



Informatica™

Informatica® Data Quality  
10.1.1

# Guia do Acelerador

Este software e a documentação são fornecidos somente sob um contrato de licença separado, contendo restrições sobre uso e divulgação. 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 meios) sem o consentimento prévio da Informatica LLC.

DIREITOS DO GOVERNO DOS ESTADOS UNIDOS Programas, softwares, bancos de dados, bem como a documentação e os dados técnicos relacionados, distribuídos a clientes do Governo dos EUA são "softwares de computador comerciais" ou "dados técnicos comerciais", de acordo com o Regulamento de Aquisição Federal aplicável e os regulamentos suplementares específicos da agência. Como tal, a utilização, duplicação, divulgação, modificação e adaptação estão sujeitas às restrições e aos termos de licença estabelecidos no contrato governamental aplicável e, na medida do que for aplicável pelos termos do contrato governamental, aos direitos adicionais estabelecidos no FAR 52.227-19, Licença de Software de Computador Comercial.

Informatica, o logotipo Informatica e PowerCenter são marcas comerciais ou marcas registradas da Informatica LLC nos Estados Unidos e em muitas jurisdições por todo o mundo. Uma lista atual das marcas comerciais da Informatica está disponível na Internet em <https://www.informatica.com/trademarks.html>. Os 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](mailto: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](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/hsqllicense.html>, <http://httpunit.sourceforge.net/doc/license.html>, <http://jung.sourceforge.net/license.txt>, [http://www.gzip.org/zlib/zlib\\_license.html](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://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/javaservice/>, <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://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/>, <http://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 QUAISQUER 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.

As informações contidas neste documento estão sujeitas a alteração sem aviso prévio. Se você encontrar problemas nesta documentação, informe-nos por escrito e envie para Informatica LLC 2100 Seaport Blvd. Redwood City, CA 94063.

Os produtos Informatica apresentam garantias segundo os termos e condições dos acordos em que são fornecidos. A INFORMATICA FORNECE AS INFORMAÇÕES NESTE DOCUMENTO "COMO ESTÃO" SEM GARANTIA DE QUALQUER TIPO, EXPRESSA OU IMPLÍCITA, INCLUINDO, SEM QUAISQUER GARANTIAS DE COMERCIALIZAÇÃO, ADEQUAÇÃO A UM DETERMINADO FIM E QUALQUER GARANTIA OU CONDIÇÃO DE NÃO-VIOLAÇÃO.

Data da Publicação: 2018-06-09

# Conteúdo

<b>Prefácio.....</b>	<b>8</b>
Recursos da Informatica. . . . .	8
Rede da Informatica. . . . .	8
Base de Dados de Conhecimento da Informatica. . . . .	8
Documentação da Informatica. . . . .	8
Matrizes de Disponibilidade de Produto Informatica. . . . .	9
Informatica Velocity. . . . .	9
Informatica Marketplace. . . . .	9
Suporte global a clientes Informatica. . . . .	9
 <b>Capítulo 1: Introdução aos Aceleradores.....</b>	 <b>10</b>
Visão Geral dos Aceleradores. . . . .	10
Estrutura do Acelerador. . . . .	10
Estrutura do Acelerador Geral. . . . .	11
Estrutura do Acelerador do Domínio de Dados. . . . .	11
Instalação do Acelerador. . . . .	12
Regras e Diretrizes da Instalação do Acelerador. . . . .	13
Importando Regras e Mapeamentos. . . . .	14
Importando Domínios de Dados e Grupos de Domínios de Dados. . . . .	14
Componentes do Acelerador. . . . .	15
Regras. . . . .	17
Mapeamentos Demonstrativos. . . . .	18
Domínios de Dados. . . . .	18
Tabelas de Referência. . . . .	18
Conjuntos de Conteúdo. . . . .	18
Marcas e Regras. . . . .	19
Uso do Acelerador no PowerCenter. . . . .	19
 <b>Capítulo 2: Acelerador Principal.....</b>	 <b>20</b>
Visão Geral do Acelerador Principal. . . . .	20
Regras de Limpeza de Dados de Endereço Principal. . . . .	21
Regras de Limpeza de Dados do Contato Principal. . . . .	22
Regras de Limpeza de Dados Corporativos Principais. . . . .	23
Regras de Limpeza de Dados Gerais Principais. . . . .	23
Regras Principais de Correspondência e de Eliminação de Duplicação. . . . .	29
Regras de Limpeza de Dados de Produto Principal. . . . .	30
Mapeamentos Demonstrativos Principais. . . . .	30
 <b>Capítulo 3: Acelerador dos Domínios de Dados.....</b>	 <b>31</b>
Visão Geral do acelerador dos Domínios de Dados. . . . .	31

Domínios de dados no acelerador dos Domínios de Dados. . . . .	32
Regras de nome de coluna no acelerador dos Domínios de Dados. . . . .	40
Regras de dados no acelerador dos Domínios de Dados. . . . .	44

## **Capítulo 4: Acelerador da Austrália/Nova Zelândia. . . . . 51**

Visão Geral do Acelerador da Austrália/Nova Zelândia. . . . .	51
Regras de Limpeza de Dados de Endereço da Austrália/Nova Zelândia. . . . .	52
Regras de Composição da Austrália/Nova Zelândia. . . . .	53
Regras de Limpeza de Dados de Contato da Austrália/Nova Zelândia. . . . .	55
Regras de Limpeza de Dados Corporativos da Austrália/Nova Zelândia. . . . .	58
Regras de Limpeza de Dados Gerais da Austrália/Nova Zelândia. . . . .	59
Regras de Eliminação de Duplicação e Correspondência da Austrália/Nova Zelândia. . . . .	60
Mapeamentos Demonstrativos da Austrália/Nova Zelândia. . . . .	62

## **Capítulo 5: Acelerador do Brasil. . . . . 64**

Visão Geral do Acelerador do Brasil. . . . .	64
Regras de Limpeza de Dados de Endereço do Brasil. . . . .	64
Regras de Composição do Brasil. . . . .	65
Regras de Limpeza de Dados de Contato do Brasil. . . . .	66
Regras de Limpeza de Dados Corporativos do Brasil. . . . .	68
Regras de Limpeza de Dados Gerais do Brasil. . . . .	68
Regras de Eliminação de Duplicação e Correspondência do Brasil. . . . .	69
Mapeamentos Demonstrativos do Brasil. . . . .	71

## **Capítulo 6: Acelerador dos Serviços Financeiros. . . . . 72**

Visão Geral do Acelerador dos Serviços Financeiros. . . . .	72
Regras de Limpeza de Dados de Contato dos Serviços Financeiros. . . . .	72
Regras de Limpeza de Dados Financeiros dos Serviços Financeiros. . . . .	73
Regras de Limpeza de Dados Gerais dos Serviços Financeiros. . . . .	76
Regras de Eliminação de Duplicação e Correspondência dos Serviços Financeiros. . . . .	76

## **Capítulo 7: Acelerador da França. . . . . 78**

Visão Geral do Acelerador da França. . . . .	78
Regras de Limpeza de Dados de Endereço da França. . . . .	78
Regras de Composição da França. . . . .	79
Regras de Limpeza de Dados de Contato da França. . . . .	80
Regras de Limpeza de Dados Corporativos da França. . . . .	82
Regras de Limpeza de Dados Gerais da França. . . . .	83
Regras de Eliminação de Duplicação e Correspondência da França. . . . .	83
Mapeamentos Demonstrativos da França. . . . .	85

## **Capítulo 8: Acelerador da Alemanha. . . . . 87**

Visão Geral do Acelerador da Alemanha. . . . .	87
--	----

Regras de Limpeza de Dados de Endereço da Alemanha. . . . .	87
Regras de Composição da Alemanha. . . . .	88
Regras de Limpeza de Dados de Contato da Alemanha. . . . .	89
Regras de Limpeza de Dados Corporativos da Alemanha. . . . .	91
Regras de Limpeza de Dados Gerais da Alemanha. . . . .	91
Regras de Eliminação de Duplicação e Correspondência da Alemanha. . . . .	92
Mapeamentos Demonstrativos da Alemanha. . . . .	95
<b>Capítulo 9: Acelerador de Portugal. . . . .</b>	<b>96</b>
Visão Geral do Acelerador de Portugal. . . . .	96
Regras de Limpeza de Dados de Endereço de Portugal. . . . .	96
Regras de Composição de Portugal. . . . .	97
Regras de Limpeza de Dados de Contato de Portugal. . . . .	98
Regras Limpeza de Dados Corporativos de Portugal. . . . .	100
Regras de Limpeza de Dados Gerais de Portugal. . . . .	100
Regras de Eliminação de Duplicação e Correspondência de Portugal. . . . .	101
Mapeamentos Demonstrativos de Portugal. . . . .	103
<b>Capítulo 10: Acelerador da Espanha. . . . .</b>	<b>104</b>
Visão Geral do Acelerador da Espanha. . . . .	104
Regras de Limpeza de Dados de Endereço da Espanha. . . . .	104
Regras de Limpeza de Dados de Contato da Espanha. . . . .	106
Regras de Limpeza de Dados Corporativos da Espanha. . . . .	107
Regras de Limpeza de Dados Gerais da Espanha . . . . .	108
Regras de Correspondência e de Eliminação de Duplicação da Espanha. . . . .	108
Mapeamentos Demonstrativos da Espanha. . . . .	111
<b>Capítulo 11: Acelerador do Reino Unido. . . . .</b>	<b>112</b>
Visão Geral do Acelerador do Reino Unido. . . . .	112
Regras de Limpeza de Dados de Endereço do Reino Unido. . . . .	112
Regras de Composição do Reino Unido. . . . .	114
Regras de Limpeza de Dados de Contato do Reino Unido. . . . .	115
Regras de limpeza de dados corporativos do Reino Unido. . . . .	118
Regras de Limpeza de Dados Financeiros do Reino Unido. . . . .	118
Regras de Limpeza de Dados Gerais do Reino Unido. . . . .	119
Regras de Eliminação de Duplicação e Correspondência do Reino Unido. . . . .	119
Mapeamentos Demonstrativos do Reino Unido. . . . .	122
<b>Capítulo 12: Acelerador dos EUA/Canadá. . . . .</b>	<b>123</b>
Visão Geral do Acelerador dos EUA/Canadá. . . . .	123
Regras de Limpeza de Dados de Endereço dos EUA/Canadá. . . . .	123
Regras de Composição dos EUA/Canadá. . . . .	126
Regras de Limpeza de Dados de Contato dos EUA/Canadá. . . . .	128

Dependências de Limpeza de Dados Corporativos dos EUA/Canadá. . . . .	133
Regras de Limpeza de Dados Gerais dos EUA/Canadá. . . . .	133
Regras de Eliminação de Duplicação e Correspondência dos EUA/Canadá. . . . .	134
Mapeamentos Demonstrativos dos EUA/Canadá. . . . .	137

# Prefácio

O *Guia do Acelerador* do Informatica Data Quality foi escrito para desenvolvedores de qualidade de dados. Este guia pressupõe que você compreende os conceitos de qualidade de dados, como padronização, análise, rotulagem e validação.

## 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](mailto: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](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](mailto: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](mailto: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

## Introdução aos Aceleradores

Este capítulo inclui os seguintes tópicos:

- [Visão Geral dos Aceleradores, 10](#)
- [Estrutura do Acelerador, 10](#)
- [Instalação do Acelerador, 12](#)
- [Componentes do Acelerador, 15](#)
- [Marcas e Regras, 19](#)
- [Uso do Acelerador no PowerCenter, 19](#)

## Visão Geral dos Aceleradores

Os aceleradores são pacotes de conteúdo que abordam problemas de qualidade de dados comuns em um país, uma região ou um setor. Um acelerador pode conter mapplets que você pode usar para analisar e aprimorar os dados em uma organização. Um acelerador também pode conter domínios de dados que você pode usar para descobrir os tipos de informações contidas nos dados.

Adicione os mapplets e os domínios de dados ao repositório do Modelo. A Informatica configura os mapplets e os domínios de dados para responder às regras comerciais que você pode definir para os dados da organização. Os aceleradores usam os termos *mapplet* e *regra* para identificar os mapplets. Quando você importa os mapplets para o repositório do Modelo, a ferramenta Developer cria os objetos de mapplet em uma pasta denominada *Regras*.

O Informatica Data Quality inclui um acelerador Principal e um acelerador dos Domínios de Dados Principais. Você pode comprar e fazer download de outros aceleradores da Informatica.

## Estrutura do Acelerador

Um acelerador é um arquivo compactado que contém os arquivos de metadados do repositório e outros arquivos em uma estrutura de diretório. A estrutura de diretório depende do tipo de acelerador. Os aceleradores gerais contêm regras, objetos de dados de referência, mapeamentos demonstrativos e fontes de dados demonstrativas. Os aceleradores de Domínios de Dados contêm regras, objetos de dados de referência, domínios de dados e grupos de domínios de dados.

## Estrutura do Acelerador Geral

Os aceleradores gerais incluem regras que analisam e aprimoram os dados da organização e os mapeamentos de amostra que demonstram as operações da regra. Os aceleradores gerais também contêm os arquivos de dados de referência e os arquivos de dados de origem usados pelos mapeamentos e pelas regras.

Um acelerador geral contém os seguintes diretórios:

- Accelerator\_Content
- Accelerator\_Sources

### Diretório Accelerator\_Content

O diretório Accelerator\_Content contém os seguintes componentes:

#### Arquivo XML do acelerador

Contém os metadados das regras, dos mapeamentos demonstrativos, das tabelas de referência e dos objetos de dados.

#### Arquivo de dados de referência

Contém os dados de referência que as regras e os mapeamentos usam para identificar as diferentes formas de valores de dados. O arquivo de dados de referência é um arquivo compactado que contém arquivos de dicionário em vários diretórios. Especifique o arquivo compactado quando você importar o arquivo XML correspondente. O processo de importação copia os dados de referência para as tabelas no banco de dados de referência.

**Nota:** Se você exportar um mapeamento que contém uma regra para o PowerCenter®, copie os arquivos de dicionário para um diretório que o Serviço de Integração do PowerCenter possa ler.

### Diretório Accelerator\_Sources

O diretório Accelerator\_Sources contém o arquivo de dados demonstrativos. O arquivo de dados demonstrativos é um arquivo compactado que contém os dados de origem dos mapeamentos demonstrativos. Copie o arquivo de dados de origem para o sistema de arquivos.

## Estrutura do Acelerador do Domínio de Dados

Os aceleradores do domínio de dados incluem os domínios de dados que determinam os tipos de informações em um conjunto de dados e as regras que definem a lógica do domínio de dados. Os aceleradores também contêm os arquivos de dados de referência que os domínios de dados e regras usam.

Um acelerador de domínio de dados contém os seguintes arquivos:

#### Arquivo de metadados do domínio de dados

Contém os metadados dos domínios de dados e dos grupos de domínios de dados que você adiciona ao glossário do domínio de dados.

#### Arquivo de metadados de regra

Contém os metadados das regras que definem a lógica do domínio de dados e dos objetos de dados de referência que os domínios de dados usam.

#### Arquivo de dados de referência dos domínios de dados

Contém os dados de referência que um domínio de dados usa quando você executa um perfil que contém o domínio de dados. O arquivo de dados de referência é um arquivo compactado que contém arquivos de dicionário em vários diretórios. Especifique o arquivo compactado quando você importar o arquivo XML correspondente. O processo de importação copia os dados de referência para as tabelas no banco de dados de referência.

