezza.

Objets extraits

Enterprise Data Catalog extrait les métadonnées des objets suivants d'une ressource IBM Netezza :

- Colonne
- Base de données
- Type de données
- Clé de distribution
- Colonne Clé de distribution
- Expression
- Table externe
- Clé étrangère
- Vue matérialisée
- Clé primaire
- Schéma
- Séquence

- Procédure stockée
- Synonyme
- Table
- Contrainte de clé unique
- Vue

Autorisations de configuration de la ressource

Configurez le privilège LIST de l'utilisateur Netezza afin d'extraire les métadonnées d'une base de données Netezza. Le privilège LIST accorde l'accès utilisateur Netezza à tous les objets de la base de données dans la base de données Netezza via Enterprise Data Catalog. Si nécessaire, vous pouvez attribuer les autorisations d'accès à la base de données afin de limiter l'accès aux objets individuels de la base de données.

Configurez les autorisations SELECT du compte utilisateur Netezza dans les vues système suivantes :

- _V_TABLE_DIST_MAP
- _V_TABLE
- _V_DATABASE
- _v_relation_keydata
- _v_relation_column
- _v_view
- _v_synonym
- _v_datatype
- _v_sequence

Prérequis

Téléchargez les fichiers JAR du pilote JDBC depuis Internet et copiez-les dans le répertoire `NetezzaScanner` du répertoire `<INFA_HOME>/services/CatalogService/ScannerJars/externalDependencies`. Il n'est pas nécessaire de redémarrer le service de catalogue.

Pour respecter les prérequis, vous pouvez également effectuer les étapes suivantes :

1. Téléchargez le fichier du pilote JDBC et copiez-le dans le répertoire `<INFA_HOME>/services/CatalogService/ScannerBinaries`.
2. Ouvrez le fichier `<INFA_HOME>/services/CatalogService/ScannerBinaries/CustomDeployer/scannerDeployer.xml` et ajoutez-y les lignes suivantes :

```
<ExecutionContextProperty>
  <ExecutionContextProperty isLocationProperty="true"
    dependencyToUnpack="netezza.zip">
    <PropertyName>NetezzaScanner_DriverLocation</PropertyName>
    <PropertyValue>scanner_agents/netezza/Drivers</PropertyValue>
  </ExecutionContextProperty>
```

3. Enregistrez le fichier `scannerDeployer.xml`.
4. Redémarrez le service de catalogue.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de connexion du type de ressource IBM Netezza :

Propriété	Description
Hôte	Nom d'hôte ou adresse IP de la machine sur laquelle le serveur de gestion de base de données s'exécute.
Port	Numéro de port de la base de données Netezza.
Utilisateur	Nom du compte utilisateur utilisé pour se connecter à la base de données Netezza.
Mot de passe	Mot de passe du compte utilisateur utilisé pour se connecter à la base de données Netezza.
Base de données	Chaîne de connexion de source de données ODBC pour une base de données Netezza. Entrez le nom de la source de données du DSN Netezza si vous en avez créé un.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Schéma	Spécifie une liste de schémas de base de données séparés par des points-virgules.
Filtre de métadonnées source	<p>Vous pouvez inclure ou exclure des tables et des vues depuis l'exécution de la ressource. Utilisez des points-virgules (;) pour séparer les noms des tables et les noms des vues.</p> <p>Pour plus d'informations sur le champ de filtre, voir "Filtre de profil de données et métadonnées source" à la page 12.</p>
Sensible à la casse	<p>Spécifie que la ressource est configurée pour l'insensibilité à la casse. Sélectionnez l'une des valeurs suivantes :</p> <ul style="list-style-type: none">- True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse.- False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. <p>La valeur par défaut est False.</p>

Propriété	Description
Mémoire	<p>Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanner.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque pmem est activé. La valeur par défaut est 1. - <code>-DskipExternalDatabaseObjects=true</code>. Exclut tous les objets externes tels que le schéma, la table, la base de données et la vue.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource IBM Netezza. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource IBM Netezza. Pour plus d'informations sur l'activation de la découverte de domaines de données, consultez la section [Composite Data Domain Discovery](#).

JDBC

Vous pouvez utiliser une connexion JDBC pour accéder aux tables dans une base de données.

Objets extraits

JDBC avec MySQL

Enterprise Data Catalog extrait les métadonnées des objets de schéma suivants de bases de données MySQL et AS400 qui utilisent JDBC pour la connectivité :

- Tables
- Contraintes (NOT NULL , UNIQUE , CHECK , PRIMARY KEY , FOREIGN KEY , DEFAULT)
- Table de clones
- Index
- Vues
- Déclencheurs
- Fonction
- Procédure
- Vue matérialisée
- Déclencheur de MV

JDBC avec AS400

- Tables
- Contraintes (NOT NULL , UNIQUE , CHECK , PRIMARY KEY , FOREIGN KEY , DEFAULT)
- Alias
- Table dupliquée
- Séquences
- Vues
- Déclencheurs
- Fonction
- Procédure
- Index
- Types de données définis par l'utilisateur

Autorisations de configuration de la ressource

Vous devez configurer les autorisations CONNECT et SELECT_CATALOG_ROLE pour le compte utilisateur de la base de données que vous utilisez pour accéder aux bases de données MySQL ou AS400.

Prérequis

Téléchargez les fichiers JAR du pilote JDBC depuis Internet et copiez-les dans le répertoire `JDBCScanner` du répertoire `<INFA_HOME>/services/CatalogService/ScannerJars/externalDependencies`. Il n'est pas nécessaire de redémarrer le service de catalogue.

Pour respecter les prérequis, vous pouvez également effectuer les étapes suivantes :

1. Téléchargez le fichier du pilote JDBC et copiez-le dans le répertoire `<INFA_HOME>/services/CatalogService/ScannerBinaries`.

2. Ouvrez le fichier `<INFA_HOME>/services/CatalogService/ScannerBinaries/CustomDeployer/scannerDeployer.xml` et ajoutez-y les lignes suivantes :

```
</ExecutionContextProperty>
<ExecutionContextProperty isLocationProperty="true"
dependencyToUnpack="genericJDBC.zip">
  <PropertyName>JDBCScanner_DriverLocation</PropertyName>
  <PropertyValue>scanner_miti/genericJDBC/Drivers</PropertyValue>
</ExecutionContextProperty>
```

3. Enregistrez le fichier `scannerDeployer.xml`.
4. Redémarrez le service de catalogue.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de connexion de la ressource JDBC :

Propriété	Description
Classe de pilote	Nom de classe du pilote JDBC.
URL	Chaîne de connexion permettant de se connecter à la base de données.
Utilisateur	Nom d'utilisateur de la base de données.
Mot de passe	Mot de passe pour le nom d'utilisateur de la base de données.
URL d'agent	Facultatif. URL de l'agent Enterprise Data Catalog qui s'exécute sur un serveur Microsoft Windows. Remarque: assurez-vous de spécifier l'URL au format suivant : <code>http://<hostname>:<connector_port>/MIMBWebServices</code> .

Remarque: Pour extraire des métadonnées de plusieurs schémas d'une source à l'aide du type de ressource JDBC, vous pouvez spécifier des noms de schémas séparés par des points-virgules lorsque vous créez la ressource. Vous pouvez entrer plusieurs noms de schémas dans le champ **Schéma** de la page **Paramètres de chargement des métadonnées**.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Catalogue	Nom de catalogue. Remarque: Vous ne pouvez pas utiliser l'option Catalogue pour les sources JDBC ou ODBC.
Schéma	Spécifie une liste de schémas à importer.
Filtre de métadonnées source	Vous pouvez inclure ou exclure des tables et des vues depuis l'exécution de la ressource. Utilisez des points-virgules (;) pour séparer les noms des tables et les noms des vues. Pour plus d'informations sur le champ de filtre, voir " Filtre de profil de données et métadonnées source " à la page 12.

Propriété	Description
Respect de la casse	Sélectionnez l'une des options suivantes pour spécifier si la base de données doit être sensible à la casse : <ul style="list-style-type: none"> - Auto. Si vous sélectionnez cette option, la ressource utilise l'API JDBC et tente de trouver si la base de données est configurée pour la sensibilité à la casse. Si l'API JDBC n'est pas disponible pour la ressource, la ressource utilise le mode insensible à la casse. - Sensible à la casse. Sélectionnez cette option pour spécifier que la base de données est configurée pour la sensibilité à la casse. - Insensible à la casse. Sélectionnez cette option pour spécifier que la base de données est configurée comme étant insensible à la casse.
Afficher le SQL d'extraction de la définition	Spécifie la requête SQL propre à la base de données pour récupérer le texte de définition de la vue.
SQL de lignage des synonymes	Spécifie la requête SQL propre à la base de données pour récupérer le lignage synonyme. Voici les deux colonnes renvoyées par la requête : <ul style="list-style-type: none"> - Nom complet du synonyme - Nom complet de la table
Étendue facultative	Spécifie les types d'objets de base de données à importer, tels que les tables et les vues, les index et les procédures. Spécifiez la liste des types d'objets de base de données facultatifs à importer. La liste peut ne contenir aucun type d'objet de base de données ou en contenir plusieurs qui seront alors séparés par des points-virgules. Par exemple, les clés, les index et les procédures stockées.
Importer des procédures stockées	Spécifie les procédures stockées à importer. La valeur par défaut est True ou False quel que soit le cas.
Mémoire	Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanneur. Spécifiez l'une des valeurs de mémoire suivantes : <ul style="list-style-type: none"> - Faible - Moyenne - Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none"> -Dscannerloglevel=<DEBUG/INFO/ERROR>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. -Dscanner.container.core=<No. of core>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. -Dscanner.yarn.app.environment=<key=value>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. -Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0>. Augmente la mémoire du conteneur du scanneur lorsque pmem est activé. La valeur par défaut est 1. -DskipExternalDatabaseObjects=true. Exclut tous les objets externes tels que le schéma, la table, la base de données et la vue.

Vous pouvez activer la découverte de données pour une ressource JDBC. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource JDBC. Pour plus d'informations sur l'activation de la découverte de domaines de données, consultez la section [Composite Data Domain Discovery](#).

Microsoft SQL Server

Vous pouvez configurer un type de ressource Microsoft SQL Server pour extraire des métadonnées des bases de données Microsoft SQL Server. Assurez-vous de configurer l'autorisation VIEW DEFINITION pour la base de données SQL et l'autorisation SELECT pour l'élément *sys.sql_expression_dependencies* de la base de données.

Objets extraits

Enterprise Data Catalog extrait les métadonnées des objets de schéma suivants d'une ressource Microsoft SQL Server :

- Tables
- Vues
- Procédure stockée étendue
- Table (définie par l'utilisateur)
- Synonyme
- Contrainte UNIQUE
- Déclencheurs
- Procédures
- Index
- Fonction table SQL
- Utilisateurs
- Schéma de partition
- Fonction de partition
- Collections de schémas XML
- Type de données défini par l'utilisateur
- Règles
- Valeurs par défaut
- Contraintes
- SQL_INLINE_TABLE_VALUED_FUNCTION
- SQL_SCALAR_FUNCTION
- SEQUENCE

Autorisations de configuration de la ressource

Avant de créer une ressource Microsoft SQL Server, configurez les autorisations du compte utilisateur de la base de données Microsoft SQL Server que vous utilisez pour vous connecter à Microsoft SQL Server.

Enterprise Data Catalog utilise l'authentification SQL Server pour se connecter à la base de données Microsoft SQL Server. Le compte utilisateur que vous utilisez pour vous connecter à Microsoft SQL Server doit être un compte de connexion SQL Server. Configurez l'autorisation VIEW DEFINITION pour le compte utilisateur. La configuration de cette autorisation répertorie tous les schémas à partir desquels l'utilisateur peut charger des métadonnées. Vous devez également configurer l'autorisation SELECT pour la valeur *sys.sql_expression_dependencies* de la base de données.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de la ressource Microsoft SQL Server :

Propriété	Description
URL d'agent	URL de l'agent Enterprise Data Catalog qui s'exécute sur un serveur Microsoft Windows. Remarque : assurez-vous de spécifier l'URL au format suivant : <code>http://<hostname>:<connector_port>/MIMBWebServices</code> .
Utilisateur	Nom du compte utilisateur Microsoft SQL Server qui se connecte à la base de données Microsoft SQL Server. Le service de catalogue utilise l'authentification Microsoft SQL Server pour se connecter à la base de données Microsoft SQL Server.
Mot de passe	Mot de passe du compte d'utilisateur qui se connecte à la base de données Microsoft SQL Server.
Hôte	Nom d'hôte de la machine sur laquelle Microsoft SQL Server s'exécute.
Port	Numéro de port pour le service du moteur de base de données Microsoft SQL Server.
Base de données	Nom de la base de données Microsoft SQL Server. Remarque: Le nom de la base de données et le nom de la ressource doivent être différents.
Instance	Nom d'instance de Microsoft SQL Server.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Importer des objets système	Spécifie les objets système à importer. La valeur par défaut est True ou False quel que soit le cas.
Schéma	Spécifie une liste de schémas de base de données séparés par des points-virgules.
Filtre de métadonnées source	Vous pouvez inclure ou exclure des tables et des vues depuis l'exécution de la ressource. Utilisez des points-virgules (;) pour séparer les noms des tables et les noms des vues. Pour plus d'informations sur le champ de filtre, voir " Filtre de profil de données et métadonnées source " à la page 12.
Importer des procédures stockées	Spécifie les procédures stockées à importer. La valeur par défaut est True ou False quel que soit le cas.
Importer les synonymes	Spécifie les synonymes à importer. Cochez cette case pour importer les synonymes. Décochez cette case pour ne pas importer les synonymes. Par défaut, la ressource est configurée pour importer les synonymes.

