



Informatica®

9.6.1 HotFix 2

Guide des données de référence

© Copyright Informatica LLC 2013, 2018

Ce logiciel et sa documentation contiennent des informations appartenant à Informatica Corporation, protégées par la loi sur le droit d'auteur et fournies dans le cadre d'un accord de licence prévoyant des restrictions d'utilisation et de divulgation. Toute ingénierie inverse du logiciel est interdite. Il est interdit de reproduire ou transmettre sous quelque forme et par quelque moyen que ce soit (électronique, photocopie, enregistrement ou autre) tout ou partie de ce document sans le consentement préalable d'Informatica Corporation. Ce logiciel peut être protégé par des brevets américains et/ou internationaux, ainsi que par d'autres brevets en attente.

L'utilisation, la duplication ou la divulgation du Logiciel par le gouvernement américain est sujette aux restrictions décrites dans l'accord de licence applicable du logiciel conformément aux documents DFARS 227.7202-1(a) et 227.7702-3(a) (1995), DFARS 252.227-7013[©](1)(ii) (OCT 1988), FAR 12.212(a) (1995), FAR 52.227-19 ou FAR 52.227-14 (ALT III) le cas échéant.

Les informations dans ce produit ou cette documentation sont sujettes à modification sans préavis. Si vous rencontrez des problèmes dans ce produit ou la documentation, veuillez nous en informer par écrit.

Informatica, Informatica Platform, Informatica Data Services, PowerCenter, PowerCenterRT, PowerCenter Connect, PowerCenter Data Analyzer, PowerExchange, PowerMart, Metadata Manager, Informatica Data Quality, Informatica Data Explorer, Informatica B2B Data Transformation, Informatica B2B Data Exchange Informatica On Demand, Informatica Identity Resolution, Informatica Application Information Lifecycle Management, Informatica Complex Event Processing, Ultra Messaging et Informatica Master Data Management sont des marques de commerce ou des marques déposées d'Informatica Corporation aux États-Unis et dans d'autres juridictions du monde. Tous les autres noms de société ou de produit peuvent être des marques de commerce ou des marques déposées de leurs détenteurs respectifs.

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 © Université de Toronto. Tous droits réservés. Copyright © Daniel Veillard. Tous droits réservés. Copyright © Uniconic, 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.

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 L'ÉTAT », 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://hsqldb.org/web/hsqldbLicense.html>, <http://httpunit.sourceforge.net/doc/license.html>, <http://jung.sourceforge.net/license.txt>, http://www.gzip.org/zlib/zlib_license.html, <http://www.openldap.org/software/release/license.html>, <http://www.libssh2.org>, <http://slf4j.org/license.html>, <http://www.sente.ch/software/OpenSourceLicense.html>, <http://fusesource.com/downloads/license-agreements/fuse-message-broker-v-5-3-license-agreement>, <http://antlr.org/license.html>, <http://aopalliance.sourceforge.net/>, <http://www.bouncycastle.org/licence.html>, <http://www.jgraph.com/jgraphdownload.html>, <http://www.jcraft.com/jsch/LICENSE.txt>, http://jotm.objectweb.org/bsd_license.html, <http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231>, <http://www.slf4j.org/license.html>, <http://nanoxml.sourceforge.net/orig/copyright.html>, <http://www.json.org/license.html>, <http://forge.ow2.org/projects/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>, et <https://code.google.com/p/lz4/>.

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>), la 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.

Ce logiciel est protégé par des brevets américains (5,794,246; 6,014,670; 6,016,501; 6,029,178; 6,032,158; 6,035,307; 6,044,374; 6,092,086; 6,208,990; 6,339,775; 6,640,226; 6,789,096; 6,823,373; 6,850,947; 6,895,471; 7,117,215; 7,162,643; 7,243,110; 7,254,590; 7,281,001; 7,421,458; 7,496,588; 7,523,121; 7,584,422; 7,676,516; 7,720,842; 7,721,270; 7,774,791; 8,065,266; 8,150,803; 8,166,048; 8,166,071; 8,200,622; 8,224,873; 8,271,477; 8,327,419; 8,386,435; 8,392,460; 8,453,159; 8,458,230; 8,707,336; 8,886,617 et RE44,478), des brevets internationaux ainsi que par d'autres brevets en attente.

EXCLUSION DE RESPONSABILITÉ : Informatica Corporation 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 Corporation 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.

Date de publication: 2018-06-08

Sommaire

Préface.....	7
Ressources Informatica.	7
Portail Mon support Informatica.	7
Documentation Informatica.	7
Matrices de disponibilité de produit Informatica.	7
Site Web Informatica.	8
Bibliothèque de procédures Informatica.	8
Base de connaissances Informatica.	8
Canal YouTube du support Informatica.	8
Informatica Marketplace.	8
Informatica Velocity.	8
Support client international Informatica.	9
 Chapitre 1: Introduction aux données de référence.....	 10
Présentation des données de référence.	10
Données de référence Informatica.	11
Données de référence définies par l'utilisateur.	12
Tables de référence.	12
Structure des tables de référence.	13
Tables de référence gérées et non gérées.	13
 Chapitre 2: Tables de référence dans l'outil Analyst.....	 14
Présentation des tables de référence de l'outil Analyst.	14
Propriétés des tables de référence.	15
Propriétés de la colonne Table de référence.	15
Création d'une table de référence dans l'éditeur de table de référence.	16
Créer une table de référence à partir des données de profil.	16
Créer une table de référence depuis les colonnes de profil.	17
Créer une table de référence depuis les valeurs de colonne.	18
Créer une table de référence depuis les modèles de colonnes.	19
Créer une table de référence à partir d'un fichier plat.	20
Propriétés de fichier plat de l'outil Analyst.	20
Création d'une table de référence à partir d'un fichier simple.	20
Créer une table de référence à partir d'une table de base de données.	21
Création d'une table de référence à partir d'une table de base de données.	21
Mises à jour de la table de référence.	22
Gérer les colonnes.	23
Gérer les lignes.	23
Rechercher et remplacer les valeurs.	24
Exporter une table de référence.	24

Activer et désactiver les éditions d'une table de référence non gérée.	25
Actualiser les valeurs de la table de référence.	25
Événements suivi d'audit.	26
Voir les événements de suivi d'audit.	26
Règles et instructions pour les tables de référence.	27

Chapitre 3: Données de référence dans l'outil Developer..... 28

Présentation des données de référence de l'outil Developer.	28
Données de référence et transformations.	29
Tables de référence.	29
Propriétés des données de la table de référence.	29
Création d'un objet de table de référence.	30
Création d'une table de référence à partir d'un fichier plat.	31
Création d'une table de référence à partir d'une source relationnelle	32
Ensembles de contenu.	33
Ensembles de caractères.	34
Modèles de classeur.	34
Ensembles de modèles.	35
Modèles probabilistes.	35
Expressions régulières.	36
Ensembles de jetons.	36
Règles et directives pour les modèles probabilistes et les modèles de classeur.	38
Création d'un ensemble de contenu.	39
Création d'un objet de données de référence dans un ensemble de contenu.	39
Génération de données de référence depuis un profil intermédiaire	40

Chapitre 4: Modèles de classeur..... 41

Présentation des modèles de classeur.	41
Exemple de transformation Classeur.	42
Structure du modèle de classeur.	42
Vues du modèle de classeur.	42
Données de référence du modèle de classeur.	43
Ajout de valeurs de données et de libellé à un modèle de classeur.	44
Suppression de valeurs de données d'un modèle de classeur.	45
Données de libellé du modèle de classeur.	45
Ajout d'un libellé à un modèle de classeur.	45
Attribution de libellés aux données du modèle de classeur.	45
Suppression d'un libellé d'un modèle de classeur.	46
Compilation d'un modèle de classeur.	46
Scores de classeur.	46
Filtres de modèle de classeur.	47
Recherche de valeurs dans les données de référence	47
Recherche de lignes de données sans libellés.	47

Recherche d'une valeur dans une ligne de données.	47
Création d'un modèle de classeur à partir d'un objet de données.	48
Opérations de copier/coller.	49
Copie d'un modèle de classeur dans un autre ensemble de contenu.	49
Importation d'un modèle de classeur depuis un autre ensemble de contenu.	49

Chapitre 5: Modèles probabilistes..... 50

Présentation des modèles probabilistes.	50
Structure du modèle probabiliste.	51
Vues du modèle probabiliste.	51
Exemple de transformation Libellé.	52
Exemple de transformation Analyseur.	53
Données de référence du modèle probabiliste.	54
Ajout d'une chaîne de données de référence à un modèle probabiliste.	54
Recherche de valeurs de données dans un modèle probabiliste.	54
Données de libellé du modèle probabiliste.	55
Libellé de dépassement.	55
Attribution de libellés à des données de modèle probabiliste.	56
Ajout d'un libellé à un modèle probabiliste.	56
Suppression d'un libellé d'un modèle probabiliste.	56
Compilation du modèle probabiliste.	57
Propriétés du modèle probabiliste.	57
Création d'un modèle probabiliste à partir d'un objet de données.	57
Opérations de copier/coller.	59
Copie d'un modèle probabiliste dans un autre ensemble de contenu.	59
Importation d'un modèle probabiliste depuis un autre ensemble de contenu.	59

Index..... 60

Préface

Le Guide des données de référence comprend des informations sur les objets de données de référence dans Informatica Data Quality. Il est rédigé pour les utilisateurs professionnels tels que les analystes et les gestionnaires de données qui travaillent sur les projets au sein d'une organisation. Le guide décrit des objets de données de référence dans l'outil Developer, l'outil Analyst, et le référentiel modèle.

Ressources Informatica

Portail Mon support Informatica

En tant que client Informatica, vous pouvez accéder au portail Mon support Informatica sur <http://mysupport.informatica.com>.

Ce site contient des informations sur les produits et les groupes d'utilisateurs, des bulletins d'information, un lien vers le système de gestion des dossiers d'assistance à la clientèle d'Informatica (ATLAS), une bibliothèque de procédures Informatica, une base de connaissances Informatica, ainsi que la documentation nécessaire sur les produits Informatica et l'accès à sa communauté d'utilisateurs.

Documentation Informatica

L'équipe Documentation d'Informatica s'efforce de fournir une documentation précise et utilisable. N'hésitez pas à contacter l'équipe Documentation d'Informatica par courriel à l'adresse infa_documentation@informatica.com pour lui faire part de vos questions, commentaires ou suggestions concernant cette documentation. Ces commentaires et suggestions nous permettront d'améliorer notre documentation. Veuillez préciser si vous acceptez d'être contacté au sujet de ces commentaires.

L'équipe Documentation met à jour la documentation chaque fois que nécessaire. Pour obtenir la toute dernière version de la documentation concernant votre produit, consultez la Documentation de produit sur <http://mysupport.informatica.com>.

Matrices de disponibilité de produit Informatica

Les matrices de disponibilité de produit (PAM) indiquent les versions des systèmes d'exploitation, les bases de données et les autres types de sources et cibles de données pris en charge par une version d'un produit. Vous pouvez consulter les PAM sur le portail Mon Support Informatica à l'adresse <https://mysupport.informatica.com/community/my-support/product-availability-matrices>.

Site Web Informatica

Vous pouvez accéder au site Web d'entreprise Informatica sur <http://www.informatica.com>. Le site contient des informations sur Informatica, son expertise, les événements à venir et les bureaux de vente. Vous y trouverez aussi des informations sur ses produits et ses partenaires. Les rubriques de service du site fournissent des informations importantes sur le support technique, la formation et l'éducation, ainsi que les services d'implémentation.

Bibliothèque de procédures Informatica

En tant que client Informatica, vous avez accès à la bibliothèque de procédures Informatica sur <http://mysupport.informatica.com>. La bibliothèque de procédures Informatica est une collection de ressources destinée à vous familiariser avec les produits Informatica et leurs fonctionnalités. Elle regroupe des articles et des démonstrations interactives qui permettent de résoudre des problèmes courants et de comparer les fonctionnalités et les comportements, et qui vous guident lors de la réalisation de tâches concrètes spécifiques.

Base de connaissances Informatica

En tant que client Informatica, vous avez accès à la base de connaissances Informatica sur <http://mysupport.informatica.com>. Utilisez la base de connaissances pour rechercher des solutions documentées aux problèmes techniques connus concernant les produits Informatica. Vous y trouverez également la réponse aux questions les plus fréquentes, des livres blancs et des conseils techniques. N'hésitez pas à contacter l'équipe Base de connaissances Informatica par courriel à l'adresse KB_Feedback@informatica.com pour lui faire part de vos questions, commentaires et suggestions concernant la base de connaissances.

Canal YouTube du support Informatica

Vous pouvez accéder au canal YouTube du support Informatica sur <http://www.youtube.com/user/INFASupport>. Le canal YouTube du support Informatica contient des vidéos concernant les solutions qui vous guident dans l'exécution de tâches spécifiques. Si vous avez des questions, commentaires ou suggestions concernant le canal YouTube du support Informatica, contactez l'équipe de support YouTube par courriel à l'adresse supportvideos@informatica.com ou envoyez un tweet à @INFASupport.

Informatica Marketplace

Informatica Marketplace est un forum où développeurs et partenaires peuvent partager des solutions qui permettent d'augmenter, d'étendre ou d'améliorer les implémentations d'intégration de données. En tirant profit des centaines de solutions disponibles sur Marketplace, vous pouvez améliorer votre productivité et accélérer le temps d'implémentation de vos projets. Vous pouvez accéder à Informatica Marketplace à l'adresse <http://www.informaticamarketplace.com>.

Informatica Velocity

Vous pouvez accéder à Informatica Velocity à l'adresse <http://mysupport.informatica.com>. Développé à partir de l'expérience concrète de centaines de projets de gestion de données, Informatica Velocity représente le savoir collectif de nos consultants, qui ont travaillé avec des entreprises du monde entier pour planifier, développer, déployer et tenir à jour des solutions de gestion des données efficaces. Si vous avez des questions, des commentaires et des suggestions sur Informatica Velocity, contactez le support des services professionnels Informatica à l'adresse ips@informatica.com.

Support client international Informatica

Vous pouvez contacter un centre de support client par téléphone ou via l'assistance en ligne.

L'assistance en ligne requiert un nom d'utilisateur et un mot de passe. Vous pouvez demander un nom d'utilisateur et un mot de passe sur <http://mysupport.informatica.com>.

Les numéros de téléphone du support client international Informatica sont disponibles sur le site Web Informatica à l'adresse

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

CHAPITRE 1

Introduction aux données de référence

Ce chapitre comprend les rubriques suivantes :

- [Présentation des données de référence, 10](#)
- [Données de référence Informatica, 11](#)
- [Données de référence définies par l'utilisateur, 12](#)
- [Tables de référence, 12](#)

Présentation des données de référence

Un objet de données de référence identifie un ensemble de valeurs de données que vous pouvez sélectionner lorsque vous configurez les transformations pour les opérations de qualité des données. Vous pouvez créer des objets de données de référence dans l'outil Developer et dans l'outil Analyst. Vous pouvez également importer des objets de données de référence dans le référentiel modèle et dans le système de fichiers. Vous pouvez utiliser le programme d'installation de contenu Data Quality pour importer des objets de données de référence et installer les données de référence dans le système de fichiers.

Vous pouvez créer et modifier les types suivants de données de référence :

Tables de référence

Une table de référence contient la version standard et les versions alternatives d'un ensemble de valeurs de données. Vous pouvez ajouter une table de référence à une transformation dans l'outil Developer afin de vérifier que les valeurs de données sources sont précises et correctement mises en forme.

Une table de base de données contient au moins deux colonnes. Une colonne contient la version standard ou souhaitée d'une chaîne. Les autres colonnes contiennent des versions alternatives. Lorsque vous ajoutez une table de référence à une transformation, la transformation recherche les données de port d'entrée qui apparaissent également dans la table. Vous pouvez créer des tables avec n'importe quelles données utiles au projet de données sur lequel vous travaillez.

Ensembles de contenu

Un ensemble de contenu est un objet du référentiel modèle qui indique les valeurs de données de référence dans le référentiel ou dans un fichier. Lorsque vous ajoutez un ensemble de contenu à une transformation, la transformation recherche les données d'entrée pour les valeurs qui correspondent aux modèles de données dans l'ensemble de contenu.

Le programme d'installation de contenu Data Quality peut installer les types de données de référence suivants :

Tables de référence Informatica

Objets de référentiel et fichiers de données développés par Informatica. Vous importez les tables de référence Informatica lorsque vous importez des objets accélérateurs dans le référentiel modèle. Les types d'informations de référence incluent des indicatifs régionaux téléphoniques, des formats de codes postaux, des prénoms, des formats de numéros de sécurité sociale, des métiers et des acronymes. Vous pouvez modifier les tables de référence Informatica.