### Arquivo de dados de referência das regras de domínio de dados

Contém os dados de referência que uma regra usa quando você executa um domínio de dados que a contém. O arquivo de dados de referência é um arquivo compactado que contém arquivos de dicionário em vários diretórios. Especifique o arquivo compactado quando você importar o arquivo XML correspondente. O processo de importação copia os dados de referência para as tabelas no banco de dados de referência.

## Instalação do Acelerador

Para instalar um acelerador, importe os metadados do objeto de repositório para um projeto de repositório do Modelo e copie os arquivos de dados demonstrativos para o sistema de arquivos. Use a Developer tool para importar os objetos de repositório.

Quando você importar regras e mapeamentos demonstrativos, selecione o projeto de repositório no Object Explorer. Quando você importar domínios de dados, selecione o projeto de repositório na caixa de diálogo **Preferências**. Em cada caso, a operação de importação solicita que você selecione o arquivo compactado que contém os dados de referência especificados pelo arquivo XML.

### Exemplo de Acelerador Geral

Você pode importar o seguinte arquivo de metadados para o acelerador Principal:

`Informatica_Core_Accelerator_1011.xml`

Quando você importar o arquivo de metadados, selecione o seguinte arquivo de dados de referência:

`Informatica_Core_Accelerator_1011.zip`

### Exemplo de Acelerador dos Domínios de Dados

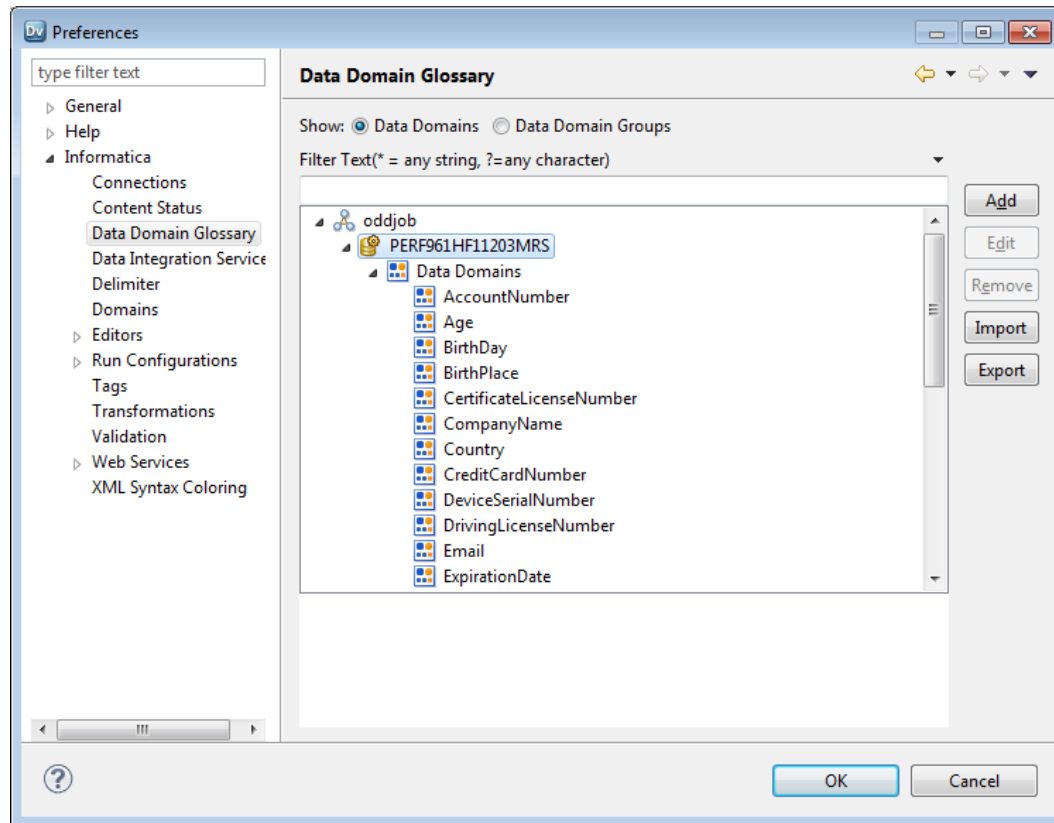
Você pode importar o seguinte arquivo de metadados para o acelerador dos Domínios de Dados Principais:

`Informatica_IDE_DataDomain_1011.xml`

Quando você importar o arquivo de metadados, selecione o seguinte arquivo de dados de referência:

`Informatica_IDE_DataDomain_1011.zip`

A seguinte imagem mostra os domínios de dados na caixa de diálogo **Preferências**:



### Dados de Origem para Mapeamentos de Amostra

Quando você importar um acelerador geral, copie os arquivos de dados demonstrativos para o seguinte diretório na máquina do Serviço de Integração de Dados:

<Diretório de Instalação do Informatica>\services\DQContent\INFA\_Content\demos\source\_data

## Regras e Diretrizes da Instalação do Acelerador

Os objetos do repositório e os arquivos de dados em um acelerador funcionam da mesma maneira que outros objetos e arquivos no sistema da Informatica. Algumas regras e diretrizes se aplicam ao conteúdo do acelerador.

Considere as seguintes regras e diretrizes quando instalar um acelerador:

- Antes de importar ou copiar arquivos, verifique se você tem todos os privilégios no Serviço de Integração de Dados, no Serviço do Gerenciamento de Conteúdo e no Serviço Analyst.
- Importe os aceleradores para um único projeto do repositório do Modelo. Crie o projeto antes de importar os aceleradores.
- Instale o acelerador Principal antes de instalar outro acelerador.
- Instale o acelerador do Domínio de Dados Principal antes de instalar o acelerador dos Domínios de Dados.
- Se você importar um arquivo de metadados que contém um objeto em comum com um acelerador importado anteriormente, substitua o objeto no repositório.

- Para usar as regras de acelerador que executam a validação de endereço, baixe e instale os arquivos de dados de referência de endereço do país especificado pelo acelerador. Para usar as regras de acelerador que executam a análise de correspondência de identidade, baixe e instale os arquivos de preenchimento de identidade do país especificado pelo acelerador. Adquirir os arquivos de dados de referência de endereço e os arquivos de preenchimento de identidade da Informatica.

## Importando Regras e Mapeamentos

Use a Developer tool para importar metadados para regras, mapeamentos de demonstração e fontes de dados de mapeamento. Durante a operação de importação, selecione o arquivo de dados de referência usado pelas regras e pelos mapeamentos.

1. Na ferramenta Developer, conecte-se ao repositório do Modelo que contém o projeto de destino dos metadados.
2. No Object Explorer, selecione o projeto de destino.  
Por exemplo, selecione o projeto *Informatica\_DQ\_Content*. Se necessário, crie um projeto no repositório do Modelo.
3. Selecione **Arquivo > Importar**.
4. Na caixa de diálogo **Importar**, selecione **Informatica > Importar Arquivo de Metadados de Objeto (Avançado)**.
5. Clique em **Avançar**.
6. Navegue até o arquivo de metadados XML na estrutura de diretórios do acelerador e selecione-o.
7. Clique em **Abrir** e em **Avançar**.
8. No painel **Origem**, selecione os itens exibidos no nó do projeto.
9. No painel **Destino**, selecione o projeto de destino.
10. Clique em **Adicionar ao Destino**.
  - Se o projeto de repositório contiver um objeto que você deseja adicionar, a ferramenta Developer solicitará que você mescle o objeto com o objeto atual. Clique em **Sim** para mesclar os objetos.
  - Se a ferramenta Developer solicitar que você renomeie os objetos, clique em **Não**.
  - Se algum objeto permanecer no painel **Origem**, use o ponteiro para mover o objeto para o projeto de destino.
11. Clique em **Avançar**.
12. Navegue até o arquivo de dados de referência compactado na estrutura de diretórios do acelerador e selecione-o.
13. Clique em **Abrir**.
14. Verifique se a página de código é UTF-8 e clique em **Avançar**.
15. No campo **Conexão de Destino**, selecione o banco de dados de referência.
16. Clique em **Concluir**.

## Importando Domínios de Dados e Grupos de Domínios de Dados

Use a caixa de diálogo Preferências para importar metadados para domínios de dados e grupos de domínios de dados. Durante a operação de importação, selecione o arquivo de dados de referência usado pelos domínios de dados.

1. Na ferramenta Developer, conecte-se ao repositório do Modelo que contém o projeto de destino dos metadados.

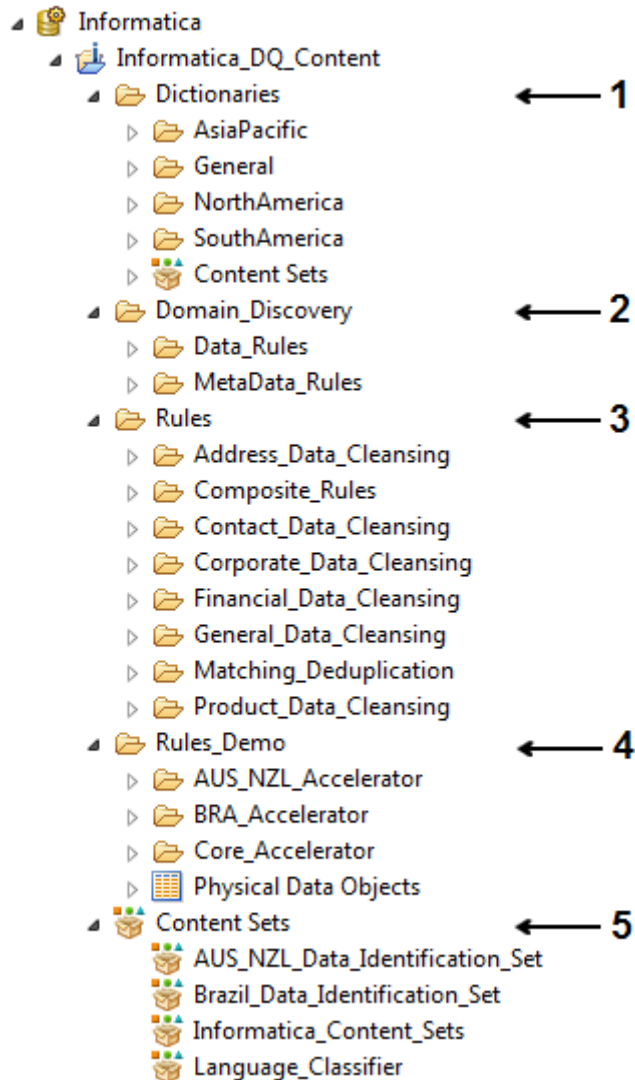
2. Selecione **Janela > Preferências**.
3. Na caixa de diálogo **Preferências**, expanda o nó **Informatica** e selecione **Glossário do Domínio de Dados**.
4. No painel do repositório, selecione o nó de nível superior dos domínios de dados ou dos grupos de domínios de dados.
5. Clique em **Importar**.
6. Navegue até o arquivo de metadados XML na estrutura de diretórios do acelerador e selecione-o.
7. Clique em **Abrir** e em **Avançar**.
8. No painel **Origem**, selecione o projeto do glossário do domínio de dados.
9. No painel **Destino**, selecione o projeto de destino.
10. Selecione a seguinte opção no campo **Resolução**:  
*Substituir opção no destino*
11. Clique em **Adicionar Conteúdo ao Destino**.
  - Se a ferramenta Developer solicitar que você adicione os objetos, clique em **Sim**.
  - Se a ferramenta Developer solicitar que você renomeie os objetos, clique em **Não**.
12. Clique em **Avançar**.
13. Se a operação de importação identificar dependências, copie os objetos dependentes do projeto de origem para o projeto de destino.
14. Clique em **Avançar**.
15. Navegue até o arquivo de dados de referência compactado na estrutura de diretórios do acelerador e selecione-o.
16. Clique em **Abrir**.
17. Verifique se a página de código é UTF-8 e clique em **Avançar**.
18. No campo **Conexão de Destino**, selecione o banco de dados de referência.
19. Clique em **Concluir**.

## Componentes do Acelerador

Quando você importa um acelerador, a ferramenta Developer cria pastas para as regras, os domínios de dados e os outros objetos especificados pelo acelerador. Cada pasta contém subpastas que organizam os objetos por país e pelo tipo de operação de qualidade de dados realizada.

Use o acelerador Principal para criar as pastas em um projeto de repositório. Ao importar mais aceleradores, você adiciona objetos e pastas ao projeto.

A seguinte imagem mostra a estrutura da pasta de projeto Informatica\_DQ\_Content quando você importa vários aceleradores para o projeto:



1. Pasta Dictionaries
2. Pasta Domain\_Discovery
3. Pasta Rules
4. Pasta Rules\_Demo
5. Pasta Conjuntos de Conteúdo

O projeto contém as seguintes pastas de nível superior:

#### **Dictionaries**

A pasta Dictionaries contém objetos de tabela de referência. Cada objeto se refere a uma tabela no banco de dados de referência.

#### **Domain\_Discovery**

A pasta Domain\_Discovery contém as regras que definem os domínios de dados nos aceleradores instalados. A pasta contém uma pasta Data\_Rules e uma pasta Metadata\_Rules. As regras na pasta Data\_Rules correspondem aos domínios de dados que analisam os valores de dados de coluna. As



regras na pasta Metadata\_Rules correspondem aos domínios de dados que analisam os nomes de coluna.

#### **Rules**

A pasta Regras contém as regras usadas para analisar e aprimorar os dados.

#### **Rules\_Demo**

A pasta Rules\_Demo contém os mapeamentos demonstrativos e as fontes de dados demonstrativas.

#### **Conjuntos de Conteúdo**

A pasta Conjuntos de Conteúdo contém os objetos de dados de referência que não especificam os dados no banco de dados de referência.

## **Regras**

As regras do acelerador definem um intervalo de operações de análise e transformação de dados. Você pode adicionar uma única regra ou uma série de regras a um mapeamento.

Use as regras do acelerador para executar as seguintes tarefas de qualidade de dados:

#### **Validação de endereço**

Valide e aprimore os dados nos registros de endereço postal. As regras exigem arquivos de dados de referência de endereço.

#### **Análise de dados**

Informações de análise de registros. As regras de análise podem extrair vários tipos de informações, incluindo nomes de pessoas, nomes de organização, números de telefone, datas e números de identificação.

#### **Padronização de dados**

Padroniza a ortografia e o formato dos valores de dados. As regras de padronização podem identificar e corrigir vários tipos de informações, incluindo nomes de pessoas, nomes de organização, números de telefone, datas e números de identificação.

#### **Análise de duplicação**

Localiza registros duplicados em um conjunto de dados. As regras de análise de duplicação comparam os registros em um conjunto de dados e geram uma pontuação numérica que representa o grau de semelhança entre os registros.

As regras de análise de duplicação podem ler registros que contêm dados corporativos gerais e registros que contêm dados de identidade. As regras de dados de identidade exigem arquivos de dados de preenchimento de identidade.

A operação de importação adiciona as regras à seguinte pasta de repositório:

```
[Informatica_DQ_Content]\Rules
```

Localize as regras que executam as operações de validação de endereço, análise de dados e padronização de dados nas subpastas *Limpeza de Dados* no projeto do acelerador. Localize as regras que executam a análise de duplicação na subpasta *Matching\_Deduplication* no projeto do acelerador.