Propriété	Description
Sensible à la casse	<p>Spécifie que la ressource est configurée pour l'insensibilité à la casse. Sélectionnez l'une des valeurs suivantes :</p> <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. <p>La valeur par défaut est False.</p>
Mémoire	<p>Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanneur.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1. - <code>-DskipExternalDatabaseObjects=true</code>. Exclut tous les objets externes tels que le schéma, la table, la base de données et la vue.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.
Attribution automatique des connexions	Spécifie si les connexions doivent être attribuées automatiquement.
Activer les ressources de référence	Extrait les métadonnées des actifs non inclus dans cette ressource, mais qui y sont référencés. Par exemple, des tables source et cible dans des mappages PowerCenter, et des tables et fichiers source des rapports Tableau.
Conserver les actifs de référence non résolus	Conserve les actifs de référence non résolus dans le catalogue après l'attribution des connexions. La conservation des actifs de référence non résolus vous permet de visualiser l'intégralité du lignage. Les actifs non résolus incluent les fichiers supprimés, les tables temporaires et d'autres actifs qui ne figurent pas dans la ressource principale.
Options d'agent	Spécifiez les options de l'agent Enterprise Data Catalog pour l'exécution de la tâche de scanneur.

Vous pouvez activer la découverte de données pour une ressource Microsoft SQL Server. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#). Vous pouvez activer la découverte de domaines de données composites pour une ressource Microsoft SQL Server. Pour plus d'informations sur l'activation de la découverte de domaines de données, consultez la section [Composite Data Domain Discovery](#).

Configuration de l'agent Enterprise Data Catalog pour la ressource sur un serveur authentifié par Windows

Pour configurer l'agent Enterprise Data Catalog pour Microsoft SQL Server, vous pouvez procéder comme suit :

1. Téléchargez le pilote JDBC SQL Server (version 4.1) dans un emplacement temporaire.
2. Copiez le fichier `sqljdbc<version>.jar` dans l'emplacement `<Enterprise_Data_Catalog_Agent_location>\java\Jdbc\mssql`.
3. Copiez le fichier `\sqljdbc_4.1\enu\auth\x64\sqljdbc_auth.dll` extrait du fichier `sqljdbc_4.1.jar` dans l'emplacement `C:\Windows\System32\`.
4. Configurez l'agent Enterprise Data Catalog pour qu'il s'exécute en tant que service. Pour configurer le service, vous pouvez spécifier l'ID Active Directory utilisé pour accéder à la base de données Microsoft SQL Server comme ID de connexion.
5. Exécutez le fichier `RestartServerService.bat`.
6. Dans Catalog Administrator, créez une ressource Microsoft SQL Server avec l'URL de l'agent Enterprise Data Catalog suivante : `http://<HostName>:<Port>/MIMBWebServices/`
7. Exécutez la ressource.

Affichage du lignage des objets de schéma et de base de données externes

Vous pouvez afficher le lignage détaillé des objets de schéma et de base de données externes dans Enterprise Data Catalog. Pour afficher le lignage des objets de schéma et de base de données externes, utilisez l'option **Attribution de connexion**.

Remarque: Pour l'attribution de connexion, vérifiez que les sources de données cible et de destination qui représentent les objets de schéma et de base de données externes sont disponibles dans Enterprise Data Catalog.

Oracle

Configurez une ressource Oracle pour extraire les métadonnées des bases de données Oracle.

Objets extraits

Enterprise Data Catalog extrait les métadonnées des objets de schéma suivants d'une ressource Oracle :

- Table
- Synonyme
- Afficher
- Types de données différents
- Informations de précision et d'échelle pour tous les types de données
- Commentaires
- Déclencheur
- Fonction
- Procédure
- Condition

- Contraintes d'intégrité
- Table organisée en index
- Vue matérialisée

Autorisations de configuration de la ressource

Vous devez configurer les autorisations CONNECT et SELECT_CATALOG_ROLE pour le compte utilisateur de la base de données que vous utilisez pour accéder à la base de données Oracle.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés du type de ressource Oracle :

Propriété	Description
Utilisateur	Nom du compte utilisateur qui se connecte à la base de données Oracle.
Mot de passe	Mot de passe du compte utilisateur qui se connecte à la base de données Oracle.
Hôte	Nom d'hôte complet de la machine sur laquelle la base de données Oracle est hébergée.
Port	Numéro de port du service de moteur de base de données Oracle.
Service	Identificateur unique du serveur de base de données Oracle. Remarque: Vous ne devez pas utiliser l'ID système (SID) comme identificateur unique du serveur de base de données Oracle.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Importer des objets système	Spécifie les objets système à importer. La valeur par défaut est True ou False quel que soit le cas.
Schéma	Spécifie une liste de schémas de base de données séparés par des points-virgules.
Filtre de métadonnées source	Vous pouvez inclure ou exclure des tables et des vues depuis l'exécution de la ressource. Utilisez des points-virgules (;) pour séparer les noms des tables et les noms des vues. Pour plus d'informations sur le champ de filtre, voir "Filtre de profil de données et métadonnées source" à la page 12.
Importer des procédures stockées	Spécifie les procédures stockées à importer. La valeur par défaut est True ou False quel que soit le cas.
Importer les synonymes privés et publics	Spécifie les synonymes privés et publics à importer. Sélectionnez l'une des valeurs suivantes : - True. Cochez cette case pour importer des synonymes privés et publics. - False. Décochez cette case pour ne pas importer les synonymes privés et publics. La valeur par défaut est False.

Propriété	Description
Sensible à la casse	<p>Spécifie que la ressource est configurée pour l'insensibilité à la casse. Sélectionnez l'une des valeurs suivantes :</p> <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. <p>La valeur par défaut est False.</p>
Importer le nom de la base de données	<p>Extrait le nom de la base de données de la source de données à partir de la variable <code>v\$database</code>. Activez cette option lorsque vous souhaitez extraire des métadonnées de la source de données Oracle Real Application Clusters (Oracle RAC).</p>
Mémoire	<p>Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanneur.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1. - <code>-DskipExternalDatabaseObjects=true</code>. Exclut tous les objets externes tels que le schéma, la table, la base de données et la vue.
Suivre les modifications de la source de données	<p>Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.</p>

Vous pouvez activer la découverte de données pour une ressource Oracle. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource Oracle. Pour plus d'informations sur l'activation de la découverte de domaines de données, consultez la section [Composite Data Domain Discovery](#).

SAP HANA

Créez une ressource selon le type de ressource SAP HANA afin d'extraire les métadonnées d'une source de données SAP HANA. Vous devez spécifier les informations de connexion lorsque vous configurez la ressource.

Objets extraits

Enterprise Data Catalog extrait les métadonnées des objets suivants dans une source de données SAP HANA :

- Table
- Synonyme
- Colonne
- Attribut
- Vue
- Vue Calcul graphique
- Colonne de vue
- Schéma
- Nombre d'objets de fonctions et de procédures stockées

Remarque: La ressource n'extrait pas les informations de lignage pour les procédures stockées et les fonctions.

- Conditions SQL
- Mesures

Autorisations de configuration de la ressource

Vérifiez que vous avez configuré les autorisations suivantes pour le nom d'utilisateur utilisé pour vous connecter à la base de données :

- **SELECT METADATA** dans les schémas pour lesquels vous voulez importer des métadonnées.
- **SELECT** dans la table `_SYS_REPO.ACTIVE_OBJECT` et les schémas système `_SYS_BI` et `_SYS_BIC`.

Prérequis

Téléchargez les fichiers JAR du pilote `ngdbc.jar` depuis Internet et copiez-les dans le répertoire `SAPHANAScanner` du répertoire `<INFA_HOME>/services/CatalogService/ScannerJars/externalDependencies`. Il n'est pas nécessaire de redémarrer le service de catalogue.

Pour respecter les prérequis, vous pouvez également effectuer les étapes suivantes :

1. Téléchargez les fichiers JAR `ngdbc.jar` depuis Internet et incluez les pilotes dans le fichier `saphanaJars.zip`.
2. Copiez le fichier `saphanaJars.zip` dans le répertoire `<INFA_HOME>/services/CatalogService/ScannerBinaries`.
3. Ouvrez le fichier `<INFA_HOME>/services/CatalogService/ScannerBinaries/CustomDeployer/scannerDeployer.xml` et ajoutez-y les lignes suivantes :

```
</ExecutionContextProperty>
  <ExecutionContextProperty isLocationProperty="true"
    dependencyToUnpack="saphanaJars.zip">
    <PropertyName>SAPHANAScanner_DriverLocation</PropertyName>
    <PropertyValue>scanner_miti/SAPHANA/Drivers</PropertyValue>
  </ExecutionContextProperty>
```

4. Enregistrez le fichier `scannerDeployer.xml`.
5. Redémarrez le service de catalogue.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de connexion à la ressource :

Propriété	Description
Utilisateur	Nom d'utilisateur de connexion à la base de données.
Mot de passe	Mot de passe de connexion à la base de données.
Hôte	Nom d'hôte ou adresse IP de la machine qui héberge la base de données.
Port	Numéro de port de la machine qui héberge la base de données. La valeur par défaut est 39013.
Base de données	Nom de l'instance de base de données.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Importer des objets système	Importe les objets système.
Schéma	Spécifiez les schémas pour lesquels vous voulez importer des métadonnées. Si vous ne spécifiez aucun schéma, Enterprise Data Catalog importe tous les schémas dans la base de données. Pour sélectionner les schémas requis, effectuez les étapes suivantes : 1. Cliquez sur Sélectionner . La boîte de dialogue Sélectionner le schéma s'affiche. 2. Sélectionnez la liste de schémas dans la liste Schéma . 3. Cliquez sur Sélectionner à l'aide d'une expression régulière pour spécifier une requête SQL afin d'identifier les noms de schémas qui correspondent au modèle dans la requête.
Filtre de métadonnées source	Spécifiez une liste de tables à inclure ou exclure lorsque la ressource collecte des métadonnées. La valeur par défaut est Tout , ce qui inclut toutes les tables de la base de données. Vous pouvez spécifier une liste de tables à inclure ou exclure en séparant les noms de tables par un point-virgule (;). Vous pouvez utiliser des expressions semblables aux expressions SQL avec des opérateurs standard, tels qu'AND, OR et NOT afin d'inclure ou d'exclure des tables de la liste de tables.
Importer des procédures stockées	Sélectionnez cette option pour importer des objets de la base de données, tels que des procédures stockées et des fonctions.
Sensible à la casse	Sélectionnez cette option pour conserver la sensibilité à la casse pour les métadonnées dans le diagramme de lignage.

Propriété	Description
Package	<p>Spécifiez une liste de noms de packages à importer. Les noms de packages sont sensibles à la casse. Veillez à séparer les noms de packages par un point virgule (;). Pour importer tous les packages, vous pouvez laisser l'option vide.</p> <p>Pour spécifier les noms de sous-packages, veillez à spécifier un nom de sous-package au format suivant : <package name>.<sub package name>.</p>
Mémoire	<p>Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanner.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journal du scanner sur des valeurs spécifiques, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement YARN. Utilisez une virgule pour séparer plusieurs valeurs de paire de clés. - <code>Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1. - <code>-DINFA_TRUSTSTORE=<trust store path></code>. Chemin d'accès au fichier Informatica truststore. - <code>-DINFA_TRUSTSTORE_PASSWORD=<trust store encrypted key></code>. Mot de passe chiffré du truststore.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Sybase

Vous pouvez configurer une ressource Sybase pour extraire les métadonnées des bases de données Sybase.

Objets extraits

Enterprise Data Catalog extrait les métadonnées des objets de schéma suivants à partir d'une ressource Sybase :

- Table
- Afficher
- Index en cluster
- Procédure

- Déclencheurs
- Règle
- Index unique
- Index non en cluster
- Index
- Fonction
- Type de données défini par l'utilisateur
- Schéma
- Colonne
- ForeignKey
- Procédure stockée
- PrimaryKeyConstraint

Autorisations de configuration de la ressource

Avant de créer une ressource Sybase, configurez les autorisations du compte d'utilisateur de la base de données Sybase ASE que vous utilisez pour vous connecter à la base de données Sybase ASE.

Configurez les autorisations SELECT dans les objets de schéma suivants :

- Tables
- Vues
- Colonnes
- Procédures

Configurez les autorisations SELECT dans les tables système suivantes :

- sysobjects
- syscomments
- sysdepends
- syscolumns
- sysindexes
- sysconstraints
- sysreferences
- sysusers
- systypes
- master.dbo.sys.servers
- master.dbo.sys.configures
- @@<servername

Prérequis

Pour respecter les prérequis afin de créer une ressource Sybase, téléchargez le fichier JAR du pilote JDBC depuis Internet, puis copiez-le dans le répertoire `SybaseScanner` du répertoire `<INFA_HOME>/services/CatalogService/ScannerJars/externalDependencies`. Il n'est pas nécessaire de redémarrer le service de catalogue.

