



Informatica® Data Quality Integration for
PowerCenter

10.1.1

Guia do Usuário

© Copyright Informatica LLC 2009, 2018

Este software e documentação contêm informações de propriedade da Informatica LLC, são fornecidos sob um contrato de licença que contém restrições quanto a seu uso e divulgação, e são protegidos por leis de copyright. A engenharia reversa do software é proibida. Não está permitida de forma alguma a reprodução ou a transmissão de qualquer parte deste documento (seja por meio eletrônico, fotocópia, gravação ou quaisquer outros) sem o consentimento prévio da Informatica LLC. Este Software pode estar protegido por patentes dos EUA e/ou internacionais e outras patentes pendentes.

O uso, duplicação ou divulgação do Software pelo Governo dos Estados Unidos estão sujeitos às restrições estipuladas no contrato de licença de software aplicável e como estabelecido em DFARS 227.7202-1(a) e 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), conforme aplicável.

As informações contidas neste produto ou documentação estão sujeitas a alteração sem aviso prévio. Informe-nos por escrito caso encontre quaisquer problemas neste produto ou documentação.

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, Informatica Master Data Management e Live Data Map são marcas comerciais ou marcas registradas da Informatica LLC nos Estados Unidos e em jurisdições pelo mundo. Todos os outros nomes de outras companhias e produtos podem ser nomes ou marcas comerciais de seus respectivos proprietários.

Partes desta documentação e/ou software estão sujeitas a copyright de terceiros, incluindo sem limitação: Copyright DataDirect Technologies. Todos os direitos reservados. Copyright © Sun Microsystems. Todos os direitos reservados. Copyright © RSA Security Inc. Todos os direitos reservados. Copyright © Ordinal Technology Corp. Todos os direitos reservados. Copyright © Aandacht c.v. Todos os direitos reservados. Copyright Genivia, Inc. Todos os direitos reservados. Copyright Isomorphic Software. Todos os direitos reservados. Copyright © Meta Integration Technology, Inc. Todos os direitos reservados. Copyright © Intalio. Todos os direitos reservados. Copyright © Oracle. Todos os direitos reservados. Copyright © Adobe Systems Incorporated. Todos os direitos reservados. Copyright © DataArt, Inc. Todos os direitos reservados. Copyright © ComponentSource. Todos os direitos reservados. Copyright © Microsoft Corporation. Todos os direitos reservados. Copyright © Rogue Wave Software, Inc. Todos os direitos reservados. Copyright © Teradata Corporation. Todos os direitos reservados. Copyright © Yahoo! Inc. Todos os direitos reservados. Copyright © Glyph & Cog, LLC. Todos os direitos reservados. Copyright © Thinkmap, Inc. Todos os direitos reservados. Copyright © Clearpace Software Limited. Todos os direitos reservados. Copyright © Information Builders, Inc. Todos os direitos reservados. Copyright © OSS Nokalva, Inc. Todos os direitos reservados. Copyright Edifecs, Inc. Todos os direitos reservados. Copyright Cleo Communications, Inc. Todos os direitos reservados. Copyright © International Organization for Standardization 1986. Todos os direitos reservados. Copyright © ej-technologies GmbH. Todos os direitos reservados. Copyright © Jaspersoft Corporation. Todos os direitos reservados. Copyright © International Business Machines Corporation. Todos os direitos reservados. Copyright © yWorks GmbH. Todos os direitos reservados. Copyright © Lucent Technologies. Todos os direitos reservados. Copyright © University of Toronto. Todos os direitos reservados. Copyright © Daniel Veillard. Todos os direitos reservados. Copyright © Unicode, Inc. Copyright IBM Corp. Todos os direitos reservados. Copyright © MicroQuill Software Publishing, Inc. Todos os direitos reservados. Copyright © PassMark Software Pty Ltd. Todos os direitos reservados. Copyright © LogiXML, Inc. Todos os direitos reservados. Copyright © 2003-2010 Lorenzi Davide, todos os direitos reservados. Copyright © Red Hat, Inc. Todos os direitos reservados. Copyright © The Board of Trustees of the Leland Stanford Junior University. Todos os direitos reservados. Copyright © EMC Corporation. Todos os direitos reservados. Copyright © Flexera Software. Todos os direitos reservados. Copyright © Jinfonet Software. Todos os direitos reservados. Copyright © Apple Inc. Todos os direitos reservados. Copyright © Telerik Inc. Todos os direitos reservados. Copyright © BEA Systems. Todos os direitos reservados. Copyright © PDFlib GmbH. Todos os direitos reservados. Copyright © Orientation in Objects GmbH. Todos os direitos reservados. Copyright © Tanuki Software, Ltd. Todos os direitos reservados. Copyright © Ricebridge. Todos os direitos reservados. Copyright © Sencha, Inc. Todos os direitos reservados. Copyright © Scalable Systems, Inc. Todos os direitos reservados. Copyright © jqWidgets. Todos os direitos reservados. Copyright © Tableau Software, Inc. Todos os direitos reservados. Copyright © MaxMind, Inc. Todos os direitos reservados. Copyright © TMate Software s.r.o. Todos os direitos reservados. Copyright © MapR Technologies Inc. Todos os direitos reservados. Copyright © Amazon Corporate LLC. Todos os direitos reservados. Copyright © Highsoft. Todos os direitos reservados. Copyright © Python Software Foundation. Todos os direitos reservados. Copyright © BeOpen.com. Todos os direitos reservados. Copyright © CNRI. Todos os direitos reservados.

Este produto inclui software desenvolvido pela Apache Software Foundation (<http://www.apache.org/>) e/ou outros softwares licenciados nas várias versões da Licença Apache (a "Licença"). Você pode obter uma cópia dessas Licenças em <http://www.apache.org/licenses/>. A menos que exigido pela legislação aplicável ou concordado por escrito, o software distribuído em conformidade com estas Licenças é fornecido "NO ESTADO EM QUE SE ENCONTRA", SEM GARANTIA OU CONDIÇÃO DE QUALQUER TIPO, seja expressa ou implícita. Consulte as Licenças para conhecer as limitações e as permissões que regulam o idioma específico de acordo com as Licenças.

Este produto inclui software desenvolvido pela Mozilla (<http://www.mozilla.org/>), direitos autorais de software de The JBoss Group, LLC; todos os direitos reservados; software copyright © 1999-2006 de Bruno Lowagie e Paulo Soares e outros produtos de software licenciados sob a Licença Pública GNU Lesser General Public License Agreement, que pode ser encontrada em <http://www.gnu.org/licenses/lgpl.html>. Os materiais são fornecidos gratuitamente pela Informatica, no estado em que se encontram, sem garantia de qualquer tipo, explícita nem implícita, incluindo, mas não limitando-se, as garantias implicadas de comerciabilidade e adequação a um determinado propósito.

O produto inclui software ACE(TM) e TAO(TM) com copyright de Douglas C. Schmidt e seu grupo de pesquisa na Washington University, University of California, Irvine e Vanderbilt University, Copyright (©) 1993-2006, todos os direitos reservados.

Este produto inclui o software desenvolvido pelo OpenSSL Project para ser usado no kit de ferramentas OpenSSL (copyright The OpenSSL Project. Todos os direitos reservados) e a redistribuição deste software está sujeita aos termos disponíveis em <http://www.openssl.org> e <http://www.openssl.org/source/license.html>.

Este produto inclui o software Curl com o Copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://curl.haxx.se/docs/copyright.html>. É permitido usar, copiar, modificar e distribuir este software com qualquer objetivo, com ou sem taxa, desde que a nota de direitos autorais acima e esta nota de permissão apareçam em todas as cópias.