Se você importar as regras de um país ou região, adicione uma subpasta para as regras de composição. Uma regra de composição combina várias regras em um formato aninhado de uma única regra.

## Mapeamentos Demonstrativos

Os mapeamentos demonstrativos são objetos de tempo de execução que aplicam a uma ou mais regras a uma fonte de dados e gravam os resultados em outra fonte de dados. Você pode usar os mapeamentos demonstrativos como modelos para outros mapeamentos.

A operação de importação adiciona os mapeamentos e os dados de objetos de fonte de dados à seguinte pasta de repositório:

```
[Informatica_DQ_Content]\Rules_Demo
```

Quando você importa um acelerador, a operação de importação adiciona a fonte de dados dos mapeamentos demonstrativos à pasta `Rules_Demo`. Copie os arquivos de fonte de dados do diretório `Accelerator_Sources` para o sistema de arquivos.

## Domínios de Dados

Um domínio de dados descreve os valores de dados que podem representar um único tipo de informações comerciais em uma coluna. Use os domínios de dados para determinar o tipo de informações em uma coluna e localizar as informações de um tipo especificado em uma coluna. Os aceleradores incluem domínios de dados para um intervalo de tipos de informações, incluindo os números de Previdência Social, os números de cartão de crédito, os endereços de e-mail e os cargos.

Por exemplo, uma tabela de banco de dados pode conter números de Previdência Social em uma coluna Comentários que qualquer usuário pode ler. Você deve identificar os registros que contêm os números de Previdência Social e excluí-los ou movê-los. Adicione o domínio de dados `SSN` a um perfil e execute-o coluna Comentários.

Você pode atribuir um domínio de dados a um ou mais grupos de domínios de dados. Use os grupos de domínios de dados para organizar os domínios de dados com base no tipo de análise comercial que os domínios de dados realizam. O glossário de domínios de dados lista os domínios de dados e os grupos de domínios de dados que você adiciona ao repositório do Modelo. Use o menu **Preferências** na ferramenta Developer para adicionar domínios de dados ao glossário de domínios de dados. Para atualizar as definições de dados em um domínio de dados, use as regras no acelerador do domínio de dados.

**Nota:** Você não pode exibir os objetos de domínio de dados no Object Explorer.

## Tabelas de Referência

Uma tabela de referência contém as versões padrão e alternativa de um conjunto de valores de dados. As regras usam as tabelas de referência para verificar se os valores de dados estão precisos e formatados corretamente.

A operação de importação adiciona as tabelas de referência à seguinte pasta de repositório:

```
[Informatica_DQ_Content]\Dictionaries
```

## Conjuntos de Conteúdo

Um conjunto de conteúdo é um objeto de dados de referência que não armazena dados em tabelas de banco de dados. Os conjuntos de conteúdo incluem conjuntos de caracteres, conjuntos de padrão, expressões regulares, conjuntos de tokens, modelos probabilísticos e modelos do classificador.

A operação de importação adiciona as regras à seguinte pasta de repositório:

```
[Informatica_DQ_Content]\Content Sets
```

**Nota:** Para exibir uma lista dos elementos em um conjunto de conteúdo, abra o conjunto de conteúdo na ferramenta Developer e selecione a guia **Marcas**.

# Marcas e Regras

As regras de acelerador incluem marcas que indicam o tipo de dados que a regra pode ler e o tipo de operação que pode executar a regra.

Para exibir as marcas que se aplicam a uma regra, abra a regra na ferramenta Developer e clique na guia Marcas. Você pode usar as opções de Pesquisa na ferramenta Developer para localizar os aceleradores que contêm uma marca especificada por você.

## Uso do Acelerador no PowerCenter

Você pode exportar regras e mapeamentos do repositório do Modelo para o sistema de arquivos e para o repositório do PowerCenter. Quando você exportar os objetos, selecione as tabelas de referência, os objetos de dados e as outras dependências nos objetos exportados.

A operação de exportação copia os dados da tabela de referência para o sistema de arquivos. Copie os arquivos para a máquina host do Serviço de Integração do PowerCenter. As localizações do arquivo de dados de referência na estrutura de diretório do PowerCenter devem corresponder às localizações das tabelas de referência na estrutura de pasta do repositório do Modelo.

O seguinte caminho descreve uma estrutura de diretório de amostra dos objetos de dados de referência em uma instalação do PowerCenter:

```
<Informatica_installation_directory>\services\<Model_repository_project_name>\<Model_repository_project_folder_name>
```

**Nota:** Se a versão do produto PowerCenter não corresponder à versão da ferramenta Developer, verifique se o ambiente do PowerCenter inclui o Plug-in de Integração do Data Quality.

Para obter mais informações sobre a integração do Data Quality com o PowerCenter, leia o Guia do Usuário de Integração de Informatica Data Quality com o PowerCenter.

## CAPÍTULO 2

# Acelerador Principal

Este capítulo inclui os seguintes tópicos:

- [Visão Geral do Acelerador Principal, 20](#)
- [Regras de Limpeza de Dados de Endereço Principal, 21](#)
- [Regras de Limpeza de Dados do Contato Principal, 22](#)
- [Regras de Limpeza de Dados Corporativos Principais, 23](#)
- [Regras de Limpeza de Dados Gerais Principais, 23](#)
- [Regras Principais de Correspondência e de Eliminação de Duplicação, 29](#)
- [Regras de Limpeza de Dados de Produto Principal, 30](#)
- [Mapeamentos Demonstrativos Principais, 30](#)

## Visão Geral do Acelerador Principal

Use as regras no acelerador Principal para verificar e aprimorar os dados comerciais em qualquer país ou região.

O acelerador Principal inclui regras que executam os seguintes processos de qualidade de dados:

- Limpeza de dados de endereço
- Limpeza de dados de contato
- Limpeza de dados corporativos
- Limpeza de dados gerais
- Limpeza de dados de correspondência e de eliminação de duplicação
- Limpeza de dados de produto

O acelerador Principal contém mapplets e objetos de dados de referência que outros aceleradores podem reutilizar. Instale o acelerador Principal antes de instalar qualquer outro acelerador.

# Regras de Limpeza de Dados de Endereço Principal

Use as regras de limpeza de dados de endereço para analisar, padronizar e validar dados de endereço.

As regras de limpeza de dados de endereço se encontram na seguinte localização do repositório:

[Informatica\_DQ\_Content]\Rules\Address\_Data\_Cleansing

A tabela a seguir descreve as regras de limpeza de dados de endereço do acelerador Principal:

Nome	Descrição
mplt_Global_AddressValidation5_v2_Discrete_Webservice	Valida os endereços postais de vários países. Use o mapplet quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.  O mapplet chama um serviço da Web de validação de endereço. Use o mapplet como um exemplo quando você estabelecer outros mapplets de serviço da Web.
mplt_Global_AddressValidation5_v2_Hybrid_Webservice	Valida os endereços postais de vários países. Use o mapplet quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.  O mapplet chama um serviço da Web de validação de endereço. Use o mapplet como um exemplo quando você estabelecer outros mapplets de serviço da Web.
mplt_Global_AddressValidation5_v2_Multiline_Webservice	Valida os endereços postais de vários países. Use o mapplet quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.  O mapplet chama um serviço da Web de validação de endereço. Use o mapplet como um exemplo quando você estabelecer outros mapplets de serviço da Web.
rule_Calc_Distance_Between_Geocoordinates	Calcula a distância entre dois conjuntos de coordenadas geográficas.
rule_Country_Identification	Identifica um país.
rule_Country_Name_Standardization	Padroniza nomes de países. A regra retorna um nome de país e os códigos de país ISO de dois e três caracteres.
rule_Geoocordinate_In_Polygon	Verifica a presença de pontos de coordenadas geográficas em uma área definida por três ou mais pontos de coordenadas geográficas.
rule_Global_Address_Parse_Hybrid	Analisa os endereços não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_Global_Address_Parse_Multiline	Analisa os endereços não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_Global_Address_Validation_Discrete	Valida a capacidade de entrega dos registros de endereço de vários países. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.

Nome	Descrição
rule_Global_Address_Validation_Discrete_w_Geocoding	Valida a capacidade de entrega dos registros de endereço de vários países e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_Global_Address_Validation_Hybrid	Valida a capacidade de entrega dos registros de endereço de vários países. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_Global_Address_Validation_Hybrid_w_Geocoding	Valida a capacidade de entrega dos registros de endereço de vários países e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_Global_Address_Validation_Multiline	Valida a capacidade de entrega dos registros de endereço de vários países. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_Global_Address_Validation_Multiline_w_Geocoding	Valida a capacidade de entrega dos registros de endereço de vários países e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.

## Regras de Limpeza de Dados do Contato Principal

Use as regras de limpeza de dados de contato para analisar e validar dados sobre contatos comerciais e indivíduos.

Localize as regras de limpeza de endereço de contato se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados de contato no acelerador Principal:

Nome	Descrição
rule_Email_Parse	Analisa endereços de e-mail dos campos de dados.
rule_Email_Parse_and_Validate	Analisa os endereços de e-mail de campos de dados e valida o formato de cada endereço de e-mail.

Nome	Descrição
rule_Email_Parse_Into_Mailbox_Domain	Analisa endereços de e-mail de caixa de correio, domínio e portas de subdomínio. Por exemplo, a regra analisa <i>info@informatica.com</i> da seguinte maneira: <ul style="list-style-type: none"> <li>- Caixa de correio: info</li> <li>- Subdomínio: informatica</li> <li>- Domínio: com</li> </ul>
rule_Email_Validation	Valida o formato de endereços de e-mail. A regra não verifica se os endereços de e-mail estão corretos ou ativos. A regra retorna Válido ou Inválido.
rule_Identify_Suspect_Names	Identifica os nomes que podem não ser nomes de pessoa autênticos. A regra compara os valores de entrada com uma tabela de referência de nomes que provavelmente não são autênticos. Por exemplo, a tabela de referência inclui os nomes de personagens fictícios.

## Regras de Limpeza de Dados Corporativos Principais

Use as regras de limpeza de dados corporativos no acelerador Principal para padronizar os dados corporativos.

As regras de limpeza de dados corporativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados corporativos no acelerador Principal:

Nome	Descrição
rule_Company_Name_Standardization	Usa tabelas de referência para padronizar os nomes de empresa.

## Regras de Limpeza de Dados Gerais Principais

Use as regras de limpeza de dados gerais para analisar, padronizar e validar dados.

As regras de limpeza de dados gerais se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados gerais no acelerador Principal:

Nome	Descrição
mplt_Parse_Tokens_Into_Single_Field	Analisa cada palavra em uma cadeia separada por espaços para uma porta separada.
rule_Add_Leading_Zero	Adiciona o numeral "0" ao início de uma cadeia.

Nome	Descrição
rule_Add_Parentheses_At_Start_End_ofLine	Adiciona símbolos parentéticos ao início e ao fim de uma cadeia.
rule_Add_Plus_To_Start_of_Line	Adiciona o símbolo de adição no início de uma cadeia.
rule_Add_Space_Around_Ampersand	Adiciona um espaço antes e depois de todos os 'e' comerciais em uma cadeia.
rule_Add_Space_Around_Hyphen	Adiciona um espaço antes e depois de todos os traços e hífen em uma cadeia.
rule_Add_Space_Between_Number_Letter	Adiciona um espaço entre um par de caracteres composto por um numeral e um caractere alfabético. Lendo da esquerda para a direita, o maplet adiciona um espaço ao primeiro par de caracteres numeral-alfabético nos dados.
rule_Add_Spaces_Around_Period	Adiciona um espaço antes e depois de todos os períodos de uma cadeia.
rule_AllTrim	Remove todos os espaços à esquerda e à direita dos campos de dados de entrada.
rule_Assign_DQ_AddressResolutionCode_Description	Atribui uma descrição à saída do Código de Resolução de Endereço da transformação do Validador de Endereço.
rule_Assign_DQ_ElementInputStatus_Description	Atribui uma descrição à saída do Status de Entrada do Elemento da transformação do Validador de Endereço. A descrição corresponde à saída das transformações do Data Quality em versões anteriores ao Data Quality 9.0.
rule_Assign_DQ_ElementRelevance_Description	Atribui uma descrição à saída da Relevância do Elemento da transformação do Validador de Endereço. A descrição corresponde à saída das transformações do Data Quality em versões anteriores ao Data Quality 9.0.
rule_Assign_DQ_ElementResultStatus_Description	Atribui uma descrição à saída do Status do Resultado do Elemento da transformação do Validador de Endereço. A descrição corresponde à saída das transformações do Data Quality em versões anteriores ao Data Quality 9.0.
rule_Assign_DQ_ExtendedElementStatus_Description	Atribui uma descrição à saída do Status do Resultado do Elemento Estendido da transformação do Validador de Endereço.
rule_Assign_DQ_GeocodingStatus_Description	Atribui uma descrição à saída do Status de Codificação Geográfica da transformação do Validador de Endereço. A descrição corresponde à saída das transformações do Data Quality em versões anteriores ao Data Quality 9.0.
rule_Assign_DQ_Mailability_Score_Description	Atribui uma descrição à saída da Pontuação de Capacidade de Envio por Correio da transformação do Validador de Endereço. A descrição corresponde à saída das transformações do Data Quality em versões anteriores ao Data Quality 9.0.
rule_Assign_DQ_Match_Code_Description	Atribui uma descrição à saída do Código de Correspondência da transformação do Validador de Endereço. A descrição corresponde à saída das transformações do Data Quality em versões anteriores ao Data Quality 9.0.