Pour respecter les prérequis, vous pouvez également effectuer les étapes suivantes :

1. Téléchargez le fichier du pilote JDBC et copiez-le dans le répertoire <INFA_HOME>/services/CatalogService/ScannerBinaries.
2. Ouvrez le fichier <INFA_HOME>/services/CatalogService/ScannerBinaries/CustomDeployer/scannerDeployer.xml et ajoutez-y les lignes suivantes :

```
</ExecutionContextProperty>
<ExecutionContextProperty isLocationProperty="true"
dependencyToUnpack="sybase_jars.zip">
  <PropertyName>SybaseScanner_DriverLocation</PropertyName>
  <PropertyValue>scanner_miti/sybase_jars/Drivers</PropertyValue>
</ExecutionContextProperty>
```

3. Enregistrez le fichier scannerDeployer.xml.
4. Redémarrez le service de catalogue.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de la ressource Sybase :

Propriété	Description
Hôte	Nom d'hôte de la machine sur laquelle la base de données Sybase est hébergée.
Port	Numéro de port du service de moteur de base de données Sybase.
Utilisateur	Nom d'utilisateur de la base de données.
Mot de passe	Mot de passe du nom d'utilisateur de la base de données.
Base de données	Nom de la base de données.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Schéma	Spécifiez une liste de base de données ou un schéma à importer.
Procédures stockées importées	Spécifie les procédures stockées à importer. La valeur par défaut est True ou False quel que soit le cas.
Sensible à la casse	Spécifie que la ressource est configurée pour l'insensibilité à la casse. Sélectionnez l'une des valeurs suivantes : <ul style="list-style-type: none">- True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse.- False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. La valeur par défaut est False.

Propriété	Description
Mémoire	<p>Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanner.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1. - <code>-DskipExternalDatabaseObjects=true</code>. Exclut tous les objets externes tels que le schéma, la table, la base de données et la vue.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource Sybase. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource Sybase. Pour plus d'informations sur l'activation de la découverte de domaines de données, consultez la section [Composite Data Domain Discovery](#).

Teradata

Teradata est l'un des types de ressources ETL dans Enterprise Data Catalog. Configurez un type de ressource Teradata pour extraire les métadonnées des bases de données Teradata.

Objets extraits

- Utilisateur
- Tables
- Colonne
- Type de données
- Vues
- Index
- IndexColumn

- PrimaryKeyConstraint
- ForeignKey
- UniqueConstraint
- CheckConstraint
- Procédure
- TableTrigger
- Base de données
- Fonctions
- Macros
- Schéma
- Procédures stockées

Autorisations de configuration de la ressource

Configurez les autorisations du compte utilisateur de la base de données Teradata que vous utilisez pour vous connecter à la base de données Teradata.

Le tableau suivant répertorie les objets de schéma et les tables système dans lesquels vous devez configurer les autorisations SELECT pour le compte utilisateur de la base de données Teradata :

Objets de schéma	Tables système
<ul style="list-style-type: none"> - Tables - Vues - Index - Packages - Procédures - Fonctions - Séquences - Déclencheurs - Synonymes - Base de données 	<ul style="list-style-type: none"> - DBC.DBCINFO - DBC.TABLETEXT - DBC.HOSTSINFO - DBC.DATABASES - DBC.TABLES - DBC.COLUMNS - DBC.ALL_RI_CHILDREN - DBC.INDICES - DBC.TRIGGERS - DBC.SHOWTBLCHECKS - DBC.TVFIELDS - DBC.TVM - DBC.UDTINFO - TDQCD.STATSRECS - DBC.DBASE - DBC.IDCOL

Prérequis

Pour respecter les prérequis afin de créer une ressource Teradata, téléchargez le fichier JAR du pilote JDBC depuis Internet, puis copiez-le dans le répertoire `TeradataScanner` du répertoire `<INFA_HOME>/services/CatalogService/ScannerJars/externalDependencies`. Il n'est pas nécessaire de redémarrer le service de catalogue.

Pour respecter les prérequis, vous pouvez également effectuer les étapes suivantes :

1. Téléchargez le fichier du pilote JDBC et copiez-le dans le répertoire `<INFA_HOME>/services/CatalogService/ScannerBinaries`.

2. Ouvrez le fichier `<INFA_HOME>/services/CatalogService/ScannerBinaries/CustomDeployer/scannerDeployer.xml` et ajoutez-y les lignes suivantes :

```
</ExecutionContextProperty>
<ExecutionContextProperty isLocationProperty="true"
dependencyToUnpack="teradataJars.zip">
  <PropertyName>TeradataScanner_DriverLocation</PropertyName>
  <PropertyValue>scanner_miti/teradata/Drivers</PropertyValue>
</ExecutionContextProperty>
```

3. Enregistrez le fichier `scannerDeployer.xml`.
4. Redémarrez le service de catalogue.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés du type de ressource Teradata :

Propriété	Description
Utilisateur	Nom du compte utilisateur permettant de se connecter à la base de données Teradata.
Mot de passe	Mot de passe du compte utilisateur permettant de se connecter à la base de données Teradata.
Hôte	Nom d'hôte complet de la machine sur laquelle la base de données Teradata est hébergée. Remarque: Pour vous connecter à une ressource Teradata via un serveur LDAP, vous devez spécifier <code>/LOGMECH=LDAP</code> sur le nom d'hôte.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées de la source de données.
Importer des objets système	Spécifie les objets système à importer. La valeur par défaut est True ou False quel que soit le cas.
Schéma	Spécifie une liste de schémas de base de données séparés par des points-virgules.
Filtre de métadonnées source	Vous pouvez inclure ou exclure des tables et des vues depuis l'exécution de la ressource. Utilisez des points-virgules (;) pour séparer les noms des tables et les noms des vues. Pour plus d'informations sur le champ de filtre, voir "Filtre de profil de données et métadonnées source" à la page 12 .
Importer des procédures stockées	Spécifie les procédures stockées à importer. La valeur par défaut est True ou False quel que soit le cas.
Sensible à la casse	Spécifie que la ressource est configurée pour l'insensibilité à la casse. Sélectionnez l'une des valeurs suivantes : <ul style="list-style-type: none">- True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse.- False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. La valeur par défaut est False.

Propriété	Description
Récupérer les types de données des vues	Spécifie l'importation du type de données Views.
Mémoire	<p>Spécifiez la valeur de mémoire requise pour exécuter une tâche de scanneur.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1. - <code>-DskipExternalDatabaseObjects=true</code>. Exclut tous les objets externes tels que le schéma, la table, la base de données et la vue.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource Teradata. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource Teradata. Pour plus d'informations sur l'activation de la découverte de domaines de données, consultez la section [Composite Data Domain Discovery](#).

CHAPITRE 10

Configuration des ressources du gestionnaire de fichiers à parties multiples Hadoop

Ce chapitre comprend les rubriques suivantes :

- [Présentation des ressources du gestionnaire de fichiers à parties multiples Hadoop , 182](#)
- [Gestionnaire de fichiers à parties multiples Hadoop, 182](#)

Présentation des ressources du gestionnaire de fichiers à parties multiples Hadoop

Le groupe de ressources du gestionnaire de fichiers à parties multiples Hadoop inclut la ressource du gestionnaire de fichiers à parties multiples Hadoop. La ressource du gestionnaire de fichiers à parties multiples Hadoop gère le lien entre les fichiers à parties multiples Hadoop et leurs cibles correspondantes. Lorsque vous exécutez la ressource, le scanner du gestionnaire de fichiers à parties multiples Hadoop extrait et migre les résultats de métadonnées dans la ressource associée vers le catalogue. Vous pouvez afficher les résultats migrés dans Enterprise Data Catalog.

Gestionnaire de fichiers à parties multiples Hadoop

Utilisez la ressource du gestionnaire de fichiers à parties multiples Hadoop pour extraire le lignage d'un fichier à parties multiples combinées vers une cible relationnelle.

Prérequis

Pour respecter les prérequis, effectuez les étapes suivantes :

- Dans le service d'intégration de données où les mappages Data Engineering sont déployés, définissez la valeur de la propriété personnalisée **HDFSRetainOriginalTargetFile** sur **True**.
- Configurez et exécutez les sources de métadonnées relationnelles et HDFS qui sont utilisées pour créer les mappages Data Engineering.

- Configurez et exécutez la ressource de métadonnées Informatica Platform.
- S'agissant des sources de métadonnées relationnelles et HDFS requises, attribuez les connexions à l'aide de l'option **Attribution de connexion** dans Catalog.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Nom de la ressource HDFS	Nom de la ressource HDFS.
Nom de la ressource relationnelle source	Nom de la source relationnelle.
Nom de la ressource relationnelle cible	Nom de la source relationnelle cible.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Extrait les métadonnées de la source de données.
Mémoire	<p>Sélectionnez la mémoire requise pour exécuter le travail du scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres de JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.

CHAPITRE 11

Configuration des ressources de gestion de fichiers

Ce chapitre comprend les rubriques suivantes :

- [Présentation des ressources de gestion des fichiers, 184](#)
- [Système de fichiers, 184](#)
- [OneDrive, 188](#)
- [Microsoft SharePoint, 192](#)

Présentation des ressources de gestion des fichiers

Le groupe de ressources de gestion des fichiers inclut toutes les ressources associées à la gestion des fichiers. Lorsque vous créez et exécutez une ressource de gestion des fichiers, le scanner de gestion des fichiers extrait et migre les résultats de métadonnées de la ressource associée vers le catalogue. Vous pouvez afficher les résultats migrés dans Enterprise Data Catalog.

Vous pouvez créer les ressources de gestion des fichiers suivantes dans Enterprise Data Catalog :

- Système de fichiers
- OneDrive
- Microsoft SharePoint

Système de fichiers

Vous pouvez choisir la ressource **Système de fichiers** pour importer les métadonnées des fichiers situés sur les systèmes de fichiers Windows et Linux dans le catalogue. Vous pouvez choisir l'option **Fichier local** pour importer les métadonnées des fichiers situés sur une machine locale. Vous pouvez choisir l'option **SFTP** ou Secure File Transfer Protocol pour importer les fichiers situés sur une machine Linux distante et choisir l'option **SMB/CIFS** ou Server Message Block/Common Internet File System pour importer les fichiers situés sur une machine Windows distante.

Propriétés de connexion à la ressource

Dans l'onglet **Général**, vous pouvez entrer un nom et une brève description de la ressource. Dans la section **Propriétés supplémentaires**, choisissez **Système de fichiers** comme ressource.

Choisissez les options suivantes dans la section **Propriétés de connexion** pour importer les métadonnées des fichiers situés sur la machine locale :

Propriété	Description
Protocole de fichier	Choisissez Fichier local pour analyser les fichiers situés sur la machine Windows ou Linux locale.
Chemin d'accès	Indiquez le chemin d'accès absolu du fichier à partir duquel vous voulez importer les métadonnées dans le catalogue. Assurez-vous que le chemin absolu que vous entrez dans ce champ est disponible sur la machine où vous avez installé Enterprise Data Catalog et tous les nœuds de cluster.
Tester la connexion	Cliquez sur Tester la connexion pour vérifier la connexion à l'emplacement spécifié.

Choisissez les options suivantes dans la section **Propriétés de connexion** pour importer les métadonnées des fichiers situés sur une machine Linux :

Propriété	Description
Protocole de fichier	Choisissez SFTP pour analyser les fichiers sur une machine Linux distante.
Nom d'utilisateur	Entrez le nom d'utilisateur pour accéder à la machine Linux.
Mot de passe	Entrez le mot de passe pour accéder à la machine Linux.
Hôte	Spécifiez le nom d'hôte ou l'adresse IP de la machine Linux.
Port	Spécifiez le numéro de port sur lequel le protocole de fichier SFTP est configuré sur la machine. La valeur par défaut est 22.
Chemin d'accès	Spécifiez le réseau ou le chemin partagé du fichier à partir duquel vous voulez importer les métadonnées dans le catalogue.
Tester la connexion	Cliquez sur Tester la connexion pour vérifier la connexion à l'emplacement spécifié.

Choisissez les options suivantes dans la section **Propriétés de connexion** pour importer les métadonnées des fichiers situés sur une machine Windows :

Propriété	Description
Protocole de fichier	Choisissez SMB/CIFS pour analyser les fichiers situés sur une machine Windows distante.
Nom d'utilisateur	Entrez le nom d'utilisateur pour accéder à la machine Windows.
Mot de passe	Entrez le mot de passe pour accéder à la machine Windows.