Ensembles de contenu Informatica

Objets de référentiel et fichiers de données développés par Informatica. Vous importez des ensembles de contenu lorsque vous importez des objets accélérateurs dans le référentiel modèle. Un ensemble de contenu contient différents types de données de référence que vous pouvez utiliser pour effectuer des opérations de recherche à l'aide des transformations de qualité des données.

Fichiers de données de référence pour adresses

Fichiers de données de référence qui contiennent des données pour les adresses livrables d'un pays. La transformation Outil de validation des adresses lit les données de référence. Vous ne pouvez pas créer ni modifier des fichiers de données de référence pour adresses.

Les données de référence pour adresses sont valides pour une période définie et vous devez actualiser vos données de manière régulière, par exemple chaque trimestre. Vous ne pouvez pas afficher ni modifier des données de référence pour adresses.

Fichiers de population d'identité

Fichiers de données de référence qui contiennent des informations sur les identités personnelles, familiales et professionnelles. La transformation Correspondance et la transformation Comparaison utilisent des fichiers de population pour trouver des identités potentielles dans les données d'entrée. Vous ne pouvez pas créer ni modifier des fichiers de données de population d'identité.

Données de référence Informatica

Vous pouvez acheter et télécharger des données de référence d'adresses et des données de population d'identité à partir d'Informatica.

Vous pouvez acheter un abonnement annuel pour les données d'adresse d'un pays et vous pouvez télécharger les données d'adresses les plus récentes à partir d'Informatica à tout moment pendant la période d'inscription.

L'utilisateur de l'installateur de contenu doit télécharger et installer séparément les données de référence et les applications. Pour des informations sur les données de référence installées dans votre système, contactez votre administrateur.

Données de référence définies par l'utilisateur

Vous pouvez utiliser les valeurs d'un objet de données pour créer un objet de données de référence.

Vous pouvez par exemple sélectionner un objet de données ou une colonne de profil qui contient des valeurs spécifiques à un projet ou à une organisation. Créez les objets de données de référence personnalisés depuis les valeurs de colonne.

Vous pouvez construire un objet de données de référence depuis une colonne de données pour vérifier les éléments suivants :

- Les lignes de données des colonnes contiennent le même type d'informations.
- Une valeur source est valide. L'objet de référence peut contenir une liste de valeurs valides ou qui ne le sont pas.

Le tableau suivant répertorie des exemples courants de colonnes de données de projets qui peuvent contenir des données de référence :

Information	Exemple de données de référence
Codes d'unités de conservation de stock (SKU)	Utilisez une colonne SKU pour créer une table de référence de codes SKU valides pour une organisation. Utilisez la table de référence pour trouver des codes SKU corrects ou incorrects dans un jeu de données.
Codes Employé	Utilisez un code Employé ou une colonne d'identifiant Employé pour créer une table de référence de codes Employé valides. Utilisez la table de référence pour rechercher des erreurs dans les données Employé.
Numéros de compte Client	Exécutez un profil sur une colonne de compte Client pour identifier un modèle de numéro de compte. Utilisez le profil pour créer un jeu de jetons de modèles de données incorrects. Utilisez le jeu de jetons pour trouver des numéros de compte qui ne sont pas conformes à la structure de numéro de compte correcte.
Noms Client	Lorsqu'un nom Client contient un prénom, un second prénom et un nom de famille, vous pouvez créer un modèle probabiliste qui définit la structure attendue des chaînes de la colonne. Utilisez le modèle probabiliste pour trouver des chaînes de données qui n'appartiennent pas à la colonne.

Tables de référence

Vous pouvez ajouter une table de référence à une transformation dans l'outil Developer. Vous configurez la transformation pour trouver des valeurs de la table de référence dans les données d'entrée et pour écrire en sortie les valeurs valides correspondantes dans la table de référence.

Pour créer une table de référence dans l'outil Developer, utilisez l'une des méthodes suivantes :

- Créer une table de référence vide et entrer les valeurs de données.
- Créer une table de référence à partir des données d'un fichier plat.
- Créer une table de référence à partir des données d'une autre table de base de données.

Structure des tables de référence

La plupart des tables de référence contiennent au moins deux colonnes. Une colonne contient les versions correctes ou requises des valeurs de données. Les autres colonnes contiennent des versions différentes des valeurs, y compris des versions alternatives qui peuvent s'afficher dans les données source.

La colonne qui contient les valeurs correctes ou requises est appelée la colonne valide. Lorsqu'une transformation lit une table de référence dans un mappage, elle cherche les valeurs dans les colonnes non valides. Lorsque la transformation trouve une valeur non valide, elle renvoie la valeur correspondante de la colonne valide. Vous pouvez également configurer une transformation pour renvoyer une seule valeur commune à la place des valeurs valides.

La colonne valide peut contenir des données en bonne et due forme, par exemple des codes postaux. Elle peut contenir des données relatives à un projet, comme par exemple le nombre d'unités de conservation de stock (SKU) propres à une organisation. Vous pouvez également créer une colonne valide à partir de données incorrectes, comme des valeurs qui contiennent des erreurs de données connues que vous voulez rechercher.

Par exemple, vous créez une table de référence qui contient une liste de numéros SKU valides dans une organisation de vente au détail. Vous ajoutez la table de référence à une transformation Libellé et créez un mappage avec la transformation. Vous exécutez le mappage avec une table de base de données de produits. Lorsque le mappage est exécuté, la transformation Libellé crée une colonne qui identifie les enregistrements de produit ne contenant pas de nombres de SKU valides.

Tables de référence et transformation Analyseur

Vous créez une table de référence avec une colonne unique lorsque vous voulez utiliser les données de la table dans une opération d'analyse basée sur un modèle. Vous configurez la transformation Analyseur pour effectuer une analyse basée sur un modèle et vous importez les données vers la configuration de la transformation.

Tables de référence gérées et non gérées

Les tables de référence stockent les métadonnées dans le référentiel modèle. Les tables de référence peuvent stocker des données de colonne dans l'entrepôt de données de référence ou dans une autre base de données.

Lorsque l'entrepôt de données de référence stocke les données de colonne, les services Informatica identifient la table comme une table de référence gérée. Lorsqu'une autre base de données stocke les données de colonne, les services Informatica identifient la table comme une table de référence non gérée. Le service de gestion de contenu utilise la connexion de la base de données pour l'entrepôt de données de référence.

Pour tenir à jour les tables de référence, vous devez posséder les privilèges suivants dans le service de gestion de contenu :

- Créer des tables de référence
- Éditer les données de la table de référence
- Éditer les métadonnées de la table de référence

Utilisez les options de sécurité de l'outil Administrator pour vérifier ou mettre à jour les privilèges du service. Pour éditer les données d'une table de référence non gérée, vérifiez également que vous avez configuré l'objet de la table de référence pour qu'elle autorise les éditions.

Remarque: Si vous éditez les métadonnées pour une table de référence non gérée dans une application de base de données, utilisez l'outil Analyst pour synchroniser le référentiel modèle avec la table de base de données. Vous devez synchroniser le référentiel modèle et la table de base de données avant d'utiliser la table de référence non gérée dans l'outil Developer.

CHAPITRE 2

Tables de référence dans l'outil Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des tables de référence de l'outil Analyst, 14](#)
- [Propriétés des tables de référence, 15](#)
- [Création d'une table de référence dans l'éditeur de table de référence, 16](#)
- [Créer une table de référence à partir des données de profil, 16](#)
- [Créer une table de référence à partir d'un fichier plat, 20](#)
- [Créer une table de référence à partir d'une table de base de données, 21](#)
- [Mises à jour de la table de référence, 22](#)
- [Événements suivi d'audit, 26](#)
- [Règles et instructions pour les tables de référence, 27](#)

Présentation des tables de référence de l'outil Analyst

Créez des tables de référence dans l'espace de travail Conception de l'outil Analyst.

Vous pouvez créer une table de référence à partir d'un fichier plat, à partir d'une source de données du référentiel modèle et à partir d'une table d'une autre base de données.

Vous pouvez créer une table de référence depuis une colonne de profil ou un sous-ensemble de données dans une colonne de profil. Vous pouvez également créer une table de référence depuis les modèles de colonne que vous choisissez à partir d'un profil.

Lorsque vous créez ou mettez à jour une table de référence, vous configurez les propriétés de la table et des colonnes de données qu'elle contient.

Propriétés des tables de référence

Vous pouvez afficher et éditer les propriétés de table de référence dans l'outil Analyst. Une table de référence affiche des propriétés générales décrivant l'objet de référentiel et les propriétés de colonne décrivant les données de colonne.

Pour afficher les propriétés, ouvrez la table de référence et sélectionnez le panneau **Propriétés**.

Pour modifier les propriétés, ouvrez la table de référence et cliquez sur **Actions > Éditer la table**.

Propriétés de la colonne Table de référence

Les propriétés de colonne contiennent des informations sur les métadonnées de la colonne.

Le tableau suivant décrit les propriétés de colonne :

Propriété	Description
Nom	Nom de chaque colonne.
Type de données	Type de données pour les données de chaque colonne. Vous pouvez sélectionner l'un des types de données suivants : <ul style="list-style-type: none">- bigint- date/heure- décimal- double- entier- chaîne <p>Vous ne pouvez pas sélectionner un type de données double lorsque vous créez une table de référence vide ou créez une table de référence à partir d'un fichier plat.</p>
Précision	Précision pour chaque colonne. La précision est le nombre maximal de chiffres ou de caractères que la colonne peut contenir. Les valeurs de précision que vous configurez dépendent du type de données.
Échelle	Échelle pour chaque colonne. L'échelle est le nombre maximal de chiffres que la colonne peut contenir à droite de la virgule. S'applique aux colonnes décimales. Les valeurs d'échelle que vous configurez dépendent du type de données.
Description	Description facultative pour chaque colonne.
Valeur Null autorisée	Indique si la colonne peut contenir des valeurs Null.
Clé	Identifie une colonne de clé. L'outil Analyst peut identifier une colonne de clé si vous importez les données de référence à partir d'une table qui spécifie une colonne de clé.

Création d'une table de référence dans l'éditeur de table de référence

Définissez la structure de la table et ajoutez des données à une table de référence dans l'éditeur de table de référence.

1. Cliquez sur **Conception > Table de référence**.
2. Sélectionnez l'option pour **Utiliser l'éditeur de table de référence**.
3. Cliquez sur **Suivant**.
4. Cliquez sur **Ajouter une colonne**.
L'outil Analyst ajoute une ligne pour entrer les attributs de colonne.
5. Pour chaque colonne que vous voulez inclure dans la table de référence, cliquez sur l'icône **Ajouter une colonne** et configurez les propriétés de chaque colonne.
Remarque: Vous pouvez réorganiser des colonnes ou en supprimer.
6. Éventuellement, saisissez une note d'audit pour la table.
La note d'audit apparaît dans le journal de suivi d'audit.
7. Entrez le nom et l'emplacement de la table de référence.
8. Cliquez sur **Terminer**.

Créer une table de référence à partir des données de profil

Vous pouvez utiliser les données de profil pour créer des tables de référence liées aux données source contenues dans le profil. Utilisez les tables de référence pour trouver les différents types d'informations dans les données source.

Vous pouvez utiliser un profil pour créer ou mettre à jour une table de référence selon les méthodes suivantes :

- Sélectionnez une colonne dans le profil et ajoutez-la à une table de référence.
- Parcourez une colonne de profil et ajoutez un sous-ensemble des données de colonne à une table de référence.
- Sélectionnez une colonne dans le profil et ajoutez les valeurs de forme pour cette colonne à une table de référence.

LIENS CONNEXES :

- [“Génération de données de référence depuis un profil intermédiaire ” à la page 40](#)

Créer une table de référence depuis les colonnes de profil

Vous pouvez créer une table de référence depuis une colonne de profil. Vous pouvez ajouter une colonne de profil à une table de référence existante. L'assistant **Nouvelle table de référence** ajoute la colonne à la table de référence.

1. Ouvrez l'espace de travail **Bibliothèque** dans l'outil Analyst.
2. Sélectionnez l'actif **Profils**.
Une liste de profils s'affiche.
3. Recherchez le profil avec la colonne que vous voulez ajouter à une table de référence.
4. Cliquez sur **Actions > Ouvrir**.
5. Sélectionnez la colonne que vous voulez ajouter à une table de référence.
6. Cliquez sur **Actions > Ajouter à la table de référence**.
7. Sélectionnez l'option pour **Créer une nouvelle table de référence**.
Éventuellement, sélectionnez **Ajouter à la table de référence existante**.
8. Cliquez sur **Suivant**.
Le nom de la colonne apparaît par défaut comme nom de la table.
9. Éventuellement, saisissez un autre nom de table, une description et une valeur par défaut.
L'outil Analyst utilise la valeur par défaut pour tout enregistrement de la table qui ne contient pas de valeur.
10. Cliquez sur **Suivant**.
11. Dans le panneau **Attributs des colonnes**, configurez les propriétés de colonne.
12. Éventuellement, choisissez de créer une colonne de description pour les lignes dans la table de référence.
Saisissez le nom et la précision pour la colonne.
13. Prévisualisez les valeurs de la colonne dans le panneau **Prévisualiser**.
14. Cliquez sur **Suivant**.
Le nom de la colonne apparaît comme nom de table par défaut. Éventuellement, saisissez un autre nom de table et une description.
15. Dans le panneau **Enregistrer dans**, sélectionnez l'emplacement où vous voulez créer la table de référence.
Le panneau **Tables de référence** répertorie les tables de référence dans l'emplacement que vous sélectionnez.
16. Éventuellement, saisissez une note d'audit.
17. Cliquez sur **Terminer**.

Créer une table de référence depuis les valeurs de colonne

Vous pouvez créer une table de référence depuis les valeurs de colonnes dans une colonne de profil. Sélectionnez une colonne dans un profil et sélectionnez les valeurs de la colonne à rajouter dans la table de référence ou créez une table de référence à rajouter aux valeurs de la colonne.

1. Ouvrez l'espace de travail **Bibliothèque** dans l'outil Analyst.
2. Sélectionnez l'actif **Profils**.
Une liste de profils s'affiche.
3. Sélectionnez un profil pour l'ouvrir et cliquez sur **Actions > Ouvrir**.
4. Sélectionnez la colonne que vous voulez ajouter à une table de référence.
5. Dans le panneau **Valeurs**, sélectionnez les valeurs de colonne que vous voulez ajouter. Utiliser les touches CTRL ou SHIFT pour sélectionner plusieurs valeurs.
6. Cliquez sur **Actions > Ajouter à la table de référence**.
L'assistant **Nouvelle table de référence** s'affiche.
7. Sélectionnez l'option pour **Créer une nouvelle table de référence**.
En option, sélectionnez **Ajouter à la table de référence existante** et cliquez sur **Suivant**. Naviguez jusqu'à la table de référence dans le dossier ou projet, prévisualisez les données de la table de référence et cliquez sur **Suivant**. Sélectionnez la colonne à ajouter et cliquez sur **Terminer**.
8. Cliquez sur **Suivant**.
9. Le nom de la colonne apparaît par défaut comme nom de la table. Éventuellement, saisissez un autre nom de table, une description et une valeur par défaut.
L'outil Analyst utilise la valeur par défaut pour tout enregistrement de la table qui ne contient pas de valeur.
10. Cliquez sur **Suivant**.
11. Dans le panneau **Attributs des colonnes**, configurez les propriétés de colonne.
12. Éventuellement, choisissez de créer une colonne de description pour les lignes dans la table de référence.
Saisissez le nom et la précision pour la colonne.
13. Prévisualisez les valeurs de la colonne dans le panneau **Prévisualiser**.
14. Cliquez sur **Suivant**.
15. Le nom de la colonne apparaît comme nom de table par défaut. Éventuellement, saisissez un autre nom de table et une description.
16. Dans le panneau **Enregistrer dans**, sélectionnez l'emplacement où vous voulez créer la table de référence.
Le panneau **Tables de référence** liste les tables de référence dans l'emplacement que vous sélectionnez.
17. Éventuellement, saisissez une note d'audit.
18. Cliquez sur **Terminer**.

Créer une table de référence depuis les modèles de colonnes

Vous pouvez créer une table de référence depuis les modèles de colonnes dans une colonne de profil. Sélectionnez une colonne dans le profil et sélectionnez les valeurs modèles à rajouter dans la table de référence ou créez une table de référence à rajouter aux valeurs modèles.