Nome	Descrição
rule_Classify_Language	<p>Classifica uma cadeia como um dos seguintes idiomas: árabe, holandês, inglês, francês, alemão, italiano, português, russo, espanhol ou turco. A regra usa o conjunto de conteúdo Language_Classifier para identificar os idiomas.</p> <p><b>Nota:</b> A regra retorna um idioma para cada cadeia que ele analisa. Se uma cadeia pertencer a um idioma que a regra não reconhece, a regra retornará o idioma mais parecido com o texto na cadeia.</p>
rule_Compare_Dates	<p>Calcula a diferença entre duas datas. O mapplet usa as seguintes unidades de medida:</p> <ul style="list-style-type: none"> <li>- Horas</li> <li>- Dias</li> <li>- Meses</li> <li>- Anos</li> </ul> <p>Cada valor de saída é exclusivo de outros valores. As saídas não podem ser adicionadas para representar a diferença entre os valores de dados.</p>
rule_Completeness	Verifica uma única porta para valores NULL. Retorna "Concluído" se a porta contém dados. Retorna "Incompleto" se a porta está vazia ou contém um valor NULL.
rule_Completeness_Multi_Port	Verifica várias portas para valores NULL. Retorna "Concluído" se todas as portas contém dados. Retorna "Incompleto" se qualquer porta está vazia ou contém um valor NULL.
rule_Concatenate_Words	Concatena dois campos. Usa um caractere de espaço como um separador.
rule_Convert_Match_Codes_to_Legacy_Values	Converte a saída da porta de Código de Correspondência em uma transformação do Validador de Endereço no código de correspondência de validação de endereço equivalente no Data Quality 8.6.
rule_CreditCard_Number_Validation	<p>Valida números de cartão de crédito dos cartões de crédito que usam o algoritmo de Luhn. A validação inclui, mas não se limita, aos seguintes cartões de crédito:</p> <ul style="list-style-type: none"> <li>- American Express</li> <li>- Diners Club Carte Blanche</li> <li>- Diners Club International</li> <li>- Diners Club EUA e Canadá</li> <li>- Discover Card</li> <li>- JCB</li> <li>- Maestro</li> <li>- Master Card</li> <li>- Solo</li> <li>- Switch</li> <li>- Visa</li> <li>- Visa Electron</li> </ul> <p>A regra retorna "Válido" ou "Inválido".</p>
rule_Date_Complete	<p>Verifica se a cadeia de entrada está em conformidade com um formato de data reconhecido pela regra. A regra lê o seguinte objeto de dados de referência:</p> <ul style="list-style-type: none"> <li>- user_defined_dates_inf</li> </ul>

Nome	Descrição
rule_Date_of_Birth_Validation	Verifica o número de anos entre uma data de nascimento e a data atual. Retorna "Adulto" ou "Menor de idade" além de "Válido" se a idade é menor ou igual a 120. Retorna "Inválido" se a idade é maior que 120.
rule_Date_Parse	Analisa os dados de data de uma cadeia para uma porta que a regra especifica. A regra reconhece datas nos seguintes formatos: <ul style="list-style-type: none"> <li>- dd/mm/yyyy</li> <li>- mm/dd/yyyy</li> <li>- yyyy/dd/mm</li> </ul> A regra retorna uma data e também retorna uma cadeia que contém o texto de entrada sem a data.
rule_Date_Standardization	Padroniza as cadeias de data em um formato de saída especificado por você. Para definir o formato de saída, abra a transformação de Expressão dq_FormatDate na regra e atualize a variável de expressão Output_Date_Format e a variável de expressão Delimitador. Se os dados de entrada não descreverem uma data válida, a regra retornará o dígito 0 para cada caractere de entrada.
rule_Date_Validation	Valida as cadeias de data exibidas em um único formato em uma coluna de dados. Para configurar o formato de data que a regra usa para validação, abra a transformação de Expressão dq_ValidateDate na regra e atualize a variável de expressão In_Date_Format. O formato padrão é "MM/DD/YYYY". A regra retorna "Válido" ou "Inválido".
rule_Date_Validation_Variable_Format	Valida as cadeias de data exibidas em vários formatos em uma coluna de dados. Use a regra quando uma fonte de dados incluir as seguintes colunas: <ul style="list-style-type: none"> <li>- Uma coluna que contém valores de data em vários formatos.</li> <li>- Uma coluna que identifica o formato do valor de data em cada linha. Se a coluna não identificar um formato de data de uma linha, a regra aplicará o formato "MM/DD/YYYY" ao valor de data.</li> </ul> A regra lê todos os valores de dados que a função <i>is_date()</i> reconhece. A regra retorna "Válido" ou "Inválido".
rule_Days_Between_Dates	Calcula o número de dias entre duas datas.
rule_Days_From_Current_Date	Calcula o número de dias entre uma data especificada e a data atual.
rule_EAN13_Algorithm	Valida um Número do Artigo Internacional. A regra retornará "Válido" se o dígito de verificação estiver correto para o número e "Inválido" se o dígito de verificação estiver incorreto.
rule_GTIN_Validation	Valida um Global Trade Item Number (GTIN). A regra valida números de oito, doze, treze e quatorze dígitos. A regra retornará "Válido" se o dígito de verificação estiver correto para o número e "Inválido" se o dígito de verificação estiver incorreto.
rule_IsNumeric	Verifica se os dados de entrada são numéricos. A regra retorna "True" ou "False".
rule_LowerCase	Retorna todos os caracteres alfabéticos em letras minúsculas.

Nome	Descrição
rule_Luhn_Algorithm	Aplica o algoritmo de Luhn a uma cadeia numérica. A regra pode validar cadeias numéricas, como números de cartão de crédito.
rule_Mask_Profanity	Verifica se há profanação nos dados de entrada. Mascara a profanação como "CENSORED" nos dados de saída.
rule_Negative_Number_Validation	Verifica se os dados de entrada são um número negativo.
rule_Numeric_Completeness	Verifica valores NULL nas entradas numéricas.
rule_Parse_Alpha_Chars_from_Non_Alpha_Chars	Identifica os caracteres alfabéticos e não alfabéticos em uma cadeia de entrada e grava cada conjunto de caracteres em diferentes portas de saída. Por exemplo, a regra analisa os seguintes valores da cadeia de entrada <code>teststring_123</code> : <code>teststring</code> <code>_123</code>
rule_Parse_First_Word	Analisa a primeira palavra em uma cadeia de entrada para uma porta que a regra especifica.
rule_Parse_Number_At_End_Of_Line	Analisa qualquer número que ocorre no final de uma cadeia de entrada para uma porta que a regra especifica. A regra lê cadeias da esquerda para direita.
rule_Parse_Number_At_Start_Of_Line	Analisa qualquer número que ocorre no início de uma cadeia de entrada para uma porta que a regra especifica. A regra lê cadeias da esquerda para direita.
rule_Parse_Profanity	Compara as cadeias de uma tabela de referência de termos profanos e analisa qualquer termo que corresponda a um valor da tabela de referência para uma porta que a regra especifica.
rule_Parse_Text_Between_Parentheses	Analisa cadeias que estão entre parênteses para uma porta que a regra especifica. A regra contém uma porta de saída das cadeias analisadas e outra para o texto de entrada sem as cadeias analisadas.
rule_Parse_Text_in_Single_Quotes	Analisa cadeias que estão entre aspas duplas para uma porta que a regra especifica. Quando os dados de entrada contêm vários elementos entre aspas, a regra analisa o elemento final. A regra lê as cadeias de entrada da esquerda para a direita. A regra contém uma porta de saída das cadeias analisadas e outra para o texto de entrada sem as cadeias analisadas.
rule_Past_Date_Label	Determina se uma data de entrada é anterior ou posterior à data do sistema.
rule_Personal_Company_Identification	Analisa nomes de pessoa e de empresa para portas diferentes que a regra especifica. A regra tem as seguintes saídas: <ul style="list-style-type: none"> <li>- Nome da pessoa</li> <li>- Nome da empresa</li> <li>- Categoria de dados, como nome da pessoa ou nome da empresa</li> <li>- Dados que a regra não pode analisar</li> </ul>
rule_Positive_Number_Validation	Verifica se os dados de entrada são um número positivo.

Nome	Descrição
rule_Prepnd_Zero_to_Single_Digit	Insere o numeral "0" para caracteres numéricos únicos.
rule_Remove_All_Leading_Zeros	Remove todas as instâncias do caractere numérico "0" do início de uma cadeia.
rule_Remove_Apostrophe	Remove apóstrofes. A regra mescla as cadeias de texto nos dois lados da apóstrofe.
rule_Remove_Control_Characters	Remove os caracteres de controle das cadeias de texto. A regra retorna uma cadeia que contém os caracteres de controle e uma cadeia que contém o texto de entrada sem os caracteres de controle.
rule_Remove_Extra_Spaces	Substitui todos os espaços consecutivos por um único espaço e retira os espaços à esquerda e à direita.
rule_Remove_Hyphen	Remove hífen.
rule_Remove_Leading_Zero	Remove uma única instância do caractere numérico "0" do início de uma cadeia.
rule_Remove_Limited_Punctuation	Remove caracteres externos. Os caracteres externos incluem barras, barras invertidas, pontos, pontos de exclamação, sublinhados e vários espaços consecutivos.
rule_Remove_Non_Numbers	Remove todos os caracteres que não são numéricos.
rule_Remove_Parentheses	Remove os símbolos de parênteses direito e esquerdo.
rule_Remove_Period	Remove pontos.
rule_Remove_Period_Parentheses	Remove os seguintes caracteres: - Parênteses esquerdo e direito - Pontos
rule_Remove_Punctuation	Remove símbolos de pontuação.
rule_Remove_Punctuation_and_Space	Remove todas as pontuações e todos os caracteres de espaço.
rule_Remove_Quotation	Remove aspas.
rule_Remove_Slashes	Remove barras diagonais e barras invertidas.
rule_Remove_Space	Remove todos os caracteres de espaço.
rule_Replace_Ampersand_With_Space	Substitui "&" comerciais por espaços.
rule_Replace_Hyphen_Underscore_with_Space	Substitui hífen e sublinhados por espaços.
rule_Replace_Hyphen_with_Space	Substitui hífen por espaços.
rule_Replace_Limited_Punct_with_Space	Substitui o seguinte caractere de pontuação por um único espaço: traço, barra invertida, ponto, ponto de exclamação e sublinhado. A regra também substitui dois, três e quatro espaços consecutivos por um único espaço.

Nome	Descrição
rule_Replace_Non_Alphabetic_with_Space	Substitui caracteres numéricos e de pontuação por um único espaço.
rule_Replace_Period_With_Space	Substitui pontos por um único espaço.
rule_Replace_Punctuation_with_Space	Substitui todas as pontuações por espaços.
rule_Replace_Slashes_With_Space	Substitui barras diagonais e barras invertidas por espaços.
rule_Reverse_String_Input	Reverte a ordem de caracteres em cadeias de entrada.
rule_String_Completeness	Verifica se uma cadeia está completa. A regra também pesquisa se há valores nas cadeias de entrada da tabela de referência string_default_values_infa. A tabela de referência contém valores como NA, DEFAULT e XX. Se uma cadeia de entrada tiver um valor na tabela de referência, a regra identificará a cadeia como incompleta.
rule_TitleCase	Converte cadeias em capitalização de título. Em cadeias de capitalização de título, a primeira letra de cada palavra fica em maiúscula.
rule_Translate_Diacritic_Characters	Substitui caracteres diacríticos por equivalentes em ASCII. Por exemplo, a regra converte "ä" para "a".
rule_UpperCase	Retorna todos os caracteres alfabéticos em letras maiúsculas.
rule_URL_Validation	Valida o formato e a estrutura de uma URL.
rule_Years_Since_Date_of_Birth	Calcula o número de anos desde a data de entrada.

## Regras Principais de Correspondência e de Eliminação de Duplicação

Use as regras de correspondência e de eliminação de duplicação para identificar os registros duplicados.

As regras de correspondência e de eliminação de duplicação se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

A tabela a seguir descreve as regras de correspondência e de eliminação de duplicação no acelerador Principal:

Nome	Descrição
mplt_Consolidate_and_Remove_Duplicate_Rows	Consolida os clusters de registros duplicados em um único registro e remove os registros duplicados redundantes.

# Regras de Limpeza de Dados de Produto Principal

Use as regras de limpeza de dados de produto para analisar, padronizar e validar dados de produto.

As regras de limpeza de dados de produto se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Product_Data_Cleansing
```

A seguinte tabela descreve as regras de limpeza de dados de produto no acelerador Principal:

Nome	Descrição
rule_Color_Parse	Analisa valores de cor para uma porta que a regra especifica.
rule_Parse_Quantity_And_UOM	Analisa a primeira ocorrência de uma quantidade e uma unidade de medida de uma cadeia para uma porta que a regra especifica. A regra é a cadeia da esquerda para direita e retorna os seguintes dados: <ul style="list-style-type: none"><li>- Quantidade.</li><li>- Unidade de medida.</li><li>- A cadeia de entrada sem a quantidade e a unidade de valores de medida.</li></ul>
rule_UOM_Standardization	Padroniza uma unidade de medida. A regra retorna valores padronizados e não padronizados de quantidade e unidade de medida. Ela também retorna uma cadeia que contém o texto de entrada com uma unidade de medida padronizada.
rule_UPC_Validation	Valida um Código de Produto Universal e retorna um Código de Produto Universal padronizado.

## Mapeamentos Demonstrativos Principais

Os mapeamentos demonstrativos do acelerador Principal usam várias regras para demonstrar processos de qualidade de dados.

Os mapeamentos demonstrativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules_Demo\Core_Accelerator
```

O acelerador contém os seguintes mapeamentos demonstrativos:

### **m\_customer\_data\_demo**

Analisa, padroniza e valida dados dos Estados Unidos e do Canadá.

### **m\_product\_demo**

Analisa descrições de produto e valida a qualidade das descrições.

## CAPÍTULO 3

# Acelerador dos Domínios de Dados

Este capítulo inclui os seguintes tópicos:

- [Visão Geral do acelerador dos Domínios de Dados, 31](#)
- [Domínios de dados no acelerador dos Domínios de Dados, 32](#)
- [Regras de nome de coluna no acelerador dos Domínios de Dados, 40](#)
- [Regras de dados no acelerador dos Domínios de Dados, 44](#)

## Visão Geral do acelerador dos Domínios de Dados

Um domínio de dados é um objeto do repositório do Modelo predefinido ou definido pelo usuário que usa regras para descobrir o significado funcional dos dados de coluna ou dos nomes de coluna. As regras de domínio de dados definem os padrões de dados e de nome de coluna que correspondem aos dados e aos metadados de origem. Você pode usar as regras de domínio de dados para atualizar a lógica do domínio de dados.

Use os domínios de dados no acelerador dos Domínios de Dados para descobrir o significado funcional dos dados de origem com base nos valores de nomes de coluna ou dados de coluna.

O acelerador dos Domínios de Dados inclui os seguintes tipos de regra:

- Regra de dados. Localiza colunas com dados que correspondem à lógica que a regra define.
- Regra de nome de coluna. Localiza colunas com os nomes de coluna que correspondem à lógica coluna-nome que a regra define.

As regras de domínio de dados retornam valores Booleanos que indicam se os dados de coluna ou nome de coluna atende aos critérios de regra. As regras de domínio de dados usam expressões regulares ou tabelas de referência para procurar valores ou padrões específicos. Por exemplo, você pode usar uma expressão de regra de nove dígitos para localizar valores de dados no formato do número de Previdência Social.

Quando você usa expressões em regras de domínio de dados, alguns valores de dados não relacionados também podem atender aos critérios da expressão de regra. Por exemplo, os CEPs dos Estados Unidos nos dados de origem podem atender ao formato do número de Previdência Social. Para fazer com que a inferência de domínio de dados seja eficiente, analise os resultados da descoberta do domínio de dados em busca de discrepâncias. Depois de analisar e verificar os resultados da descoberta do domínio de dados, você pode decidir associar um domínio de dados a uma coluna de dados.

# Domínios de dados no acelerador dos Domínios de Dados

Use os domínios de dados predefinidos em perfis para realizar a descoberta do domínio de dados e identificar as características de dados críticos em uma empresa.