O produto inclui software copyright 2001-2005 (©) MetaStuff, Ltd. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://www.dom4j.org/license.html>.

O produto inclui o copyright de software © 2004-2007, The Dojo Foundation. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://dojotoolkit.org/license>.

Este produto inclui o software ICU com o copyright International Business Machines Corporation e outros. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://source.icu-project.org/repos/icu/icu/trunk/license.html>.

Este produto inclui o copyright de software © 1996-2006 Per Bothner. Todos os direitos reservados. O direito de usar tais materiais é estabelecido na licença que pode ser encontrada em <http://www.gnu.org/software/kawa/Software-License.html>.

Este produto inclui o software OSSP UUID com Copyright © 2002 Ralf S. Engelschall, Copyright © 2002 e OSSP Project Copyright © 2002 Cable & Wireless Deutschland. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://www.opensource.org/licenses/mit-license.php>.

Este produto inclui software desenvolvido pela Boost (<http://www.boost.org/>) ou sob a licença de software Boost. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em http://www.boost.org/LICENSE_1_0.txt.

Este produto inclui software copyright © 1997-2007 University of Cambridge. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://www.pcre.org/license.txt>.

Este produto inclui o copyright de software © 2007 The Eclipse Foundation. Todos os direitos reservados. As permissões e as limitações relativas a este software estão sujeitas aos termos disponíveis em <http://www.eclipse.org/org/documents/epl-v10.php> e em <http://www.eclipse.org/org/documents/edl-v10.php>.

Este produto inclui softwares licenciados de acordo com os termos disponíveis em <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/license.html>, <http://www.jgraph.com/jgraphdownload.html>, <http://www.jcraft.com/jsch/LICENSE.txt>, http://jotm.objectweb.org/bsd_license.html, <http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231>, <http://www.slf4j.org/license.html>, <http://nanoxml.sourceforge.net/orig/copyright.html>, <http://www.json.org/license.html>, <http://forge.ow2.org/projects/jaservice/>, <http://www.postgresql.org/about/license.html>, <http://www.sqlite.org/copyright.html>, <http://www.tcl.tk/software/tcltk/license.html>, <http://www.jaxen.org/faq.html>, <http://www.jdom.org/docs/faq.html>, <http://www.slf4j.org/license.html>, <http://www.iodbc.org/dataspace/iodbc/wiki/iodbc/License>, <http://www.keplerproject.org/md5/license.html>, <http://www.toedter.com/en/jcalendar/license.html>, <http://www.edankert.com/bounce/index.html>, <http://www.net-snmp.org/about/license.html>, <http://www.openmdx.org/#FAQ>, http://www.php.net/license/3_01.txt, <http://srp.stanford.edu/license.txt>, <http://www.schneier.com/blowfish.html>, <http://www.jmock.org/license.html>, <http://xsom.java.net>, <http://benalman.com/about/license/>, <https://github.com/CreateJS/EaselJS/blob/master/src/easeljs/display/Bitmap.js>, <http://www.h2database.com/html/license.html#summary>, <http://jsoncpp.sourceforge.net/LICENSE>, <http://jdbc.postgresql.org/license.html>, <http://protobuf.googlecode.com/svn/trunk/src/google/protobuf/descriptor.proto>, <https://github.com/rantav/hector/blob/master/LICENSE>, <http://web.mit.edu/Kerberos/krb5-current/doc/mitK5license.html>, <http://jibx.sourceforge.net/jibx-license.html>, <https://github.com/lyokato/libgeohash/blob/master/LICENSE>, <https://github.com/hjiang/jsonxx/blob/master/LICENSE>, <https://code.google.com/p/lz4/>, <https://github.com/jedisct1/libsodium/blob/master/LICENSE>, <http://one-jar.sourceforge.net/index.php?page=documents&file=license>, <https://github.com/EsotericSoftware/kryo/blob/master/license.txt>, <http://www.scala-lang.org/license.html>, <https://github.com/tinkerpop/blueprints/blob/master/LICENSE.txt>, <http://gee.cs.oswego.edu/dl/classes/EDU/oswego/cs/dl/util/concurrent/intro.html>, <https://aws.amazon.com/asl/>, <https://github.com/twbs/bootstrap/blob/master/LICENSE>, <https://sourceforge.net/p/xmlunit/code/HEAD/tree/trunk/LICENSE.txt>.

Este produto inclui software licenciado de acordo com a Academic Free License (<http://www.opensource.org/licenses/afl-3.0.php>), a Common Development and Distribution License (<http://www.opensource.org/licenses/cddl1.php>), a Common Public License (<http://www.opensource.org/licenses/cpl1.0.php>), a Sun Binary Code License Agreement Supplemental License Terms, a BSD License (<http://www.opensource.org/licenses/bsd-license.php>), a nova BSD License (<http://opensource.org/licenses/BSD-3-Clause>), a MIT License (<http://www.opensource.org/licenses/mit-license.php>), a Artistic License (<http://www.opensource.org/licenses/artistic-license-1.0>) e a Initial Developer's Public License Version 1.0 (<http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/>).

Este produto inclui copyright do software © 2003-2006 Joe Walnes, 2006-2007 XStream Committers. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://xstream.codehaus.org/license.html>. Este produto inclui software desenvolvido pelo Indiana University Extreme! Lab. Para obter mais informações, visite <http://www.extreme.indiana.edu/>.

Este produto inclui software Copyright © 2013 Frank Balluffi e Markus Moeller. Todos os direitos reservados. As permissões e limitações relativas a este software estão sujeitas aos termos da licença MIT.

Consulte as patentes em <https://www.informatica.com/legal/patents.html>.

ISENÇÃO DE RESPONSABILIDADE: a Informatica LLC fornece esta documentação no estado em que se encontra, sem garantia de qualquer tipo, expressa ou implícita, incluindo, mas não limitando-se, as garantias implícitas de não infração, comercialização ou uso para um determinado propósito. A Informatica LLC não garante que este software ou documentação não contenha erros. As informações fornecidas neste software ou documentação podem incluir imprecisões técnicas ou erros tipográficos. As informações deste software e documentação estão sujeitas a alterações a qualquer momento sem aviso prévio.

AVISOS

Este produto da Informatica (o "Software") traz determinados drivers (os "drivers da DataDirect") da DataDirect Technologies, uma empresa em funcionamento da Progress Software Corporation ("DataDirect"), que estão sujeitos aos seguintes termos e condições:

1. OS DRIVERS DA DATADIRECT SÃO FORNECIDOS NO ESTADO EM QUE SE ENCONTRAM, SEM GARANTIA DE QUALQUER TIPO, EXPRESSA OU IMPLÍCITA, INCLUINDO, MAS NÃO LIMITANDO-SE, AS GARANTIAS IMPLÍCITAS DE COMERCIALIZAÇÃO, ADEQUAÇÃO A UMA FINALIDADE ESPECÍFICA E NÃO INFRAÇÃO.
2. EM NENHUM CASO, A DATADIRECT OU SEUS FORNECEDORES TERCEIRIZADOS SERÃO RESPONSÁVEIS, EM RELAÇÃO AO CLIENTE FINAL, POR QUALQUER DANOS DIRETOS, INDIRETOS, INCIDENTAIS, ESPECIAIS, CONSEQUENCIAIS OU DEMAIS QUE POSSAM ADVIR DO USO DE DRIVERS ODBC, SENDO OU NÃO ANTERIORMENTE INFORMADOS DAS POSSIBILIDADES DE TAIS DANOS. ESTAS LIMITAÇÕES SE APLICAM A TODAS AS CAUSAS DE AÇÃO, INCLUINDO, SEM LIMITAÇÕES, QUEBRA DE CONTRATO, QUEBRA DE GARANTIA, NEGLIGÊNCIA, RESPONSABILIDADE RIGOROSA, DETURPAÇÃO E OUTROS ATOS ILÍCITOS.