1. Ouvrez l'espace de travail **Bibliothèque** dans l'outil Analyst.
2. Sélectionnez l'actif **Profils**.
Une liste de profils s'affiche.
3. Sélectionnez la colonne que vous voulez ajouter à une table de référence.
4. Dans le panneau **Modèles**, sélectionnez les modèles de colonnes que vous voulez ajouter. Utiliser les touches CTRL ou SHIFT pour sélectionner plusieurs valeurs.
5. Cliquez sur **Actions > Ajouter à la table de référence**.
L'assistant **Nouvelle table de référence** s'affiche.
6. Sélectionnez l'option pour **Créer une nouvelle table de référence**.
En option, sélectionnez **Ajouter à la table de référence existante** et cliquez sur **Suivant**. Naviguez jusqu'à la table de référence dans le dossier ou projet, prévisualisez les données de la table de référence et cliquez sur **Suivant**. Sélectionnez la colonne à ajouter et cliquez sur **Terminer**.
7. Cliquez sur **Suivant**.
8. Le nom de la colonne apparaît par défaut comme nom de la table. Éventuellement, saisissez un autre nom de table, une description et une valeur par défaut.
L'outil Analyst utilise la valeur par défaut pour tout enregistrement de la table qui ne contient pas de valeur.
9. Cliquez sur **Suivant**.
10. Dans le panneau **Attributs des colonnes**, configurez les propriétés de colonne.
11. Éventuellement, choisissez de créer une colonne de description pour les lignes dans la table de référence.
Saisissez le nom et la précision pour la colonne.
12. Prévisualisez les valeurs de la colonne dans le panneau **Prévisualiser**.
13. Cliquez sur **Suivant**.
14. Le nom de la colonne apparaît comme nom de table par défaut. Éventuellement, saisissez un autre nom de table et une description.
15. Dans le panneau **Enregistrer dans**, sélectionnez l'emplacement où vous voulez créer la table de référence.
Le panneau **Tables de référence** liste les tables de référence dans l'emplacement que vous sélectionnez.
16. Éventuellement, saisissez une note d'audit.
17. Cliquez sur **Terminer**.

Créer une table de référence à partir d'un fichier plat

Vous pouvez importer des données de référence depuis un fichier CSV. Utilisez l'assistant **Nouvelle table de référence** pour importer les données du fichier.

Vous devez configurer les propriétés de chaque fichier plat que vous utilisez pour créer une table de référence.

Propriétés de fichier plat de l'outil Analyst

Lorsque vous importez un fichier plat comme table de référence, vous devez configurer les propriétés pour chaque colonne de ce fichier. Les options que vous configurez déterminent la manière dont l'outil Analyst lit les données du fichier.

Le tableau suivant décrit les propriétés que vous pouvez configurer lorsque vous importez le fichier de données pour une table de référence :

Propriétés	Description
Délimiteurs	Caractères utilisés pour séparer les colonnes de données. Utilisez le champ Autre pour entrer un séparateur différent. Les délimiteurs doivent être des caractères imprimables et doivent être différents du caractère d'échappement et des guillemets s'ils sont sélectionnés. Vous ne pouvez pas sélectionner des caractères multi-octets non imprimables comme délimiteurs.
Qualificateur de texte	Guillemet qui définit les limites des chaînes de texte. Choisissez Aucun guillemet, Guillemet simple ou Guillemets doubles. Si vous sélectionnez un guillemet, l'assistant ignore les délimiteurs au sein des paires de guillemets.
Noms des colonnes	Importe les noms de colonne depuis la première ligne. Sélectionnez cette option si les noms de colonne apparaissent dans la première ligne. L'assistant utilise les données de la première ligne dans l'aperçu des noms de colonne. La valeur par défaut n'est pas activée.
Valeurs	Option pour démarrer l'importation d'une valeur à partir d'une ligne. Indique dans la prévisualisation le numéro de la ligne à partir duquel l'assistant démarre la lecture lorsqu'il importe le fichier.

Création d'une table de référence à partir d'un fichier simple

Lorsque vous créez des données de table de référence à partir d'un fichier simple, cette table utilise la structure de colonne du fichier et importe les données du fichier.

1. Cliquez sur **Conception > Table de référence** dans l'outil Analyst.
L'assistant **Nouvelle table de référence** s'affiche.
2. Sélectionnez l'option pour **Importer un fichier simple**.
3. Cliquez sur **Suivant**.
4. Cliquez sur **Parcourir** pour sélectionner un fichier simple.
5. Cliquez sur **Charger** pour envoyer le fichier vers un répertoire dans le répertoire d'installation des services Informatica accessible par l'outil Analyst.

6. Entrez le nom de table. Entrez éventuellement une description et une valeur par défaut.
L'outil Analyst utilise la valeur par défaut pour tout enregistrement de la table qui ne contient pas de valeur.
7. Sélectionnez une page de code qui correspond aux données dans le fichier simple.
8. Prévisualisez les données dans le panneau **Prévisualisation du fichier**.
9. Cliquez sur **Suivant**.
10. Configurez les propriétés du fichier simple.
11. Dans le panneau **Prévisualisation**, cliquez sur **Afficher** pour actualiser la prévisualisation.
12. Cliquez sur **Suivant**.
13. Dans le panneau **Attributs des colonnes**, vérifiez ou modifiez les propriétés de colonne pour chaque colonne.
14. Créez éventuellement une colonne de description pour les lignes dans la table de référence. Saisissez le nom et la précision pour la colonne.
15. Saisissez éventuellement une note d'audit pour la table.
16. Cliquez sur **Terminer**.

Créer une table de référence à partir d'une table de base de données

Lorsque vous créez une table de référence à partir d'une table de base de données, vous vous connectez à la base de données et importez les données de la table.

Avant d'importer les données de la table de référence depuis une table de base de données, vérifiez que le domaine Informatica contient une connexion à la base de données. Si le domaine ne contient pas de connexion à la base de données, vous pouvez en définir une dans l'outil Analyst.

Pour créer une connexion à la base de données, cliquez sur **Gérer > les connexions**.

Création d'une table de référence à partir d'une table de base de données

Pour créer la table de référence, connectez-vous à une base de données et importez les données.

1. Cliquez sur **Conception > Table de référence** dans l'outil Analyst.
L'assistant **Nouvelle table de référence** s'affiche.
2. Sélectionnez l'option **Connecter à une table relationnelle**.
3. Sélectionnez **Table non gérée** pour créer une table qui ne stocke pas de données dans l'entrepôt de données de référence.
Pour effectuer des opérations d'édition sur une table de référence non gérée, sélectionnez l'option **Editable**.
Cliquez sur **Suivant**.
4. Sélectionnez la connexion de base de données dans la liste des connexions établies.
Cliquez sur **Suivant**.

5. Dans le panneau **Tables**, sélectionnez une table.
6. Vérifiez les propriétés de la table dans le panneau **Propriétés**.
Cliquez sur **Prévisualisation des données** pour prévisualiser les données de la table.
Cliquez sur **Suivant**.
7. Dans le panneau **Attributs des colonnes**, sélectionnez la colonne valide.
Éventuellement, spécifiez un commentaire d'audit à écrire dans le suivi d'audit lorsqu'un utilisateur met à jour la table de référence.
Si vous créez une table de référence gérée, vous pouvez effectuer les actions suivantes dans le panneau **Attributs des colonnes** :
 - Éditer les noms des colonnes de la table de référence.
 - Ajouter une colonne de métadonnées pour les descriptions de niveau de ligne.
8. Cliquez sur **Suivant**.
9. Entrez un nom et éventuellement une description pour la table de référence.
10. Dans le panneau **Dossiers**, indiquez le projet ou le dossier dans lequel la table de référence doit être stockée.
Le panneau **Tables de référence** répertorie les tables de référence dans le dossier que vous sélectionnez.
11. Cliquez sur **Terminer**.

Mises à jour de la table de référence

Les données métier que contient une table de référence peuvent changer avec le temps. Examinez et mettez à jour les données et métadonnées d'une table de référence pour vérifier que la table contient des informations correctes. Mettez les tables de référence à jour dans l'outil Analyst. Vous pouvez mettre à jour les données et métadonnées d'une table de référence gérée ou non gérée.

Vous pouvez effectuer les opérations suivantes sur les données et métadonnées de table de référence :

Gérer les colonnes

Vous pouvez ajouter et supprimer des colonnes, et éditer les propriétés de colonne.

Gérer les lignes

Vous pouvez ajouter des lignes de données à une table de référence.

Éditer les valeurs de données de référence

Vous pouvez éditer une valeur de données de référence.

Remplacer les valeurs de données

Utilisez l'option **Rechercher et remplacer** pour remplacer les valeurs de données qui ne sont plus correctes ou pertinentes pour l'organisation. Vous pouvez trouver une valeur dans une colonne et la remplacer par une autre valeur. Vous pouvez remplacer toutes les valeurs d'une colonne par une valeur unique.

Exporter une table de référence

Exportez une table de référence vers un fichier de valeurs séparées par des virgules (CSV), un fichier dictionnaire ou un fichier Excel.

Activer ou désactiver les éditions sur une table non gérée

Mettez à jour une table de référence non gérée pour activer ou désactiver les éditions des données et métadonnées d'une table.

Actualiser les données de la table de référence

Rechargez les données de la table de référence dans l'outil Analyst pour afficher les dernières modifications dans les données.

Gérer les colonnes

Vous pouvez changer les colonnes dans une table de référence. Vous pouvez également définir les propriétés des tables de référence.

1. Cliquez sur **Conception > Ouvrir**.
2. Choisissez l'actif Tables de référence.
3. Cliquez sur le nom de la table de référence pour l'ouvrir dans le panneau **Table de référence**.
4. Cliquez sur **Actions > Éditer la table**.

La boîte de dialogue **Editer les propriétés des colonnes** s'affiche. Utilisez les options de cette boîte de dialogue pour effectuer les opérations suivantes :

- Modifier la colonne valide de la table.
 - Supprimer une colonne de la table.
 - Modifier un nom de colonne.
 - Mettre à jour le texte descriptif d'une colonne.
 - Mettre à jour le statut d'édition de la table de référence.
 - Mettre à jour la note d'audit de la table. La note d'audit apparaît dans le fichier journal d'audit pour toute action que vous effectuez dans la fenêtre **Editer les propriétés des colonnes**.
5. Lorsque vous avez terminé les opérations à effectuer, cliquez sur **OK**.

Gérer les lignes

Vous pouvez ajouter, éditer ou supprimer les lignes dans une table de référence.

1. Cliquez sur **Conception > Ouvrir**.
2. Choisissez l'actif Tables de référence.
3. Cliquez sur le nom de la table de référence pour l'ouvrir dans le panneau **Table de référence**.
4. Éditez les lignes de données. Vous pouvez éditer les lignes de données en procédant comme suit :

- Pour ajouter une ligne, sélectionnez **Actions > Ajouter une ligne**.

Dans la fenêtre **Ajouter une ligne**, entrez une valeur pour chaque colonne. Éventuellement, saisissez une note d'audit.

Cliquez sur **OK** pour appliquer les modifications.

- Pour éditer une valeur de données, double-cliquez sur la valeur dans la table de référence et mettez-la à jour.
Après avoir édité les données, utilisez les options de ligne pour accepter ou rejeter l'édition.
- Pour éditer plusieurs lignes, sélectionnez les lignes concernées et cliquez sur **Actions > Editer**.

Dans la fenêtre **Modifier plusieurs lignes**, entrez une valeur pour chaque colonne de la ligne. Éventuellement, saisissez une note d'audit.

Cliquez sur **OK** pour appliquer les modifications pour chaque ligne.

- Pour supprimer des lignes, sélectionnez les lignes concernées et cliquez sur **Actions > Supprimer**.

Dans la fenêtre **Supprimer les lignes**, vous pouvez saisir une note d'audit.

Cliquez sur **OK** pour supprimer les données.

Remarque: Utilisez l'outil Developer pour éditer des données de ligne dans une grande table de référence. Par exemple, si une table de référence contient plus de 500 lignes, éditez-la dans l'outil Developer.

Rechercher et remplacer les valeurs

Vous pouvez rechercher et remplacer des valeurs de données dans une table de référence. Utilisez les options de recherche et remplacement lorsqu'une table contient une ou plusieurs instances d'une valeur de données que vous devez mettre à jour.

1. Cliquez sur **Conception > Ouvrir**.
2. Choisissez l'actif Tables de référence.
3. Cliquez sur le nom de la table de référence pour l'ouvrir dans le panneau **Table de référence**.
4. Cliquez sur **Actions > Rechercher et remplacer**.
La barre d'outils **Rechercher et remplacer** apparaît.
5. Entrez les critères de recherche dans la barre d'outils :
 - Entrez une valeur de données dans le champ **Rechercher**.
 - Sélectionnez les colonnes dans lesquelles effectuer la recherche. Par défaut, la recherche porte sur toutes les colonnes.
 - Entrez une valeur de données dans le champ **Remplacer par**.
6. Recherchez dans les colonnes sélectionnées la valeur de données indiquée dans le champ **Rechercher**.
Utilisez les options suivantes pour effectuer un remplacement individuel ou global des valeurs :
 - Utilisez les options **Suivant** et **Précédent** pour rechercher les valeurs une par une.
 - Pour remplacer une valeur, sélectionnez **Remplacer**.
 - Utilisez l'option **Tout mettre en surbrillance** pour afficher toutes les instances de la valeur.
 - Pour remplacer toutes les instances de la valeur, sélectionnez **Remplacer tout**.

Exporter une table de référence

Exportez une table de référence vers un fichier de valeurs séparées par des virgules, un fichier dictionnaire ou un fichier Microsoft Excel.

1. Cliquez sur **Conception > Ouvrir**.
2. Choisissez l'actif Tables de référence.
3. Cliquez sur le nom de la table de référence pour l'ouvrir dans le panneau **Table de référence**.
4. Cliquez sur **Actions > Exporter les données**.
La fenêtre **Exporter les données dans un fichier** apparaît.

Le tableau suivant décrit les options de la fenêtre **Exporter les données vers un fichier** :

Option	Description
Nom de fichier	Nom de fichier pour les données exportées.
Format de fichier	Format du fichier exporté. Vous pouvez sélectionner les formats suivants : <ul style="list-style-type: none">• csv. Fichier de valeurs séparées par des virgules.• xls. Fichier Microsoft Excel.• dic. Fichier dictionnaire Informatica.
Exporter noms champs comme première ligne	Option des noms des colonnes. Sélectionnez cette option pour indiquer que la première ligne du fichier contient les noms des colonnes.
Page de code	Page de code des données de référence. La page de code par défaut est UTF-8.

5. Cliquez sur **OK** pour exporter le fichier.

Activer et désactiver les éditions d'une table de référence non gérée

Vous pouvez configurer une table de référence non gérée de sorte à activer ou désactiver la mise à jour de ses valeurs de données et de ses colonnes.

Enregistrez la table de référence avant de modifier son statut d'édition.

1. Cliquez sur **Conception > Ouvrir**.
2. Choisissez l'actif Tables de référence.
3. Cliquez sur le nom de la table de référence pour l'ouvrir dans le panneau **Table de référence**.
4. Cliquez sur **Actions > Éditer la table**.

La fenêtre **Editer les propriétés des colonnes** apparaît.

5. Sélectionnez ou désélectionnez l'option **Editable**.

Lorsque vous changez le statut d'édition de la table de référence, la boîte de dialogue des propriétés se ferme.

Actualiser les valeurs de la table de référence

Utilisez l'option **Actualiser** pour mettre à jour la table de référence en fonction des dernières valeurs dans l'outil Analyst.

Vous devrez peut-être actualiser les valeurs de la table de référence qui s'affichent dans l'outil Analyst. Par exemple, il est possible que quelqu'un modifie la table de référence dans l'outil Developer et que vous deviez afficher les dernières modifications. Rechargez les données de la table de référence avec l'option **Actualiser**.

Pour recharger les valeurs de la table de référence, cliquez sur **Actions > Actualiser**.

Événements suivi d'audit

Vous pouvez afficher un suivi d'audit des modifications que vous apportez aux tables de référence. L'outil Analyst crée des événements de journal de suivi d'audit quand vous apportez un changement à une table de référence et entrez une note de suivi d'audit.

Les événements de journal de suivi d'audit fournissent des informations sur les tables de référence que vous gérez. Vous pouvez configurer des options de requête dans l'onglet Suivi d'audit pour filtrer les événements de journal que vous voyez. Vous pouvez configurer des filtres sur la plage de date, le type, le nom d'utilisateur et le statut.

Le tableau suivant décrit les options de filtre que vous configurez lorsque vous affichez les événements de journal de suivi d'audit :

Option	Description
Date	Dates de début et de fin des événements de journal à rechercher. Utilisez le calendrier pour définir les dates.
Type	Type d'événement de suivi d'audit. Vous pouvez filtrer et afficher les types d'événements suivants : <ul style="list-style-type: none">- Données. Événements associés aux données de la table de référence. Ils comprennent les opérations d'ajout, de suppression et de mise à jour de ligne.- Métadonnées. Événements associés aux métadonnées de la table de référence. Ils comprennent les opérations de création, d'ajout ou de suppression de colonne, ainsi que la sélection d'une colonne valide.
Utilisateur	Utilisateur qui a modifié la table de référence et entré le commentaire relatif au suivi d'audit. Le filtre affiche le nom complet et le nom de connexion de l'utilisateur.
Statut	Statut des événements du journal de suivi d'audit. Le statut correspond à l'action que vous avez effectuée dans l'éditeur de la table de référence.