Propriété	Description
Emplacement physique	<p>S'applique aux fichiers plats. Entrez le chemin d'accès absolu du répertoire partagé sur la machine Windows.</p> <p>Supposons, par exemple, que les fichiers se trouvent dans les dossiers <code>c:\user1\1021\SMB\test_files\test1</code> et <code>c:\user1\1021\SMB\test_files\test2</code>. Vous mappez l'emplacement <code>c:\user1\1021\SMB</code> au répertoire partagé <code>\\SMB\</code>. Dans ce scénario, entrez <code>\\SMB\</code> dans le champ Emplacement physique.</p> <p>Pour plus d'informations sur la façon dont Enterprise Data Catalog utilise le champ Emplacement physique pour le lignage, voir l'Annexe F - Lignage entre les ressources PowerCenter et les fichiers plats.</p>
Hôte	Spécifiez le nom d'hôte ou l'adresse IP de la machine.
Chemin d'accès	<p>Spécifiez le réseau ou le chemin partagé du fichier à partir duquel vous voulez importer les métadonnées dans le catalogue.</p> <p>Par exemple, entrez <code>\\SMB\test_files\</code> dans le champ Chemin.</p>
Tester la connexion	Cliquez sur Tester la connexion pour vérifier la connexion à l'emplacement spécifié.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Extrait et ingère les métadonnées de la source de données.
Types de fichiers	<p>Sélectionnez un ou tous les types de fichiers suivants à partir desquels vous voulez extraire les métadonnées :</p> <ul style="list-style-type: none"> - Tous. Utilisez cette option pour spécifier si vous voulez extraire des métadonnées de tous les types de fichiers. - Sélectionner. Utilisez cette option pour spécifier que vous voulez extraire des métadonnées à partir de types de fichiers spécifiques. Effectuez les étapes suivantes pour spécifier les types de fichiers : <ol style="list-style-type: none"> 1. Cliquez sur Sélectionner. La boîte de dialogue Sélectionner des types de fichiers spécifiques s'affiche. 2. Sélectionnez les fichiers requis parmi les options suivantes : <ul style="list-style-type: none"> - Formats non structurés étendus Utilisez cette option pour extraire les métadonnées des types de fichiers tels que les fichiers audio, vidéo, image et les ebooks. - Types de fichiers structurés. Utilisez cette option pour extraire les métadonnées des types de fichiers tels que JSON, XML, texte et les fichiers délimités. - Types de fichiers non structurés. Utilisez cette option pour extraire des métadonnées à partir de types de fichiers tels que Microsoft Excel, Microsoft PowerPoint, Microsoft Word, pages Web, fichiers compressés, e-mails et PDF. 3. Cliquez sur Sélectionner. <p>Remarque: Vous pouvez sélectionner l'option Types de fichiers spécifiques dans la boîte de dialogue pour sélectionner les fichiers de toutes les catégories.</p>
Traiter les fichiers sans extension en tant que	<p>Sélectionnez l'une des options suivantes pour identifier les fichiers sans extension :</p> <ul style="list-style-type: none"> - Aucun - Avro - Parquet

Propriété	Description
Entrer le délimiteur de fichier	<p>Vous pouvez spécifier l'un des délimiteurs suivants :</p> <ul style="list-style-type: none"> - Délimiteur à caractère unique. Spécifiez le délimiteur de fichier si le fichier à partir duquel vous extrayez les métadonnées utilise un délimiteur autre que les délimiteurs suivants : <ul style="list-style-type: none"> - Virgule (,) - Tabulation horizontale (\t) - Point-virgule (;) - Deux points (:) - Barre verticale () <p>Assurez-vous de placer le délimiteur entre des guillemets simples. Par exemple : '\$ '. Utilisez une virgule pour séparer plusieurs délimiteurs. Par exemple : '\$ ', '% ', '& '</p> <ul style="list-style-type: none"> - Délimiteur à plusieurs caractères. Spécifiez plusieurs caractères comme délimiteur en les plaçant entre des guillemets simples pour les fichiers CSV. Par exemple : ' \$# '. <p>Si vous spécifiez un délimiteur à plusieurs caractères, n'indiquez aucun autre délimiteur pour le scanneur. Lorsque vous exécutez une ressource de système de fichiers avec un délimiteur à plusieurs caractères, le scanneur considère les fichiers contenant le délimiteur comme des fichiers CSV valides et les fichiers restants comme des fichiers non structurés lors du traitement des fichiers.</p> <p>Les ressources suivantes ne prennent pas en charge le délimiteur à plusieurs caractères pour les fichiers CSV dans Enterprise Data Catalog :</p> <ul style="list-style-type: none"> - Ressource de système de fichiers qui utilise le protocole SFTP ou SMB/CIFS. - Amazon S3 - OneDrive - SharePoint - Microsoft Azure Blob Storage - Azure Data Lake Store
Autres types de fichiers	Extrait les métadonnées de fichier de base telles que la taille du fichier, le chemin d'accès au fichier et les informations d'horodatage provenant d'autres types de fichiers.
Répertoire de premier niveau	<p>Spécifiez un répertoire ou une liste de répertoires sous le répertoire source. Si vous ne spécifiez pas cette option, Enterprise Data Catalog importe tous les fichiers du répertoire source spécifié.</p> <p>Pour spécifier un répertoire ou une liste de répertoires, vous pouvez effectuer les étapes suivantes :</p> <ol style="list-style-type: none"> 1. Cliquez sur Sélectionner.... La boîte de dialogue Sélectionner répertoire de premier niveau s'affiche. 2. Sélectionnez les répertoires requis à l'aide de l'une des options suivantes : <ul style="list-style-type: none"> - Sélectionner dans la liste : sélectionnez les répertoires requis à partir d'une liste de répertoires. - Sélectionner à l'aide d'une expression régulière : fournissez une expression régulière SQL pour sélectionner des schémas qui correspondent à l'expression. <p>Remarque: Si vous sélectionnez plusieurs répertoires, vous devez les séparer à l'aide d'un point-virgule (;).</p>
Inclure le sous-répertoire	Sélectionnez cette option pour importer tous les fichiers des sous-répertoires du répertoire source.
Sensible à la casse	<p>Spécifie que la ressource est configurée pour la sensibilité à la casse. Sélectionnez l'une des valeurs suivantes :</p> <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. <p>La valeur par défaut est True.</p>

Propriété	Description
Mémoire	<p>Sélectionnez la mémoire requise pour exécuter le travail du scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource du **système de fichiers**. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource du système de fichiers. Pour plus d'informations sur l'activation de la découverte de domaines de données, consultez la section [Composite Data Domain Discovery](#).

OneDrive

OneDrive est un service d'hébergement et de synchronisation de fichiers géré par Microsoft dans le cadre de sa suite de services bureautiques en ligne.

Objets extraits

Enterprise Data Catalog n'extrait que les fichiers de la source OneDrive.

Autorisations de configuration de la ressource

Configurez l'autorisation `read` dans la source de données Microsoft OneDrive pour le compte utilisateur permettant d'accéder à la source de données.

Types de fichiers pris en charge

La ressource OneDrive permet d'extraire les métadonnées de fichiers structurés, non structurés et non structurés étendus.

Fichiers structurés pris en charge :

- Fichiers AVRO
- Fichiers délimités
- Fichiers texte
- Fichiers JSON
- Fichiers Parquet
- Fichiers XML

Fichiers non structurés pris en charge :

- Fichiers Apple
- Fichiers compressés
- E-mail

Fichiers non structurés étendus :

- Fichiers VB
- Fichiers ASP
- Fichiers TIF
- Fichiers LOG
- Fichiers CSS
- Fichiers ASPX
- Fichiers DLL
- Fichiers GIF
- Fichiers SQL

Attribuez l'autorisation `read` aux fichiers pour l'extraction des métadonnées.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de la connexion :

Propriété	Description
URL OneDrive	URL permettant d'accéder à OneDrive.
Nom d'utilisateur	Nom d'utilisateur pour la connexion à OneDrive.
Mot de passe	Mot de passe pour la connexion à OneDrive
Répertoire source	Répertoire source à partir duquel tous les fichiers sont analysés.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Extrait et ingère les métadonnées de la source de données.
Types de fichiers	<p>Sélectionnez un ou tous les types de fichiers suivants à partir desquels vous voulez extraire les métadonnées :</p> <ul style="list-style-type: none"> - Tous. Utilisez cette option pour spécifier si vous voulez extraire des métadonnées de tous les types de fichiers. - Sélectionner. Utilisez cette option pour spécifier que vous voulez extraire des métadonnées à partir de types de fichiers spécifiques. Effectuez les étapes suivantes pour spécifier les types de fichiers : <ul style="list-style-type: none"> 1. Cliquez sur Sélectionner. La boîte de dialogue Sélectionner des types de fichiers spécifiques s'affiche. 2. Sélectionnez les fichiers requis parmi les options suivantes : <ul style="list-style-type: none"> - Formats non structurés étendus Utilisez cette option pour extraire les métadonnées des types de fichiers tels que les fichiers audio, vidéo, image et les ebooks. - Types de fichiers structurés. Utilisez cette option pour extraire des métadonnées à partir de types de fichiers tels que Avro, Parquet, JSON, XML, fichiers texte et fichiers délimités. - Types de fichiers non structurés. Utilisez cette option pour extraire des métadonnées à partir de types de fichiers tels que Apple, Microsoft Excel, Microsoft PowerPoint, Microsoft Word, pages Web, fichiers compressés, e-mails et PDF. 3. Cliquez sur Sélectionner. <p>Remarque: Vous pouvez sélectionner l'option Types de fichiers spécifiques dans la boîte de dialogue pour sélectionner les fichiers de toutes les catégories.</p>
Traiter les fichiers sans extension As	<p>Sélectionnez l'une des options suivantes pour identifier les fichiers sans extension :</p> <ul style="list-style-type: none"> - Aucun - Avro - Parquet
Entrer le délimiteur de fichier	<p>Spécifiez le délimiteur de fichier si le fichier à partir duquel vous extrayez des métadonnées utilise un délimiteur autre qu'un délimiteur de la liste suivante :</p> <ul style="list-style-type: none"> - Virgule (,) - Tabulation horizontale (\t) - Point-virgule (;) - Deux points (:) - Barre verticale () <p>Assurez-vous de placer le délimiteur entre des guillemets simples. Par exemple : '\$ '. Utilisez une virgule pour séparer plusieurs délimiteurs. Par exemple : '\$ ', '% ', '& '</p>
Autres types de fichiers	Extrait les métadonnées de fichier de base telles que la taille du fichier, le chemin d'accès au fichier et les informations d'horodatage provenant d'autres types de fichiers.

Propriété	Description
Répertoire de premier niveau	<p>Spécifiez un répertoire ou une liste de répertoires sous le répertoire source. Si vous ne spécifiez pas cette option, Enterprise Data Catalog importe tous les fichiers du répertoire source spécifié. Pour spécifier un répertoire ou une liste de répertoires, vous pouvez effectuer les étapes suivantes :</p> <ol style="list-style-type: none"> 1. Cliquez sur Sélectionner.... La boîte de dialogue Sélectionner répertoire de premier niveau s'affiche. 2. Sélectionnez les répertoires requis à l'aide de l'une des options suivantes : <ul style="list-style-type: none"> - Sélectionner dans la liste : sélectionnez les répertoires requis à partir d'une liste de répertoires. - Sélectionner à l'aide d'une expression régulière : fournissez une expression régulière SQL pour sélectionner des schémas qui correspondent à l'expression. <p>Remarque: Si vous sélectionnez plusieurs répertoires, vous devez les séparer à l'aide d'un point-virgule (;).</p>
Inclure le sous-répertoire	Sélectionnez cette option pour importer tous les fichiers des sous-répertoires du répertoire source.
Sensible à la casse	<p>Spécifie que la ressource est configurée pour la sensibilité à la casse. Sélectionnez l'une des valeurs suivantes :</p> <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. <p>La valeur par défaut est True.</p>
Mémoire	<p>Sélectionnez la mémoire requise pour exécuter le travail du scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. - <code>Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanner lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource OneDrive. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource OneDrive. Pour plus d'informations sur l'activation de la découverte de domaines de données, consultez la section [Composite Data Domain Discovery](#).

Microsoft SharePoint

SharePoint est une plate-forme collaborative Web qui s'intègre à Microsoft Office. SharePoint est essentiellement un système de gestion et de stockage des documents configurable, le cas échéant.

Objets extraits

Enterprise Data Catalog extrait uniquement les fichiers de la source de données Microsoft SharePoint.

Autorisations de configuration de la ressource

Configurez l'autorisation `read` dans la source de données Microsoft SharePoint pour le compte utilisateur permettant d'accéder à la source de données.

Types de fichiers pris en charge

La ressource Microsoft SharePoint permet d'extraire les métadonnées de fichiers structurés, non structurés et non structurés étendus.

Fichiers structurés pris en charge :

- Fichiers AVRO
- Fichiers délimités
- Fichiers texte
- Fichiers JSON
- Fichiers Parquet
- Fichiers XML

Fichiers non structurés pris en charge :

- Fichiers Apple
- Fichiers compressés
- E-mail

Fichiers non structurés étendus :

- Fichiers VB
- Fichiers ASP
- Fichiers TIF
- Fichiers LOG
- Fichiers CSS
- Fichiers ASPX
- Fichiers DLL
- Fichiers GIF
- Fichiers SQL

Attribuez l'autorisation `read` aux fichiers pour l'extraction des métadonnées.

Propriétés de connexion à la ressource

Le tableau suivant décrit les propriétés de la connexion :

Propriété	Description
URL SharePoint	URL permettant d'accéder à SharePoint.
Nom d'utilisateur	Nom d'utilisateur permettant d'accéder à SharePoint.
Mot de passe	Mot de passe permettant d'accéder à SharePoint.
Type de contenu SharePoint	Spécifie le type de contenu SharePoint. Sélectionnez l'un des types de contenu SharePoint suivants : <ul style="list-style-type: none">- Tous. Analyse le contenu des listes et des bibliothèques SharePoint.- Liste SharePoint. Analyse le contenu des listes SharePoint.- Bibliothèque SharePoint. Analyse le contenu des bibliothèques SharePoint.
Activer l'analyse du sous-site	Sélectionnez cette option pour analyser des sous-sites à partir du site SharePoint.
Inclure les sous-sites imbriqués	Spécifiez cette option pour analyser les sous-sites imbriqués dans le sous-site de niveau supérieur. Cette option s'applique lorsque vous sélectionnez l'option Activer l'analyse du sous-site.