Data da Publicação: 2018-07-03

Conteúdo

Prefácio.....	5
Recursos da Informatica.	5
Rede da Informatica.	5
Base de Dados de Conhecimento da Informatica.	5
Documentação da Informatica.	5
Matrizes de Disponibilidade de Produto Informatica.	6
Informatica Velocity.	6
Informatica Marketplace.	6
Suporte global a clientes Informatica.	6
 Capítulo 1: Integração do Informatica Data Quality.....	 7
Visão Geral de Integração da Qualidade de Dados.	7
Recursos e Funcionalidade da Qualidade de Dados.	8
Instaladores de Integração.	9
Exportação de Objetos para o PowerCenter.	9
Instalação de Dados de Referência.	9
Matriz de Integração.	10
Regras e Diretrizes para a Integração de Produtos.	10
Comportamento das Transformações de Qualidade de Dados no PowerCenter.	11
 Capítulo 2: Instalação	 14
Visão Geral da Instalação.	14
Requisitos de Sistema.	14
Antes da Instalação.	15
Instalando o Plug-in de Servidor.	15
Instalando no Windows.	15
Instalando no UNIX.	16
Instalando o Plug-in Cliente.	16
Registrando o Plug-In no PowerCenter.	17
 Capítulo 3: Requisitos de Dados de Referência.....	 18
Visão Geral dos Dados de Referência.	18
Requisitos do Arquivo de Dados de Referência.	19
Requisitos de Dados de Referência de Endereço.	20
Requisitos de Dados de Preenchimento de Identidade.	21
Requisitos do Modelo de Probabilidade e do Modelo de Classificador.	23
Localização de Arquivos do Modelo Probabilístico e do Modelo de Classificador.	23
Alocação de Memória do Serviço de Integração do PowerCenter.	23

Prefácio

Os usuários do PowerCenter podem trabalhar com mapeamentos e objetos de dados de referência que eles importam de um ambiente do Informatica Data Quality. O *Guia do Usuário de Integração de Informatica Data Quality para PowerCenter* foi elaborado para os usuários de PowerCenter que executam mapeamentos de qualidade de dados e gerenciam os objetos de dados de referência associados.

Recursos da Informatica

Rede da Informatica

A Rede da Informatica hospeda o Suporte Global a Clientes da Informatica, a Base de Dados de Conhecimento da Informatica e outros recursos de produtos. Para acessar a Rede da Informatica, visite <https://network.informatica.com>.

Como membro, você pode:

- Acessar todos os seus recursos Informatica em um só lugar.
- Pesquisar a Base de Dados de Conhecimento em busca de recursos de produtos, incluindo documentações, perguntas frequentes e práticas recomendadas.
- Visualizar informações sobre disponibilidade de produtos.
- Revisar seus casos de suporte.
- Encontrar a sua Rede de Grupo de Usuários da Informatica local e colaborar com seus colegas.

Base de Dados de Conhecimento da Informatica

Use a Base de Dados de Conhecimento da Informatica para pesquisar a Rede da Informatica em busca de recursos de produtos, como documentações, artigos de instruções, práticas recomendadas e PAMs.

Para acessar a Base de Dados de Conhecimento, visite <https://kb.informatica.com>. Em caso de dúvidas, comentários ou ideias sobre a Base de Dados de Conhecimento, entre em contato com a equipe da Base de Dados de Conhecimento da Informatica em KB_Feedback@informatica.com.

Documentação da Informatica

Para obter a documentação mais recente do seu produto, navegue pela Base de Dados de Conhecimento da Informatica em https://kb.informatica.com/_layouts/ProductDocumentation/Page/ProductDocumentSearch.aspx

Em caso de dúvidas, comentários ou ideias sobre esta documentação, entre em contato com a equipe de Documentação da Informatica pelo e-mail infa_documentation@informatica.com.

Matrizes de Disponibilidade de Produto Informatica

As Matrizes de Disponibilidade de Produto (PAMs) indicam as versões dos sistemas operacionais, os bancos de dados e outros tipos de fontes e destinos de dados com os quais uma versão de produto é compatível. Se você for membro da Rede da Informatica, poderá acessar PAMs em <https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Velocity

O Informatica Velocity é uma coleção de dicas e práticas recomendadas desenvolvidas pelos Serviços Profissionais da Informatica. Desenvolvido com base na experiência no mundo real de centenas de projetos de gerenciamento de dados, o Informatica Velocity representa o conhecimento coletivo de nossos consultores, que trabalharam com organizações de todo o mundo para planejar, desenvolver, implantar e manter soluções de gerenciamento de dados bem-sucedidas.

Se você for membro da Rede da Informatica, poderá acessar os recursos do Informatica Velocity em <http://velocity.informatica.com>.

Se você tiver dúvidas, comentários ou ideias sobre o Informatica Velocity, entre em contato com os Serviços Profissionais da Informatica em ips@informatica.com.

Informatica Marketplace

O Informatica Marketplace é um fórum onde você pode encontrar soluções que aumentam, ampliam ou aprimoram suas implementações da Informatica. Aproveitando qualquer uma das centenas de soluções fornecidas por desenvolvedores e parceiros da Informatica, você pode melhorar sua produtividade e agilizar o tempo de implementação nos seus projetos. Você pode acessar o Informatica Marketplace através do link <https://marketplace.informatica.com>.

Suporte global a clientes Informatica

Você pode entrar em contato com um Centro de Suporte Global por telefone ou via Suporte Online na Rede da Informatica.

Para descobrir o número de telefone local do Suporte Global a Clientes da Informatica, visite o site da Informatica no seguinte link:

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

Se você for membro da Rede da Informatica, poderá usar o Suporte Online em <http://network.informatica.com>.

CAPÍTULO 1

Integração do Informatica Data Quality

Este capítulo inclui os seguintes tópicos:

- [Visão Geral de Integração da Qualidade de Dados, 7](#)
- [Recursos e Funcionalidade da Qualidade de Dados, 8](#)
- [Instaladores de Integração, 9](#)
- [Exportação de Objetos para o PowerCenter, 9](#)
- [Instalação de Dados de Referência, 9](#)
- [Matriz de Integração, 10](#)
- [Regras e Diretrizes para a Integração de Produtos, 10](#)
- [Comportamento das Transformações de Qualidade de Dados no PowerCenter, 11](#)

Visão Geral de Integração da Qualidade de Dados

Os usuários do Informatica podem criar e executar mapeamentos no Informatica Data Quality. Os usuários salvam os mapeamentos e outros objetos no repositório do Modelo. Os usuários podem exportar os objetos para uso no PowerCenter. Como um usuário do PowerCenter, você pode importar os objetos para o repositório do PowerCenter e adicionar os objetos a sessões.

Use os objetos para analisar e aprimorar o conteúdo e a estrutura dos seus dados.

Você pode usar os objetos que os usuários criaram no Data Quality 10.1.1 com as seguintes versões do PowerCenter:

- PowerCenter 10.1.1.
- PowerCenter 10.1.
- PowerCenter 9.6.1 com o hotfix mais recente instalado.

Se você usar a versão atual do Data Quality e do PowerCenter, poderá importar objetos que têm origem no repositório do Modelo sem nenhuma instalações adicionais. Se você usar uma versão anterior do PowerCenter, instale e registre a versão atual do plug-in de Integração de Informatica Data Quality.

Recursos e Funcionalidade da Qualidade de Dados

O Informatica Data Quality usa aplicativos Informatica para criar objetos, salvar objetos no repositório do Modelo e executar mapeamentos.