Les événements de journal de suivi d'audit comprennent également les commentaires de suivi d'audit et les valeurs de colonnes que vous avez insérées, actualisées ou supprimées.

Voir les événements de suivi d'audit

Voir les événements de journal de suivi d'audit pour obtenir plus d'informations sur les changements apportés à la table de référence.

1. Cliquez sur **Conception > Ouvrir** et sélectionnez **Tables de référence**.
2. Cliquez sur un nom de table de référence pour l'ouvrir dans un panneau. L'onglet **Table de référence** apparaît.
3. Cliquez sur **Suivi d'audit**.
4. Configurez les options des filtres.
5. Cliquez sur **Afficher**.

Les événements de journal pour les options de requêtes spécifiées apparaissent.

Règles et instructions pour les tables de référence

Utilisez les règles et directives suivantes lorsque vous travaillez avec les tables de référence de l'outil Analyst :

- Lorsque vous importez une table de référence depuis une base de données Oracle, IBM DB2, IBM DB2/zOS, IBM DB2/iOS ou Microsoft SQL Server, l'outil Analyst ne peut pas afficher l'aperçu si les noms de table, d'affichage, de schéma, de synonyme et de colonne contiennent des caractères à casse mixte ou des caractères en minuscule.

Si la table de référence que vous affichez change dans un autre client, utilisez l'option **Actualiser** pour recharger la table de référence et afficher les modifications dans l'outil Analyst.

Pour prévisualiser les données des tables qui résident dans des bases de données sensibles à la casse, définissez l'attribut Prise en charge des identifiants à casse mixte sur True. Définissez l'attribut sur True dans les connexions pour les bases de données Oracle, IBM DB2, IBM DB2/zOS, IBM DB2/iOS et Microsoft SQL Server dans l'outil Developer ou Administrator.

- Quand vous créez une table de référence à partir de modèles de colonnes extrapolés dans un format, l'outil Analyst renseigne la table de référence avec des modèles de colonnes dans un format différent.

Par exemple, quand vous créez une table de référence pour la colonne modèle X(5), l'outil Analyst affiche le format suivant pour le modèle de colonne dans la table de référence : xxxxx.

- Lorsque vous importez une table de base de données Oracle, vérifiez la longueur de la colonne VARCHAR2 de la table. L'outil Analyst ne peut pas importer une table de base de données Oracle qui contient une colonne VARCHAR2 dont la longueur est supérieure à 1000.
- Pour lire une table de référence, vous devez exécuter les autorisations pour la connexion à la base de données qui stocke les valeurs des données de la table. Par exemple, si l'entrepôt de données de référence stocke les valeurs de données, vous devez disposer d'autorisations d'exécution sur la connexion à l'entrepôt de données de référence. Vous devez disposer d'autorisations d'exécution pour accéder en lecture ou en écriture à la table de référence. Les autorisations de connexion à la base de données s'appliquent à toutes les données de référence dans la base de données.

CHAPITRE 3

Données de référence dans l'outil Developer

Ce chapitre comprend les rubriques suivantes :

- [Présentation des données de référence de l'outil Developer, 28](#)
- [Données de référence et transformations, 29](#)
- [Tables de référence, 29](#)
- [Ensembles de contenu, 33](#)

Présentation des données de référence de l'outil Developer

Vous pouvez créer, mettre à jour et afficher les propriétés de configuration pour les objets de données de référence dans l'outil Developer.

Utilisez l'outil Developer pour créer et mettre à jour les types d'objets suivants :

Tables de référence

Une table de référence contient la version standard et les versions alternatives d'un ensemble de valeurs de données. Vous pouvez ajouter une table de référence à une transformation dans l'outil Developer afin de vérifier que les valeurs de données sources sont précises et correctement mises en forme.

Ensembles de contenu

Un ensemble de contenu est un objet du référentiel modèle qui indique les valeurs de données de référence dans le référentiel ou dans un fichier. Un ensemble de contenu contient différents types de données de référence que vous pouvez utiliser pour effectuer des opérations de recherche dans les transformations de qualité des données.

Vous pouvez également travailler avec des fichiers de données de référence d'adresse et des fichiers de population d'identité dans l'outil Developer. Vous sélectionnez les fichiers de données de référence d'adresse lorsque vous configurez une transformation Outil de validation des adresses. Vous sélectionnez les fichiers de population d'identité lorsque vous configurez une transformation Correspondance pour l'analyse de correspondance d'identité.

Données de référence et transformations

Plusieurs transformations lisent des données de référence pour effectuer des tâches de qualité des données.

Les transformations suivantes peuvent lire les données de référence :

- Outil de validation des adresses. Lit les données de référence d'adresses afin de vérifier l'exactitude des adresses.
- Convertisseur de casse. Lit des tables de données de référence pour identifier les chaînes dont la casse doit être modifiée.
- Classeur Lit des données d'ensemble de contenu pour identifier le type d'information d'une chaîne.
- Comparaison. Lit des données de la population d'identité lors de l'analyse de doublons.
- Libellé. Lit des données d'ensemble de contenu pour identifier et libeller des chaînes.
- Correspondance. Lit des données de la population d'identité lors de l'analyse de doublons.
- Analyseur. Lit des données d'ensemble de contenu pour analyser des chaînes en se basant sur les informations qu'elles contiennent.
- Normalisation. Lit les tables de données de référence pour normaliser des chaînes dans un format commun.

Le programme d'installation de Data Quality Content contient des objets de données de référence Informatica que vous pouvez importer.

Tables de référence

Vous pouvez ajouter une table de référence à une transformation dans l'outil Developer. Vous configurez la transformation pour trouver des valeurs de la table de référence dans les données d'entrée et pour écrire en sortie les valeurs valides correspondantes dans la table de référence.

Pour créer une table de référence dans l'outil Developer, utilisez l'une des méthodes suivantes :

- Créer une table de référence vide et entrer les valeurs de données.
- Créer une table de référence à partir des données d'un fichier plat.
- Créer une table de référence à partir des données d'une autre table de base de données.

Propriétés des données de la table de référence

Vous pouvez afficher les propriétés des données et métadonnées de la table de référence dans l'outil Developer. L'outil Developer affiche les propriétés lorsque vous ouvrez la table de référence depuis le référentiel modèle.

Une table de référence affiche des propriétés générales et des propriétés de colonne. Vous pouvez afficher les propriétés de table de référence dans l'outil Developer. Vous pouvez afficher et modifier les propriétés de table de référence dans l'outil Analyst.

Le tableau suivant décrit les propriétés générales d'une table de référence :

Propriété	Description
Nom	Nom de la table de référence.
Description	Description facultative de la table de référence.

Le tableau suivant décrit les propriétés de colonne d'une table de référence :

Propriété	Description
Valide	Identifie la colonne contenant les données de référence valides.
Nom	Nom de chaque colonne.
Type de données	Type de données de chaque colonne.
Précision	Précision de chaque colonne.
Échelle	Échelle de chaque colonne.
Description	Description du contenu de la colonne. Vous pouvez éventuellement ajouter une description lorsque vous créez la table de référence.
Inclure une colonne pour les descriptions de bas niveau	Indique que la table de référence contient une colonne pour les descriptions des données de colonne.
Valeur par défaut	Valeur par défaut des champs de la colonne. Vous pouvez éventuellement ajouter une valeur par défaut lorsque vous créez la table de référence.
Nom de connexion	Nom de la connexion à la base de données contenant les valeurs de données de la table de référence.

Création d'un objet de table de référence

Sélectionnez cette option lorsque vous voulez créer une table de référence vide et ajouter des valeurs manuellement.

1. Sélectionnez **Fichier > Nouveau > Table de référence** dans le menu de l'outil Developer.
2. Dans l'assistant Nouvelle table, sélectionnez **Table de référence vide**.
3. Saisissez le nom de la table.
4. Sélectionnez un projet dans lequel stocker les métadonnées de la table.
Cliquez sur **Parcourir** lorsque vous êtes sur le champ Emplacement. La boîte de dialogue **Sélectionner un emplacement** s'affiche à l'écran et affiche les projets du référentiel. Sélectionnez le projet dont vous avez besoin.
Cliquez sur **Suivant**.
5. Ajoutez deux colonnes ou plus à la table. Cliquez sur l'option **Nouveau** pour créer une colonne.

Le tableau suivant décrit les propriétés de chaque colonne :

Propriété	Valeur par défaut
Nom	colonne
Type de données	chaîne
Précision	10
Échelle	0
Description	Vide. Propriété facultative.

- Sélectionnez la colonne qui contient les valeurs valides. Vous pouvez modifier l'ordre des colonnes que vous créez.
- Le tableau suivant décrit les propriétés facultatives :

Propriété	Valeur par défaut
Vous pouvez inclure une colonne pour les descriptions de niveau de ligne.	Effacé
Note d'audit	Vide
Valeur par défaut	Vide

Cliquez sur **Terminer**.

La table de référence s'ouvre dans l'espace de travail de l'outil Developer.

Création d'une table de référence à partir d'un fichier plat

Vous pouvez créer une table de référence à partir des données stockées dans un fichier plat.

- Sélectionnez **Fichier > Nouveau > Table de référence** dans le menu de l'outil Developer.
- Dans l'assistant Nouvelle table, sélectionnez **Table de référence à partir d'un fichier plat**.
- Parcourez les répertoires pour accéder au fichier à utiliser comme source de données pour la table.
- Saisissez le nom de la table.
- Sélectionnez un projet dans lequel stocker les métadonnées de la table.
Cliquez sur **Parcourir** lorsque vous êtes sur le champ Emplacement. La boîte de dialogue **Sélectionner un emplacement** s'affiche à l'écran et affiche les projets du référentiel. Sélectionnez le projet dont vous avez besoin.
Cliquez sur **Suivant**.
- Définissez UTF-8 comme page de code.
- Précisez les délimiteurs utilisés par le fichier plat.
- Si le fichier plat contient des noms de colonnes, sélectionnez l'option d'importation des noms de colonnes à partir de la première ligne du fichier.

9. Le tableau suivant décrit les propriétés facultatives de la table :

Propriété	Valeur par défaut
Qualificateur de texte	Pas de guillemets
Débuter l'importation à la ligne	Ligne 1
Séparateur de lignes	\012 LF (\n)
Traiter les délimiteurs consécutifs comme uniques	Effacé
Caractère d'échappement	Vide
Conserver le caractère d'échappement dans les données	Effacé
Nombre maximum de lignes à consulter	500

Cliquez sur **Suivant**.

10. Sélectionnez la colonne qui contient les valeurs valides.
11. Le tableau suivant décrit les propriétés facultatives :

Propriété	Valeur par défaut
Vous pouvez inclure une colonne pour les descriptions de niveau de ligne.	Effacé
Note d'audit	Vide
Valeur par défaut	Vide
Nombre maximum de lignes à consulter	500

Cliquez sur **Terminer**.

La table de référence s'ouvre dans l'espace de travail de l'outil Developer.

Création d'une table de référence à partir d'une source relationnelle

Vous pouvez utiliser une source de base de données pour créer une table de référence gérée ou non gérée.

Remarque: Vous pouvez configurer une connexion de base de données dans l'Explorateur de connexion. Si l'outil Developer n'affiche pas l'Explorateur de connexion, sélectionnez **Fenêtre > Afficher aperçu > Explorateur de connexion** à partir du menu de l'outil Developer.

1. Sélectionnez **Fichier > Nouveau > Table de référence** dans le menu de l'outil Developer.
2. Dans l'assistant Nouvelle table, sélectionnez **Table de référence à partir d'une source relationnelle**. Cliquez sur **Suivant**.
3. Sélectionnez la connexion à une base de données.

Cliquez sur **Parcourir** lorsque vous êtes sur le champ Connexion. La boîte de dialogue **Sélectionner une connexion** s'ouvre et affiche les connexions de base de données disponibles.

Pour créer une table de référence gérée, connectez-vous à l'entrepôt de données de référence. Pour créer une table de référence non gérée, connectez-vous à une autre base de données.

Cliquez sur **OK** lorsque vous avez sélectionné une connexion.

4. Si la connexion de base de données que vous sélectionnez ne spécifie pas l'entrepôt de données de référence, sélectionnez **Table non gérée**.

Si vous voulez effectuer des opérations d'édition sur une table de référence non gérée, sélectionnez l'option **Editable**.

5. Sélectionnez une table de base de données.

Cliquez sur **Parcourir** lorsque vous êtes sur le champ Ressource. La boîte de dialogue **Sélectionner une ressource** s'ouvre et affiche les ressources sur la connexion de base de données. Explorez la base de données et sélectionnez une table de base de données.

6. Saisissez le nom de la table.

7. Sélectionnez un projet dans lequel stocker l'objet de la table de référence.

Cliquez sur **Parcourir** lorsque vous êtes sur le champ Emplacement. La boîte de dialogue **Sélectionner un emplacement** s'affiche et présente les projets du référentiel.

Sélectionnez un projet et cliquez sur **Suivant**.

8. Sélectionnez la colonne qui contient les valeurs valides.

9. Le tableau suivant décrit les propriétés facultatives que vous pouvez modifier :

Propriété	Valeur par défaut
Vous pouvez inclure une colonne pour les descriptions de niveau de ligne.	Effacé
Note d'audit	Vide
Valeur par défaut	Vide
Nombre maximum de lignes à consulter	500

Cliquez sur **Terminer**.

Ensembles de contenu

Un ensemble de contenu est un objet de référentiel modèle qui stocke les données ou métadonnées pour d'autres objets de données de référence. Un ensemble de contenu peut inclure des jeux de caractères, des ensembles de modèles, les jeux de jetons, des expressions régulières, des modèles probabilistes et des modèles de classeur. Utilisez un jeu de contenu pour définir et organiser les objets de données de référence qui sont associés à un seul projet, type d'information ou objectif métier.

L'outil Developer inclut les jeux de caractères définis par le système et des jeux de jetons qui ne s'affichent pas dans le référentiel modèle. Pour afficher et utiliser les objets définis par le système, configurez une stratégie dans la transformation Libellé, la transformation Analyseur ou la transformation Normalisation.

Ensembles de caractères

Un jeu de caractères contient des expressions qui identifient certains caractères et certaines plages de caractères. Vous pouvez utiliser des jeux de caractères dans les transformations Libellé qui utilisent le mode de libellé de caractère.

Les plages de caractère spécifient une plage séquentielle de codes de caractère. Par exemple, la plage de caractère « [A-C] » correspond aux caractères majuscules « A, » « B, » et « C ». Cette plage de caractère ne correspond pas aux caractères minuscules « a, » « b, » ou « c ».

Utilisez les jeux de caractères pour identifier un caractère spécifique ou une plage de caractères comme faisant partie des opérations de libellé. Par exemple, vous pouvez libeller tous les numéros d'une colonne qui contient des numéros de téléphone. Après avoir libellé les nombres, vous pouvez identifier les modèles avec une transformation Analyseur et écrire les modèles problématiques dans des ports de sortie séparés.

Propriétés de jeu de caractères

Configurez les propriétés qui déterminent les opérations de libellé de caractères pour un jeu de caractères.

Le tableau suivant décrit les propriétés d'un jeu de caractères défini par l'utilisateur :

Propriété	Description
Étiquette	Définit l'étiquette appliquée par une transformation Libellé aux données qui correspondent au jeu de caractères.
Mode standard	Active une vue d'édition simple qui inclut des champs pour la plage de début et la plage de fin.
Plage de début	Spécifie le premier caractère dans une plage de caractères.
Plage de fin	Spécifie le dernier caractère dans une plage de caractères. Pour une plage contenant un caractère unique, ne remplissez pas ce champ.
Mode avancé	Active une vue d'édition avancée dans laquelle vous pouvez saisir manuellement des plages de caractères à l'aide de caractères de plages et de caractères délimiteurs.
Caractère de plage	Modifie de manière temporaire le symbole représentant une plage de caractères. Le caractère de plage rétablit le caractère par défaut lorsque vous fermez le jeu de caractères.
Caractère délimiteur	Modifie de manière temporaire le symbole qui sépare les plages de caractères. Le caractère délimiteur rétablit le caractère par défaut lorsque vous fermez le jeu de caractères.

Modèles de classeur

Un modèle de classeur analyse les chaînes d'entrée et détermine les types d'information que les chaînes sont le plus susceptibles de contenir. Vous pouvez utiliser un modèle de classeur dans une transformation Classeur. Utilisez un modèle de classeur lorsque les chaînes d'entrée contiennent une grande quantité de données.