**Nota:** Na tabela, o símbolo de asterisco (\*) é um caractere curinga.

A seguinte tabela descreve os domínios de dados disponíveis no acelerador dos Domínios de Dados:

Nome	Descrição	Tipo de Regra Dependente	Grupo de Domínios de Dados
Account_Status	Descobre os dados de coluna que correspondem aos valores de status da conta em uma tabela de referência.	Regra de dados	Account_Bank
AccountNumber	Descobre os nomes de coluna que contêm a cadeia "a*c*num" ou "acc".	Regra de nome de coluna	Account_Bank, PCI, PHI
Age	Descobre os nomes de coluna que contêm a cadeia "age" ou "dob", além dos dados de coluna com valores de 1 a 120.	Regra de nome de coluna Regra de dados	PII
AlphaNumeric_SpecialCharacter	Descobre os dados de coluna que contêm dados alfanuméricos não formatados e dados de caracteres especiais.	Regra de dados	Geral
Austria_NationalID	Descobre os dados de coluna que correspondem ao ID nacional da Áustria.	Regra de dados	NationalID
BIC_SwiftCode	Descobre os dados de coluna que correspondem ao código de Bank Identifier Code (BIC) ou Society for Worldwide Interbank Financial Telecommunication (SWIFT) pelo reconhecimento de padrão e pelo código de país.	Regra de dados	Account_Bank
BinayValue	Descobre os dados de coluna que contêm valores binários.	Regra de dados	Geral
BirthDay	Descobre os nomes de coluna que contêm a cadeia "dob", "date*of*bir*" ou "birth*da*" e identifica os dados de coluna que correspondem a datas de nascimento válidas.	Regra de nome de coluna Regra de dados	PII
BirthPlace	Descobre os nomes de coluna que contêm a cadeia "birth*place" ou "location*birth".	Regra de nome de coluna	PII



Nome	Descrição	Tipo de Regra Dependente	Grupo de Domínios de Dados
Brazil_IDDoc	Descobre os dados de coluna que correspondem ao formato do número do cartão de ID Brasileiro chamado Registro Geral.	Regra de dados	NationalID
Brazil_NationalID	Descobre os dados de coluna que correspondem ao ID nacional do Brasil.	Regra de dados	NationalID
Bulgaria_NationalID	Descobre os dados de coluna que correspondem ao ID nacional da Bulgária.	Regra de dados	NationalID
Canada_SIN	Descobre os dados de coluna que correspondem ao formato do Número da Previdência Social do Canadá.	Regra de dados	NationalID
CertificateLicenseNumber	Descobre os nomes de coluna que contêm a cadeia "cert*lic*number", "cert*lic*no*", "lic* nu*" ou "lic*no*".	Regra de nome de coluna	PHI
China_NationalID	Descobre os dados de coluna que correspondem ao ID nacional da China.	Regra de dados	NationalID
CompanyName	Descobre os nomes de coluna que contêm a cadeia "company" e identifica os dados de coluna que correspondem aos valores de organização-nome em uma tabela de referência.	Regra de nome de coluna Regra de dados	Contact, PII
ComputerAddress	Descobre os dados de coluna que correspondem aos formatos de endereço de computador.	Regra de dados	Geral
Country	Descobre os nomes de coluna que contêm a cadeia "iso*countr*code", "iso*country" ou "countr*", além de identificar os dados de coluna que correspondem aos nomes de países.	Regra de nome de coluna Regra de dados	Address, PII
CountryCode_Phone	Descobre os números de telefone com base nos códigos de discagem internacional.	Regra de dados	Contact

Nome	Descrição	Tipo de Regra Dependente	Grupo de Domínios de Dados
CreditCardNumber	Descobre os nomes de coluna que contêm a cadeia "ccn", "cr*ca*nu" ou "credit*no*", além de identificar os dados de coluna que correspondem ao formato do número de cartão de crédito de várias organizações de cartão de crédito.	Regra de nome de coluna Regra de dados	Account_Bank, PCI, PII
CreditCard_AMEX	Descobre os dados de coluna que correspondem ao formato do número de cartão de crédito American Express.	Regra de dados	Account_Bank
CreditCard_DinersCard	Descobre os dados de coluna que correspondem ao formato do número de cartão de crédito Diners Club International.	Regra de dados	Account_Bank
CreditCard_Discover Card	Descobre os dados de coluna que correspondem ao formato do número de cartão de crédito Discover.	Regra de dados	Account_Bank
CreditCard_JCB	Descobre os dados de coluna que correspondem ao formato do número de cartão de crédito JCB International.	Regra de dados	Account_Bank
CreditCard_Master Card	Descobre os dados de coluna que correspondem ao formato do número de cartão de crédito MasterCard.	Regra de dados	Account_Bank
CreditCard_Visa	Descobre os dados de coluna que correspondem ao formato do número de cartão de crédito Visa.	Regra de dados	Account_Bank
Croatia_NationalID	Descobre os dados de coluna que correspondem ao ID nacional da Croácia.	Regra de dados	NationalID
Date_AllFormats	Descobre os valores de data nos dados de coluna.	Regra de dados	Geral
Date_MM_DD_YYYY	Descobre as cadeias de data nos dados de origem exibidas em um único formato em uma coluna de dados. O formato padrão é "MM/DD/YYYY".	Regra de dados	Geral
Denmark_NationalID	Descobre os dados de coluna que correspondem ao ID nacional da Dinamarca.	Regra de dados	NationalID

Nome	Descrição	Tipo de Regra Dependente	Grupo de Domínios de Dados
DeviceSerialNumber	Descobre os nomes de coluna que contêm a cadeia "device*number", "device*no*", "serial*number", "serial*no*" ou "device*identi*".	Regra de nome de coluna	PHI
DriverLicense_Canada	Descobre os dados de coluna que correspondem aos números da carteira de motorista do Canadá, exceto para as províncias Columbia Britânica, Quebec, Manitoba e Ilha do Príncipe Eduardo.	Regra de dados	NationalID
DriverLicense_GBR	Descobre os dados de coluna que correspondem aos números da carteira de motorista do Reino Unido.	Regra de dados	NationalID
DriverLicense_USA	Descobre os dados de coluna que correspondem aos números da carteira de motorista da maioria dos estados dos Estados Unidos.	Regra de dados	NationalID
DrivingLicenseNumber	Descobre os nomes de coluna que contêm a cadeia "license" ou "driver*license". Identifica os dados de coluna que correspondem aos números de carteira de motorista no Reino Unido, nos Estados Unidos e em algumas províncias canadenses com base nos requisitos de tamanho e padrão.	Regra de nome de coluna Regra de dados	PII
Email	Descobre os nomes de coluna que contêm a cadeia "email" e identifica os dados de coluna que correspondem a um formato de ID de e-mail predefinido.	Regra de nome de coluna Regra de dados	Contact, PHI
ExpirationDate	Descobre os nomes de coluna que contêm a cadeia "exp*da*" ou "cr*exp*", além de identificar os dados de coluna que correspondem a datas de cartão de crédito expirados.	Regra de nome de coluna Regra de dados	PCI
Finland_NationalID	Descobre os dados de coluna que correspondem ao ID nacional da Finlândia.	Regra de dados	NationalID

Nome	Descrição	Tipo de Regra Dependente	Grupo de Domínios de Dados
FirstName	Descobre os nomes de coluna que contêm a cadeia "f*nam*", além de identificar os dados de coluna que correspondem aos valores em uma tabela de referência de nomes.	Regra de nome de coluna Regra de dados	Contact, PCI, PII
France_INSEE	Descobre os dados de coluna que correspondem ao formato do número do French Institut National de la Statistique et des Études Économiques (INSEE).	Regra de dados	NationalID
Gender	Descobre os nomes de coluna que contêm a cadeia "gender" ou cadeias como "female" e "male", além de identificar os dados de coluna que correspondem aos valores de sexo em uma tabela de referência.	Regra de nome de coluna Regra de dados	Contact, PII
Geocode_Latitude	Descobre os nomes de coluna que contêm a cadeia "latitude", além de identificar os dados de coluna que correspondem às coordenadas de latitude válidas.	Regra de nome de coluna Regra de dados	Address, Geral
Geocode_LatitudeLongitude	Descobre os nomes de coluna que contêm cadeias como "latitude", "longitude" e "geocode", além de identificar os dados de coluna que correspondem às coordenadas de latitude ou longitude válidas.	Regra de nome de coluna Regra de dados	Address, Geral
Geocode_Longitude	Descobre os nomes de coluna que contêm a cadeia "longitude", além de identificar os dados de coluna que correspondem às coordenadas de longitude válidas.	Regra de nome de coluna Regra de dados	Address, Geral
Grade	Descobre os nomes de coluna que contêm a cadeia "grade".	Regra de nome de coluna	PII
GreatBritain_NINO	Descobre os dados de coluna que correspondem ao formato do número de Previdência Social da Grã-Bretanha.	Regra de dados	NationalID
Hostname	Descobre os dados de coluna que correspondem aos nomes de host válidos.	Regra de dados	Geral

Nome	Descrição	Tipo de Regra Dependente	Grupo de Domínios de Dados
IBAN	Descobre os dados de coluna que correspondem ao formato do International Bank Account Number de vários países europeus.	Regra de dados	Account_Bank
India_NationalID	Descobre os dados de coluna que correspondem ao formato do Número de Conta Permanente da Índia.	Regra de dados	NationalID
IPAddress	Descobre os nomes de coluna que contêm a cadeia "ip" ou "inter*port*add", além de identificar os dados de coluna que correspondem a um formato de endereço IP predefinido.	Regra de nome de coluna Regra de dados	PII
ISBN	Descobre os dados de coluna que correspondem ao formato do Número Padrão Internacional de Livro.	Regra de dados	Geral
Italy_FiscalCode	Descobre os dados de coluna que correspondem ao formato do ID nacional da Itália.	Regra de dados	NationalID
JobPosition	Descobre os nomes de coluna que contêm a cadeia "title", "position" ou "designation".	Regra de nome de coluna	PII
Korea_NationalID	Descobre os dados de coluna que correspondem ao ID nacional da Coreia.	Regra de dados	NationalID
LastName	Descobre os nomes de coluna que contêm a cadeia "lname", "su*name" ou "last*name", além de identificar os dados de coluna que correspondem aos valores em uma tabela de referência de sobrenomes.	Regra de nome de coluna Regra de dados	Contact, PCI, PII
Norway_NationalID	Descobre os dados de coluna que correspondem ao ID nacional da Noruega.	Regra de dados	NationalID
Passport_DEU_MR	Descobre os dados de coluna que correspondem ao formato legível por máquina do número de passaporte a Alemanha.	Regra de dados	NationalID
Passport_GBR	Descobre os dados de coluna que correspondem ao formato do número de passaporte do Reino Unido.	Regra de dados	NationalID

Nome	Descrição	Tipo de Regra Dependente	Grupo de Domínios de Dados
Passport_India	Descobre os dados de coluna que correspondem ao formato do número de passaporte da Índia.	Regra de dados	NationalID
Passport_MachineReadable	Descobre os dados de coluna que correspondem aos números de passaporte legíveis por máquina de todos os países.	Regra de dados	NationalID
Passport_USA_MR	Descobre os dados de coluna que correspondem ao formato legível por máquina de número de passaporte dos Estados Unidos.	Regra de dados	NationalID
PhoneNumber	Descobre os nomes de coluna que contêm a cadeia "phone" ou "fax", além de identificar os dados de coluna que correspondem ao formato do número de telefone dos Estados Unidos.	Regra de nome de coluna Regra de dados	Contact, PHI
CEP	Descobre os dados de coluna que correspondem aos códigos postais de vários países.	Regra de dados	Address
Romania_NationalID	Descobre os dados de coluna que correspondem ao ID nacional da Romênia.	Regra de dados	NationalID
Salary	Descobre os nomes de coluna que contêm a cadeia "compensation", "salary" ou "wages".	Regra de nome de coluna	PII
SouthAfrica_NationalID	Descobre os dados de coluna que correspondem ao ID nacional da África do Sul.	Regra de dados	NationalID
SSN	Descobre os nomes de coluna que contêm a cadeia "SSN", "social*sec*no" ou "social*sec*num*", além de identificar os dados de coluna que correspondem ao formato do número de Previdência Social.	Regra de nome de coluna Regra de dados	NationalID, PHI
SSN_General	Descobre as colunas de dados que correspondem ao formato do número de Previdência Social.	Regra de dados	NationalID, PII

Nome	Descrição	Tipo de Regra Dependente	Grupo de Domínios de Dados
State	Descobre os nomes de coluna que contêm a cadeia "add*sta", "state" ou "us*sta*", além de identificar os dados de coluna que correspondem aos nomes de estado dos Estados Unidos.	Regra de nome de coluna Regra de dados	PII
Street	Descobre os nomes de coluna que contêm uma das seguintes cadeias: - street - road - lane - court - avenue - way - blvd - boule*ard	Regra de nome de coluna	PII
Sweden_NationalID	Descobre os dados de coluna que correspondem ao ID nacional da Suécia.	Regra de dados	NationalID
Taiwan_NationalID	Descobre os dados de coluna que correspondem ao ID nacional de Taiwan.	Regra de dados	NationalID
UniquelidentifyingNumber	Descobre os nomes de coluna que contêm a cadeia "unique*iden*number" ou "iden*num".	Regra de nome de coluna	PHI
UPC	Descobre os dados de coluna que correspondem a um Código de produto Universal.	Regra de dados	Geral
URL	Descobre os nomes de coluna que contêm a cadeia "uni*res*loc", "URL" ou "web", além de identificar os dados de coluna que correspondem aos formatos de URL predefinidos.	Regra de nome de coluna Regra de dados	PHI
USZip_5digit	Descobre os dados de coluna que correspondem aos CEPs dos Estados Unidos.	Regra de dados	Address

Nome	Descrição	Tipo de Regra Dependente	Grupo de Domínios de Dados
VehicleRegPlateNumber	Descobre os nomes de coluna que contêm a cadeia "registration", "number*plate", "license*plate" ou "vehicle*registration".	Regra de nome de coluna	PII
ZipCode	Descobre os nomes de coluna que contêm a cadeia "zip" ou "pin", além de identificar os dados de coluna que correspondem aos CEPs dos Estados Unidos.	Regra de nome de coluna Regra de dados	PII

## Regras de nome de coluna no acelerador dos Domínios de Dados

Use as regras de nome de coluna do domínio de dados para identificar colunas de dados com nomes que correspondem à lógica de nome de coluna que as regras definem. Cada regra usa uma ou mais expressões regulares para pesquisar strings comuns que o nome da coluna pode incluir.

Por exemplo, a regra *dataDomain\_MetaDataRule\_BIC\_SWIFTCode* contém uma transformação de Rotulador que pesquisa usando as seguintes expressões regulares:

```
^*[iI][sS][oO].*[9][3][6][2].*$
^*[sS][wW][iI][fF][tT]*[bB][iI][cC]$
^*[bB][iI][cC].*[cC][oO][dD][eE].*$
```

As regras de nome de coluna analisam os caracteres nos nomes de coluna. As regras de nome de coluna não analisam os valores de dados nas colunas.