Talvez seja necessário instalar o Plug-in do Informatica Data Quality no PowerCenter se você não usou as versões de produto atuais.

Use o Informatica Data Quality para projetar e executar mapeamentos que obtêm os seguintes objetivos:

- Crie dados de perfil de dados. A criação de perfil de dados revela o conteúdo e a estrutura de seus dados. A criação de perfil de dados é uma etapa essencial em qualquer projeto de dados, pois é capaz de identificar pontos fortes e fracos nos dados e ajudar você a definir seu plano de projeto.
- Crie scorecards para analisar a qualidade dos dados. Um scorecard é uma representação gráfica das medidas de qualidade em um perfil.
- Padronize os valores de dados. Padronize os dados para remover erros e inconsistências encontrados na execução de um perfil. Você pode padronizar variações em pontuação, formatação e ortografia. Por exemplo, você pode garantir que os valores de cidade, estado e CEP sejam consistentes.
- Analise os registros. Analise registros de dados para aprimorar a estrutura de registro e gerar mais informações a partir de seus dados. É possível dividir um único campo de dados de formulário livre em campos que contêm tipos de informações diferentes. Você também pode adicionar informações aos seus registros. Por exemplo, você pode sinalizar registros do cliente como clientes pessoais ou comerciais.
- Valide endereços postais. A validação de endereço avalia e aprimora a precisão e a capacidade de entrega dos seus dados de endereço postal. A validação de endereço corrige erros em endereços e completa endereços parciais comparando registros de endereço com dados de referência de transportadoras postais nacionais. A validação de endereço pode ainda adicionar informações postais que aceleram a entrega de correspondências e reduz os custos de postagem.
- Localize registros duplicados. A análise de registros duplicados compara um conjunto de registros com outro para encontrar valores semelhantes ou correspondentes nas colunas de dados selecionadas. Você define o nível de similaridade que indica uma boa correspondência entre os valores do campo. Além disso, você também pode definir o peso relativo fixo para cada coluna em cálculos de correspondência. Por exemplo, você pode priorizar informações de sobrenome em relação às de nome.
- Crie e execute regras de qualidade de dados. A Informatica fornece regras pré-criadas que você pode executar ou editar para atender os objetivos do seu projeto. Você pode criar regras na ferramenta Desenvolvedor.
- Colabore com os usuários da Informatica. As tabelas de dados de referência e regras que você adiciona ao repositório do Modelo estão disponíveis para usuários na ferramenta Desenvolvedor e na ferramenta Analyst. Os usuários podem colaborar em projetos, e usuários diferentes podem assumir a propriedade de objetos em estágios diferentes de um projeto.
- Exporte mapeamentos para o PowerCenter. Você pode exportar mapeamentos para o PowerCenter reutilizar os metadados para integração física de dados ou para criar serviços Web.

Instaladores de Integração

Se você usar versões diferentes da Qualidade de Dados e do PowerCenter para criar e armazenar objetos, instale o Plug-in da Qualidade de Dados da Informatica para o PowerCenter. Baixe os instaladores de plug-in da Informatica.

Instale a versão de plug-in para a versão do PowerCenter que você deseja usar. O plug-in tem instaladores de Cliente e Servidor.

Execute o instalador do Cliente na máquina cliente do PowerCenter para exibir mapplets e mapeamentos no repositório e adicioná-los às sessões. Execute o instalador do Servidor na máquina com o Serviço de Integração do PowerCenter para executar as sessões em fluxos de trabalho.

Após instalar o plug-in na máquina do Serviço de Integração, registre o plug-in com o repositório do PowerCenter.

Exportação de Objetos para o PowerCenter

Use a ferramenta Desenvolvedor para exportar objetos, como mapplets e mapeamentos, do repositório do Modelo do Informatica. É possível exportar objetos do repositório do Modelo para o repositório do PowerCenter, bem como exportar objetos do repositório do Modelo para o sistema de arquivos. Use o PowerCenter Repository Manager para importar objetos XML para o repositório do PowerCenter.

- Para obter informações sobre como exportar objetos do repositório do Modelo, consulte o *Guia de Mapeamento do Informatica Developer* e o *Guia da Informatica Developer Tool*.
- Para obter informações sobre como importar objetos para o repositório do PowerCenter, consulte o *Guia do Repositório do PowerCenter*.

Instalação de Dados de Referência

Se um mapplet ou mapeamento implantado lê arquivos de dados de referência, instale esses arquivos na máquina com o Serviço de Integração do PowerCenter.

Você pode usar a ferramenta Desenvolvedor para exportar os dados de referência do repositório do Modelo quando exporta um mapplet ou mapeamento. O Desenvolvedor exporta os dados de referência como um arquivo compactado. Você extrai o arquivo e copia os dados de referência para a máquina do Serviço de Integração.

Você também pode usar o Instalador de Conteúdo da Qualidade de Dados para adicionar dados de referência à máquina do Serviço de Integração. Execute o Instalador de Conteúdo para instalar os arquivos de dados de referência de endereço e os arquivos de dados de preenchimento de identidade.

Matriz de Integração

O nível de integração entre o Data Quality e o PowerCenter depende das respectivas versões dos produtos. Você pode exportar objetos e dados de referência de várias versões do Data Quality para várias versões do PowerCenter. Não é possível exportar objetos ou dados de referência do PowerCenter para o Data Quality.

A seguinte tabela mostra as versões de produto que podem ser integradas:

Nome do Produto	PowerCenter 10.1.1	PowerCenter 10.1	PowerCenter 10.0	PowerCenter 9.6.1
Data Quality 10.1.1	Sim	Instalar plug-ins	Não	Instalar plug-ins
Data Quality 10.1	Não	Sim	Não	Instalar plug-ins
Data Quality 10.0	Não	Não	Sim	Instalar plug-ins
Data Quality 9.6.1	Não	Não	Não	Sim

Os valores da tabela representam os seguintes tipos de integração:

Sim

É possível exportar objetos e arquivos de um ambiente do Data Quality para um ambiente do PowerCenter. Não é necessário instalar ou registrar um plug-in. É possível executar objetos de qualidade de dados em um repositório do PowerCenter que você atualizar para a versão especificada.

Instalar plug-ins

É possível exportar objetos e arquivos de um ambiente do Data Quality para um ambiente do PowerCenter. Instale e registre os plug-ins de cliente e servidor para a versão do PowerCenter que você utiliza. É possível executar objetos de qualidade de dados em um repositório do PowerCenter que você atualizar para a versão especificada.

Não

Não é possível usar os objetos e arquivos definidos em um ambiente do Data Quality para um ambiente do PowerCenter.

Regras e Diretrizes para a Integração de Produtos

Você pode exportar objetos de qualidade de dados de diferentes versões do Informatica Data Quality para diferentes versões do PowerCenter. Instale o plug-in de Integração na versão do Informatica Data Quality da qual você exporta os objetos.

Considere as seguintes regras e diretrizes ao integrar o Informatica Data Quality e o PowerCenter:

- Não instale o plug-in de Integração se a versão de produto do repositório do PowerCenter corresponder à versão de produto do repositório do Modelo.
- Não é possível executar dois plug-ins de Integração do Data Quality simultaneamente em uma única máquina do PowerCenter. O PowerCenter usa o plug-in que você instalou mais recentemente. Por exemplo, se você instalar o plug-in de Integração do Data Quality 9.6.1 em uma máquina do PowerCenter 9.5.1, não poderá mais importar e executar objetos de qualidade de dados de um repositório do Modelo da versão 9.5.1. Você pode continuar a executar qualquer objeto de qualidade de dados importado para o repositório do PowerCenter antes da instalação do plug-in 9.6.1.