Un modèle de classeur contient des lignes de données et des valeurs de libellé. L'ensemble de lignes de données représente les données source que vous pouvez sélectionner dans la transformation Classeur. Les valeurs des libellés décrivent les types d'information que les lignes de données sont susceptibles de contenir. Vous attribuez un libellé à chaque ligne de données du modèle.

Pour lier les lignes de données aux libellés dans un modèle de classeur, vous devez compiler le modèle. Le processus de compilation génère une série d'associations logiques entre les lignes de données et les libellés. Lorsque vous exécutez un mappage qui lit le modèle, le service d'intégration de données applique la logique

du modèle aux données d'entrée de la transformation. Le service d'intégration de données renvoie le libellé qui décrit le plus précisément les informations dans chaque ligne de données.

Vous pouvez créer un modèle de classeur dans l'outil Developer. Le référentiel modèle stocke l'objet de modèle de classeur. L'outil Developer écrit les lignes de données, les libellés et les données de compilation dans un fichier de la structure de répertoires Informatica.

Ensembles de modèles

Un ensemble de modèles contient des expressions qui identifient les modèles de données dans la sortie d'une opération de libellé de jeton. Vous pouvez utiliser les ensembles de modèles pour analyser le port de sortie des données à jeton et écrire des chaînes correspondantes sur un ou plusieurs ports de sortie. Utilisez les ensembles de modèles dans les transformations Analyseur qui utilisent le mode d'analyse de modèles.

Par exemple, vous pouvez configurer une transformation Analyseur pour utiliser les ensembles de modèles qui identifient les noms et les initiales. Cette transformation utilise les ensembles de modèles pour analyser la sortie d'une transformation Libellé en mode libellé de jeton. Vous pouvez configurer la transformation Analyseur pour écrire les noms et les initiales dans la sortie pour séparer les ports.

Propriétés de l'ensemble de modèles

Configurez les propriétés qui déterminent les modèles dans un ensemble de modèles.

Le tableau suivant décrit la propriété pour un ensemble de modèles défini par l'utilisateur :

Propriété	Description
Patron	Définit les modèles que l'analyseur de modèles recherche. Vous pouvez entrer plusieurs modèles pour un ensemble de modèles. Vous pouvez entrer des modèles construits à partir d'une combinaison de caractères génériques, de caractères et de chaînes.

Modèles probabilistes

Un modèle probabiliste analyse les valeurs de données d'entrée et détermine les types d'informations que les valeurs sont le plus susceptibles de contenir. Utilisez un modèle probabiliste dans une transformation Libellé et une transformation Analyseur.

Un modèle probabiliste contient des valeurs de données et des valeurs de libellé. L'ensemble de valeurs de données représente les données source que vous pouvez sélectionner dans la transformation. Les valeurs de libellé décrivent les types d'informations que les valeurs de données contiennent. Vous attribuez un libellé à chaque valeur de données dans le modèle.

Pour lier les valeurs de données aux libellés dans un modèle probabiliste, vous devez compiler le modèle. Le processus de compilation génère une série d'associations logiques entre les valeurs de données et les libellés. Lorsque vous exécutez un mappage qui lit le modèle, le service d'intégration de données applique la logique du modèle aux données d'entrée de la transformation. Le service d'intégration de données renvoie le libellé qui décrit le plus précisément les valeurs de données d'entrée.

Créez un modèle probabiliste dans l'outil Developer. Le référentiel modèle stocke l'objet de modèle probabiliste. L'outil Developer écrit les valeurs de données, les libellés et les données de compilation dans un fichier dans la structure de répertoires Informatica.

Expressions régulières

Dans le contexte d'ensembles de contenu, une expression régulière est une expression que vous pouvez utiliser dans des opérations d'analyse et de libellé. Utilisez les expressions régulières pour identifier une ou plusieurs chaînes dans les données d'entrée. Vous pouvez utiliser des expressions régulières dans les transformations Analyseur qui utilisent le mode d'analyse de jeton. Vous pouvez également utiliser des expressions régulières dans les transformations Libellé qui utilisent le mode libellé de jeton.

Les transformations Analyseur utilisent des expressions régulières pour faire correspondre les modèles dans les données d'entrée et analyser toutes les chaînes de correspondance vers une ou plusieurs sorties. Par exemple, vous pouvez utiliser une expression régulière pour identifier toutes les adresses électroniques dans les données d'entrée et analyser chaque composant d'adresse électronique vers une sortie différente.

Les transformations Libellé utilisent des expressions régulières pour faire correspondre un modèle d'entrée et créer un libellé unique. Les expressions régulières qui ont plusieurs sorties ne génèrent pas plusieurs libellés.

Propriétés de l'expression régulière

Configurez les propriétés qui permettent de déterminer la méthode d'identification et d'écriture de chaînes de sortie d'une expression régulière.

Le tableau suivant décrit les propriétés d'une expression régulière définie par l'utilisateur :

Propriété	Description
Nombre de sorties	Définit le nombre de ports de sortie que l'expression régulière écrit.
Expression régulière	Définit un modèle que la transformation Analyseur utilise pour faire correspondre des chaînes.
Expression de test	Contient des données que vous entrez pour tester l'expression régulière. Dès que vous entrez des données dans ce champ, celui-ci met en évidence les chaînes qui correspondent à l'expression régulière.
Expression suivante	Déplace vers la chaîne suivante qui correspond à l'expression régulière et modifie la police de cette chaîne en gras.
Expression précédente	Déplace la chaîne précédente qui correspond à l'expression régulière et modifie la police de cette chaîne en gras.

Ensembles de jetons

Un jeu de jetons contient des expressions identifiant les jetons spécifiques. Vous pouvez utiliser des jeux de jetons dans les transformations Libellé qui utilisent le mode libellé de jeton. Vous pouvez également utiliser des jeux de jetons dans les transformations Analyseur qui utilisent le mode d'analyse de jeton.

Utilisez les jeux de jetons pour identifier des jetons dans le cadre des opérations de libellé et d'analyse. Par exemple, vous pouvez utiliser un jeu de jetons pour étiqueter toutes les adresses électroniques qui utilisent un format « NomCompte@NomDomaine ». Après avoir étiqueté les jetons, vous pouvez utiliser la transformation Analyseur pour écrire les adresses électroniques dans les ports de sortie que vous indiquez.

Propriétés des jeux de jetons

Configurez les propriétés qui déterminent les opérations de libellé pour un jeu de jetons.

Le tableau suivant décrit les propriétés d'un jeu de caractères défini par l'utilisateur :

Propriété	Mode de jeu de jetons	Description
Nom	N/A	Définit le nom du jeu de jetons.
Description	N/A	Décrit le jeu de jetons.
Options de jeu de jetons	N/A	Définit si le jeu de jetons utilise le mode d'expression régulière ou de caractère.
Étiquette	Expression régulière	Définit le libellé qu'une transformation Libellé applique aux données qui correspondent au jeu de jetons.
Expression régulière	Expression régulière	Définit un modèle que la transformation Libellé utilise pour établir des correspondances avec les chaînes.
Expression de test	Expression régulière	Contient des données que vous entrez pour tester l'expression régulière. Lorsque vous saisissez des données dans ce champ, il met en surbrillance les chaînes qui correspondent à l'expression régulière.
Expression suivante	Expression régulière	Déplace vers la chaîne suivante qui correspond à l'expression régulière et modifie la police de cette chaîne en gras.
Expression précédente	Expression régulière	Déplace la chaîne précédente qui correspond à l'expression régulière et modifie la police de cette chaîne en gras.
Étiquette	Caractère	Définit l'étiquette appliquée par une transformation Libellé aux données qui correspondent au jeu de caractères.
Mode standard	Caractère	Active une vue d'édition simple qui inclut des champs pour la plage de début et la plage de fin.
Plage de début	Caractère	Spécifie le premier caractère dans une plage de caractères.

Propriété	Mode de jeu de jetons	Description
Plage de fin	Caractère	Spécifie le dernier caractère dans une plage de caractères. Pour des plages à caractère unique, ne remplissez pas ce champ.
Mode avancé	Caractère	Active une vue d'édition avancée dans laquelle vous pouvez saisir manuellement des plages de caractères à l'aide de caractères de plages et de caractères délimiteurs.
Caractère de plage	Caractère	Modifie de manière temporaire le symbole représentant une plage de caractères. Le caractère de plage rétablit le caractère par défaut lorsque vous fermez le jeu de caractères.
Caractère délimiteur	Caractère	Modifie de manière temporaire le symbole qui sépare les plages de caractères. Le caractère délimiteur rétablit le caractère par défaut lorsque vous fermez le jeu de caractères.

Règles et directives pour les modèles probabilistes et les modèles de classeur

Chaque modèle probabiliste et de classeur du référentiel modèle identifie un fichier dans la structure de répertoires Informatica. Les fichiers contiennent les valeurs de données et les libellés que vous ajoutez au modèle dans l'outil Developer. Les fichiers contiennent également la logique de compilation qui définit les associations entre les valeurs de données et les libellés.

Envisagez les règles et directives suivantes lorsque vous travaillez avec des modèles probabilistes ou des modèles de classeur :

- Lorsque vous exécutez un mappage qui inclut un modèle, le service d'intégration de données applique la logique du modèle compilé aux données d'entrée de la transformation. Le service d'intégration de données ne lit pas les valeurs de données ou les libellés du modèle pendant l'exécution du mappage.
- Vous pouvez éventuellement supprimer les valeurs de données et les libellés d'un modèle probabiliste ou d'un modèle de classeur. Par exemple, vous pouvez décider de supprimer des données sensibles ou privées d'un modèle. Vous pouvez supprimer les valeurs de données individuelles et les libellés dans l'outil Developer. Vous pouvez supprimer toutes les valeurs de données et les libellés lorsque vous exportez un modèle depuis le référentiel modèle.
Remarque: Si vous supprimez toutes les valeurs de données et les libellés d'un modèle, vous ne pouvez pas compiler le modèle.
- Lorsque vous supprimez une ou plusieurs valeurs de données ou des libellés d'un modèle, la logique du modèle compilé ne représente plus les données actuelles du fichier modèle. Pour synchroniser la logique

du modèle, les valeurs de données et les libellés, compilez à nouveau le modèle. Ne compilez pas le modèle si vous souhaitez conserver la logique du modèle actuelle.

- Pour protéger les données dans un modèle de classeur ou un modèle probabiliste, sauvegardez le fichier de modèle dans la structure de répertoires Informatica. Sauvegardez le fichier avant de supprimer toutes les valeurs de données et les libellés d'un modèle.
- Recherchez les fichiers de modèle dans la machine hôte du service de gestion du contenu.

Par défaut, les fichiers de modèle probabiliste se trouvent à l'emplacement suivant et portent l'extension de nom de fichier suivante :

```
<Informatica_Installation_Directory>/tomcat/bin/ner/<filename>.ner
```

Par défaut, les fichiers de modèle de classeur se trouvent à l'emplacement suivant et portent l'extension de nom de fichier suivante :

```
<Informatica_Installation_Directory>/tomcat/bin/classifier/<filename>.classifier
```

- Si vous procédez à une mise à niveau de l'installation Informatica, il est possible que vous deviez compiler les modèles probabilistes et les modèles de classeur avant de les utiliser dans un mappage. Si un modèle ne contient pas de données, remplacez le fichier actuel dans la structure de répertoires Informatica par le fichier de sauvegarde qui contient les données.

Création d'un ensemble de contenu

Créez un jeu de contenu pour gérer des objets de données de référence qui font référence à un seul projet, type d'information ou objectif métier.

1. Dans la vue **Explorateur d'objets**, sélectionnez un projet ou un dossier dans lequel stocker le jeu de contenu.
2. Cliquez sur **Fichier > Nouveau > Ensemble de contenu**.
3. Entrez un nom pour l'ensemble de contenu.
4. Vous pouvez éventuellement sélectionner **Parcourir** pour changer l'emplacement du référentiel modèle pour l'ensemble de contenu.
5. Cliquez sur **Terminer**.

Création d'un objet de données de référence dans un ensemble de contenu

Vous pouvez créer un jeu de caractères, un ensemble de formes, un jeu de jetons, une expression régulière, un modèle probabiliste et un modèle de classeur dans un ensemble de contenu.

1. Ouvrez un ensemble de contenu dans l'éditeur et sélectionnez la vue **Contenu**.
2. Sélectionnez un type d'objet de données de référence.
3. Cliquez sur **Ajouter**.
4. Entrez un nom pour l'objet de données de référence.
Entrez éventuellement une description de l'objet.
5. Configurez les propriétés de l'objet de données de référence.
6. Cliquez sur **Terminer**.

Astuce: Vous pouvez copier des objets de données de référence d'un ensemble de contenu vers un autre. Utilisez les options **Copier dans** et **Coller depuis** pour créer une copie d'un objet dans un ensemble de contenu. Utilisez la touche **CTRL** pour sélectionner plusieurs objets de l'ensemble de contenu.

Génération de données de référence depuis un profil intermédiaire

Un profil intermédiaire est un profil que vous exécutez sur une transformation dans un mappage. Vous pouvez exécuter un profil intermédiaire pour générer des données pour un objet de données de référence.

Par exemple, exécutez un profil sur une transformation que vous connectez aux ports d'entrée d'une transformation Libellé ou d'une transformation Analyseur. Ajoutez les données de profil à un modèle probabiliste et appliquez le modèle probabiliste à la transformation Libellé ou la transformation Analyseur.

1. Ouvrez le mappage qui contient la transformation devant faire l'objet du profil.
2. Sélectionnez la transformation.

Par exemple, sélectionnez une transformation que vous connectez aux ports d'entrée d'une transformation Libellé ou d'une transformation Analyseur.

3. Cliquez sur **Profilier maintenant**.
4. Sélectionnez l'onglet Résultats dans le profil, puis vérifiez les résultats du profil.
5. Dans Profilage de colonnes, sélectionnez une colonne.
6. Sous Détails, sélectionnez l'option pour afficher les valeurs du profil.

L'éditeur affiche les valeurs de données dans la colonne que vous avez sélectionnée. Vous pouvez sélectionner toutes les valeurs de la colonne ou un sous-ensemble de valeurs.

7. Exportez les données de colonne dans un fichier.
 - Pour exporter toutes les valeurs de colonne, cliquez sur l'option **Exporter les fréquences de valeur vers le fichier**.
 - Pour exporter un sous-ensemble de valeurs de colonne, cliquez avec le bouton droit de la souris sur les valeurs et sélectionnez **Envoyer à > Exporter les résultats vers le fichier**.
8. Enregistrez le fichier qui contient les données de colonne. Vous pouvez sauvegarder le fichier sur la machine des services Informatica ou sur la machine de l'outil Developer.

Vous pouvez utiliser le fichier comme source de données pour un objet de données de référence. Par exemple, créez un modèle probabiliste depuis le fichier.

LIENS CONNEXES :

- ["Créer une table de référence à partir des données de profil" à la page 16](#)

CHAPITRE 4

Modèles de classeur

Ce chapitre comprend les rubriques suivantes :

- [Présentation des modèles de classeur, 41](#)
- [Structure du modèle de classeur, 42](#)
- [Vues du modèle de classeur, 42](#)
- [Données de référence du modèle de classeur, 43](#)
- [Données de libellé du modèle de classeur, 45](#)
- [Scores de classeur, 46](#)
- [Filtres de modèle de classeur, 47](#)
- [Création d'un modèle de classeur à partir d'un objet de données, 48](#)
- [Opérations de copier/coller, 49](#)

Présentation des modèles de classeur

Un modèle de classeur est un objet de données de référence dans un ensemble de contenu. Utilisez un modèle de classeur pour analyser les chaînes de texte longues qui contiennent plusieurs valeurs. Un modèle de classeur identifie le type d'informations le plus courant dans chaque chaîne.

Ajoutez un modèle de classeur à une transformation Classeur. La transformation recherche les valeurs communes des données du modèle de classeur et des données de chaque ligne d'entrée. La transformation utilise ces valeurs communes pour classer le type d'informations que représente chaque ligne.

Vous utilisez un modèle de classeur lorsque les données d'entrée présentent les caractéristiques suivantes :

- Les données d'entrée contiennent du texte. Les modèles de classeur appliquent des processus en langage naturel aux données textuelles pour identifier les types d'informations contenus dans le texte. Les processus en langage naturel détectent les mots pertinents contenus dans la chaîne d'entrée. Les processus en langage naturel ignorent les mots non pertinents.
- Les chaînes de données d'entrée contiennent plusieurs valeurs. Par exemple, vous pouvez créer une colonne de données qui inclut le contenu d'un e-mail dans chaque champ.

La transformation Classeur lit les types de données Chaîne. La transformation n'impose aucune limite quant à la longueur des chaînes d'entrée.