Você pode localizar as regras de nome de coluna na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Domain_Discovery_MetaData_Rules
```

A seguinte tabela descreve as regras de nome de coluna no acelerador dos Domínios de Dados:

Nome	Descrição
dataDomain_MetaDataRule_ABARoutingNumber	Procura nomes que possam descrever uma coluna de números de roteamento da American Banking Association.
dataDomain_MetaDataRule_AccountNumber	Procura nomes que possam descrever uma coluna de números de conta.
dataDomain_MetaDataRule_AccountStatus	Procura nomes que possam descrever uma coluna de informações de status da conta.
dataDomain_MetaDataRule_Address	Procura nomes que possam descrever uma coluna de informações de endereço.



Nome	Descrição
dataDomain_MetaDataRule_AdmissionDate	Procura nomes que possam descrever uma coluna de informações de data de admissão.
dataDomain_MetaDataRule_Age	Procura nomes que possam descrever uma coluna de idade ou data das informações de nascimento.
dataDomain_MetaDataRule_BankAccount	Procura nomes que possam descrever uma coluna de informações da conta bancária.
dataDomain_MetaDataRule_BIC_SwiftCode	Procura nomes que possam descrever uma coluna de Códigos de Identificador de Negócios. Os Códigos de Identificador de Negócios também são chamados de códigos Swift e códigos ISO 9362.
dataDomain_MetaDataRule_BirthDay	Procura nomes que possam descrever uma coluna de data de nascimento ou de informações de aniversário.
dataDomain_MetaDataRule_BirthPlace	Procura nomes que possam descrever uma coluna de local ou localização das informações de nascimento.
dataDomain_MetaDataRule_CertificateLicenseNumber	Procura nomes que possam descrever uma coluna das informações de número de licença de certificado.
dataDomain_MetaDataRule_City	Procura nomes que possam descrever uma coluna de informações da cidade.
dataDomain_MetaDataRule_CompanyName	Procura nomes que possam descrever uma coluna de informações de nome da empresa.
dataDomain_MetaDataRule_Computer_Address	Procura nomes que possam descrever uma coluna de dados de endereço MAC ou do computador.
dataDomain_MetaDataRule_Country	Procura nomes que possam descrever uma coluna de informações do país, incluindo informações de código de país ISO.
dataDomain_MetaDataRule_CountryCode_Phone	Procura nomes que possam descrever uma coluna de códigos telefônicos de país.
dataDomain_MetaDataRule_County	Procura nomes que possam descrever uma coluna de informações do município.
dataDomain_MetaDataRule_CreditCardNumber	Procura nomes que possam descrever uma coluna de números de cartão de crédito.
dataDomain_MetaDataRule_CreditCardTrack1FormatB	Procura nomes que possam descrever uma coluna de informações de faixa 1 formato b de um cartão de crédito.
dataDomain_MetaDataRule_Currency	Procura nomes que possam descrever uma coluna de informações de moeda.
dataDomain_MetaDataRule_DeviceSerialNumber	Procura nomes que possam descrever uma coluna de informações de número de dispositivo ou de número de série.
dataDomain_MetaDataRule_DischargeDate	Procura nomes que possam descrever uma coluna de informações de data de quitação.

Nome	Descrição
dataDomain_MetaDataRule_DrivingLicenseNumber	Procura nomes que possam descrever uma coluna de informações de carteira de motorista.
dataDomain_MetaDataRule_Email	Procura nomes que possam descrever uma coluna de informações de e-mail.
dataDomain_MetaDataRule_ExpirationDate	Procura nomes que possam descrever uma coluna de informações de data de expiração, por exemplo, informações de data de expiração de cartões de crédito.
dataDomain_MetaDataRule_FirstName	Procura nomes que possam descrever uma coluna de informações de nome.
dataDomain_MetaDataRule_FullName	Procura nomes que possam descrever uma coluna de informações de nome completo.
dataDomain_MetaDataRule_Gender	Procura nomes que possam descrever uma coluna de informações de gênero.
dataDomain_MetaDataRule_Grade	Procura nomes que possam descrever uma coluna de informações de grau.
dataDomain_MetaDataRule_HealthCareBeneficiaryNumber	Procura nomes que possam descrever uma coluna de números de beneficiários de assistência médica.
dataDomain_MetaDataRule_Height	Procura nomes que possam descrever uma coluna de informações de altura.
dataDomain_MetaDataRule_Hostname	Procura nomes que possam descrever uma coluna de informações de nome do host do computador.
dataDomain_MetaDataRule_IBAN	Procura nomes que possam descrever uma coluna de International Bank Account Numbers.
dataDomain_MetaDataRule_ICD10	Procura nomes que possam descrever uma coluna de valores da décima revisão da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde.
dataDomain_MetaDataRule_ICD9	Procura nomes que possam descrever uma coluna de valores da nona revisão da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde.
dataDomain_MetaDataRule_IPAddress	Procura nomes que possam descrever uma coluna de informações de endereço IP do computador.
dataDomain_MetaDataRule_ISBN	Procura nomes que possam descrever uma coluna de International Standard Book Numbers.
dataDomain_MetaDataRule_ITIN_USA	Procura nomes que possam descrever uma coluna de Individual Taxpayer Identification Numbers.
dataDomain_MetaDataRule_JobPosition	Procura nomes que possam descrever uma coluna de informações de cargo, função ou designação.
dataDomain_MetaDataRule_LastName	Procura nomes que possam descrever uma coluna de informações de sobrenome.

Nome	Descrição
dataDomain_MetaDataRule_Latitude	Procura nomes que possam descrever uma coluna de informações de latitude.
dataDomain_MetaDataRule_LatitudeLongitude	Procura nomes que possam descrever uma coluna de informações de latitude, longitude ou coordenadas geográficas.
dataDomain_MetaDataRule_Longitude	Procura nomes que possam descrever uma coluna de informações de longitude.
dataDomain_MetaDataRule_MaidenName	Procura nomes que possam descrever uma coluna de informações de nome de solteira.
dataDomain_MetaDataRule_MiddleName	Procura nomes que possam descrever uma coluna de informações de nome do meio.
dataDomain_MetaDataRule_NationalId	Procura nomes que possam descrever uma coluna de números de identidade nacional.
dataDomain_MetaDataRule_NDC_USA	Procura nomes que possam descrever uma coluna de informações de National Drug Code.
dataDomain_MetaDataRule_NPI_USA	Procura nomes que possam descrever uma coluna de números de National Provider Identifier.
dataDomain_MetaDataRule_Passport	Procura nomes que possam descrever uma coluna de informações de passaporte.
dataDomain_MetaDataRule_PhoneNumber	Procura nomes que possam descrever uma coluna de números de telefone ou números de fax.
dataDomain_MetaDataRule_Quantity	Procura nomes que possam descrever uma coluna de informações de quantidade.
dataDomain_MetaDataRule_Race	Procura nomes que possam descrever uma coluna de informações de etnia ou cor da pele.
dataDomain_MetaDataRule_Religion	Procura nomes que possam descrever uma coluna de informações sobre religião, fé ou crença.
dataDomain_MetaDataRule_Salary	Procura nomes que possam descrever uma coluna de informações de salários, rendimentos ou remunerações.
dataDomain_MetaDataRule_SSN	Procura nomes que possam descrever uma coluna de números de Social Security.
dataDomain_MetaDataRule_State	Procura nomes que possam descrever uma coluna de informações de estado dos Estados Unidos.
dataDomain_MetaDataRule_Street	Procura nomes que possam descrever uma coluna de informações de endereço.
dataDomain_MetaDataRule_UniqueIdentifying Number	Procura nomes que possam descrever uma coluna de números de identificação exclusivos.

Nome	Descrição
dataDomain_MetaDataRule_UPC_EAN	Procura nomes que possam descrever uma coluna de Códigos de Produto Universais ou Números de Artigos Europeus.
dataDomain_MetaDataRule_URL	Procura nomes que possam descrever uma coluna de informações de URL (Uniform Resource Locator) ou endereço da Web.
dataDomain_MetaDataRule_VehicleRegPlate Number	Procura nomes que possam descrever uma coluna de números de certificado de registro do veículo ou placa do veículo.
dataDomain_MetaDataRule_Weight	Procura nomes que possam descrever uma coluna de informações de peso.
dataDomain_MetaDataRule_ZipCode	Procura nomes que possam descrever uma coluna de CEPs.

## Regras de dados no acelerador dos Domínios de Dados

Use as regras de dados do domínio de dados para identificar as colunas que contêm dados que correspondem aos critérios da regra.

As regras de dados se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Domain_Discovery\Data_Rules
```

A seguinte tabela descreve as regras de dados no acelerador dos Domínios de Dados:

Nome	Descrição
dataDomain_DataRule_ABARoutingNumber	Identifica os dados de coluna que correspondem ao formato de um número de roteamento da American Banking Association. O número de roteamento identifica uma instituição financeira em uma transação financeira.
dataDomain_DataRule_Account_Status	Identifica os dados de coluna que correspondem aos valores de status de conta nos dados de referência.
dataDomain_DataRule_Address_Data	Identifica dados de coluna que representam as informações de endereço. A regra reconhece globalmente os dados de endereço de vários países.
dataDomain_DataRule_Age	Identifica os dados de coluna com valores de 1 a 120.
dataDomain_DataRule_Alphanumeric_Special Character	Identifica os dados de coluna que contêm dados alfanuméricos não formatados e dados de caracteres especiais.
dataDomain_DataRule_Amount	Identifica os dados de coluna que representam uma quantidade física.
dataDomain_DataRule_AUT_NATID	Identifica os dados de coluna que correspondem ao formato do ID nacional da Áustria.

Nome	Descrição
dataDomain_DataRule_BankAccount_USA	Identifica os dados de coluna que correspondem a um formato de número de conta bancária dos Estados Unidos.
dataDomain_DataRule_BGR_NATID	Identifica os dados de coluna que correspondem ao formato do ID nacional da Bulgária.
dataDomain_DataRule_BIC_SWIFTCode	Identifica os dados de coluna que correspondem ao código de Bank Identifier Code (BIC) ou Society for Worldwide Interbank Financial Telecommunication (SWIFT) pelo reconhecimento de padrão e pelo código de país.
dataDomain_DataRule_BinaryValues	Identifica os dados de coluna que contêm valores binários.
dataDomain_DataRule_BirthDay	Identifica os dados de coluna que correspondem às datas de nascimento válidas. A regra verifica o número de anos entre a data de entrada e a data atual. A regra retorna "Adulto", "Menor" ou "Válido" com base nos valores de 1 a 120. A regra retorna "Inválido" para todos os outros valores.
dataDomain_DataRule_BRA_IDDoc	Identifica os dados de coluna que correspondem ao formato de número do cartão de identificação de <i>Registro Geral</i> no Brasil.
dataDomain_DataRule_BRA_Personal_ID	Identifica os dados de coluna que correspondem ao formato do ID pessoal do Brasil.
dataDomain_DataRule_CAN_SIN	Identifica os dados de coluna que correspondem ao formato de número de Social Insurance do Canadá.
dataDomain_DataRule_CHN_NATID	Identifica os dados de coluna que correspondem ao formato do ID nacional da China.
dataDomain_DataRule_City	Identifica os dados de coluna que contêm um nome de cidade válido. A regra lê os dados de referência que contêm nomes de cidades internacionais.
dataDomain_DataRule_CompanyName	Identifica os dados de coluna que correspondem aos valores de nome da organização nos dados de referência.
dataDomain_DataRule_Computer_Address	Identifica os dados de coluna que correspondem ao formato de endereços IP e MAC.
dataDomain_DataRule_Country	Identifica os dados de coluna que correspondem a um nome de país ISO.
dataDomain_DataRule_CountryCode_Phone	Identifica os dados de coluna que correspondem aos números de telefone baseados em códigos de discagem internacionais.
dataDomain_DataRule_County	Identifica os dados de coluna que correspondem a um nome de município dos Estados Unidos.
dataDomain_DataRule_CreditCard_AMEX	Identifica os dados de coluna que correspondem ao formato do número de cartão de crédito American Express.
dataDomain_DataRule_CreditCard_DinersCard	Identifica os dados de coluna que correspondem ao formato do número de cartão de crédito Diners Club International.

Nome	Descrição
dataDomain_DataRule_CreditCard_DiscoverCard	Identifica os dados de coluna que correspondem ao formato do número de cartão de crédito Discover.
dataDomain_DataRule_CreditCard_JCB	Identifica os dados de coluna que correspondem ao formato do número de cartão de crédito JCB International.
dataDomain_DataRule_CreditCard_MasterCard	Identifica os dados de coluna que correspondem ao formato do número de cartão de crédito MasterCard.
dataDomain_DataRule_CreditCard_Visa	Identifica os dados de coluna que correspondem ao formato do número de cartão de crédito Visa.
dataDomain_DataRule_CreditCardNumber	Identifica os dados de coluna que correspondem ao formato do número de cartão de crédito das maiores organizações de cartão de crédito, como a American Express, a Diners Club International e a Maestro.
dataDomain_DataRule_CreditCardTrack1FormatB	Identifica os dados de coluna que correspondem às informações de cartão de crédito de Faixa 1 Formato B.
dataDomain_DataRule_Currency	Identifica os dados de coluna que correspondem a um termo de moeda nos dados de referência.
dataDomain_DataRule_Date_Validation	Identifica as cadeias de data nos dados de origem exibidas em um único formato em uma coluna de dados. Para configurar o formato de data que a regra usa para validação, abra a transformação de Expressão dq_ValidateDate na regra e atualize a variável de expressão In_Date_Format. O formato padrão é "MM/DD/YYYY". A regra retorna "Válido" ou "Inválido".
dataDomain_DataRule_Date_Validation_All_Formats	Identifica os valores de data nos dados de coluna e padroniza os dados de coluna em um formato de data específico.
dataDomain_DataRule_DEU_Machine_Readable_Passport	Identifica os dados de coluna que correspondem ao formato legível por máquina do número de passaporte da Alemanha.
dataDomain_DataRule_DNK_NATID	Identifica os dados de coluna que correspondem ao formato do ID nacional da Dinamarca.
dataDomain_DataRule_DriversLicense	Identifica os dados de coluna que correspondem aos números de carteira de motorista do Canadá, do Reino Unido e dos Estados Unidos com base no comprimento e no padrão dos valores de dados.
dataDomain_DataRule_DriversLicense_Canada	Identifica os dados de coluna que correspondem aos números da carteira de motorista do Canadá, exceto para as províncias Columbia Britânica, Quebec, Manitoba e Ilha do Príncipe Eduardo.
dataDomain_DataRule_DriversLicense_Canada_narrow	Identifica os dados de coluna que correspondem aos números da carteira de motorista do Canadá, exceto para as províncias Columbia Britânica, Quebec, Manitoba e Ilha do Príncipe Eduardo. A regra é semelhante à regra dataDomain_DataRule_DriversLicense_Canada. No entanto, dataDomain_DataRule_DriversLicense_Canada_narrow executa uma análise mais restrita para reduzir a probabilidade de falsos positivos.

Nome	Descrição
dataDomain_DataRule_DriversLicense_GBR	Identifica os dados de coluna que correspondem aos números de carteira de motorista do Reino Unido.
dataDomain_DataRule_DriversLicense_narrow	Identifica os dados de coluna que correspondem aos números de carteira de motorista do Reino Unido e de muitos estados e províncias do Canadá e dos Estados Unidos. A regra não valida os números das províncias de Columbia Britânica, Quebec, Manitoba e Ilha do Príncipe Eduardo. Para reduzir a probabilidade de falsos positivos, a regra não valida números que contêm entre quatro e oito dígitos.
dataDomain_DataRule_DriversLicense_USA	Identifica os dados de coluna que correspondem aos números da carteira de motorista da maioria dos estados dos Estados Unidos.
dataDomain_DataRule_DriversLicense_USA_narrow	Identifica os dados de coluna que correspondem aos números da carteira de motorista da maioria dos estados dos Estados Unidos. Para reduzir a probabilidade de falsos positivos, a regra exclui os valores de dados que contêm entre seis e oito dígitos. Por exemplo, a regra exclui um valor como 01012017.
dataDomain_DataRule_Email	Identifica os dados de coluna que correspondem a um formato de ID de e-mail predefinido.
dataDomain_DataRule_ExpirationDate	Identifica os dados de coluna que correspondem aos dados de cartão de crédito expirados. A regra compara a data de entrada com a data do sistema para realizar a validação.
dataDomain_DataRule_FIN_NATID	Identifica os dados de coluna que correspondem ao formato do ID nacional da Finlândia.
dataDomain_DataRule_FirstName	Identifica os dados de coluna que correspondem aos valores em um conjunto de dados de referência de nomes.
dataDomain_DataRule_FRA_INSEE	Identifica os dados de coluna que correspondem ao formato do número do French Institut National de la Statistique et des Études Économiques (INSEE).
dataDomain_DataRule_FullName	Identifica as strings em uma coluna de dados que contêm nomes, nomes do meio e sobrenomes. A regra compara as palavras em cada string com os dados de referência.
dataDomain_DataRule_GBR_NINO	Identifica os dados de coluna que correspondem ao formato do número da Previdência Social do Reino Unido.
dataDomain_DataRule_GBR_Passport_Number	Identifica os dados de coluna que correspondem ao formato do número de passaporte do Reino Unido.
dataDomain_DataRule_Gender	Identifica os dados de coluna que correspondem aos valores de gênero nos dados de referência.
dataDomain_DataRule_Height	Identifica os dados de coluna com os valores de 1 a 8, onde 8 representa a altura em pés.
dataDomain_DataRule_HostName	Identifica os dados de coluna que correspondem aos nomes de host válidos.