Le tableau suivant décrit les propriétés **supplémentaires** et **avancées** pour les paramètres de métadonnées source dans l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Extrait et ingère les métadonnées de la source de données.
Types de fichiers	<p>Sélectionnez un ou tous les types de fichiers suivants à partir desquels vous voulez extraire les métadonnées :</p> <ul style="list-style-type: none">- Tous. Utilisez cette option pour spécifier si vous voulez extraire des métadonnées de tous les types de fichiers.- Sélectionner. Utilisez cette option pour spécifier que vous voulez extraire des métadonnées à partir de types de fichiers spécifiques. Effectuez les étapes suivantes pour spécifier les types de fichiers :<ol style="list-style-type: none">1. Cliquez sur Sélectionner. La boîte de dialogue Sélectionner des types de fichiers spécifiques s'affiche.2. Sélectionnez les fichiers requis parmi les options suivantes :<ul style="list-style-type: none">- Formats non structurés étendus Utilisez cette option pour extraire les métadonnées des types de fichiers tels que les fichiers audio, vidéo, image et les ebooks.- Types de fichiers structurés. Utilisez cette option pour extraire des métadonnées à partir de types de fichiers tels que Avro, Parquer, JSON, XML, fichiers texte et fichiers délimités.- Types de fichiers non structurés. Utilisez cette option pour extraire des métadonnées à partir de types de fichiers tels que Apple, Microsoft Excel, Microsoft PowerPoint, Microsoft Word, pages Web, fichiers compressés, e-mails et PDF.3. Cliquez sur Sélectionner. <p>Remarque: Vous pouvez sélectionner l'option Types de fichiers spécifiques dans la boîte de dialogue pour sélectionner les fichiers de toutes les catégories.</p>

Propriété	Description
Traiter les fichiers sans extension As	Sélectionnez l'une des options suivantes pour identifier les fichiers sans extension : <ul style="list-style-type: none"> - Aucun - Avro - Parquet
Entrer le délimiteur de fichier	Spécifiez le délimiteur de fichier si le fichier à partir duquel vous extrayez des métadonnées utilise un délimiteur autre qu'un délimiteur de la liste suivante : <ul style="list-style-type: none"> - Virgule (,) - Tabulation horizontale (\t) - Point-virgule (;) - Deux points (:) - Barre verticale () Assurez-vous de placer le délimiteur entre des guillemets simples. Par exemple : '\$ '. Utilisez une virgule pour séparer plusieurs délimiteurs. Par exemple : '\$ ', '% ', '& '
Autres types de fichiers	Extrait les métadonnées de fichier de base telles que la taille du fichier, le chemin d'accès au fichier et les informations d'horodatage provenant d'autres types de fichiers.
Répertoire de premier niveau	Spécifiez un répertoire ou une liste de répertoires sous le répertoire source. Si vous ne spécifiez pas cette option, Enterprise Data Catalog importe tous les fichiers du répertoire source spécifié. Pour spécifier un répertoire ou une liste de répertoires, vous pouvez effectuer les étapes suivantes : 1. Cliquez sur Sélectionner.... La boîte de dialogue Sélectionner répertoire de premier niveau s'affiche. 2. Sélectionnez les répertoires requis à l'aide de l'une des options suivantes : <ul style="list-style-type: none"> - Sélectionner dans la liste : sélectionnez les répertoires requis à partir d'une liste de répertoires. - Sélectionner à l'aide d'une expression régulière : fournissez une expression régulière SQL pour sélectionner des schémas qui correspondent à l'expression. Remarque: Si vous sélectionnez plusieurs répertoires, vous devez les séparer à l'aide d'un point-virgule (;).
Inclure le sous-répertoire	Sélectionnez cette option pour importer tous les fichiers des sous-répertoires du répertoire source. Remarque: Cette option est obligatoire pour analyser des fichiers à partir de SharePoint.
Sensible à la casse	Spécifie que la ressource est configurée pour la sensibilité à la casse. Sélectionnez l'une des valeurs suivantes : <ul style="list-style-type: none"> - True. Cochez cette case pour spécifier que la ressource est configurée comme étant sensible à la casse. - False. Décochez cette case pour spécifier que la ressource est configurée comme étant insensible à la casse. La valeur par défaut est True.
Mémoire	Sélectionnez la mémoire requise pour exécuter le travail du scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé : <ul style="list-style-type: none"> - Faible - Moyenne - Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .

Propriété	Description
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, ERROR ou INFO. La valeur par défaut est INFO. - <code>Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur. - <code>Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de données pour une ressource SharePoint. Pour plus d'informations sur l'activation de la découverte de données, consultez la section [Enable Data Discovery](#).

Vous pouvez activer la découverte de domaines de données composites pour une ressource SharePoint. La ressource Microsoft SharePoint opère sur un site par ressource. Pour plus d'informations sur l'activation de la découverte de domaines de données composites, consultez la section [Composite Data Domain Discovery](#).

CHAPITRE 12

Configuration des ressources de source de diffusion

Ce chapitre comprend les rubriques suivantes :

- [, 196](#)
- [Apache Kafka, 196](#)

Le groupe de la source de diffusion inclut la ressource Apache Kafka. Lorsque vous créez et exécutez la ressource, le scanneur Apache Kafka extrait et migre les résultats de la source de données Apache Kafka vers le catalogue. Vous pouvez afficher les résultats migrés dans Enterprise Data Catalog.

Apache Kafka

Apache Kafka est une plate-forme de diffusion des données distribuée open-source qui fournit un débit élevé et unifié ainsi qu'un environnement à latence faible pour le traitement des données en temps réel.

Objets extraits

La ressource Apache Kafka extrait les métadonnées des détails du schéma des messages publiés dans les rubriques Kafka d'une source de données Apache Kafka.

Connexion à une source de données Apache Kafka compatible SSL

Pour connecter une source de données Apache Kafka à un registre de schéma compatible SSL, procédez comme suit :

1. Téléchargez les certificats SSL à partir du registre de schéma pour Apache Kafka à l'aide d'un navigateur Web.
Remarque: Veillez à importer le certificat compatible SSL du registre de schéma pour Apache Kafka dans le répertoire Certificats.
2. Copiez les certificats dans le répertoire `<INFA_HOME>/services/shared/security/`.

3. Accédez au répertoire `<INFA_HOME>/source/java/jre/bin`, puis exécutez la commande `keytool` pour importer chaque certificat copié en tant que certificat approuvé dans le fichier keystore du domaine Informatica :

```
keytool -import -file <INFA_HOME>/services/shared/security/<certificate>.cer -alias  
<alias name> -keystore <INFA_HOME>/services/shared/security/infa_truststore.jks -  
storepass <Informatica domain keystore password>
```

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Plate-forme Confluent	Indique si la plate-forme sélectionnée est une plate-forme de données de flux.
URL du Registre de schéma	Une URL pour accéder au Registre de schéma. La syntaxe de l'URL est : <code>http://host1:port1</code> Remarque: Si le Registre de schéma est activé pour SSL, le certificat public du Registre de schéma Confluent doit être importé dans le truststore Informatica.
Liste de brokers	Spécifiez la liste de brokers au format suivant : <code><host number>:<port number></code> . Pour spécifier plusieurs brokers, utilisez une virgule comme délimiteur.
Protocole de sécurité	Protocole de sécurité pour les communications avec les brokers. Les valeurs valides sont notamment : PLAINTEXT, SSL, SASL_PLAINTEXT et SASL_SSL.
Chemin du fichier truststore SSL	Emplacement du fichier <code>infa_truststore</code> par défaut avec le nom du fichier, à partir du nœud de domaine. Chemin par défaut : <code>INFA_HOME/services/shared/security/infa_truststore.jks</code>
Mot de passe du truststore SSL	Mot de passe par défaut du truststore SSL.
Chemin du fichier keystore SSL	Emplacement du fichier <code>infa_keystore</code> par défaut avec le nom du fichier, à partir du nœud de domaine. Chemin par défaut : <code>INFA_HOME/services/shared/security/infa_keystore.jks</code>
Mot de passe du keystore SSL	Mot de passe par défaut du keystore SSL.
Mot de passe de la clé SSL	Mot de passe de la clé SSL.
Mécanisme d'authentification SASL	Mécanisme SASL utilisé pour les connexions client. Spécifiez GSSAPI pour Kerberos et PLAIN pour une authentification utilisateur simple.
Configuration JAAS de SASL	Spécifiez les paramètres de contexte de connexion JAAS pour les connexions SASL au format pris en charge par le fichier de configuration JAAS. Le format utilisé est <code>loginModuleClass controlFlag (optionName=optionValue)*;</code> Pour les brokers, le préfixe d'écoute doit être utilisé comme préfixe de la configuration JAAS et le nom du mécanisme SASL doit être en minuscules.

Propriété	Description
Chemin du fichier de configuration Kerberos	Chemin complet du fichier de configuration Kerberos si vous utilisez l'authentification Kerberos pour la ressource.
Délai d'expiration de l'API en millisecondes	Durée d'expiration pour l'API Apache Kafka. Spécifiez le délai d'expiration en millisecondes.

Le tableau suivant décrit les propriétés que vous pouvez configurer dans la section **Métadonnées source** de l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Active l'extraction des métadonnées..
Noms ou modèles de rubriques	<p>Liste des noms de rubrique ou des modèles de caractère générique qui doivent être inclus ou exclus de l'analyse des métadonnées. Une liste vide indique que toutes les rubriques disponibles sont incluses dans l'analyse des métadonnées. Utilisez une virgule pour séparer les noms de rubrique.</p> <p>Par exemple, vous pouvez utiliser les expressions régulières ou les modèles de métadonnées suivants pour inclure ou exclure les rubriques de l'analyse des métadonnées :</p> <ul style="list-style-type: none"> - <code>^(abc).*</code> Sélectionne les rubriques ou les modèles qui commencent par abc. - <code><Topic name>_[^de]<Topic name></code>. Sélectionne les rubriques qui ne contiennent ni d ni e dans l'exemple de chaîne. - <code><Topic name>[iJK]<Topic name>\d</code>. Sélectionne les rubriques qui comprennent les caractères i, J et K, qui se terminent par un chiffre dans l'exemple de chaîne. <p>Remarque:</p> <ul style="list-style-type: none"> - Si les rubriques sélectionnées incluent du contenu XML ou JSON incorrect, le schéma inféré par la ressource peut être incorrect. - Si la rubrique sélectionnée inclut tous les types de messages (XML, JSON, CSV, etc.), le schéma inféré par la ressource peut être incorrect. - Si la rubrique sélectionnée n'a aucun contenu, il est impossible d'afficher les détails du schéma dans la ressource.
Stratégie d'interrogation	<p>Sélectionnez l'une des stratégies d'interrogation :</p> <ul style="list-style-type: none"> - Interroger à partir du début - Interroger à partir de la fin
Mémoire	<p>Valeur de mémoire requise pour exécuter une tâche de scanneur.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>

Propriété	Description
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanner. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> -Dscannerloglevel=<DEBUG/INFO/ERROR>. Définit le niveau de journalisation du scanner sur des valeurs, telles que DEBUG, INFO ou ERROR. La valeur par défaut est INFO. -Dscanner.container.core=<No. of core>. Augmente le nombre de cœurs disponibles pour le conteneur du scanner. La valeur doit être un nombre. -Dscanner.yarn.app.environment=<key=value>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur pour les paires clé-valeur multiples. -Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0>. Augmente la mémoire du conteneur du scanner lorsque pmem est activé. La valeur par défaut est 1. -Djava.security.krb5.conf=<Path of the Kerberos configuration file>. Extraie les métadonnées de la source de données activée pour Kerberos lorsque le protocole de sécurité est SASL.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Remarque:

- La ressource Apache Kafka dispose d'une taille d'échantillonnage de 30 messages. La taille d'échantillonnage ne peut pas être modifiée.
- La ressource Apache Kafka prend en charge un lignage externe avec des mappages de diffusion.

Exécution d'une ressource Apache Kafka sur un cluster compatible Kerberos

Si vous voulez exécuter une ressource Apache Kafka sur un cluster compatible Kerberos, procédez comme suit :

1. Copiez le fichier krb.conf à l'emplacement suivant : <Install Directory>/data/ldmbcmev/Informatica/LDM20_309/source/services/shared/security/krb5.conf
2. Copiez le fichier krb.conf à l'emplacement /etc sur tous les clusters où le service de catalogue est en cours d'exécution.
3. Copiez le fichier keytab dans le répertoire /opt dans les emplacements suivants :
 - Emplacement commun pour tous les clusters sur lesquels le service de catalogue est en cours d'exécution.
 - Machine du domaine.
 - Machine du cluster Kerberos.
4. Ajoutez les informations de la machine de l'hôte KDC à l'emplacement etc/hosts de la machine du domaine et de celle du cluster sur laquelle le service de catalogue est en cours d'exécution.

CHAPITRE 13

Configuration des ressources NoSQL

Ce chapitre comprend les rubriques suivantes :

- [Présentation des ressources NoSQL, 200](#)
- [Apache Cassandra, 200](#)
- [MongoDB, 203](#)

Présentation des ressources NoSQL

Le groupe de ressources NoSQL inclut les ressources NoSQL. Lorsque vous créez et exécutez une ressource NoSQL, le scanneur NoSQL extrait et migre les métadonnées de la ressource associée vers le catalogue. Vous pouvez afficher les résultats migrés dans Enterprise Data Catalog.

Apache Cassandra

Apache Cassandra est un espace de stockage en colonnes open source distribué et un système de gestion de base de données NoSQL conçu pour gérer de grandes quantités de données sur plusieurs serveurs de commodités. Il fournit une haute disponibilité des données sans défaillance.

Objets extraits

La ressource Apache Cassandra extrait les métadonnées des actifs suivants dans une source de données Apache Cassandra :

- Espace de clés
- Tables
- Vues
- Fonction
- Champs

Autorisations de configuration de la ressource

Configurez l'autorisation `read` dans la source de données Apache Cassandra pour le compte utilisateur permettant d'accéder à la source de données.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Hôte	Nom d'hôte ou adresse IP du serveur Apache Cassandra.
Mot de passe du keystore	Mot de passe pour accéder au keystore Apache Cassandra
Chemin du keystore	Emplacement du keystore Apache Cassandra.
Centre de données local	Nom du centre de données qui contient le nœud requis.
Mot de passe	Mot de passe du compte utilisateur pour accéder au serveur Apache Cassandra.
Port	Numéro de port du serveur Apache Cassandra. Le numéro de port par défaut est 9042.
SSL activé	Option permettant d'activer SSL
Mot de passe du fichier truststore	Mot de passe pour accéder au truststore Apache Cassandra.
Chemin du truststore	Emplacement du trustore Apache Cassandra.
Nom d'utilisateur	Nom d'utilisateur pour accéder au serveur Apache Cassandra.