- Não é possível editar mapplets ou mapeamentos de qualidade de dados em um repositório do PowerCenter.

Comportamento das Transformações de Qualidade de Dados no PowerCenter

Durante a importação de um mapeamento ou mapplet de qualidade de dados do repositório do Modelo para o repositório do PowerCenter, o comportamento das transformações de qualidade de dados pode mudar.

Considerações sobre a Transformação de Consolidação

O Informatica Data Quality atualizou a estratégia padrão na transformação de Consolidação da versão 9.5.1. Na versão 9.5.1 e posteriores, a transformação de Consolidação seleciona o registro com o ID de linha mais alto como o registro sobrevivente. Na versão 9.5 e anteriores, a transformação de Consolidação usava o valor não em branco mais frequente em uma coluna que você especifica para identificar o registro sobrevivente.

Se você fizer upgrade de um repositório do PowerCenter que contém uma transformação de Consolidação para a versão 9.5.1 ou posterior, o upgrade manterá o valor não em branco mais frequente como a configuração padrão. Se você importar um mapeamento contendo uma transformação de Consolidação com configurações padrão de um repositório do Modelo da versão 9.5.1 ou posterior, essa transformação usará o ID de linha mais alto para identificar o registro sobrevivente.

Conversão em Mapplets

As seguintes transformações podem ser convertidas em mapplets na importação para o repositório do PowerCenter:

- Consolidação
- Exceção
- Gerador de Chave
- Correspondência

Durante a importação de um mapeamento que contenha uma ou mais das transformações, o PowerCenter converte cada uma delas em um mapplet que contém a configuração expandida.

Quando você importa um mapplet que contém uma ou mais das transformações, o PowerCenter expande a configuração de cada transformação em várias transformações.

O PowerCenter pode adicionar transformações à transformação de qualidade de dados ou à expansão de mapplet para concluir a lógica de transformação. Por exemplo, o PowerCenter expande a transformação Correspondência com outras transformações.

Considerações da transformação de Decisão

Você não poderá exportar um mapeamento ou mapplet que contém uma transformação de Decisão para o PowerCenter se o script da transformação incluir um parâmetro do sistema. A operação de exportação não pode converter o parâmetro do sistema em um valor que pode ser usado pelo PowerCenter. Antes de exportar um mapeamento ou mapplet com uma transformação de Decisão que utiliza um parâmetro do sistema, substitua o parâmetro por um valor apropriado.

Considerações sobre a Transformação de Exceção

O PowerCenter não pode usar uma transformação de Exceção que se origina em um repositório do PowerCenter anterior à versão 9.5. Se você atualizar um repositório do PowerCenter para a versão 9.5 ou posterior e o repositório tiver um mapeamento com uma transformação de Exceção, exporte o mapeamento novamente. Crie o mapeamento novamente na Developer tool, se necessário.

Ao executar um mapeamento que inclui uma transformação de Exceção, você poderá atualizar as tabelas de saída de registros duplicados ou registros inválidos na ferramenta Analyst. Se você planeja atualizar uma tabela de registros inválidos na ferramenta Analyst, será necessário habilitar a alta precisão na sessão que executa o mapeamento.

Considerações sobre a Transformação de Correspondência

Você não poderá executar um mapeamento de correspondência de identidade no PowerCenter se esse mapeamento fizer a leitura ou a gravação de dados de índice de identidade em tabelas de banco de dados. Para executar a análise de correspondência de identidade no PowerCenter, configure a transformação de Correspondência na Developer tool de forma a gravar dados de índice de identidade em arquivos temporários.

Transformações de Várias Estratégias

Você pode definir várias estratégias de transformação de dados em cada uma das transformações a seguir na Developer tool:

- Caso
- Classificador
- Decisão
- Gerador de Chave
- Rotulador
- Correspondência
- Mesclar
- Analisador
- Padronizador

Durante a importação de um mapplet ou de um mapeamento que contenha uma transformação de várias instâncias, cada estratégia é convertida em uma transformação separada.

Configurações de partição

Antes de executar uma tarefa de sessão que contém um mapeamento de qualidade de dados em uma grade, identifique a versão do PowerCenter na qual você criou os nós da grade. Se os nós tiverem sido criados no PowerCenter 9.5.1 HotFix 2 ou em uma versão anterior do PowerCenter, o mapeamento poderá falhar.

Para verificar se mapeamentos de qualidade de dados podem ser executados em todos os nós da grade, realize uma das seguintes tarefas:

- Analise as propriedades avançadas em qualquer Serviço de Integração do PowerCenter que distribua tarefas para nós na grade. Configure cada Serviço de Integração do PowerCenter para ignorar requisitos de recursos quando ele distribui tarefas. Por padrão, as propriedades avançadas especificam que o Serviço de Integração do PowerCenter ignora requisitos de recursos.
- Recrie qualquer nó que você criou no PowerCenter 9.5.1 HotFix 2 ou versão anterior. Quando você recriar o nó, o PowerCenter aplicará os arquivos de configuração do nó atual a esse nó.

Quando uma tarefa de sessão é executada em uma grade, o processo do Data Transformation Manager (DTM) distribui os segmentos de sessão entre os nós da grade. Algumas transformações não são

particionáveis em uma grade. Quando uma transformação não é particionável em uma grade, o DTM cria um único grupo de partições para os segmentos da transformação e executa esses segmentos em um único nó.

A seguinte tabela descreve as configurações de partição para transformações de qualidade de dados:

Transformação	Configuração de Partição
Validação de Endereço	Através da grade
Associação	Não pode ser particionado
Relatório de AV	Particionável localmente
Conversor de Maiúsculas/Minúsculas	Através da grade
Classificador	Particionável localmente
Cluster	Não pode ser particionado
Comparação	Através da grade
Consolidação	Não pode ser particionado
Decisão	Através da grade
Exceção	Não pode ser particionado
Ponto de Execução	Através da grade
Gerador de Chave	Não pode ser particionado
Armazenamento de Chaves	Não pode ser particionado
Rotulador	Através da grade
Mesclar	Através da grade
Gerador de Par	Não pode ser particionado
Analisador (análise de probabilidade)	Particionável localmente
Analisador	Através da grade
Padronizador de Token	Através da grade
Analisador Baseado no Peso	Através da grade

CAPÍTULO 2

Instalação

Este capítulo inclui os seguintes tópicos:

- [Visão Geral da Instalação, 14](#)
- [Requisitos de Sistema, 14](#)
- [Antes da Instalação, 15](#)
- [Instalando o Plug-in de Servidor, 15](#)
- [Instalando o Plug-in Cliente, 16](#)
- [Registrando o Plug-In no PowerCenter, 17](#)

Visão Geral da Instalação

O plug-in Data Quality Integration tem um componente de servidor e um componente cliente.

Execute as seguintes etapas para instalar e registrar o plug-in:

1. Execute o instalador do Servidor na máquina com o Serviço de Integração do PowerCenter.
2. Execute o instalador do Cliente em uma máquina do cliente do PowerCenter.
3. Registre o arquivo `IDQZIntegration.xml` com o repositório do PowerCenter.

Requisitos de Sistema

O plug-in de Integração do Informatica Data Quality tem os mesmos requisitos de sistema dos aplicativos e serviços do PowerCenter nos quais é instalado.

Para obter informações sobre os pré-requisitos do PowerCenter e os requisitos de sistema, consulte a documentação de instalação do PowerCenter.

Antes da Instalação

Considere o comportamento do instalador a seguir se você usar dados de referência no PowerCenter.

O instalador do plug-in de Integração grava arquivos de propriedades na máquina do Serviço de Integração do PowerCenter. Os arquivos de propriedades especificam as localizações dos arquivos de dados de referência que os mapeamentos de qualidade de dados podem ler.