Nome	Descrição
dataDomain_DataRule_HRV_NATID	Identifica os dados de coluna que correspondem ao formato do ID nacional da Croácia.
dataDomain_DataRule_IBAN	Identifica os dados de coluna que correspondem ao formato do International Bank Account Number de vários países europeus.
dataDomain_DataRule_ICD_10	Identifica os dados de coluna que correspondem aos nomes das condições na décima revisão da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde (ICD). A Organização Mundial da Saúde (OMS) mantém a classificação.
dataDomain_DataRule_ICD_9	Identifica os dados de coluna que correspondem aos nomes das condições na nona revisão da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde (ICD). A Organização Mundial da Saúde (OMS) mantém a classificação.
dataDomain_DataRule_IND_NATID	Identifica os dados de coluna que correspondem ao formato do Número de Conta Permanente da Índia.
dataDomain_DataRule_IND_Passport	Identifica os dados de coluna que correspondem ao formato do número de passaporte da Índia.
dataDomain_DataRule_IPAddress	Identifica os dados de coluna que correspondem ao formato de endereço IP predefinido.
dataDomain_DataRule_ISBN	Identifica os dados de coluna que correspondem ao formato do International Standard Book Number.
dataDomain_DataRule_ISIN	Identifica os dados de coluna que correspondem ao International Securities Identification Number (ISIN). Um ISIN identifica uma segurança como uma ação ou um título.
dataDomain_DataRule_ItalyFiscalCode	Identifica os dados de coluna que correspondem ao formato do ID nacional da Itália.
dataDomain_DataRule_ITIN_USA	Identifica os dados de coluna que correspondem ao formato de um Individual Taxpayer Identification Number (ITIN) dos Estados Unidos. O Internal Revenue Service emite os números de identificação.
dataDomain_DataRule_JobPosition	Identifica os dados de coluna que correspondem aos nomes de cargo nos dados de referência.
dataDomain_DataRule_KOR_NATID	Identifica os dados de coluna que correspondem ao formato do ID nacional da Coreia.
dataDomain_DataRule_LastName	Identifica os dados de coluna que correspondem aos valores em um conjunto de dados de referência de sobrenomes.
dataDomain_DataRule_Latitude	Identifica os dados de coluna que correspondem às coordenadas de latitude válidas.
dataDomain_DataRule_LatitudeLongitude	Identifica os dados de coluna que correspondem aos pares válidos das coordenadas de latitude e longitude. Cada par é separado por um ponto e vírgula.



Nome	Descrição
dataDomain_DataRule_Longitude	Identifica os dados de coluna que correspondem às coordenadas de longitude válidas.
dataDomain_DataRule_Machine_Readable_Passport	Identifica os dados de coluna que correspondem a números de passaportes legíveis por máquina de todos os países.
dataDomain_DataRule_NDC_USA	Identifica os dados de coluna que correspondem a um valor de National Drug Code (NDC) no diretório de National Drug Code dos Estados Unidos. Cada código identifica exclusivamente uma droga que um fabricante desenvolveu para uso humano.
dataDomain_DataRule_NOR_NATID	Identifica os dados de coluna que correspondem ao formato do ID nacional da Noruega.
dataDomain_DataRule_NPI_USA	Identifica os dados de coluna que correspondem a um número do National Provider Identifier (NPI) nos Estados Unidos. Os centros dos serviços Medicare e Medicaid emitem os números para os fornecedores de assistência médica.
dataDomain_DataRule_PhoneNumber	Identifica os dados de coluna que correspondem ao formato do número de telefone dos Estados Unidos.
dataDomain_DataRule_PostCode	Identifica os dados de coluna que correspondem aos CEPs de vários países.
dataDomain_DataRule_Quantity	Identifica os dados de coluna que descrevem uma quantidade física e inclui unidades de medida.
dataDomain_DataRule_Race	Identifica os dados de coluna que correspondem ao nome de uma raça de pessoas nos dados de referência.
dataDomain_DataRule_Religion	Identifica os dados de coluna que correspondem ao nome de uma religião nos dados de referência.
dataDomain_DataRule_ROU_NATID	Identifica os dados de coluna que correspondem ao formato do ID nacional da Romênia.
dataDomain_DataRule_SouthAfrica_NATID	Identifica os dados de coluna que correspondem ao formato do ID nacional da África do Sul.
dataDomain_DataRule_Spanish_NIF	Identifica os dados de coluna que correspondem ao formato do número de identificação fiscal (NIF) da Espanha.
dataDomain_DataRule_SSN	Identifica os dados de coluna que correspondem ao formato de número de Social Security dos Estados Unidos.
dataDomain_DataRule_State	Identifica os dados de coluna que correspondem aos nomes de estados dos Estados Unidos.
dataDomain_DataRule_Street	Identifica as strings nos dados de coluna que descrevem as informações de endereço, por exemplo, rua, estrada, avenida. A regra usa uma expressão regular para localizar descritores de rua nos dados da coluna.
dataDomain_DataRule_SWE_NATID	Identifica os dados de coluna que correspondem ao formato do ID nacional da Suécia.

Nome	Descrição
dataDomain_DataRule_TWN_NATID	Identifica os dados de coluna que correspondem ao formato do ID nacional de Taiwan.
dataDomain_DataRule_UPC	Identifica os dados de coluna que correspondem a um Código de Produto Universal válido. Um Código de Produto Universal é um tipo de código de barras.
dataDomain_DataRule_UPC_EAN	Identifica os dados de coluna que correspondem a um Código de Produto Universal ou Número de Artigo Europeu válido. Os Códigos de Produto Universais e Números de Artigos Europeus são tipos de código de barras.
dataDomain_DataRule_URL	Identifica os dados de coluna que correspondem aos formatos de URL predefinidos.
dataDomain_DataRule_US_Zip5	Identifica os dados de coluna que correspondem aos ZIP Codes dos Estados Unidos.
dataDomain_DataRule_USA_Machine_Readable_Passport	Identifica os dados de coluna que correspondem a um formato legível por máquina do número de passaporte dos Estados Unidos.
dataDomain_DataRule_USA_SSN_post_2011June	Identifica os dados de coluna que correspondem ao formato do número de Social Security em tamanho, valores numéricos e valores mínimo e máximo das seções de área, grupo e número de série. Com base na iniciativa de Aleatorização de SSN em vigor desde 25 de junho de 2011, a regra não verifica a emissão de um número de Social Security e a combinação do grupo e do número de área.
dataDomain_DataRule_Weight	Identifica os dados de coluna que descrevem um valor de peso. A regra verifica se há um número entre 0 e 500.
dataDomain_DataRule_ZipCode	Identifica os dados de coluna que correspondem aos ZIP Codes dos Estados Unidos.

## CAPÍTULO 4

# Acelerador da Austrália/Nova Zelândia

Este capítulo inclui os seguintes tópicos:

- [Visão Geral do Acelerador da Austrália/Nova Zelândia, 51](#)
- [Regras de Limpeza de Dados de Endereço da Austrália/Nova Zelândia, 52](#)
- [Regras de Composição da Austrália/Nova Zelândia, 53](#)
- [Regras de Limpeza de Dados de Contato da Austrália/Nova Zelândia, 55](#)
- [Regras de Limpeza de Dados Corporativos da Austrália/Nova Zelândia, 58](#)
- [Regras de Limpeza de Dados Gerais da Austrália/Nova Zelândia, 59](#)
- [Regras de Eliminação de Duplicação e Correspondência da Austrália/Nova Zelândia, 60](#)
- [Mapeamentos Demonstrativos da Austrália/Nova Zelândia, 62](#)

## Visão Geral do Acelerador da Austrália/Nova Zelândia

Use as regras no acelerador da Austrália/Nova Zelândia para verificar e aprimorar os dados de organizações da Austrália/Nova Zelândia.

O acelerador da Austrália/Nova Zelândia inclui regras que executam as seguintes operações de qualidade de dados:

- Limpeza de dados de endereço
- Limpeza de dados de contato
- Limpeza de dados corporativos
- Limpeza de dados gerais
- Correspondência e eliminação de duplicação

O acelerador da Austrália/Nova Zelândia também inclui regras de composição. Uma regra de composição combina várias regras em um único objeto.

O acelerador depende de regras instaladas pelo Acelerador Principal.

# Regras de Limpeza de Dados de Endereço da Austrália/Nova Zelândia

Use as regras de limpeza de dados de endereço para analisar, padronizar e validar dados de endereço.

As regras de limpeza de dados de endereço se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados de endereço no acelerador da Austrália/Nova Zelândia:

Nome	Descrição
rule_AUS_Address_Parse_Hybrid	Analisa os endereços australianos não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_AUS_Address_Parse_Multiline	Analisa os endereços australianos não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_AUS_Address_Validation_Distinct	Valida a capacidade de entrega de endereços australianos. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_AUS_Address_Validation_Distinct_w_Geocoding	Valida a capacidade de entrega de endereços australianos e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_AUS_Address_Validation_Hybrid	Valida a capacidade de entrega de endereços australianos. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_AUS_Address_Validation_Hybrid_w_Geocoding	Valida a capacidade de entrega de endereços australianos e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_AUS_Address_Validation_Multiline	Valida a capacidade de entrega de endereços australianos. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_AUS_Address_Validation_Multiline_w_Geocoding	Valida a capacidade de entrega de endereços australianos e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.

Nome	Descrição
rule_NZL_Address_Parse_Hybrid	Analisa os endereços neozelandeses não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_NZL_Address_Parse_Multiline	Analisa os endereços neozelandeses não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_NZL_Address_Validation_Distinct	Valida a capacidade de entrega de endereços neozelandeses. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_NZL_Address_Validation_Distinct_w_Geocoding	Valida a capacidade de entrega de endereços neozelandeses e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_NZL_Address_Validation_Hybrid	Valida a capacidade de entrega de endereços neozelandeses. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_NZL_Address_Validation_Hybrid_w_Geocoding	Valida a capacidade de entrega de endereços neozelandeses e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_NZL_Address_Validation_Multiline	Valida a capacidade de entrega de endereços neozelandeses. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_NZL_Address_Validation_Multiline_w_Geocoding	Valida a capacidade de entrega de endereços neozelandeses e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.

## Regras de Composição da Austrália/Nova Zelândia

Use as regras de composição no acelerador da Austrália/Nova Zelândia para adicionar um conjunto de regras a um mapeamento como um único objeto. Uma regra de composição é uma regra que faz uso da lógica de outras regras do acelerador.

As regras de composição se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```

A seguinte tabela descreve as regras de composição no acelerador da Austrália/Nova Zelândia:

Nome	Descrição
rule_AUS_Contact_Data	Analisa, padroniza e valida os dados de contato da Austrália, como endereços, números de telefone e Tax File Numbers.
rule_NZL_Contact_Data	Analisa, padroniza e valida os dados de contato da Nova Zelândia, como endereços, números de telefone e números do Inland Revenue Department (IRD).

### Regra de Composição para Dados de Contato da Austrália

A regra `rule_AUS_Contact_Data` lê mapplets de várias pastas no repositório. A regra também inclui uma transformação não reutilizável.

A seguinte tabela lista os nomes e as localizações de repositório das regras e da transformação em `rule_AUS_Contact_Data`:

Regra	Localização
Case_Converter	Transformação não reutilizável
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_AUS_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_AUS_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_AUS_Gender_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_AUS_Multi_Person_Name_Parse	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_AUS_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_AUS_Phone_Number_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_AUS_Tax_File_Number_Standardization	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_AUS_Tax_File_Number_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_Prenome_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_Salutation_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing

### Regra de Composição para Dados de Contato da Nova Zelândia

A regra `rule_NZL_Contact_Data` lê mapplets de várias pastas no repositório. A regra também inclui uma transformação não reutilizável.

A seguinte tabela lista os nomes e as localizações de repositório das regras e da transformação em `rule_NZL_Contact_Data`:

Regra	Localização
Case_Converter	Transformação não reutilizável
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_AUS_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_AUS_Multi_Person_Name_Parse	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_NZL_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_NZL_Gender_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_NZL_IRD_Number_Standardization	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_NZL_IRD_Number_Validate	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_NZL_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_NZL_Phone_Number_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_Prenome_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_Salutation_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing

## Regras de Limpeza de Dados de Contato da Austrália/Nova Zelândia

Use as regras de limpeza de dados de contato para analisar, padronizar e validar dados sobre contatos comerciais e indivíduos.

As regras de limpeza de dados de contato se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados de contato no acelerador da Austrália/Nova Zelândia:

Nome	Descrição
rule_AUS_Driver_Licence_Number_Validation	Valida números de carteira de habilitação da Austrália com base em requisitos de tamanho e padrão.
rule_AUS_Gender_Assignment	Atribui o sexo de acordo com os nomes. A regra retorna "M" para nomes masculinos, "F" para nomes feminino e "U" quando o sexo é desconhecido. Por exemplo, a regra atribui a "John Smith" o sexo "M" de masculino.
rule_AUS_Given_Name_Standard	Gera determinados nomes com base em nomes australianos.
rule_AUS_Multi_Person_Name_Parse	<p>Analisa os valores de nome de pessoa em portas diferentes. A regra cria portas para valores como título, nome, nome do meio e sobrenome.</p> <p>A saída da regra inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p> <p>Quando os dados de nome identificam mais de uma pessoa, a regra cria uma porta de saída para cada nome completo. Por exemplo, a regra pode ler o nome "John and Jane Smith" e criar portas de saída para "John Smith" e "Jane Smith".</p>
rule_AUS_Personal_Name_Parsing_FML	<p>Analisa os valores em um nome de pessoa em portas diferentes.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Nome, nome do meio e sobrenome</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p>
rule_AUS_Personal_Name_Parsing_LFM	<p>Analisa os valores em um nome de pessoa em portas diferentes.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Sobrenome, nome e nome do meio</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p>
rule_AUS_Phone_Number_Parse	<p>Analisa um número de telefone australiano em uma cadeia. A regra analisa o primeiro número de telefone dos dados, lendo da direita para a esquerda.</p> <p>A regra reconhece números de telefone que usam zeros à esquerda, códigos de discagem internacional ou extensões que começam com o símbolo de jogo da velha. A regra processa os seguintes símbolos de pontuação: o sinal de adição, os parênteses e o símbolo de jogo da velha. Antes de executar a regra, remova todas as outras pontuações, incluindo os espaços duplos.</p> <p>A regra retorna um número de telefone e também retorna uma string que contém o texto de entrada com o número de telefone removido.</p>
rule_AUS_Phone_Number_Standardization	Padroniza os números de telefone australianos para formatos de discagem internacional e local. A regra reconhece números de telefone que usam zeros à esquerda, códigos de discagem internacional ou extensões que começam com o símbolo de jogo da velha.