Le tableau suivant décrit les propriétés que vous pouvez configurer dans la section **Métadonnées source** de l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Active l'extraction des métadonnées..
Sensible à la casse	Spécifie que la ressource est configurée pour l'insensibilité à la casse. Sélectionnez l'une des valeurs suivantes : True. Cochez cette case pour indiquer que la ressource est configurée comme étant sensible à la casse. False. Décochez cette case pour indiquer que la ressource est configurée comme étant insensible à la casse. La valeur par défaut est True.
Espace clé	Option d'importation d'un schéma de la base de données particulier lorsque vous spécifiez une liste d'espaces clés à l'aide de l'API Parcourir.

Propriété	Description
Mémoire	<p>Valeur de mémoire requise pour exécuter une tâche de scanneur.</p> <p>Spécifiez l'une des valeurs de mémoire suivantes :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	<p>Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres :</p> <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, INFO ou ERROR. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur pour les paires clé-valeur multiples. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Vous pouvez activer la découverte de domaine pour une ressource Cassandra. Pour plus d'informations, consultez les sections [Enable Data Discovery](#) et [Composite Data Domain Discovery](#).

Prérequis pour effectuer la découverte de domaine sur une ressource Cassandra

Avant d'exécuter les profils pour effectuer la découverte de domaine sur une ressource Cassandra, veillez à configurer les paramètres de connexion et du service d'intégration de données.

Paramètre de connexion :

1. Dans Informatica Administrator, cliquez sur l'onglet **Connexion** dans la section **Gérer**.
2. Dans la section **Navigateur de domaine**, sélectionnez la connexion à la ressource Cassandra que vous utilisez pour exécuter les profils.
3. Dans la section **Propriété avancée**, définissez **Caractère d'identificateur SQL** sur `" "` (guillemets).
4. Cliquez sur **OK**.
5. Dans le champ **Propriétés de connexion supplémentaires**, ajoutez la valeur suivante :
`EnableCaseSensitive=0`.
6. Cliquez sur **OK**.

Paramètre du service d'intégration de données :

1. Dans Informatica Administrator, cliquez sur l'onglet **Services et nœuds** dans la section **Gérer**.
2. Accédez au service d'intégration de données
3. Dans le champ **Propriétés avancées**, définissez la valeur de **Taille maximale des segments de mémoire** sur 4 096 Mo.
4. Cliquez sur **OK**.

5. Dans le champ **Propriétés personnalisées**, ajoutez la valeur suivante :
`ExecutionContextOptions.JVMMaxMemory=1024M`
6. Cliquez sur **OK**.

MongoDB

MongoDB est une base de données interplate-forme basée sur des documents qui fournit des performances élevées, une haute disponibilité et une évolutivité.

Objets extraits

La ressource MongoDB extrait les métadonnées des actifs suivants dans une source de données MongoDB :

- Base de données
- Collections
- Champs
- Vues

Vous pouvez afficher le lignage externe des objets, des objets imbriqués et de tous les types de données MongoDB, à l'exception du type de données en tableau, lorsque vous utilisez les mappages Informatica Intelligent Cloud Services, PowerCenter et Data Engineering Integration.

Remarque: La ressource MongoDB n'affiche pas le lignage interne ni la relation entre les vues et les collections.

Autorisations de configuration de la ressource

Configurez l'autorisation `read` sur la source de données MongoDB pour le compte utilisateur que vous utilisez pour accéder à la source de données.

Connexion à une source de données MongoDB compatible SSL

Pour vous connecter à une source de données MongoDB compatible SSL, procédez comme suit :

1. Téléchargez les certificats SSL MongoDB à l'aide d'un navigateur Web.
Remarque: Veillez à importer le certificat MongoDB Trust Services dans le répertoire Certificats.
2. Copiez les certificats dans le répertoire `<INFA_HOME>/services/shared/security/`.
3. Accédez au répertoire `<INFA_HOME>/source/java/jre/bin`, puis exécutez la commande `keytool` pour importer chaque certificat copié en tant que certificat approuvé dans le fichier keystore du domaine Informatica :

```
keytool -import -file <INFA_HOME>/services/shared/security/<certificate>.cer -alias  
<alias name> -keystore <INFA_HOME>/services/shared/security/infa_truststore.jks -  
storepass <Informatica domain keystore password>
```

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Utiliser l'URL de connexion	URL pour la connexion au serveur MongoDB. La valeur par défaut est Oui.
URL de connexion	URL pour la connexion au serveur MongoDB. La syntaxe de l'URL pour un serveur MongoDB autonome est <code>mongodb://<hostname>:<port>/<database>?retryWrites=true&w=majority</code> La syntaxe de l'URL pour MongoDB en tant que service dans le cluster est <code>mongodb+srv://<username>:<password>@<hostname>:<port>/test?retryWrites=true&w=majority</code>
Hôte	Nom d'hôte ou adresse IP du serveur MongoDB.
Port	Numéro de port du serveur MongoDB.
SSL activé	Indique si le serveur MongoDB est compatible SSL.
Authentification	Type d'authentification pour la connexion au serveur MongoDB. Sélectionnez l'une des options suivantes : <ul style="list-style-type: none">- Aucune.- Nom d'utilisateur/Mot de passe. Active l'authentification basée sur les rôles.
Nom d'utilisateur	Nom d'utilisateur du compte pour la connexion au serveur MongoDB.
Mot de passe	Mot de passe du compte pour la connexion au serveur MongoDB.
Base de données d'authentification	Base de données d'authentification dans laquelle les données utilisateur sont définies. Par défaut, les données utilisateur sont stockées dans la base de données d'administration.

Le tableau suivant décrit les propriétés que vous pouvez configurer dans la section **Métadonnées source** de l'onglet **Paramètres de chargement des métadonnées** :

Propriété	Description
Activer les métadonnées source	Active l'extraction des métadonnées.
Base de données	Importe la base de données requise pour l'exécution de la ressource. Utilisez un point-virgule pour séparer les noms de base de données. Une chaîne vide importe toutes les bases de données. Un nom de base de données importe toutes les collections qui appartiennent à la base de données.
Nombre de lignes pour l'échantillonnage	Spécifiez le nombre de lignes sur lesquelles la ressource s'exécute. Par défaut, vous pouvez choisir une taille d'échantillonnage de 20 lignes. La taille d'échantillonnage maximale dans une exécution de ressource est de 1 000 lignes. La ressource MongoDB limite la taille d'échantillonnage à 30 documents pour identifier le schéma.

Propriété	Description
Option d'échantillonnage	Spécifiez l'une des options d'échantillonnage suivantes : <ul style="list-style-type: none"> - N premières lignes - N lignes aléatoires - N dernières lignes
Filtre de métadonnées source	Spécifiez une combinaison d'expressions régulières et de caractères génériques pour inclure ou exclure des actifs spécifiques de l'exécution de la ressource. Utilisez un point-virgule pour séparer les modèles de caractère générique.
Importer les collections système	Option permettant d'indiquer si les collections MongoDB doivent être importées. Sélectionnez la valeur True si l'utilisateur dispose des privilèges de propriétaire de la base de données.
Mémoire	Valeur de mémoire requise pour exécuter une tâche de scanneur. Spécifiez l'une des valeurs de mémoire suivantes : <ul style="list-style-type: none"> - Faible - Moyenne - Élevée Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal .
Options de JVM	Paramètres JVM que vous pouvez définir pour configurer le conteneur du scanneur. Utilisez les arguments suivants pour configurer les paramètres : <ul style="list-style-type: none"> - <code>-Dscannerloglevel=<DEBUG/INFO/ERROR></code>. Définit le niveau de journalisation du scanneur sur des valeurs, telles que DEBUG, INFO ou ERROR. La valeur par défaut est INFO. - <code>-Dscanner.container.core=<No. of core></code>. Augmente le nombre de cœurs disponibles pour le conteneur du scanneur. La valeur doit être un nombre. - <code>-Dscanner.yarn.app.environment=<key=value></code>. Valeur de paire de clés que vous devez définir dans l'environnement Yarn. Utilisez une virgule comme séparateur pour les paires clé-valeur multiples. - <code>-Dscanner.pmem.enabled.container.memory.jvm.memory.ratio=<1.0/2.0></code>. Augmente la mémoire du conteneur du scanneur lorsque <code>pmem</code> est activé. La valeur par défaut est 1.
Nombre de threads pour l'exécution	Nombre de threads requis pour l'extraction des métadonnées.
Suivre les modifications de la source de données	Affichez les notifications de modification de la source de métadonnées dans Enterprise Data Catalog.

Fonctionnalités non supportées

- La ressource MongoDB ne prend pas en charge les tableaux indexés ni les tables virtuelles.
- Pour des tableaux hétérogènes, la ressource MongoDB récupère la clé pour inférer le schéma. La valeur des tableaux hétérogènes n'est pas utilisée pour inférer le schéma.

CHAPITRE 14

Configuration des ressources Informatica Platform

Ce chapitre comprend les rubriques suivantes :

- [Présentation des ressources Informatica Platform, 206](#)
- [Informatica Data Quality, 206](#)
- [Ressource Informatica Similarity Discovery, 212](#)
- [Dépannage, 214](#)

Présentation des ressources Informatica Platform

Le groupe de ressources Informatica Platform comprend les ressources Informatica Data Quality et Informatica Similarity Discovery.

Lorsque vous créez et exécutez la ressource Informatica Data Quality, le scanner extrait et migre les résultats du profil de l'entrepôt de profilage vers le catalogue. Vous pouvez ensuite afficher les résultats migrés dans Enterprise Data Catalog. Lorsque vous créez et exécutez la ressource Informatica Similarity Discovery, le scanner découvre des colonnes similaires en fonction du nom de la colonne, du modèle de données de colonne et des valeurs uniques dans les colonnes des ressources sélectionnées.

Informatica Data Quality

Lorsque vous exécutez un profil dans Informatica Analyst ou dans Informatica Developer, le service d'intégration de données stocke les résultats du profil dans l'entrepôt de profilage. Pour afficher ces résultats dans Enterprise Data Catalog, vous pouvez créer une ressource Informatica Data Quality dans Catalog Administrator et lui associer l'entrepôt de profilage. Lorsque vous exécutez la ressource, le scanner d'Informatica Data Quality extrait et migre les résultats du profil dans l'entrepôt de profilage associé vers le catalogue. Les résultats du profil migrés incluent les résultats du profil de colonne, les résultats de la découverte de domaines de données, les résultats du profilage de règle, les fiches d'évaluation et les fréquences de valeurs. Vous pouvez migrer les résultats du profil depuis les versions actuelle ou antérieure

dans tout domaine Informatica. La ressource Informatica Data Quality prend en charge les systèmes de base de données relationnelle.

Dans Catalog Administrator, vous pouvez créer plusieurs ressources d'entrepôt de profilage. Lorsque vous créez une ressource d'entrepôt de profilage, vous associez un entrepôt de profilage à la ressource. Enterprise Data Catalog analyse l'entrepôt de profilage associé et répertorie tous les schémas avec les résultats du profil. Vous pouvez choisir un ou plusieurs schémas pour lesquels le scanneur Informatica Data Quality migre les résultats vers le catalogue.

Lorsque vous exécutez le scanneur Informatica Data Quality, le catalogue compare l'horodatage des résultats du catalogue avec les résultats migrés. Enterprise Data Catalog affiche les derniers résultats du profil et les résultats de la curation en fonction de l'horodatage. Enterprise Data Catalog fusionne les résultats de l'inférence avec ceux existants dans le catalogue.

Vous pouvez afficher la fréquence de valeur d'une colonne ou d'une règle dans l'outil Developer tool ou Analyst tool. Le scanneur Informatica Data Quality migre les fréquences de valeurs des colonnes avec les résultats du profil vers le catalogue. L'entrepôt de profilage stocke un maximum de 16 000 valeurs pour une colonne. Le scanner identifie les 1 000 principales valeurs maximales d'une colonne et les migre vers Enterprise Data Catalog.

Exemple

Supposez que vous disposez de plus d'1 million de tables réparties dans 700 schémas et plusieurs bases de données dans votre entreprise. Au fil des années, vous avez exécuté des profils dans ces schémas dans l'outil Developer tool ou Analyst tool. Tous les résultats du profilage se trouvent dans une ou plusieurs bases de données de l'entrepôt de profilage. Désormais, vous voulez implémenter Enterprise Data Catalog dans votre entreprise et accéder aux résultats du profil existant et les afficher pour les schémas du catalogue. Au lieu d'exécuter les profils dans les schémas et les bases de données de Catalog Administrator, ce qui constitue une tâche longue et gourmande en ressources, vous décidez de migrer les résultats vers le catalogue. En outre, vous voulez que les développeurs et les analystes continuent à utiliser les outils Developer tool et Analyst tool.

Dans ce cas, vous pouvez exécuter les scanners de base pour extraire les métadonnées des schémas et des bases de données vers le catalogue. Ensuite, vous pouvez créer une ressource d'entrepôt de profilage, choisir les schémas et les connexions pour lesquels vous voulez migrer les résultats du profil, puis exécuter la ressource d'entrepôt de profilage. Le scanner migre les résultats du profil des schémas sélectionnés vers le catalogue. Cette action économise du temps et des efforts. Un autre avantage est que les utilisateurs peuvent continuer à utiliser les outils Developer tool et Analyst tool. Vous pouvez également migrer les résultats, le cas échéant.