Vous devez compiler les modèles de classeur dans l'outil Developer. Lorsque vous compilez un modèle, vous créez des associations entre des valeurs de données similaires dans le modèle. La transformation Classeur utilise les données compilées pour rechercher des informations dans les données d'entrée.

Exemple de transformation Classeur

Vous pouvez utiliser un modèle de classeur et une transformation Classeur pour classer les e-mails en fonction du texte qu'ils contiennent.

Par exemple, vous travaillez dans un centre d'assistance client et vous consultez les e-mails que les clients envoient à l'entreprise. L'entreprise compte des clients dans de nombreux pays et reçoit des e-mails dans de nombreuses langues. Vous voulez trier les e-mails par langue, de sorte à pouvoir envoyer chaque e-mail au centre le plus à même de répondre au client.

Effectuez les étapes suivantes pour trier les e-mails :

1. Écrivez les e-mails dans un seul fichier ou une seule table de base de données.
2. Créez un modèle de classeur qui contient l'exemple de texte pour chaque langue.

Remarque: Vous pouvez utiliser des échantillons de données des e-mails en tant que source des données pour le modèle. Copiez le texte de l'e-mail dans un fichier ou une table de base de données et créez une source de données à partir du fichier ou de la table dans le référentiel modèle.

3. Ajoutez le modèle de classeur à une transformation Classeur.
4. Ajoutez la transformation à un mappage et connectez les ports de transformation à la source de données et aux cibles de données. Créez une cible des données pour chaque langue.

Lorsque vous exécutez le mappage, la transformation Classeur analyse les e-mails et écrit leur texte dans la cible de données appropriée. Vous pouvez partager la cible de données avec les membres de l'équipe du centre d'assistance approprié.

Structure du modèle de classeur

Un modèle de classeur contient une colonne de valeurs de données de référence et une colonne qui indique un libellé pour chaque ligne de données de valeurs de référence. Lorsqu'une transformation Classeur compare les données d'entrée et les données du modèle, la transformation renvoie le libellé qui décrit le plus précisément les données d'entrée.

Un modèle de classeur contient également des données de compilation. La transformation Classeur utilise les données de compilation pour calculer les similarités entre les données de référence et les données d'entrée. Lorsque vous compilez un modèle, vous créez ou mettez à jour les données de compilation.

Le référentiel modèle stocke l'objet de modèle de classeur. Lorsque vous créez ou mettez à jour un modèle de classeur, vous écrivez les données de référence et les métadonnées compilées dans un fichier sur la machine de l'outil Developer. Le fichier est en lecture seule. Vous pouvez lire le chemin du fichier dans le modèle de classeur dans l'outil Developer.

Vues du modèle de classeur

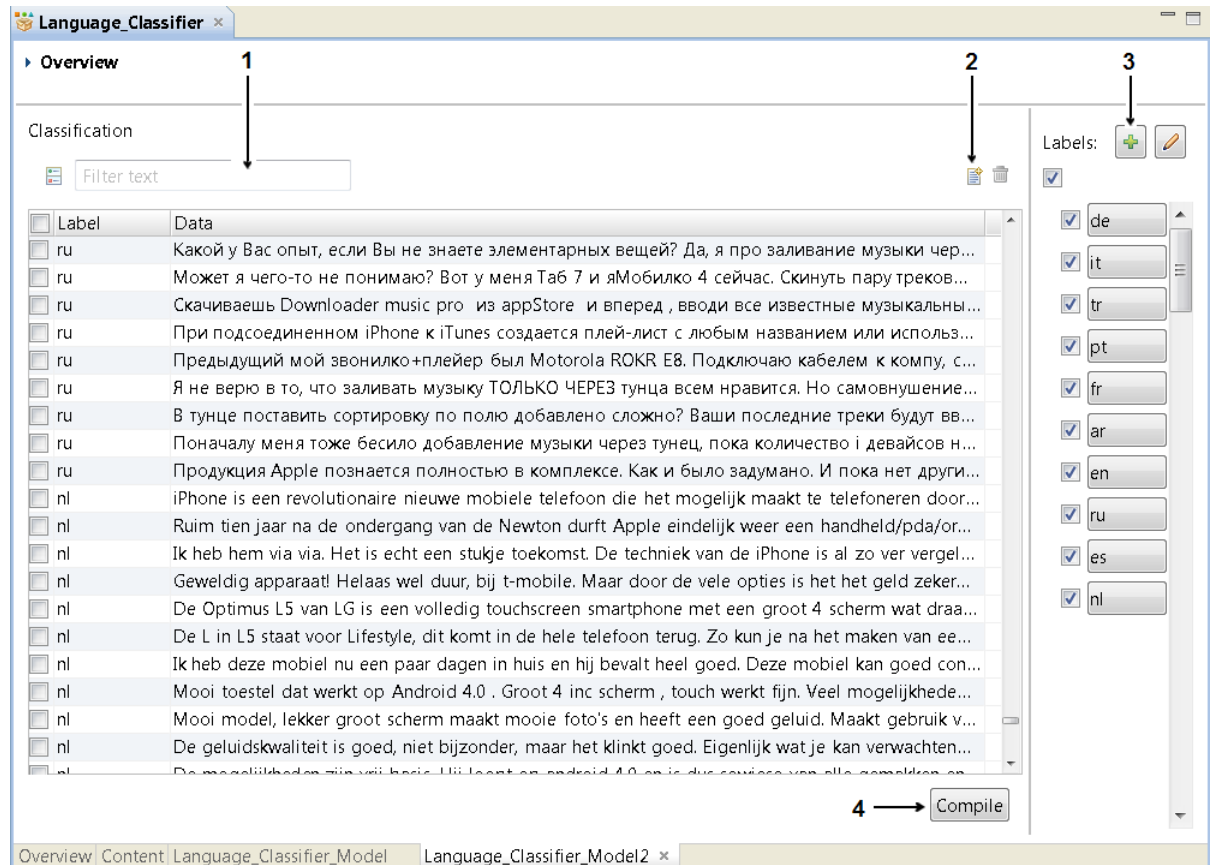
Vous pouvez utiliser la vue par défaut et la vue détaillée pour mettre à jour les données dans un modèle de classeur. La vue par défaut affiche les valeurs des libellés et les valeurs de données dans un tableau. La vue détaillée affiche les valeurs de données dans une série de zones de texte.

Utilisez la vue par défaut pour vérifier et mettre à jour les libellés sur chaque ligne. Vous pouvez sélectionner une ligne, plusieurs lignes ou toutes les lignes. La vue par défaut peut afficher environ 100 caractères de données de ligne. La vue détaillée peut afficher toutes les données de chaque ligne. Cliquez sur l'option

Données du modèle de classeur pour basculer d'une vue à l'autre. Utilisez la vue détaillée pour vérifier et mettre à jour les données dans une seule ligne.

Vous pouvez ajouter des données, filtrer les lignes de données et ajouter des libellés aux lignes dans chaque vue. Vous pouvez rechercher les valeurs de données dans une seule ligne dans la vue détaillée.

L'image suivante présente la vue par défaut d'un modèle de classeur qui contient des données pour le classement par langue :



1. Trouve une valeur dans les lignes de données.
2. Insère une ligne de données.
3. Gère les libellés.
4. Compile le modèle de classeur.

Pour attribuer un libellé à une ligne de données, sélectionnez la ligne et cliquez sur un nom de libellé.

Données de référence du modèle de classeur

Un modèle de classeur contient une colonne de données de référence qui peut inclure des phrases, des paragraphes ou des pages de texte. Les données de référence représentent les différents types d'entrée de texte qu'une transformation Classeur peut lire dans un mappage. Lorsque vous créez un modèle, vérifiez que

les données de référence comprennent les types de textes que vous vous attendez à rencontrer lorsque vous exécutez le mappage.

Vous pouvez utiliser les données source de mappage pour créer un modèle de classeur. Sélectionnez un échantillon de données source et copiez-le dans le modèle.

Tenez compte des règles et directives suivantes lorsque vous utilisez des données de référence de modèle de classeur :

- Un champ de données de référence peut être de n'importe quelle longueur. Vous pouvez entrer des pages de texte dans chaque champ de données.
- Vous devez importer des données de référence à partir d'un objet de données.
- Vous ne pouvez pas éditer des valeurs de données de référence. Cependant, vous pouvez supprimer une ligne de données.

Ajout de valeurs de données et de libellé à un modèle de classeur

Utilisez une source de données pour ajouter des données à un modèle de classeur. Vous pouvez ajouter des valeurs de données et des valeurs de libellé.

1. Ouvrez l'ensemble de contenu qui contient le modèle.
2. Sélectionnez le nom du modèle et cliquez sur **Editer**.
3. Cliquez sur **Adjonction de données**.

L'assistant Modèle de classeur s'ouvre.

4. Parcourez le référentiel modèle et sélectionnez l'objet de données que vous voulez utiliser. Cliquez sur **Suivant**.

Remarque: Ne sélectionnez pas un objet de données de média social comme source de données.

5. Vérifiez les colonnes dans l'objet de données et sélectionnez une colonne à ajouter en tant que colonne de données ou de libellé pour le modèle. Vous pouvez ajouter une colonne de données de référence et une colonne de libellé au cours de la même opération.
 - Pour utiliser une colonne de source de données en tant que colonne de données de référence dans le modèle, sélectionnez le nom de la colonne et cliquez sur **Données**.
Vous pouvez sélectionner plusieurs colonnes de données. Le modèle de classeur fusionne le contenu des colonnes que vous sélectionnez en une seule colonne.
 - Pour utiliser une colonne de source de données comme colonne de libellé pour le modèle, sélectionnez le nom de la colonne et cliquez sur **Libellé**.

Cliquez sur **Suivant**.

6. Sélectionnez le nombre de lignes à copier à partir de la source de données.
Sélectionnez toutes les lignes ou saisissez le nombre de lignes à copier. Si vous saisissez un nombre, le modèle compte les lignes depuis le début de l'ensemble de données.
7. Cliquez sur **Terminer** et enregistrez le modèle.

Après avoir ajouté les données, vérifiez que les lignes de données que vous avez ajoutées incluent des valeurs de libellé.

Suppression de valeurs de données d'un modèle de classeur

Vous pouvez supprimer des valeurs de données dans la vue par défaut et dans la vue détaillée d'un modèle de classeur. Pour supprimer toutes les valeurs de données, utilisez la vue par défaut.

1. Ouvrez le modèle de classeur dans l'outil Developer. Pour ouvrir le modèle, sélectionnez le nom du modèle dans l'ensemble de contenu et cliquez sur **Editer**.
2. Sélectionnez la ligne qui contient les données à supprimer.
Vous pouvez sélectionner une seule ligne, plusieurs lignes ou toutes les lignes.
3. Cliquez sur **Supprimer**.

Données de libellé du modèle de classeur

Un modèle de classeur contient des libellés descriptifs pour les types d'informations des champs de données de référence. Lorsque vous créez un modèle, vous attribuez un libellé à chaque champ de données.

Vous pouvez sélectionner une colonne en tant que colonne de données de libellé lorsque vous ajoutez des données de source de données à un modèle de classeur. Vous pouvez également saisir les noms de libellé dans le modèle. Vous pouvez attribuer tout libellé du modèle à tout champ de données de référence.

Les libellés sont indépendants des valeurs de données de référence qu'ils décrivent. Si vous supprimez toutes les lignes de table qui contiennent un libellé sélectionné, vous ne supprimez pas le libellé du modèle. Si vous supprimez un libellé, vous ne supprimez pas les valeurs de données de référence que vous avez associées au libellé.

Ajout d'un libellé à un modèle de classeur

Si vous ajoutez des valeurs de données sans libellés à un modèle de classeur, ajoutez les libellés séparément.

1. Ouvrez l'ensemble de contenu qui contient le modèle.
2. Sélectionnez le nom du modèle et cliquez sur **Editer**.
3. Sélectionnez l'option **Nouveau libellé** pour ajouter un libellé.
Le libellé apparaît dans le modèle.
4. Entrez un nom dans la boîte de dialogue **Nouveau libellé**.
5. Cliquez sur **OK**.

Attribution de libellés aux données du modèle de classeur

Les libellés représentent les types d'informations qu'un modèle de classeur peut identifier dans les données source. Lorsque vous créez un modèle de classeur, vérifiez que chaque ligne de données du modèle est associée à un libellé.

Vous pouvez attribuer un libellé à une ligne ou à plusieurs lignes.

1. Ouvrez l'ensemble de contenu qui contient le modèle.
2. Sélectionnez le nom du modèle et cliquez sur **Editer**.
3. Filtrez les lignes de données dans le modèle pour afficher les lignes auxquelles aucun libellé n'est associé.

Pour afficher les lignes qui ne sont associées à aucun libellé, effacez tous les noms de libellé dans le panneau **Libellés**.

4. Sélectionnez une ou plusieurs lignes de données. Vous pouvez utiliser l'option Sélectionner tout pour sélectionner toutes les lignes qui s'affichent.

Le modèle ajoute une coche aux lignes sélectionnées.

5. Parcourez les valeurs de libellé du modèle et sélectionnez un libellé à appliquer aux lignes de données.

Le modèle attribue le libellé aux lignes sélectionnées.

6. Compilez le modèle pour ajouter les noms des libellés à la logique du modèle de classeur.

Si vous attribuez un libellé que vous avez effacé de l'affichage des noms de libellé, le modèle masque les lignes. Sélectionnez le nom de libellé dans le panneau Libellés pour afficher les lignes.

Suppression d'un libellé d'un modèle de classeur

Vous pouvez supprimer un libellé dans la vue par défaut ou la vue détaillée d'un modèle de classeur.

1. Ouvrez l'ensemble de contenu qui contient le modèle.
2. Sélectionnez le nom du modèle et cliquez sur **Editer**.
3. Cliquez sur **Propriétés**.
4. Dans la boîte de dialogue **Gérer les libellés**, sélectionnez un ou plusieurs libellés à supprimer.
Vous pouvez sélectionner plusieurs libellés.
5. Cliquez sur **Supprimer**.
6. Cliquez sur **OK** pour fermer la boîte de dialogue.

Compilation d'un modèle de classeur

À chaque fois que vous éditez une valeur de libellé ou la valeur de données de référence dans un modèle de classeur, vous devez compiler le modèle. Lorsque vous compilez le modèle, vous mettez à jour les données de compilation dans le modèle.

- Pour mettre à jour les données de compilation, ouvrez le modèle dans l'outil Developer et cliquez sur **Compiler**.

Scores de classeur

Une transformation Classeur compare chaque ligne de données d'entrée à chaque ligne de données de référence dans un modèle de classeur. La transformation calcule un score pour chaque comparaison. Les scores représentent les degrés de similarité entre la ligne d'entrée et les lignes de données de référence.

Lorsque vous exécutez un mappage qui contient une transformation Classeur, le mappage renvoie le libellé qui identifie la ligne des données de référence dont le score est le plus élevé. Le score est compris entre 0 et 1. Un score élevé indique une correspondance élevée entre les données d'entrée et les données du modèle.

Vérifiez les scores du classeur pour vous assurer que la sortie du libellé décrit précisément chaque ligne de données d'entrée. Vous pouvez également vérifier les scores pour vous assurer que le modèle de classeur est approprié pour les données d'entrée. Si la sortie de transformation contient un important pourcentage de scores bas, le modèle de classeur est peut-être inapproprié. Pour améliorer les comparaisons, compilez à nouveau le modèle. Si le modèle compilé n'améliore pas les scores, remplacez le modèle dans la transformation.

Filtres de modèle de classeur

Vous pouvez appliquer des filtres aux données du modèle de classeur dans les vues par défaut et détaillée.

Vous pouvez utiliser des filtres pour effectuer les tâches suivantes :

- Recherchez les valeurs de données auxquelles aucun libellé n'est associé. Utilisez les options de libellé pour filtrer les lignes de données qui apparaissent dans un modèle de classeur. Si une ligne de données n'utilise pas de libellé, ajoutez un libellé à la ligne.
- Recherchez les valeurs de données dans les lignes de données de référence. Utilisez le filtre dans la vue par défaut pour rechercher des valeurs de données dans les données de référence. Vérifiez que les données de référence chevauchent les données source dans un mappage.
- Recherchez une valeur de données dans une ligne de données de référence. Utilisez le filtre dans la vue détaillée lorsque vous avez besoin de vérifier qu'une ligne de données de référence contient une valeur de données. Une ligne de données peut contenir un grand nombre de valeurs de données.

Recherche de valeurs dans les données de référence

Utilisez le filtre dans la vue par défaut pour vérifier que les lignes de données de référence contiennent les valeurs de données attendues.

1. Ouvrez l'ensemble de contenu qui contient le modèle.
2. Sélectionnez le nom du modèle et cliquez sur **Editer**.
3. Saisissez une valeur de texte dans le champ de filtre.

L'outil Developer affiche les lignes de données qui contiennent le texte de filtre.