O instalador grava os seguintes arquivos:

- AD50.cfg. Armazena as propriedades de configuração para dados de referência de endereço.
- CLASSIFIER.properties. Armazena as propriedades de configuração para dados do modelo de classificador.
- IDQTx.cfg. Armazena as propriedades de configuração para dados de preenchimento de identidade.
- NER.properties. Armazena as propriedades de configuração para dados do modelo de probabilidade.

O instalador adiciona os arquivos ao seguinte diretório:

```
[PowerCenter_installation]/server/bin
```

Se o instalador encontrar um arquivo de propriedades no diretório `bin`, ele criará uma cópia de backup desse arquivo e instalará um novo arquivo. Para manter a configuração atual dos dados de referência após a instalação, mescle o conteúdo dos arquivos novos e antigos.

Se você instalar o PowerCenter 9.5.1 ou uma versão mais recente, o processo de instalação do servidor adicionará os arquivos de propriedades ao diretório `bin` por padrão.

Pré-requisito de Dados de Referência de Endereço

Antes de instalar dados de referência de endereço, interrompa o Serviço de Integração do PowerCenter. Reinicie o serviço depois de instalar os dados.

Instalando o Plug-in de Servidor

O instalador do plug-in de Servidor de Integração do Informatica Data Quality adiciona arquivos de aplicativo de transformação de qualidade de dados à estrutura de diretório do Serviço de Integração do PowerCenter.

Instalando no Windows

Execute o instalador do Windows Server para adicionar o plug-in a uma máquina com o Serviço de Integração do PowerCenter em uma plataforma Windows.

Feche todos os aplicativos antes de instalar.

1. Faça logon na máquina com uma conta de sistema do usuário.
2. Extraia o arquivo do Servidor de Integração para Windows.
3. Vá para o diretório em que você extraiu os arquivos e execute `install.exe`.
4. Selecione um idioma e clique em **OK**.
5. Na tela **Bem-vindo**, clique em **Avançar**.
6. Verifique se os serviços do PowerCenter não estão em execução e clique em **OK**.
7. Especifique o caminho para o diretório do PowerCenter. O instalador fornece um caminho padrão para esse diretório. Para selecionar outro caminho, clique em **Fechar**.

8. Clique em **Avançar**.
9. Examine o resumo de pré-instalação. Clique em **Anterior** se desejar corrigir algum erro.
10. Clique em **Instalar**.
11. Examine o resumo de pós-instalação.
Para obter mais informações sobre as tarefas de instalação e para exibir as propriedades de configuração dos componentes instalados, consulte os arquivos de log de instalação.
12. Clique em **Concluído**.

Instalando no UNIX

Execute o instalador do UNIX Server para adicionar o plug-in a uma máquina com o Serviço de Integração do PowerCenter em uma plataforma UNIX.

1. Faça login na máquina com uma conta de sistema do usuário.
2. Feche todos os outros aplicativos.
3. Extraia o arquivo do Integration Server for UNIX.
4. Abra um shell UNIX e navegue para o diretório em que você extraiu os arquivos de instalação.
5. Digite `sh install.bin`.
6. Para aceitar o inglês como o idioma de instalação, pressione **Enter**.
7. Leia o texto de Boas-vindas e pressione **Enter**.
8. Digite um caminho absoluto para o diretório de instalação do PowerCenter ou pressione **Enter** para aceitar o caminho padrão.
9. Examine o resumo de pré-instalação. Para corrigir um erro, digite `back` e pressione **Enter**.
10. Pressione **Enter** para iniciar a instalação.
11. Examine o resumo de pós-instalação.
Para obter mais informações sobre as tarefas de instalação e para exibir as propriedades de configuração dos componentes instalados, consulte os arquivos de log de instalação.
12. Pressione **Enter** para concluir o processo de instalação.

Instalando o Plug-in Cliente

Execute o instalador do Data Quality Integration Client nas máquinas Cliente do PowerCenter.

1. Feche todos os outros aplicativos.
2. Extraia o arquivo de Cliente de Integração.
3. Abra o diretório no qual você extraiu os arquivos.
4. Execute `install.exe`.
5. Na tela **Bem-vindo**, clique em **Avançar**.
O instalador exibe um prompt para fechar os aplicativos cliente do PowerCenter antes de continuar a instalação.
6. Verifique se não há nenhum aplicativo cliente do PowerCenter em execução e clique em **OK**.

7. Especifique o caminho para o diretório do Cliente do PowerCenter. O instalador fornece um caminho padrão para esse diretório. Para selecionar outro caminho, clique em **Fechar**.
8. Clique em **Avançar**.
9. Examine o resumo de pré-instalação. Clique em **Anterior** se desejar corrigir algum erro.
10. Clique em **Instalar**.
11. Examine o resumo de pós-instalação.
Para obter mais informações sobre as tarefas de instalação e para exibir as propriedades de configuração dos componentes instalados, consulte os arquivos de log de instalação.
12. Clique em **Concluído**.

Registrando o Plug-In no PowerCenter

Registre o plug-in em qualquer repositório do PowerCenter que deverá conter objetos de qualidade de dados.

O instalador de Integração grava `IDQZIntegration.xml` no diretório `$INFA_HOME\server\bin\native` na máquina com o Serviço de Integração do PowerCenter. Verifique se você pode acessar esse arquivo. Se necessário, copie ou envie esse arquivo por FTP para a máquina com o Serviço de Integração do PowerCenter.

Se você não tiver os privilégios corretos para registrar o plug-in, contate o administrador do Serviço de Repositório do PowerCenter.

1. Abra o Console de Administração do PowerCenter.
2. No Navegador, selecione o repositório que você deseja usar para objetos de qualidade de dados.
3. Execute o Serviço de Repositório no modo exclusivo.
4. Clique na guia **Plug-ins**.
5. Clique no link **Registrar o plug-in de Serviço de Repositório**.
6. Clique em **Procurar** para localizar o arquivo `IDQZIntegration.xml`.
7. Para substituir o registro de um plug-in de Integração mais antigo, selecione a opção para atualizar o registro de plug-in.
8. Insira seu nome de usuário e senha para o repositório.
9. Clique em **OK**.
O Serviço de Repositório registra o plug-in no repositório. Os resultados da operação de registro são exibidos no log de atividades.
10. Execute o Serviço de Repositório no modo normal.

CAPÍTULO 3

Requisitos de Dados de Referência

Este capítulo inclui os seguintes tópicos:

- [Visão Geral dos Dados de Referência, 18](#)
- [Requisitos do Arquivo de Dados de Referência, 19](#)
- [Requisitos de Dados de Referência de Endereço, 20](#)
- [Requisitos de Dados de Preenchimento de Identidade, 21](#)
- [Requisitos do Modelo de Probabilidade e do Modelo de Classificador, 23](#)

Visão Geral dos Dados de Referência

As transformações de qualidade de dados pode ler dados de referência. Se qualquer transformação no repositório do PowerCenter ler dados de referência, verifique se os arquivos de dados de referência estão presentes no ambiente do PowerCenter.

As transformações de qualidade de dados no PowerCenter podem usar os seguintes tipos de dados de referência:

Arquivos de dados de referência

Os arquivos de texto que contêm as versões padrão de termos comerciais e versões alternativas dos termos. É possível editar os arquivos de dados de referência.

Arquivos de dados de referência de endereço

Arquivos de dados que contêm dados completos para todos os endereços postais em um país. Instale os arquivos de dados de referência de endereço na máquina do Serviço de Integração do PowerCenter. Não é possível editar arquivos de dados de referência de endereço.