Extraction des résultats de profil de règle

Vous pouvez appliquer des règles et exécuter des profils de règle dans les outils Developer tool et Analyst tool. Vous pouvez utiliser les méthodes suivantes pour créer ou appliquer des règles dans les profils de colonne :

- Dans l'outil Developer tool, créez un mapplet et validez-le en tant que règle. La règle s'affiche en tant que règle réutilisable dans l'outil Analyst tool. Vous pouvez appliquer la règle aux profils de colonnes dans les outils Analyst tool et Developer tool.
- Dans l'outil Analyst tool, créez une spécification de règle et générez un mapplet. Vous pouvez appliquer la spécification de règle aux profils de colonnes dans l'outil Analyst tool. Dans l'outil Developer tool, validez le mapplet en tant que règle. La règle s'affiche en tant que règle réutilisable que vous pouvez utiliser dans les profils de colonnes.
- Vous pouvez utiliser des règles prédéfinies dans les profils de colonnes. Informatica fournit des règles prédéfinies avec les outils Developer tool et Analyst tool.

Lorsque vous exécutez la ressource Informatica Data Quality, elle extrait les résultats du profil de règle vers le catalogue. Vous ne pouvez pas appliquer des règles d'expression au profil. Vous ne pouvez pas associer

les termes d'entreprise avec un actif de règle dans Enterprise Data Catalog. Pour en savoir plus sur la création de règles et leur application aux profils dans les outils Developer tool et Analyst tool, consultez le *Guide Informatica Data Discovery*.

Extraction des résultats de fiche d'évaluation

Vous pouvez créer et modifier une fiche d'évaluation dans l'outil Developer tool et l'outil Analyst tool. Une fiche d'évaluation comprend plusieurs composants, tels que des mesures, des groupes de mesures, des scores et des seuils. Les fiches d'évaluation permettent à une entreprise de mesurer la qualité des données au niveau de la mesure. Vous pouvez créer des groupes de mesures pour regrouper les mesures connexes. Une fois le profil exécuté, vous pouvez ajouter des colonnes à partir des résultats de profil comme mesures pour une fiche d'évaluation. Lorsque vous exécutez une fiche d'évaluation, les outils Analyst tool et Developer tool génèrent les valeurs moyennes pondérées de chaque groupe de mesures. Lorsque vous exécutez la ressource Data Quality, elle extrait les résultats de la fiche d'évaluation vers le catalogue.

Vous ne pouvez pas associer les termes d'entreprise avec un actif de fiche d'évaluation dans Enterprise Data Catalog. Pour en savoir plus sur la création, la modification et l'exécution de fiches d'évaluation dans les outils Developer tool et Analyst tool, consultez le *Guide Informatica Data Discovery*.

Objets extraits

Le scanneur Informatica Data Quality extrait les mesures de profilage suivantes de l'entrepôt de profilage vers le catalogue :

- Résultats du profil de colonne
 - Valeurs Null, uniques et non uniques.
 - Modèles
 - Types de données
 - Valeurs maximum et minimum
 - Valeur d'écart-type
 - Valeur moyenne
 - Valeur totale
- Résultats de la découverte de domaines de données
 - Résultats de l'inférence
 - Résultats de la curation
- Résultats du profil de règle
 - Valeurs Null, unique et non unique
 - Modèles
 - Valeurs maximale et minimale
 - Longueur maximale et minimale
 - Types de données
- Fiches d'évaluation
 - Groupes de mesures
 - Scores de mesure
 - Scores de qualité
- Fréquences de valeurs

Prérequis

Pour respecter les prérequis, effectuez l'étape suivante :

- Avant de migrer les résultats du profilage d'une ressource, vous devez exécuter la ressource correspondante dans Catalog Administrator afin d'extraire les métadonnées dans le catalogue. Par exemple, vous voulez migrer les résultats du profil de schéma S1 d'Oracle vers le catalogue. Avant de migrer les résultats du profil de schéma Oracle, créez et exécutez une ressource Oracle avec S1 et utilisez-la pour configurer la ressource d'entrepôt de profilage.
- Exportez les domaines de données du domaine de l'entrepôt de profilage vers le domaine Enterprise Data Catalog si l'entrepôt de profilage et Enterprise Data Catalog se trouvent sur deux domaines différents.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Nom	Entrez un nom pour la ressource d'entrepôt de profilage.
Description	Vous pouvez entrer éventuellement une brève description de la ressource.
Type d'actif	Choisissez l'option Informatica Data Quality comme type de ressource.
Type de base de données d'entrepôt de profilage	Choisissez Oracle , IBM DB2 , ou Microsoft SQL Server comme type de base de données d'entrepôt de profilage.
Utilisateur	Entrez le nom d'utilisateur pour accéder à la base de données d'entrepôt de profilage.
Mot de passe	Entrez le mot de passe pour accéder à la base de données d'entrepôt de profilage.
Hôte	Entrez le nom d'hôte ou l'adresse IP du domaine Informatica, ou la chaîne de connexion JDBC complète. Pour obtenir un exemple de chaînes de connexion JDBC, consultez la rubrique "Dépannage" à la page 214 .
Port	Entrez le numéro de port du domaine Informatica ou du service de moteur de base de données. La valeur par défaut est 1521.
Service	Entrez le nom de service Oracle. Cette option s'affiche lorsque la base de données d'entrepôt de profilage est Oracle.
Base de données	Entrez le nom de la base de données Microsoft SQL Server. Le scanner utilise la valeur comme contrainte pour importer la base de données Microsoft SQL Server. Vous pouvez entrer un ou plusieurs noms de bases de données séparés par des virgules. Lorsque vous spécifiez un nom de base de données, le scanner n'importe que les schémas et les tables dans la base de données. Si vous entrez, par exemple, AdvDW , le scanner importe les tables et les schémas de la base de données AdvDW. Cette option s'affiche lorsque la base de données d'entrepôt de profilage est Microsoft SQL Server ou IBM DB2.

Propriété	Description
Nom du schéma	Entrez le nom du schéma. Le scanner utilise la valeur comme contrainte pour spécifier le schéma non par défaut de la base de données Microsoft SQL Server. Cette option s'affiche lorsque la base de données d'entrepôt de profilage est Microsoft SQL Server.
Extraire les résultats de	Choisissez la règle, la fiche d'évaluation, le profil ou le profil et la fréquence de valeurs pour extraire les résultats de l'entrepôt de profilage.
Nom du domaine	Nom du domaine Informatica. Remarque: Cette propriété s'applique à la version 10.2.0 HotFix 2 et aux versions ultérieures.
Nom d'utilisateur du domaine	Nom d'utilisateur employé par le service d'intégration de données pour accéder au service de référentiel modèle.
Mot de passe du domaine	Mot de passe de l'utilisateur du référentiel modèle.
Domaine de sécurité	Nom du domaine de sécurité auquel l'utilisateur du domaine est rattaché.
Hôte du domaine	Nom d'hôte du nœud principal de passerelle.
Port du domaine	Numéro de port du nœud principal de passerelle.
Nom du service de référentiel	Nom du service de référentiel modèle.
Nom de connexion de l'entrepôt	Nom de l'entrepôt de profilage.

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer les métadonnées source	Sélectionnez cette option pour extraire les métadonnées des sources de données.
Connexions de schéma profilé	Sélectionnez un ou plusieurs schémas dans la boîte de dialogue Sélectionner les connexions de schéma profilé . Le scanner d'entrepôt de profilage migre les résultats du profilage des schémas sélectionnés vers le catalogue.
Cumulatif	Sélectionnez ou désélectionnez cette option, selon les besoins. - Sélectionnez cette option pour que le scanner numérise tous les résultats du profil d'une source de données afin d'extraire les derniers résultats de la colonne en fonction de l'horodatage. Pour plus d'informations sur cette option, consultez les exemples dans la rubrique "Dépannage" à la page 214 . - Désélectionnez cette option pour récupérer les derniers résultats du profil.
Pourcentage d'acceptation automatique	Entrez une valeur de comprise entre 0 et 100. Enterprise Data Catalog accepte automatiquement les domaines de données lorsque le pourcentage d'inférence dépasse le nombre configuré.

Propriété	Description
Incrémentiel	<p>Sélectionnez ou désélectionnez cette option, selon les besoins.</p> <ul style="list-style-type: none"> - Sélectionnez cette option pour migrer la valeur delta des résultats du profil dans chaque exécution. Dans la première exécution de la ressource d'entrepôt de profilage, le scanner migre les résultats du profil pour toutes les tables. Dans les exécutions suivantes de la ressource, seule la valeur delta des résultats du profil est migrée. <p>Pour plus d'informations sur cette option, consultez les exemples dans la rubrique "Dépannage" à la page 214.</p> <ul style="list-style-type: none"> - Désélectionnez cette option pour récupérer les derniers résultats du profil.
Mémoire	<p>Spécifiez la mémoire requise pour exécuter la tâche du scanner.</p> <p>Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données migré :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevée <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal.</p>
Options de JVM	Entrez les paramètres JVM pour configurer le conteneur du scanner.
Suivre les modifications de la source de données	Sélectionnez cette option pour afficher la notification de modification de la source de métadonnées dans Enterprise Data Catalog.

Tâches de ressource Informatica Data Quality

Lorsque vous exécutez la ressource Informatica Data Quality, le scanner effectue les tâches suivantes :

1. Vérifie que le catalogue contient les métadonnées source pour la ressource.
Pour migrer, par exemple, les résultats du profil Oracle, vous devez d'abord créer et exécuter la ressource d'Oracle dans Catalog Administrator.
2. Vérifie que la ressource correspondante contient les résultats du profil.
3. Vérifie que la ressource correspondante contient les décisions de curation.
4. Si les résultats de la curation n'existent pas, le scanner migre les résultats vers le catalogue.
5. S'ils existent, le scanner compare l'horodatage des décisions de curation dans le catalogue avec la décision de curation dans l'entrepôt de profilage. Le scanner migre et remplace les résultats de la curation si ceux de l'entrepôt de profilage comportent le dernier horodatage.

Par exemple, vous rejetez le domaine de données SSN pour une colonne de la table Employé dans le catalogue le 05/01/19. Vous acceptez le domaine de données SSN pour la même colonne dans l'outil Developer tool le 05/02/2019. Vous exécutez le scanner Informatica Data Quality le 05/03/19. La décision de curation du domaine de données SSN dans l'entrepôt de profilage comporte le dernier horodatage. Le scanner migre et remplace la décision de curation dans le catalogue. Le domaine de données SSN affiche l'acceptation dans le catalogue.

Remarque: Enterprise Data Catalog affiche **SYSTEM** dans la colonne **Attribué par** pour une ressource de domaine de données lorsque vous organisez un domaine de données dans l'outil Developer tool ou Analyst tool et que la décision est migrée vers le catalogue.

6. Fusionne les résultats inférés du domaine de données.
Dans l'outil Developer tool, par exemple, vous acceptez les domaines de données SSN et Âge dans la table Employé. Dans Enterprise Data Catalog, vous acceptez les domaines de données Adresse et Ville

dans la même table Employé. Lorsque vous exécutez le scanneur Informatica Data Quality, celui-ci fusionne les résultats acceptés du domaine de données et affiche *SSN*, *Âge*, *Adresse* et *Ville* comme domaines de données inférés pour la table Employé dans le catalogue.

7. Accepte automatiquement un domaine de données si le pourcentage d'inférence dépasse la valeur **Pourcentage d'acceptation automatique**.
8. Identifie les 1 000 principales valeurs maximales d'une colonne et les migre vers le catalogue.
9. Dans Catalog Administrator, lorsque vous cliquez sur l'onglet **Lien de rapport manquant** de la section **Attribution de connexion**, le fichier exporté n'affiche pas les détails des objets source et cible non liés.

Ressource Informatica Similarity Discovery

En tant qu'analyste de données ou architecte de données, vous pouvez analyser les données de votre entreprise pour trouver des colonnes similaires. Vous pouvez exécuter la ressource Informatica Similarity Discovery pour découvrir des colonnes similaires en fonction du nom de la colonne, du modèle de données de colonne et des facteurs de valeurs uniques.

Dans Catalog Administrator, vous pouvez créer plusieurs ressources Informatica Similarity Discovery. La ressource Informatica Similarity Discovery s'exécute sur plusieurs ressources pour découvrir des colonnes similaires dans le catalogue. Vous pouvez spécifier les ressources sur lesquelles vous souhaitez exécuter la découverte de similarité.

Lorsque vous créez et exécutez la ressource Informatica Similarity Discovery, la ressource analyse la base de données PostgreSQL, compare les données préparées pour déduire des colonnes similaires et conserve les résultats dans la base de données PostgreSQL. Après avoir exécuté la ressource Informatica Similarity Discovery, vous pouvez afficher des colonnes similaires dans Enterprise Data Catalog.

Prérequis

Avant de créer et d'exécuter la ressource Informatica Similarity Discovery, assurez-vous que vous avez sélectionné l'option **Exécuter le profil de similarité** pour les ressources sur lesquelles vous souhaitez exécuter la découverte de similarité dans Catalog Administrator.

Exemple

Le catalogue contient des détails client similaires répartis entre plusieurs ressources. Vous souhaitez peut-être identifier toutes les ressources qui contiennent des détails client similaires afin de pouvoir intégrer les données dans une seule ressource et supprimer les données en double. Pour atteindre cet objectif, vous pouvez activer l'option **Exécuter le profil de similarité** pour les ressources identifiées, puis créer et exécuter la ressource Informatica Similarity Discovery sur ces ressources pour identifier et afficher des colonnes similaires dans Enterprise Data Catalog.

Pour obtenir plus d'exemples de similarité de colonne, reportez-vous à la rubrique [Business Example](#) dans le *Guide d'Informatica Catalog Administrator*.