Recherche de lignes de données sans libellés

Effacez toutes les options de libellé pour afficher les lignes de données de référence qui n'ont pas de libellé.

Lorsque vous ouvrez un modèle de classeur, l'outil Developer affiche par défaut l'ensemble des lignes et des libellés.

1. Ouvrez l'ensemble de contenu qui contient le modèle.
2. Sélectionnez le nom du modèle et cliquez sur **Editer**.
3. Dans la vue par défaut, sélectionnez ou désélectionnez un libellé.

Lorsque vous sélectionnez un libellé, le modèle affiche les chaînes de données que vous avez associées à ce libellé. Lorsque vous désélectionnez un libellé, le modèle masque les chaînes de données que vous avez associées à ce libellé.

4. Pour vérifier que toutes les chaînes de données utilisent un libellé, effacez toutes les valeurs de libellé. Le modèle affiche toutes les chaînes qui n'utilisent pas de libellé.
5. Cliquez sur une valeur de libellé pour ajouter le libellé à la chaîne de données.

Recherche d'une valeur dans une ligne de données

Utilisez le filtre dans la vue détaillée pour rechercher une valeur de données dans une seule ligne de données.

1. Ouvrez l'ensemble de contenu qui contient le modèle.
2. Sélectionnez le nom du modèle et cliquez sur **Editer**.
3. Sélectionnez la vue détaillée.

4. Dans la vue par défaut, entrez une valeur dans le champ de filtre.
Le modèle affiche les lignes de données qui contiennent la valeur.
5. Sélectionnez une ligne de données de recherche.
6. Saisissez la valeur de recherche dans le champ de recherche sous la ligne de données.
Le modèle met en surbrillance la première instance de la valeur dans la ligne.
7. Cliquez sur la flèche vers le bas pour trouver l'instance suivante de la valeur dans la ligne.
Utilisez les flèches vers le haut et le bas pour passer d'une valeur à une autre dans la ligne de données.

Création d'un modèle de classeur à partir d'un objet de données

Utilisez un objet de données en tant que source pour les données du modèle de classeur.

Un modèle de classeur fonctionne de manière optimale lorsque vous utilisez les données d'entrée de la transformation Classeur en tant que source des données de référence du modèle. Par exemple, vous pouvez exécuter un profil sur l'objet de transformation dans le mappage. Créez un objet de données à partir des résultats de profil.

1. Dans l'explorateur d'objets, ouvrez et créez un jeu de contenu.
2. Sélectionnez la vue **Contenu**.
3. Sélectionnez **Modèles de classeur**, puis cliquez sur **Ajouter**.
L'assistant Modèle de classeur s'ouvre.
4. Saisissez un nom pour le modèle de classeur.
Entrez éventuellement une description textuelle du modèle.
5. Parcourez le référentiel modèle et sélectionnez l'objet de données qui contient les données de référence.
Cliquez sur **Suivant**.
6. Vérifiez les colonnes de l'objet de données et sélectionnez une colonne à ajouter en tant que valeurs de données de référence ou valeurs de libellé pour le modèle.
 - Pour ajouter une colonne de données en tant que données de référence, sélectionnez le nom de colonne et cliquez sur **Données**.
 - Pour utiliser une colonne de données en tant que source pour les valeurs de libellé, sélectionnez le nom de la colonne et cliquez sur **Libellé**.Cliquez sur **Suivant**.
7. Sélectionnez le nombre de lignes à copier à partir de la source de données.
Sélectionnez toutes les lignes ou saisissez le nombre de lignes à copier. Si vous saisissez un nombre, le modèle compte les lignes depuis le début de l'ensemble de données.
8. Cliquez sur **Terminer** et enregistrez le modèle.
Compilez le modèle après avoir créé le modèle de classeur.

Opérations de copier/coller

Vous pouvez copier un modèle de classeur d'un ensemble de contenu vers un autre dans un référentiel modèle. Copiez un modèle de classeur pour partager des ressources avec d'autres utilisateurs de l'outil Developer.

Vous pouvez copier un modèle vers un autre ensemble de contenu ou importer un modèle dans l'ensemble de contenu actuel. Vous pouvez importer plusieurs modèles de plusieurs ensembles de contenu dans le référentiel en une seule fois.

Lorsque vous copiez un modèle, le service de gestion du contenu crée une copie du fichier de données du modèle sur la machine du service. Chaque modèle utilise un fichier de données différent.

Copie d'un modèle de classeur dans un autre ensemble de contenu

Vous pouvez copier un modèle de classeur d'un ensemble de contenu vers un autre dans un référentiel modèle. Lorsque vous copiez un modèle de classeur, vous spécifiez l'objet de modèle et les ensembles de contenu source et de destination.

1. Ouvrez l'ensemble de contenu qui contient le modèle de classeur.
2. Sélectionnez un modèle de classeur et cliquez sur **Copier dans**.
3. Parcourez le référentiel modèle et sélectionnez un ensemble de contenu.

Vous pouvez copier le modèle de classeur vers un ensemble de contenu dans le projet actuel ou dans un autre projet.

4. Cliquez sur **OK**.

L'outil Developer copie le modèle de classeur dans l'ensemble de contenu sélectionné.

Importation d'un modèle de classeur depuis un autre ensemble de contenu

Vous pouvez importer un modèle de classeur d'un ensemble de contenu vers un autre dans un référentiel modèle. Lorsque vous importez un modèle de classeur, vous spécifiez un ou plusieurs objets de modèle et les ensembles de contenu source et de destination.

1. Ouvrez l'ensemble de contenu qui doit contenir le modèle de classeur.
2. Sélectionnez un modèle de classeur et cliquez sur **Coller depuis**.
3. Parcourez le référentiel modèle et sélectionnez un modèle de classeur.

Vous pouvez coller le modèle de classeur d'un ensemble de contenu dans le projet actuel ou un autre projet.

4. Cliquez sur **OK**.

L'outil Developer colle le modèle de classeur dans l'ensemble de contenu actuel.

CHAPITRE 5

Modèles probabilistes

Ce chapitre comprend les rubriques suivantes :

- [Présentation des modèles probabilistes, 50](#)
- [Structure du modèle probabiliste, 51](#)
- [Vues du modèle probabiliste, 51](#)
- [Exemple de transformation Libellé, 52](#)
- [Exemple de transformation Analyseur, 53](#)
- [Données de référence du modèle probabiliste, 54](#)
- [Données de libellé du modèle probabiliste, 55](#)
- [Propriétés du modèle probabiliste, 57](#)
- [Création d'un modèle probabiliste à partir d'un objet de données, 57](#)
- [Opérations de copier/coller, 59](#)

Présentation des modèles probabilistes

Un modèle probabiliste est un objet de données de référence que vous créez dans un ensemble de contenu. Utilisez un modèle probabiliste pour analyser une chaîne de données qui contient plusieurs valeurs de données. Un modèle probabiliste identifie les types d'informations présents dans chaque valeur de la chaîne. Vous pouvez ajouter un modèle probabiliste à une transformation Libellé et une transformation Analyseur.

Utilisez un modèle probabiliste dans une transformation Libellé pour attribuer un libellé descriptif à chaque valeur d'une chaîne d'entrée. La transformation Libellé écrit les libellés sur un seul port de sortie. Utilisez un modèle probabiliste dans une transformation Analyseur pour écrire chaque valeur d'une chaîne d'entrée sur un port qui représente les informations que contient la valeur. La transformation Analyseur crée un port de sortie pour chaque type d'information.

Vous concevez et compilez un modèle probabiliste dans l'outil Developer. Lorsque vous définissez un modèle probabiliste, vous ajoutez une série de lignes de données au modèle et vous attribuez un libellé à chaque valeur de chaque ligne. Lorsque vous compilez un modèle probabiliste, l'outil Developer crée des associations entre les valeurs de données et les libellés que vous avez ajoutés. La transformation Libellé et la transformation Analyseur utilisent des processus en langage naturel pour comparer les données du modèle probabiliste et les données du port d'entrée.

Les processus en langage naturel utilisent les techniques suivantes pour identifier les types d'informations qui figurent dans les valeurs de données :

- Les processus en langage naturel peuvent reconnaître les valeurs de données similaires et affecter le même libellé aux valeurs.
- Les processus en langage naturel peuvent comparer une valeur de données aux valeurs adjacentes dans la chaîne. Les processus en langage naturel analysent la séquence de valeurs pour comprendre l'utilisation de chaque chaîne et vérifier les types d'informations que les chaînes représentent.

Structure du modèle probabiliste

Un modèle probabiliste contient un ensemble de lignes de données et un ensemble correspondant de libellés. Les lignes de données contiennent des exemples des différentes valeurs qui peuvent s'afficher dans les données d'entrée de la transformation. Les libellés identifient les types d'informations que les valeurs d'entrée sont supposées contenir.

Lorsque vous définissez un modèle probabiliste, vous affectez chaque libellé à une ou plusieurs valeurs dans les lignes de données. Lorsque vous compilez le modèle, l'outil Developer génère un objet de données de référence qui représente les relations entre les valeurs de libellé et les valeurs de données.

L'outil Developer stocke les valeurs de libellé et les valeurs de données dans un fichier de données sur la machine hôte des services Informatica. Le fichier de données contient également les métadonnées qui définissent les associations entre les valeurs de libellé et les valeurs de données. Lorsque vous compilez un modèle probabiliste, vous actualisez les liens entre les valeurs de libellé et les valeurs de données.

Si vous supprimez le fichier de données utilisé par un modèle probabiliste, ce dernier passe en lecture seule. Vous ne pouvez pas compiler un modèle probabiliste en lecture seule.

Une ligne de données peut contenir une ou plusieurs valeurs. Chaque ligne de données peut présenter une structure différente. Vous pouvez attribuer le même libellé à plusieurs valeurs dans une ligne de données. Par ailleurs, vous pouvez attribuer un libellé différent à des valeurs identiques qui figurent à différentes positions dans une ligne. Affectez chaque libellé à une valeur de données au moins avant de compiler le modèle probabiliste.

Remarque: Pour optimiser les fonctionnalités du modèle probabiliste, vérifiez que chaque ligne de données contient plusieurs valeurs. L'ordre des valeurs sur chaque ligne doit correspondre autant que possible à l'ordre dans lequel les valeurs apparaissent dans les données source. Si les lignes de données contiennent des valeurs uniques, la transformation Libellé ou la transformation Analyseur ne peut pas appliquer les processus en langage naturel lors de l'analyse des données d'entrée.

Vues du modèle probabiliste

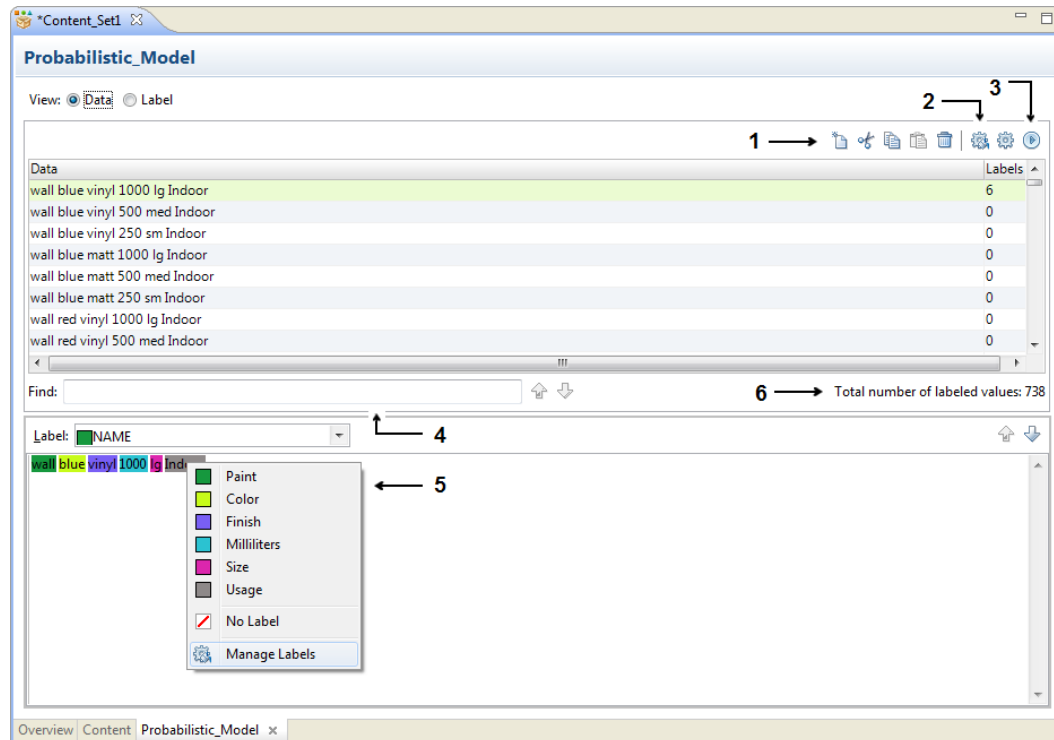
Lorsque vous modifiez un modèle probabiliste, vous pouvez travailler dans la vue Données ou Libellé. Utilisez la vue Données pour ajouter des lignes de données au modèle et attribuer des libellés aux valeurs de données dans chaque ligne. Utilisez la vue Libellé pour trouver des libellés inutilisés.

La vue Données affiche les lignes de données dans le modèle probabiliste et montre le nombre de libellés que vous assignez à chaque ligne. Lorsque vous sélectionnez une ligne, la ligne s'affiche dans un éditeur en dessous de la colonne de lignes de données. Pour attribuer un libellé à une valeur de données dans une ligne, cliquez avec le bouton droit de la souris sur la valeur dans l'éditeur et sélectionnez un libellé dans le menu. La

vue Données affiche également le nombre total de libellés que vous avez attribués aux valeurs dans le modèle probabiliste actuel.

La vue Libellé répertorie les libellés que vous définissez dans le modèle probabiliste. Lorsque vous sélectionnez un libellé, la vue Libellé affiche les lignes qui contiennent les valeurs que le libellé décrit. La vue Libellé affiche les valeurs de données que vous avez assignées au libellé sur chaque ligne.

L'image suivante affiche la vue Données d'un modèle probabiliste :



1. Insère une ligne de données.
2. Gère les libellés.
3. Compile le modèle.
4. Trouve une valeur dans les lignes de données. Utilisez les flèches pour trouver la ligne suivante ou la ligne précédente qui contient la valeur.
5. Assigne un libellé à une valeur de données.
6. Affiche le nombre total de valeurs de données qui ont un libellé.

Exemple de transformation Libellé

La base de données client d'une compagnie d'assurance contient plusieurs erreurs d'entrée de données. Vous êtes gestionnaire des données dans cette compagnie d'assurance. Vous configurez un mappage avec une transformation Libellé pour déterminer les différents types de données que contient chaque colonne.

Le tableau suivant décrit des exemples de données provenant de la base de données client :

ID ligne	Champ 1	Champ 2	Champ 3
1	19132954	AIM SECURITIES	PETRIE TAYBRO

ID ligne	Champ 1	Champ 2	Champ 3
2	10110169	JASE TRAPANI	BANK OF NEW YORK
3	10111786	WANGER ASSET MANAGEMENT, LLP	JAN SEEDORF
4	10112299	FELIX LEVENER	HARVARD MAGAZINE
5	10112036	DESCHÊNES & FILS LTÉE (QUEBEC)	RICHARD TREMBLAY
6	BERGER ASSOCIATES	10111101	DAREEN HULSMAN
7	19131385	EAGLE FINANCIAL GROUP INC	PATRICK MCKINNIE
8	LAKENYA PASKETT	WHITEHALL FINANCIAL GROUP	15954710

Lorsque vous exécutez le mappage, la transformation Libellé compare les données d'entrée aux données de référence du modèle probabiliste. La transformation Libellé attribue un libellé à chaque valeur d'entrée. La transformation écrit les libellés sur un port de sortie. Chaque ligne de sortie contient un ensemble de libellés qui définit la structure de données sur la ligne d'entrée correspondante.

Le tableau suivant décrit les libellés que la transformation Libellé ajoute au port de sortie :

ID ligne	Libellés de sortie
1	numéro du contact de l'organisation
2	contact de l'organisation, numéro
3	numéro du contact de l'organisation
4	contact de l'organisation, numéro
5	numéro du contact de l'organisation
6	organisation, numéro du contact
7	organisation, numéro du contact
8	numéro du contact, organisation

Exemple de transformation Analyseur

Un supermarché stocke des descriptions de produits dans une colonne d'une table de base de données. Ces descriptions contiennent plusieurs valeurs de données qui représentent différents types d'informations. Vous êtes gestionnaire des données dans ce supermarché. Vous voulez créer des colonnes pour les différents types d'informations figurant dans les descriptions des produits.

Vous configurez un mappage avec une transformation Analyseur pour organiser les valeurs de données dans les champs appropriés.