Arquivos de preenchimento de identidade

Arquivos de dados que definem tipos diferentes de identidade e contêm algoritmos usados pela transformação de Correspondência. Instale os arquivos de preenchimento na máquina de Serviços do PowerCenter. Não é possível editar arquivos de preenchimento.

Conjuntos de conteúdo

Os arquivos de dados que contêm dados de referência especificados por uma ou mais transformações. Quando você exporta um mapeamento que lê conjuntos de conteúdo, o processo de exportação exporta os dados do conjunto de conteúdo com os metadados de transformação. No caso dos arquivos do

modelo de classificador e de probabilidade, o processo exporta os dados da estrutura de diretório da Qualidade de Dados.

Requisitos do Arquivo de Dados de Referência

Um mapeamento de qualidade de dados pode ler objetos de dados de referência definidos no repositório do Modelo.

Quando um usuário do Informatica Data Quality exporta um mapeamento que lê um objeto de dados de referência, o processo de exportação grava os dados de referência no sistema de arquivos como um ou mais arquivos de dados de referência. Ao usar o mapeamento no PowerCenter, você deve verificar se os arquivos estão instalados em um local que o Serviço de Integração pode ler.

Especifique o local do arquivo para os arquivos de dados de referência quando exporta o mapeamento. Exporte os arquivos para a estrutura de diretório do PowerCenter, ou exporte os arquivos para o sistema de arquivos e copie os arquivos para a máquina do Serviço de Integração do PowerCenter.

Os locais do arquivo de dados de referência na estrutura de diretório do PowerCenter devem corresponder aos locais da tabela de referência na estrutura do repositório do Modelo. A estrutura de diretório deve seguir este formato:

```
<Diretório_Raiz_do_PowerCenter>\services\<Nome _do_Projeto_do_Repositório_do_Modelo>\<Nome_da_Pasta_do_Projeto_do_Repositório_do_Modelo>
```

Nota: O PowerCenter lê arquivos de dados de referência no formato UTF-8. A ferramenta Developer exporta dados de referência para o PowerCenter no formato UTF-8.

Instalando Dados de Referência do Instalador de Conteúdo da Qualidade de Dados

O conjunto de arquivos do Instalador de Conteúdo da Qualidade de Dados inclui arquivos de dados de referência para a estrutura de diretório do PowerCenter.

Identifique o diretório de instalação do PowerCenter quando você executar o Instalador de Conteúdo. Por padrão, o Instalador de Conteúdo anexa o seguinte caminho à raiz:

```
\services\DQContent\Conteúdo_do_INFA\
```

Edite o caminho para corresponder à estrutura de repositório do Modelo ou copie os arquivos de dados de referência para o caminho necessário depois de executar o Instalador de Conteúdo.

Usando uma variável de ambiente para definir o local do arquivo de dados de referência

Você pode definir uma variável de ambiente *INFA_CONTENT* na máquina do Serviço de Integração do PowerCenter para definir o caminho que o Serviço de Integração usa para ler arquivos de dados de referência.

Use essa variável de ambiente se você não puder instalar os arquivos de dados de referência no local exigido na máquina do Serviço de Integração do PowerCenter. Instale os arquivos em um local que o Serviço de Integração do PowerCenter possa ler e crie a variável de ambiente *INFA_CONTENT* com o caminho instalado.

Requisitos de Dados de Referência de Endereço

Um mapeamento que executa a validação de endereços lê dados de referência de endereço. Você adquiriu e fez download de dados de referência de endereço da Informatica. Copie os arquivos de dados para a máquina do Serviço de Integração do PowerCenter.

A tabela a seguir descreve os tipos de dados de referência de endereço que você pode instalar:

Tipo de Dados de Referência	Quando Usar
Lote	Use para operações gerais de validação de endereço.
CAMEO	Use para adicionar dados demográficos de consumidores aos registros de endereço.
Certificado	Use para verificar se os registros de endereço atendem aos padrões de certificação definidos por um serviço de correio. Os seguintes países definem padrões de certificação: <ul style="list-style-type: none">- Austrália. Certifica a correspondência de acordo com o padrão AMAS (Sistema de Aprovação de Correspondência de Endereço).- Canadá. Certifica a correspondência de acordo com o padrão SERP (Programa de Análise e Reconhecimento de Software) padrão.- França. Certifica a correspondência de acordo com o padrão SNA (Serviço Nacional de Gerenciamento de Endereços).- Nova Zelândia. Certifica a correspondência de acordo com o padrão SendRight.- Estados Unidos. Certifica a correspondência de acordo com o padrão CASS (Sistema de Suporte para Precisão de Codificação).
Lista de sugestões	Use para encontrar versões válidas de um registro de endereço incompleto. Use dados da lista de sugestões ao configurar um mapeamento de validação de endereço para processar registros de endereço simples em tempo real.
Codificação Geográfica	Use para adicionar coordenadas de latitude e de longitude aos registros de endereço.
Complementares	Use para identificar a área geográfica ou a área populacional que contém um endereço. A transformação do Validador de Endereço pode adicionar dados suplementares a registros de endereço dos seguintes países: <ul style="list-style-type: none">- Japão. Identifica o distrito urbano que contém o endereço.- Sérvia. Identifica o código postal no nível da rua para o endereço.- Reino Unido. Adiciona um identificador do ponto de entrega ao endereço.- Estados Unidos. Identifica o centro populacional que contém o endereço.

Você também pode usar o Instalador de Conteúdo da Qualidade de Dados para instalar os dados de referência de endereço. Execute o Instalador de Conteúdo do Servidor na máquina do Serviço de Integração do PowerCenter para instalar os arquivos.

O Instalador de Conteúdo do Data Quality cria um diretório /av/ na estrutura de diretório do PowerCenter e grava dados de referência de endereço nessa localização. O Serviço de Integração do PowerCenter lê o caminho para os dados de referência de endereço no arquivo de configuração AD50.cfg.

O instalador do PowerCenter grava o arquivo AD50.cfg no seguinte caminho:

```
[Instalação_do_PowerCenter]/server/bin
```

Quando você instala dados de referência de endereço, verifique se as seguintes propriedades estão definidas no AD50.cfg:

Localização dos Dados de Referência

Digite o caminho para os dados de referência de endereço na propriedade `ReferenceDataLocation`.

Chave de licença de dados

Insira a licença dados na propriedade `LicenseKey`. Se você tiver mais de uma licença, digite cada licença em uma cadeia de caracteres separados por vírgula.

Valores de pré-carregamento

Insira pelo menos uma abreviação de país como valor de pré-carregamento para cada tipo de dados de referência que um mapeamento lê. Digite ALL para aplicar uma configuração de pré-carregamento para todos os países.

O Serviço de Integração pode usar um método diferente para carregar dados para cada país. Por exemplo, você pode especificar o pré-carregamento total de dados em lotes/interativo dos Estados Unidos e o pré-carregamento parcial de dados em lotes/interativo do Reino Unido. O Serviço de Integração também pode usar um método de pré-carregamento diferente para cada tipo de dados. Por exemplo, você pode especificar o pré-carregamento total de dados em lotes/interativo dos Estados Unidos e o pré-carregamento parcial dos dados de geocodificação dos Estados Unidos.

As configurações de pré-carregamento total substituem as configurações de pré-carregamento parcial, e as configurações de pré-carregamento parcial substituem as configurações que indicam nenhum pré-carregamento de dados. Por exemplo, se você não inserir ALL para nenhum pré-carregamento de dados e inserir USA para o pré-carregamento total, o Serviço de Integração carregará todos os dados dos Estados Unidos na memória e não carregará os dados de nenhum outro país. Se você não tiver um requisito de pré-carregamento, digite ALL para o não pré-carregamento de dados de nenhum tipo de dado de referência que você planeje usar.