Propriétés de connexion à la ressource

L'onglet **Général** inclut les propriétés suivantes :

Propriété	Description
Nom	Entrez un nom pour la ressource.
Description	Vous pouvez entrer éventuellement une brève description de la ressource.
Type d'actif	Choisissez l'option Informatica Similarity Discovery comme type de ressource.
Ressources à regrouper	<p>Sélectionnez les ressources sur lesquelles vous souhaitez exécuter la découverte de similarité. Par défaut, vous pouvez sélectionner quatre ressources pour la découverte de similarité. Si vous souhaitez ajouter plus de ressources, vous pouvez configurer la propriété personnalisée suivante pour le service de catalogue dans Informatica Administrator :</p> <p><code>IdmCustomOptions.scanner.simDiscoveryScanner.maxPermittedResourceCount.</code></p> <p>Remarque: Après avoir exécuté la ressource Informatica Similarity Discovery avec un ensemble de ressources sélectionnées pour la découverte de similarité, vous ne pouvez pas sélectionner les mêmes ressources pour exécuter la découverte de similarité à l'aide d'une autre ressource Informatica Similarity Discovery. Par exemple, si vous utilisez les ressources R1, R2 et R3 pour exécuter la découverte de similarité, vous ne pouvez pas réutiliser les ressources R1, R2 et R3 avec une autre ressource Découverte de similarité.</p>

L'onglet **Paramètres de chargement des métadonnées** inclut les propriétés suivantes :

Propriété	Description
Activer la découverte de similarité	Sélectionnez l'option pour exécuter la découverte de similarité sur les ressources sélectionnées.
Fonctionnalités activées	<p>Sélectionnez ou désélectionnez les options selon vos besoins. Sélectionnez les fonctionnalités sur lesquelles vous souhaitez exécuter la découverte de similarité.</p> <p>Fonctionnalités de découverte de similarité :</p> <ul style="list-style-type: none">- Nom. Détermine les colonnes similaires dans les sources de données en fonction des noms de colonnes. La correspondance de noms utilise la correspondance approximative de chaînes pour identifier les colonnes qui portent des noms similaires.- Modèles. Détermine le modèle de données de colonnes dans les sources de données et calcule le pourcentage de chevauchement des modèles dans les paires de colonnes pour tous les types de données, sauf les types de données numériques.- Valeurs uniques. Détermine les colonnes similaires dans les sources de données en fonction des valeurs distinctes. Le chevauchement des valeurs distinctes est calculé en pourcentage des valeurs distinctes qui se chevauchent dans deux colonnes. <p>Pour plus d'informations sur les fonctionnalités de la découverte de similarité, reportez-vous à la rubrique Column Similarity.</p>
Cumulatif	Choisissez cette option pour conserver les résultats de similarité de colonne issus de l'analyse précédente dans les résultats d'analyse des ressources suivantes. Si vous ne choisissez pas cette option, les résultats de profil précédents sont supprimés.

Propriété	Description
Mémoire	<p>Sélectionnez la mémoire requise pour exécuter le travail du scanner. Sélectionnez l'une des valeurs suivantes en fonction de la taille du jeu de données importé :</p> <ul style="list-style-type: none"> - Faible - Moyenne - Élevé <p>Remarque: Pour plus d'informations sur les valeurs de mémoire, consultez l'article sur le <i>réglage des performances d'Enterprise Data Catalog</i> sous l'onglet des articles de la bibliothèque de guides pratiques du Informatica Doc Portal..</p>
Options personnalisées	Spécifiez les options personnalisées pour exécuter le travail du scanner.
Suivre les modifications de la source de données	Sélectionnez cette option pour afficher la notification de modification de la source de métadonnées dans Enterprise Data Catalog.

Tâches de ressources

Lorsque vous exécutez la ressource Informatica Similarity Discovery, le scanner effectue les tâches suivantes :

1. Vérifie si les ressources sélectionnées sont activées avec l'option **Exécuter le profil de similarité**.
2. Analyse les données préparées dans la base de données PostgreSQL, compare les données préparées pour déduire des colonnes similaires et conserve les résultats dans la base de données PostgreSQL.

La matrice suivante décrit le type de calcul de similarité effectué sur les colonnes de différents types de données déduits :

Type de colonne	Nom	Modèle	Valeurs uniques
Numérique	Oui	Non	Non
Chaîne	Oui	Oui	Oui
Date	Oui	Non	Non

Pour plus d'informations sur le processus de similarité de colonne, reportez-vous à la rubrique [Column Similarity Process](#).

Dépannage

Dans l'outil Developer tool, j'exécute un profil dans les trois premières colonnes d'une source de données dans la première exécution. Ensuite, j'exécute le profil dans les trois colonnes suivantes de la source de données. Qu'affichera Enterprise Data Catalog après l'exécution de la ressource d'entrepôt de profilage ?

Par défaut, le catalogue affiche les derniers résultats de l'exécution du profil. Pour afficher toutes les six colonnes du catalogue, choisissez l'option **Cumulatif** dans la ressource d'entrepôt de profilage et exécutez la ressource. Enterprise Data Catalog affiche les derniers résultats de chaque colonne dans la source de données en fonction de l'horodatage.

Comment l'option Cumulatif fonctionne-t-elle ?

Supposez que vous disposez d'une table Employé contenant les colonnes C1, C2, C3 et C4. Dans l'outil Developer tool, vous créez et exécutez un profil dans la table. Les exécutions du profil incluent R1, R2 et R3.

La liste suivante affiche les colonnes que vous choisissez lorsque vous exécutez le profil :

1. R1. Vous avez choisi C1 et C2.
2. R2. Vous avez choisi C3 et C4.
3. R3. Vous avez choisi C1 et C4.

Dans Catalog Administrator, vous créez et exécutez la ressource d'entrepôt de profilage avec l'option **Cumulatif**. Le scanner Informatica Data Quality extrait les derniers résultats des colonnes suivants en fonction de l'horodatage :

1. C1. Le scanner extrait les derniers résultats de R3.
2. C2. Le scanner extrait les derniers résultats de R1.
3. C3. Le scanner extrait les derniers résultats de R2.
4. C4. Le scanner extrait les derniers résultats de R3.

Comment l'option Incrémentiel fonctionne-t-elle ?

1. Dans l'outil Developer tool, vous exécutez un profil P1 dans les tables Employé et Département à 10 h le 05/02/19.
2. Dans Catalog Administrator, vous créez une ressource d'entrepôt de profilage PWH à l'aide de l'option **Incrémentiel** et exécutez la ressource à 10 h 15 le 05/02/19.
Le scanner migre les résultats du profil de P1 vers le catalogue.
3. Vous réexécutez P1 à 10 h le 05/03/19.
4. Vous exécutez PWH à 10 h 15 le 05/03/19.
Le scanner ne migre pas de résultats, car il n'y a aucun nouveau résultat, ni changement dans l'horodatage.
5. Vous créez et exécutez le profil P2 dans les tables Employé et Département à 10 h le 05/04/19.
6. Vous exécutez PWH à 10 h 15 le 05/04/19.
Le scanner migre les résultats du profil de P2 vers le catalogue.
7. Vous exécutez le profil dans la table Paie à 10 h le 05/06/19.
8. Vous exécutez PWH à 10 h 15 le 05/06/19.
Le scanner ne migre que les résultats du profil de la table Paie, car vous avez choisi l'option **Incrémentiel**. Dans ce cas, les résultats du profil de la table Paie sont également appelés delta.

Quels résultats s'affichent dans le catalogue si j'exécute la ressource de profilage et la ressource d'entrepôt de profilage dans la même table ?

Le catalogue compare l'horodatage des résultats du profil provenant de l'exécution de la ressource de profilage et les résultats du profil migrés. Il affiche les résultats avec le dernier horodatage.

Quels types de tâches sont disponibles pour la ressource d'entrepôt de profilage dans l'onglet Surveillance ?

Le type de tâche **Chargement des métadonnées** est disponible pour la ressource d'entrepôt de profilage.

Quelles opérations de statistiques et de progression du profilage s'affichent dans l'onglet Surveillance pour la ressource d'entrepôt de profilage ?

Dans l'onglet **Statistiques**, les noms de statistiques de profilage suivants et leurs valeurs s'affichent :

- Nombre d'identificateurs de résultats du profil récupérés

- Nombre de résultats du profil récupérés
- Résultats de fréquences de valeurs pour les tables extraites

Dans l'onglet **Progression**, les noms d'opérations de progression suivants s'affichent avec l'horodatage et les résultats :

- Extraction des résultats du profil
- Extraction des identificateurs de résultats du profil
- Ingestion des résultats du profil
- Extraction de VF

Si le catalogue contient plusieurs résultats pour la ressource, quels sont ceux qui sont remplacés par le scanner d'entrepôt de profilage ?

Supposez que vous disposez d'un schéma S1 d'Oracle. Vous créez des connexions C1 et C2 d'Oracle dans l'outil Administrator tool.

Dans l'outil Developer tool, vous choisissez C1 pour importer les tables Oracle de S1, créez un profil pour les tables et exécutez le profil.

Dans Catalog Administrator, vous choisissez C1 et S1 pour créer la ressource R1 d'Oracle. Vous choisissez C2 et S1 pour créer la ressource R2 d'Oracle. Vous exécutez les ressources. Le catalogue affiche les résultats de la ressource.

Lorsque vous créez une ressource d'entrepôt de profilage, vous choisissez la connexion C1 de l'entrepôt de profilage. Lorsque vous exécutez la ressource d'entrepôt de profilage, le catalogue compare et remplace les résultats de la ressource de R1.

Je vois une incompatibilité de noms de domaines de données dans le catalogue. Comment résoudre ce problème ?

Ce problème se produit lorsque l'entrepôt de profilage et Enterprise Data Catalog se trouvent dans des domaines différents. Pour synchroniser les domaines de données, exportez-les depuis le domaine de l'entrepôt de profilage et importez-les dans le domaine Enterprise Data Catalog.

Quels sont les types de bases de données pris en charge et leurs chaînes de connexion JDBC complètes que je peux utiliser pour le scanner de l'entrepôt de profilage ?

Le tableau suivant répertorie les exemples de chaînes de connexion JDBC pour tous les types de bases de données pris en charge :

Type de base de données	Valeur du nom de la classe	Valeur de la chaîne de connexion
Oracle	com.informatica.jdbc.oracle.OracleDriver	jdbc:informatica:oracle://<nom d'hôte>:<port>;SID=<sid>
DB2	com.informatica.jdbc.db2.DB2Driver	jdbc:informatica:db2://<hostname>:<port>;DatabaseName=<dbname>
SQL Server	com.informatica.jdbc.sqlserver.SQLServerDriver	jdbc:informatica:sqlserver://<host>:<port>;databaseName=<dbname>

Type de base de données	Valeur du nom de la classe	Valeur de la chaîne de connexion
Sybase	com.informatica.jdbc.sybase.SybaseDriver	jdbc:informatica:sybase:// <host>:<port>;databaseName=<dbname>
MYSQL	com.informatica.jdbc.mysql.MySQLDriver	jdbc:informatica:mysql:// <host>:<port>;databaseName=<dbname>

J'exécute un profil sur une colonne source qui comprend un caractère spécial dans son nom. Pourquoi l'exécution du profil échoue-t-elle ?

Si une table ou une colonne source contient un caractère spécial dans son nom ou que le nom commence par un nombre, l'entrepôt de profilage remplace chaque caractère spécial ou le nombre par un trait de soulignement (_). Par conséquent, la table n'est pas stockée dans l'entrepôt de profilage sous son nom d'origine et vous pouvez voir deux tables au lieu de la table source dans Enterprise Data Catalog. Lorsque vous exécutez un profil, l'entrepôt de profilage ne peut pas mettre à jour la table avec les résultats de profilage et l'exécution du profil échoue. Vous pouvez noter un comportement identique lorsqu'une ressource de référence est créée dans le scanneur Data Quality.

Avant de sélectionner les colonnes sur lesquelles vous souhaitez exécuter le profil, supprimez les caractères spéciaux et les nombres des noms de table et de colonne.

Dans l'outil Developer tool, j'exécute un profil pour la découverte de domaines de données et la migration des résultats de profil vers Enterprise Data Catalog. J'effectue ensuite la curation sur les résultats de profil, puis exécute à nouveau la ressource Informatica Data Quality pour migrer les résultats vers Enterprise Data Catalog. Les résultats de profil mis à jour sont-ils affichés dans Enterprise Data Catalog ?

Les résultats mis à jour ne sont pas affichés dans Enterprise Data Catalog. Pour afficher les résultats de profil mis à jour, assurez-vous d'exécuter un profil après la curation et de migrer les résultats de profil vers Enterprise Data Catalog.

INDEX

F

Filtre de métadonnées source
présentation [12](#)

P

propriété de type de ressource
Atlas [52](#)
Axon [76](#)
Cassandra [200](#)
Cloudera Navigator [54](#)
DatabaseScripts [124](#)
DataStage [143](#)
DIH [122](#)
Erwin [147](#)
GCS [49](#)
glossaire d'entreprise [79](#)
Glue [43](#)
Hadoop [182](#)
Hive [63](#)
IBM DB2 [155](#)
IBM DB2 for z/OS [157](#)
IBM Netezza [160](#)
IICS [125](#)
Informatica MDM [127](#)
Informatica Platform [130](#)
JDBC [164](#)
Kafka [196](#)

propriété de type de ressource (*a continué*)

MicroStrategy [86](#)
MongoDB [203](#)
Oracle [170](#)
PowerBI [114](#)
PowerCenter [133](#)
Qlik Sense [117](#)
Salesforce [39](#)
SAP BusinessObjects [95](#)
SAP BW [104](#)
SAP BW/4HANA [109](#)
SAP HANA [173](#)
SAP PowerDesigner [151](#)
SAP S/4HANA [70](#)
Snowflake [46](#)
SQL Server [167](#)
SSIS [139](#)
Sybase [175](#)
Teradata [178](#)

R

ressource

attributs [11](#)
création [11](#)
planification du scanneur [10](#)
présentation [10](#)
scanneur [9](#)