Le fragment de donnée suivant contient la description du produit « Jus d'orange » :

Jus d'orange Sunnydream sans sucre 350 ml

Le tableau suivant décrit les données de sortie que la transformation Analyseur crée à partir des données d'entrée :

Nom du produit	Type de produit	Détails du produit	Taille du produit
Sunnydream	Jus d'orange	Sans sucre	350 ml

Données de référence du modèle probabiliste

Les valeurs de données de référence d'un modèle probabiliste représentent les types de données d'entrée qu'une transformation Libellé ou Analyseur peut lire dans un mappage.

Vous pouvez ajouter, éditer et supprimer des lignes de données de référence dans l'outil Developer. Vous pouvez coller les données de référence depuis le Presse-papiers. Vous pouvez également utiliser les données source de mappage en tant que valeurs de données de référence dans un modèle probabiliste.

Une fois que vous avez ajouté les valeurs de données, créez les colonnes de libellé et attribuez un libellé à chaque valeur de données dans chaque ligne du modèle.

Remarque: Les valeurs de libellé du modèle indiquent la structure possible des lignes d'entrée lues par la transformation. Vérifiez que le modèle probabiliste contient la structure de libellé que vous prévoyez que la transformation trouvera dans les données source.

Ajout d'une chaîne de données de référence à un modèle probabiliste

Utilisez la vue Données pour ajouter une chaîne de données de référence à une ligne dans un modèle probabiliste.

1. Ouvrez l'ensemble de contenu qui contient le modèle.
Sélectionnez le nom du modèle et cliquez sur **Editer**.
2. Sélectionnez la vue Données.
3. Cliquez sur **Nouveau** pour ajouter une ligne vide au modèle.
4. Entrez une chaîne de données dans la ligne vide.
5. Attribuez un libellé à chaque valeur de la chaîne.
6. Enregistrez le modèle probabiliste.

Après avoir enregistré le modèle, vous pouvez éventuellement le compiler.

Recherche de valeurs de données dans un modèle probabiliste

Utilisez la vue Données pour trouver les lignes de données dans le modèle qui contient une valeur de données.

1. Ouvrez l'ensemble de contenu qui contient le modèle.
Sélectionnez le nom du modèle et cliquez sur **Modifier**.
2. Sélectionnez la vue Données.

3. Entrez une valeur de données dans le champ **Rechercher**.

Vous pouvez également entrer les caractères génériques suivants :

* ?

La vue Données affiche la première ligne dans le modèle contenant la valeur que vous avez saisie.

4. Utilisez la flèche vers le haut ou vers le bas pour afficher chaque ligne contenant la valeur.

Données de libellé du modèle probabiliste

Les valeurs de libellé dans un modèle probabiliste représentent les types d'informations que les valeurs de données de référence peuvent contenir. Lorsque vous ajoutez des lignes de données à un modèle, attribuez un libellé à chaque valeur de données dans chaque ligne. Les libellés que vous ajoutez au modèle s'affichent dans la vue Libellé et dans les options de menu de la vue Données.

Vous pouvez attribuer tous les libellés du modèle à toutes les valeurs de données de référence. Si la même valeur présente des significations différentes dans deux lignes de données de référence, vous pouvez attribuer des libellés différents à la valeur dans chaque ligne.

Vous pouvez définir la même combinaison de libellés pour plusieurs chaînes d'entrée. La présence de plusieurs exemples d'un libellé augmente l'exactitude du modèle probabiliste.

Libellé de dépassement

Lorsqu'une transformation ne peut pas attribuer un libellé que vous définissez à une valeur de données d'entrée, elle attribue un libellé de dépassement aux données.

La transformation Libellé attribue un libellé de dépassement à toutes les valeurs de données qu'elle ne peut pas identifier. La transformation Analyseur crée une colonne de dépassement pour les données non attribuées.

Le tableau suivant montre comment une transformation Analyseur utilise un port de dépassement pour analyser les éléments de données d'adresse qu'un modèle probabiliste ne reconnaît pas :

Données d'entrée	Port Street_Name	Port Street_Descriptor	Port de dépassement
Park Place	Park	Place	
Park Avenue	Park	Avenue	
Madison avenue	Madison	Avenue	
Central Park	Central	Park	
Washington Square Park	Washington	Square	Park
Madison Square Garden	Madison	Square	Garden

La transformation affecte les valeurs au port de débordement lorsque le nombre de valeurs de données d'entrée est supérieur au nombre de libellés figurant dans le modèle probabiliste. Avant d'utiliser un modèle dans une transformation, vérifiez les données source du mappage et assurez-vous que le modèle contient le nombre approprié de valeurs de libellé.

Attribution de libellés à des données de modèle probabiliste

Attribuez un libellé à toutes les valeurs de données dans un modèle probabiliste.

Vous pouvez attribuer des libellés différents à une même valeur de données si celle-ci apparaît à différents emplacements dans les lignes de données. Si une valeur de données utilise un libellé incorrect, vous pouvez le mettre à jour.

1. Ouvrez l'ensemble de contenu qui contient le modèle.
2. Sélectionnez le nom du modèle et cliquez sur **Editer**.
3. Sélectionnez la vue Données.
4. Recherchez la valeur de données qui n'a pas de libellé ou dont le libellé est incorrect.
5. Sélectionnez la ligne de données qui contient la valeur de données.

La ligne s'affiche dans l'éditeur.

6. Cliquez avec le bouton droit de la souris sur une valeur de données dans l'éditeur et sélectionnez un libellé dans le menu contextuel.

L'outil Developer attribue le libellé à la valeur de données.

7. Enregistrez le modèle probabiliste.

Après avoir enregistré le modèle, vous pouvez éventuellement le compiler.

Ajout d'un libellé à un modèle probabiliste

Ajoutez un libellé pour chaque type d'information que les valeurs de données du modèle représentent. Si vous utilisez le modèle probabiliste dans une transformation Analyseur, ajoutez un libellé pour chaque port de sortie qui sera créé par la transformation.

1. Ouvrez l'ensemble de contenu qui contient le modèle.
2. Sélectionnez le nom du modèle et cliquez sur **Modifier**.
3. Dans la vue Données ou Libellé, cliquez sur **Gérer les libellés**.
La boîte de dialogue **Gérer les libellés** s'affiche.
4. Dans la boîte de dialogue **Gérer les libellés**, cliquez sur **Nouveau**.
Un libellé s'affiche dans la première ligne vide de la boîte de dialogue.
5. Modifiez le nom du libellé. Éventuellement, mettez à jour la couleur du libellé.
6. Cliquez sur **OK** pour ajouter le libellé au modèle.
7. Enregistrez le modèle probabiliste.

Après avoir ajouté le libellé, vous devez l'affecter à une valeur de données au moins. Compilez le modèle avant de l'utiliser dans une transformation.

Suppression d'un libellé d'un modèle probabiliste

Lorsque vous supprimez un libellé d'un modèle, toutes les valeurs de données associées à ce libellé sont conservées dans le modèle. Attribuez un autre libellé à chaque valeur de données.

1. Ouvrez l'ensemble de contenu qui contient le modèle.
2. Sélectionnez le nom du modèle et cliquez sur **Modifier**.
3. Dans la vue Données ou Libellé, cliquez sur **Gérer les libellés**.
La boîte de dialogue **Gérer les libellés** s'affiche.

4. Dans la boîte de dialogue **Gérer les libellés**, cliquez sur **Supprimer**.
5. Cliquez sur **OK** pour supprimer le libellé.
6. Enregistrez le modèle probabiliste.

Remarque: Un libellé est un élément structurel dans un modèle probabiliste. Si vous ajoutez ou supprimez un libellé après avoir ajouté le modèle à une transformation, vous invalidez l'opération qui utilise le modèle. Pour utiliser le modèle que vous avez mis à jour, supprimez et recréez l'opération de transformation.

Compilation du modèle probabiliste

À chaque fois que vous mettez à jour un modèle probabiliste, vous pouvez compiler le modèle. Compilez le modèle pour mettre à jour sa logique avec les valeurs de données et de libellé actuelles.

Avant de compiler le modèle, vérifiez que toutes les valeurs de libellé identifient au moins une valeur de données.

- Pour compiler le modèle, ouvrez-le dans l'outil Developer et cliquez sur **Compiler**.

Propriétés du modèle probabiliste

Vous pouvez afficher et mettre à jour les propriétés générales et les propriétés avancées d'un modèle probabiliste.

Pour ouvrir l'éditeur de propriétés, sélectionnez l'option **Modifier les propriétés** dans la vue Données ou l'option **Modifier les propriétés avancées** dans la vue Libellé.

Les propriétés générales affichent le nom du modèle probabiliste, son éventuelle description et le nom du fichier de données du modèle. Les propriétés avancées affichent les propriétés de calcul que l'outil Developer utilise pour compiler le modèle probabiliste.

L'élément de base de la compilation de modèles probabilistes est le *n-gramme*. Un n-gramme est une série de lettres qui suit ou précède d'autres lettres pour former un mot. Lorsqu'un mappage s'exécute, la transformation Libellé ou Analyseur crée plusieurs n-grammes pour chaque valeur de la colonne de données de référence du modèle probabiliste. La transformation compare les valeurs des données d'entrée aux valeurs de données de référence et aux n-grammes. Les propriétés avancées d'un modèle probabiliste déterminent la manière dont le modèle probabiliste gère les n-grammes et d'autres fonctionnalités de modèle.

Remarque: Les valeurs par défaut dans les propriétés avancées représentent les paramètres souhaités pour l'analyse probabiliste et la compilation du modèle probabiliste. Si vous modifiez une propriété avancée, vous risquez d'affecter négativement la précision de l'analyse probabiliste. Ne modifiez pas les propriétés avancées sauf si vous comprenez les conséquences des modifications que vous effectuez.

Création d'un modèle probabiliste à partir d'un objet de données

Vous pouvez utiliser un objet de données en tant que source de données d'un modèle probabiliste. Par exemple, utilisez l'objet de données source à partir du mappage qui va lire le modèle probabiliste. Vous

pouvez également profiler un objet dans le mappage et créer un objet de données à partir des résultats du profil.

Un modèle probabiliste fonctionne de manière optimale lorsque vous utilisez les données d'entrée d'une transformation Libellé ou Analyseur en tant que source des données de référence du modèle. Par exemple, vous pouvez exécuter un profil sur l'objet de transformation dans le mappage. Créez un objet de données à partir des résultats de profil.

1. Ouvrez ou créez un jeu de contenu.
2. Sélectionnez la vue **Contenu**.
3. Sélectionnez **Modèles probabilistes**, puis cliquez sur **Ajouter**.
L'assistant Modèle probabiliste s'affiche à l'écran.
4. Sélectionnez l'option **Modèle probabiliste issu d'objets de données**.
Cliquez sur **Suivant**.
5. Saisissez un nom pour le modèle probabiliste.
Entrez éventuellement une description textuelle du modèle.
6. Parcourez le référentiel modèle et sélectionnez l'objet de données qui contient les données de référence.
Cliquez sur **Suivant**.
7. Vérifiez les colonnes de données dans l'objet de données et sélectionnez les colonnes à ajouter au modèle probabiliste.
 - Pour ajouter une colonne de données en tant que données de référence, sélectionnez le nom de colonne et cliquez sur **Données**. Vous pouvez ajouter plusieurs colonnes de données de référence. L'opération d'importation joint les colonnes de façon que les valeurs de la même ligne dans chaque colonne s'affichent dans la même ligne de données dans le modèle.
 - Pour utiliser une colonne de données en tant que source pour les valeurs de libellé, sélectionnez le nom de la colonne et cliquez sur **Libellé**. Vous pouvez ajouter plusieurs colonnes de données de libellé. L'opération d'importation concatène les colonnes de façon que les valeurs de chaque colonne s'affichent l'une sous l'autre dans une colonne unique dans le modèle.Cliquez sur **Suivant**.
8. Sélectionnez le nombre de lignes à copier à partir de la source de données.
Sélectionnez toutes les lignes ou saisissez le nombre de lignes à copier. Si vous saisissez un nombre, le modèle compte les lignes depuis le début de l'ensemble de données.
9. Définissez les délimiteurs à utiliser pour les valeurs de données de référence. Spécifiez un délimiteur pour identifier plusieurs valeurs qui représentent une seule information.
Le délimiteur par défaut est un espace.
10. Cliquez sur **Terminer** et enregistrez le modèle.
Le modèle probabiliste s'ouvre dans l'outil Developer.
Après avoir créé le modèle probabiliste, affectez les libellés aux valeurs de données et compilez le modèle.

Opérations de copier/coller

Vous pouvez copier un modèle probabiliste d'un ensemble de contenu vers un autre dans un référentiel modèle. Copiez un modèle probabiliste pour partager les ressources avec d'autres utilisateurs de l'outil Developer.

Vous pouvez copier un modèle vers un autre ensemble de contenu ou importer un modèle dans l'ensemble de contenu actuel. Vous pouvez importer plusieurs modèles de plusieurs ensembles de contenu dans le référentiel en une seule fois.

Lorsque vous copiez un modèle, le service de gestion du contenu crée une copie du fichier de données du modèle sur la machine hôte des services Informatica. Chaque modèle utilise un fichier de données différent.

Copie d'un modèle probabiliste dans un autre ensemble de contenu

Vous pouvez copier un modèle probabiliste d'un ensemble de contenu vers un autre dans un référentiel modèle. Lorsque vous copiez un modèle probabiliste, vous spécifiez l'objet de modèle et les ensembles de contenu source et de destination.

1. Ouvrez l'ensemble de contenu qui contient le modèle probabiliste.
2. Sélectionnez un modèle probabiliste et cliquez sur **Copier dans**.
3. Parcourez le référentiel modèle et sélectionnez un ensemble de contenu.

Vous pouvez copier le modèle probabiliste vers un ensemble de contenu dans le projet en cours ou un autre projet.

4. Cliquez sur **OK**.

L'outil Developer copie le modèle probabiliste dans l'ensemble de contenu sélectionné.

Importation d'un modèle probabiliste depuis un autre ensemble de contenu

Vous pouvez importer un modèle probabiliste d'un ensemble de contenu vers un autre dans un référentiel modèle. Lorsque vous importez un modèle probabiliste, vous spécifiez un ou plusieurs objets de modèle et les ensembles de contenu source et de destination.

1. Ouvrez l'ensemble de contenu devant contenir le modèle probabiliste.
2. Sélectionnez un modèle probabiliste et cliquez sur **Coller depuis**.
3. Parcourez le référentiel modèle et sélectionnez un modèle probabiliste.

Vous pouvez coller le modèle probabiliste d'un ensemble de contenu dans le projet actuel ou un autre projet.

4. Cliquez sur **OK**.

L'outil Developer colle le modèle probabiliste dans l'ensemble de contenu actuel.

INDEX

C

créer une table de référence depuis les colonnes de profil
tables de référence [17](#)
Créer une table de référence depuis les modèles de colonnes
tables de référence [19](#)
créer une table de référence depuis les valeurs de colonne
tables de référence [18](#)
créer une table de référence manuellement
tables de référence [16](#)

E

ensembles de caractères [34](#)
ensembles de contenu
ensembles de caractères [34](#)
ensembles de jetons [36](#)
ensembles de modèles [35](#)
expressions régulières [36](#)
modèles de classeur [34](#)
modèles probabilistes [35](#)
ensembles de jetons [36](#)
ensembles de modèles [35](#)
exporter une table de référence
tables de référence [24](#)
expressions régulières [36](#)

G

gérer les colonnes
tables de référence [23](#)
gérer les lignes
tables de référence [23](#)

I

importer une table de référence
tables de référence [20](#)

M

modèles de classeur
dans les ensembles de contenu [34](#)

modèles de classeur (*a continué*)
règles et directives [38](#)
modèles probabilistes
dans les ensembles de contenu [35](#)
règles et directives [38](#)

O

Outil Analyst
rechercher et remplacer des valeurs de données de référence [24](#)

S

service de gestion de contenu
privilèges de la table de référence [13](#)

T

tables de référence
Actualisation dans l'outil Analyst [25](#)
créer une table de référence depuis les colonnes de profil [17](#)
Créer une table de référence depuis les modèles de colonnes [19](#)
créer une table de référence depuis les valeurs de colonne [18](#)
créer une table de référence manuellement [16](#)
exporter une table de référence [24](#)
gérées et non gérées [13](#)
gérer les colonnes [23](#)
gérer les lignes [23](#)
importer une table de référence [20](#)
Présentation de l'outil Analyst [14](#)
privilèges [13](#)
propriétés dans l'outil Analyst [15](#)
propriétés dans l'outil Developer [12](#), [29](#)
rechercher et remplacer des valeurs dans l'outil Analyst [24](#)
voir les tables de suivi d'audit [26](#)

V

voir les événements de table d'audit
tables de référence [26](#)