Você não especifica um valor de pré-carregamento para dados suplementares.

Se desejar, edite outras propriedades no arquivo `AD50.cfg`.

Nota: O Informatica Data Quality 9.x não lê o arquivo `AD50.cfg`. Usuários do Informatica Data Quality 9.x inserem a localização dos dados de referência, as informações da chave de licença e os valores de pré-carregamento por meio da ferramenta Administrador.

Requisitos de Dados de Preenchimento de Identidade

Um mapeamento que executa análise de correspondência de identidade lê arquivos de dados de referência denominados preenchimentos de identidade. Adquirir e faça download dos arquivos de preenchimento de identidade da Informatica. Copie os arquivos de preenchimento na máquina do Serviço de Integração do PowerCenter ou em um diretório compartilhado em uma máquina que o Serviço de Integração do PowerCenter possa acessar.

Você também pode usar o Instalador de Conteúdo do Data Quality para instalar os arquivos de preenchimento. Execute o Instalador de Conteúdo do Servidor na máquina do Serviço de Integração do PowerCenter para instalar os arquivos.

Você deve copiar ou instalar os arquivos de preenchimento em um diretório com o nome `padrão`.

O Serviço de Integração do PowerCenter pode ler a localização do diretório do arquivo de preenchimento por meio do arquivo de configuração `IDQTx.cfg` ou da variável de ambiente `SSAPR`. O Serviço de Integração do PowerCenter procura a localização no arquivo `IDQTx.cfg`. Se o arquivo `IDQTx.cfg` não estiver presente ou não

especificar uma localização, o Serviço de Integração do PowerCenter procurará a localização na variável de ambiente SSAPR.

Nota: O arquivo IDQTx.cfg e a variável de ambiente SSAPR especificam o caminho para o diretório pai do diretório padrão. O caminho não inclui o nome de diretório.

O instalador do PowerCenter grava o arquivo IDQTx.cfg no seguinte caminho:

```
[Instalação_do_PowerCenter]/server/bin
```

O arquivo IDQTx.cfg usa a propriedade *IdentityReferenceDataLocation* para especificar a localização do diretório de arquivos de preenchimento. Atualize a propriedade com a localização atual do diretório.

Por exemplo, você pode definir o seguinte caminho na propriedade *IdentityReferenceDataLocation*:

```
[PowerCenter_Installation]/services/DQContent/INFA_Content/identity/
```

Você pode definir um caminho absoluto ou relativo para o diretório padrão na propriedade *IdentityReferenceDataLocation*. Defina um caminho absoluto para um diretório na máquina do Serviço de Integração do PowerCenter ou para um diretório compartilhado em uma máquina que o serviço possa acessar. Defina um caminho relativo para um diretório na máquina do Serviço de Integração do PowerCenter.

O Serviço de Integração do PowerCenter lê um caminho relativo no seguinte diretório:

```
[Instalação_do_PowerCenter]/server/bin
```

Por exemplo, para especificar um caminho relativo a um diretório de arquivo de preenchimento no diretório bin, defina o seguinte caminho relativo:

```
./
```

Propriedades do IDQTx.cfg

Quando você define a propriedade *IdentityReferenceDataLocation* no arquivo IDQTx.cfg, também pode definir as seguintes propriedades:

IdentityCacheDir

Caminho para o diretório que armazena os arquivos de dados temporários criados quando o mapeamento é executado. O caminho identifica um diretório pai. O Serviço de Integração do PowerCenter grava os arquivos temporários nos diretórios abaixo da localização que você especificar.

A propriedade tem o seguinte valor padrão:

```
./identityCache
```

IdentityIndexDir

Caminho para o diretório que contém os arquivos de índice temporários criados quando o mapeamento é executado. A análise de correspondência de identidade usa o índice para classificar os registros em grupos antes da análise de correspondência. O caminho identifica um diretório pai. O Serviço de Integração do PowerCenter grava os arquivos de índice nos diretórios abaixo da localização que você especificar.

A propriedade tem o seguinte valor padrão:

```
./identityIndex
```

Nota: Um usuário da ferramenta Developer pode definir o caminho do diretório de cache e o caminho do diretório de índice na transformação de Correspondência. O Serviço de Integração do PowerCenter usará o caminho no arquivo IDQTx.cfg se a transformação de Correspondência não especificar o caminho.

Requisitos do Modelo de Probabilidade e do Modelo de Classificador

Modelos probabilísticos e modelos de classificador são arquivos de dados de referência que identificam os tipos de informações em uma string de dados. Antes de executar um mapeamento que lê um modelo probabilístico ou um arquivo de modelo de classificador, você configura o PowerCenter para ler os arquivos.

Desempenhe as seguintes tarefas de configuração:

- Verifique as localizações do arquivo de modelo na estrutura de diretórios do PowerCenter.
- Atualize a alocação de memória do Java no Serviço de Integração do PowerCenter.

Localização de Arquivos do Modelo Probabilístico e do Modelo de Classificador

Ao exportar um mapeamento que grava dados de modelo probabilístico ou de modelo de classificados, você especifica a localização de um diretório para o arquivo de dados do modelo exportado. Exporte o arquivo para a estrutura de diretórios do PowerCenter ou exporte o arquivo para o sistema de arquivos e o copie para a máquina de serviços do PowerCenter.

Por padrão, o Serviço de Integração do PowerCenter lê os arquivos de modelo no seguinte diretório:

```
[PowerCenter_Installation]/server/bin
```

Se os arquivos não estiverem presentes no local padrão, o Serviço de Integração do PowerCenter lerá a localização dos arquivos nos arquivos de propriedade. Você deve atualizar os arquivos de propriedade com os caminhos para os arquivos de modelo.

A seguinte tabela descreve os arquivos de propriedades:

Nome do Arquivo	Descrição
CLASSIFIER.properties	Contém a propriedade <i>ClassifierFileLocation</i> . Atualize a propriedade para identificar o caminho para os arquivos de modelo de classificador.
NER.properties	Contém a propriedade <i>NERFileLocation</i> . Atualize a propriedade para identificar o caminho para os arquivos de modelo probabilístico.

O Serviço de Integração do PowerCenter lê os arquivos de propriedades do seguinte diretório:

```
[PowerCenter_Installation]/server/bin
```

Você pode definir um caminho absoluto ou um caminho relativo para identificar a localização de arquivos de modelo em cada arquivo de propriedades. Se você definir um caminho relativo, o Serviço de Integração do PowerCenter adicionará o caminho para o seguinte diretório:

```
[PowerCenter_Installation]/server/bin
```

Nota: Use barras invertidas para definir os caminhos de diretório para os arquivos de modelo em todas as plataformas.

Alocação de Memória do Serviço de Integração do PowerCenter

Um Serviço de Integração do PowerCenter usa grandes quantidades de memória para ler um arquivo de modelo de classificador ou um arquivo de modelo probabilístico. Se o Serviço de Integração do PowerCenter

não puder acessar a memória necessária, ele falhará ao ler o arquivo de modelo e não poderá processar os dados.

Consulte a alocação de memória no Serviço de Integração do PowerCenter. Se você executar sessões que fazem referência a um arquivo de modelo de classificador ou a um arquivo de modelo probabilístico, defina a alocação de memória como pelo menos 512 MB.

Você pode usar a propriedade de processo *Memória Máxima para Java SDK* para alocar a memória. Para analisar ou atualizar a propriedade, faça logon na ferramenta Administrator e selecione o Serviço de Integração do PowerCenter que executa as tarefas de sessão. Você encontrará a propriedade *Memória Máxima do Java SDK* na exibição **Processos**.