Nome	Descrição
rule_AUS_Phone_Number_Validation	Valida o código de área e o tamanho dos números de telefone australianos. A regra retorna a região do número de telefone, assim como códigos que indicam se o código de área e o tamanho de um número de telefone são válidos.
rule_AUS_Tax_File_Number_Parse	Analisa os Tax File Numbers (TFN) australianos.
rule_AUS_Tax_File_Number_Standardization	Padroniza os Tax File Numbers (TFN) australianos. Para configurar o formato padronizado, edite a variável de expressão TFN_Format da transformação de Expressão dq_Format_TFN. O padrão é "No_Punctuation".
rule_AUS_Tax_File_Number_Validation	Valida Tax File Numbers (TFN) australianos com base no dígito de verificação em cada número.
rule_NZL_Gender_Assignment	Atribui o sexo de acordo com os nomes da Nova Zelândia. A regra retorna "M" para nomes masculinos, "F" para nomes feminino e "U" quando o sexo é desconhecido. Por exemplo, a regra atribui a "John Smith" o sexo "M" de masculino.
rule_NZL_Given_Name_Standard	Gera determinados nomes com base em nomes da Nova Zelândia.
rule_NZL_IRD_Number_Parse	Analisa cadeias numéricas de nove dígitos como números do Inland Revenue Department (IRD) da Nova Zelândia.
rule_NZL_IRD_Number_Standardization	Padroniza números do Inland Revenue Department (IRD) da Nova Zelândia. Para configurar o formato padronizado, edite a variável de expressão IRD_Format da transformação de Expressão dq_Format_IRD. O padrão é "No_Punctuation". A regra requer que a entrada seja uma cadeia de nove dígitos.
rule_NZL_IRD_Number_Validate	Valida os números do Inland Revenue Department (IRD) da Nova Zelândia com base no dígito de verificação em cada número.
rule_NZL_Phone_Number_Parse	Analisa um número de telefone da Nova Zelândia em uma cadeia. A regra analisa o primeiro número de telefone dos dados, lendo da direita para a esquerda. A regra reconhece números de telefone que usam zeros à esquerda, códigos de discagem internacional ou extensões que começam com o símbolo de jogo da velha. A regra processa os seguintes símbolos de pontuação: o sinal de adição, os parênteses e o símbolo de jogo da velha. Antes de executar a regra, remova todas as outras pontuações, incluindo os espaços duplos. A regra retorna um número de telefone e também retorna uma string que contém o texto de entrada com o número de telefone removido.
rule_NZL_Phone_Number_Standardization	Padroniza os números de telefone da Nova Zelândia para formatos de discagem internacional e local. A regra reconhece números de telefone que usam zeros à esquerda, códigos de discagem internacional ou extensões que começam com o símbolo de jogo da velha.
rule_NZL_Phone_Number_Validation	Valida o código de área e o tamanho dos números de telefone da Nova Zelândia. A regra retorna a região do número de telefone, assim como códigos que indicam se o código de área e o tamanho de um número de telefone são válidos.

Nome	Descrição
rule_Prename_Assignment	Gera um título de acordo com o sexo. Você pode alterar a variável de expressão female_prename de "Sra." para "Srta.".
rule_Salutation_Assignment	Gera saudações formais e casuais com base em nomes e tokens de nome. Por exemplo, quando os dados de entrada contêm "Sr. John Smith", a regra gera a saudação formal "Prezado Sr. Smith," e a saudação casual "Prezado John,". É possível alterar o prefixo e a pontuação editando as variáveis na transformação de Expressão dq_Generate_Salutation.

### Dependências de Regras de Limpeza de Dados de Contato Principais

O acelerador da Austrália/Nova Zelândia depende das seguintes regras de limpeza de dados de contato do acelerador Principal:

- rule\_Email\_Validation

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados do Contato Principal” na página 22.](#)

## Regras de Limpeza de Dados Corporativos da Austrália/Nova Zelândia

Use as regras de limpeza de dados corporativos para analisar, padronizar e validar dados corporativos.

As regras de limpeza de dados corporativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados corporativos no acelerador da Austrália/Nova Zelândia:

Nome	Descrição
rule_AUS_Business_Number_Parse	Analisa cadeias numéricas de 11 dígitos como Australian Business Numbers (ABN).
rule_AUS_Business_Number_Standardize	Padroniza Australian Business Numbers (ABN) para o formato NN NNN NNN NNN. A regra requer que a entrada seja uma cadeia de 11 dígitos.
rule_AUS_Business_Number_Validation	Valida Australian Business Numbers (ABN) com base no dígito de verificação em cada número.
rule_AUS_Company_Name_Standardization	Padroniza os nomes de empresa para os valores da tabela de referência australiana.

# Regras de Limpeza de Dados Gerais da Austrália/Nova Zelândia

Use as regras de limpeza de dados gerais para identificar os tipos de informações contidas nos campos de entrada.

As regras de limpeza de dados gerais se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados gerais no acelerador da Austrália/Nova Zelândia:

Nome	Descrição
rule_AUS_NZL_NER_Field_Identification	Identifica o tipo de informações contidas em um campo de entrada. A regra pode identificar nomes, IDs Pessoais, nomes de empresa, datas e dados de endereço da Austrália e da Nova Zelândia. A regra retorna um rótulo que descreve o tipo de dados de entrada. A regra usa técnicas de correspondência probabilística para identificar os tipos de informações.

## Dependências de Regras de Limpeza de Dados Gerais Principais

O acelerador da Austrália/Nova Zelândia depende das seguintes regras de limpeza de dados gerais do acelerador Principal:

- rule\_Assign\_DQ\_GeocodingStatus\_Description
- rule\_Assign\_DQ\_Mailability\_Score\_Description
- rule\_Assign\_DQ\_Match\_Code\_Description
- rule\_Remove\_Extra\_Spaces
- rule\_Remove\_Hyphen
- rule\_Remove\_Leading\_Zero
- rule\_Remove\_Period\_Parentheses
- rule\_Remove\_Punctuation
- rule\_Remove\_Punctuation\_and\_Space
- rule\_Remove\_Space
- rule\_Replace\_Limited\_Punct\_with\_Space
- rule\_UpperCase

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados Gerais Principais” na página 23](#).

# Regras de Eliminação de Duplicação e Correspondência da Austrália/Nova Zelândia

Use as regras de correspondência e de eliminação de duplicação no acelerador da Austrália/Nova Zelândia para medir os níveis de semelhança entre os registros em conjuntos de dados.

As regras de correspondência e de eliminação de duplicação se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

A tabela a seguir descreve as regras de correspondência e de eliminação de duplicação no acelerador da Austrália/Nova Zelândia:

Nome	Descrição
mplt_AUS_Firstname_and_TFN_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados australianas com base em Tax File Numbers (TFN) e em nomes. O mapplet gera chaves de grupo com base em dados de TFN.
mplt_AUS_IMO_Company_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados australianos com base em nomes e endereços de empresas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_AUS_IMO_Familyname_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados australianos com base em nomes e endereços de família. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_AUS_IMO_Individual_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados australianos com base em nomes e endereços de pessoas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_AUS_IMO_Personal_Name_and_Data_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados australianos com base em nomes de pessoas e em dados pessoais. Os campos da coluna de dados pessoais devem conter um único tipo de dados, como o número de telefone, o e-mail ou o Tax File Number. O mapplet gera chaves de grupo com base em dados pessoais.
mplt_AUS_Individual_Name_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas e em dados de endereços australianos. O mapplet usa uma combinação de caracteres dos valores de sobrenome e dos valores de código postal para gerar chaves de grupo.
mplt_AUS_Individual_Name_and_Date_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas australianas e em datas. O mapplet gera chaves de grupo com base em dados de data.
mplt_AUS_Individual_Name_and_Email_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em endereços de e-mail e em nomes de pessoas australianas. O mapplet gera chaves de grupo com base em dados de endereço de e-mail.
mplt_AUS_Individual_Name_and_Phone_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas australianas e em números de telefone. O mapplet gera chaves de grupo dos dados de número de telefone.
mplt_AUS_Individual_Name_and_TFN_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados australianas com base em Tax File Numbers (TFN) e em nomes de pessoas. O mapplet gera chaves de grupo com base em dados de TFN.

Nome	Descrição
mplt_AUS_Individual_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas australianas. O mapplet gera códigos NYSIIS dos valores de sobrenome e usa os códigos NYSIIS como chaves de grupo.
mplt_AUS_NZL_Company_Name_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de empresa e em dados de endereço da Austrália e da Nova Zelândia. O mapplet usa uma combinação de caracteres dos valores de nome de empresa e dos valores de código postal para gerar chaves de grupo.
mplt_AUS_NZL_Familyname_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em sobrenomes e em dados de endereço da Austrália e da Nova Zelândia. O mapplet usa uma combinação de caracteres dos valores de sobrenome e dos valores de código postal para gerar chaves de grupo.
mplt_Company_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base nos nomes de empresa. O mapplet gera os códigos Soundex dos valores de nome de empresa e usa os códigos Soundex como chaves de grupo.
mplt_NZL_Firstname_and_IRD_Match	Usa estratégias de correspondência de campo para identificar as linhas duplicadas de dados da Nova Zelândia com base em números do Inland Revenue Department (IRD) e em nomes. O mapplet gera chaves de grupo com base em números do IRD.
mplt_NZL_IMO_Company_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados da Nova Zelândia com base em nomes e endereços de empresas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_NZL_IMO_Familyname_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados da Nova Zelândia com base em nomes e endereços de família. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_NZL_IMO_Individual_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados da Nova Zelândia com base em nomes e endereços de pessoas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_NZL_IMO_Personal_Name_and_Data_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados da Nova Zelândia com base em nomes de pessoas e em dados pessoais. Os campos da coluna de dados pessoais devem conter um único tipo de dados, como o número de telefone, o e-mail ou o número do Inland Revenue Department. O mapplet gera chaves de grupo com base em dados pessoais.
mplt_NZL_Individual_Name_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas e em endereços neozelandeses. O mapplet usa uma combinação de caracteres dos valores de sobrenome e dos valores de código postal para gerar chaves de grupo.
mplt_NZL_Individual_Name_and_Date_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas neozelandesas e em datas. O mapplet gera chaves de grupo com base em dados de data.
mplt_NZL_Individual_Name_and_Email_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em endereços de e-mail e em nomes de pessoas neozelandesas. O mapplet gera chaves de grupo com base em dados de endereço de e-mail.

Nome	Descrição
mplt_NZL_Individual_Name_and_IRD_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas neozelandesas e em números do Inland Revenue Department (IRD). O mapplet gera chaves de grupo com base em números do IRD.
mplt_NZL_Individual_Name_and_Phone_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas neozelandesas e em números de telefone. O mapplet gera chaves de grupo dos dados de número de telefone.
mplt_NZL_Individual_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas neozelandesas. O mapplet gera códigos NYSIIS dos valores de sobrenome e usa os códigos NYSIIS como chaves de grupo.
rule_AUS_NZL_Company_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes de empresa e em endereços da Austrália e da Nova Zelândia.
rule_AUS_NZL_Famlyname_and_Address_MatchScore	Gera uma pontuação de correspondência com base em sobrenomes e em endereços da Austrália e da Nova Zelândia.
rule_AUS_NZL_Firstname_and_PID_MatchScore	Gera uma pontuação de correspondência com base em nomes e em números de identificação pessoal.
rule_AUS_NZL_Individual_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes de pessoas e em endereços da Austrália e da Nova Zelândia.
rule_AUS_NZL_Individual_Name_and_PID_MatchScore	Gera uma pontuação de correspondência com base em nomes de pessoas e em números de identificação pessoal.
rule_Company_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de empresa.
rule_Individual_Name_and_Date_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nas datas.
rule_Individual_Name_and_Email_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos endereços de e-mail.
rule_Individual_Name_and_Phone_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos números de telefone.
rule_Individual_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas.

## Mapeamentos Demonstrativos da Austrália/Nova Zelândia

Os mapeamentos demonstrativos do acelerador da Austrália/Nova Zelândia usam várias regras para demonstrar processos de qualidade de dados.

Os mapeamentos demonstrativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules_Demo\AUS_NZL_Accelerator
```

O acelerador contém os seguintes mapeamentos demonstrativos:

**m\_AUS\_customer\_data\_demo**

Analisa, padroniza e valida dados da Austrália/Nova Zelândia.

**m\_AUS\_customer\_matching\_demo**

Analisa e padroniza os dados de identidade da Austrália e da Nova Zelândia e realiza a análise de correspondência de identidade nos dados.

O mapeamento analisa as seguintes combinações de dados e gera clusters de correspondência para cada combinação:

- Dados de nome e endereço da pessoa
- Nome e número de telefone da pessoa

## CAPÍTULO 5

# Acelerador do Brasil

Este capítulo inclui os seguintes tópicos:

- [Visão Geral do Acelerador do Brasil, 64](#)
- [Regras de Limpeza de Dados de Endereço do Brasil, 64](#)
- [Regras de Composição do Brasil, 65](#)
- [Regras de Limpeza de Dados de Contato do Brasil, 66](#)
- [Regras de Limpeza de Dados Corporativos do Brasil, 68](#)
- [Regras de Limpeza de Dados Gerais do Brasil, 68](#)
- [Regras de Eliminação de Duplicação e Correspondência do Brasil, 69](#)
- [Mapeamentos Demonstrativos do Brasil, 71](#)

## Visão Geral do Acelerador do Brasil

Use as regras no acelerador do Brasil para verificar e aprimorar os dados de organizações do Brasil.

O acelerador do Brasil inclui regras que executam as seguintes operações de qualidade de dados:

- Limpeza de dados de endereço
- Limpeza de dados de contato
- Limpeza de dados corporativos
- Limpeza de dados gerais
- Correspondência e eliminação de duplicação

O acelerador do Brasil também inclui uma regra de composição. Uma regra de composição combina várias regras em um único objeto.

O acelerador depende de regras instaladas pelo acelerador Principal.

## Regras de Limpeza de Dados de Endereço do Brasil

Use as regras de limpeza de dados de endereço para analisar, padronizar e validar dados de endereço.

As regras de limpeza de dados de endereço se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```



A tabela a seguir descreve as regras de limpeza de dados de endereço no acelerador do Brasil:

Nome	Descrição
rule_BRA_Address_Parse_Hybrid	Analisa os endereços brasileiros não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_BRA_Address_Parse_Multiline	Analisa os endereços brasileiros não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_BRA_Address_Validation_Distinct	Valida a capacidade de entrega de endereços brasileiros. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_BRA_Address_Validation_Distinct_w_Geocoding	Valida a capacidade de entrega de endereços brasileiros e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_BRA_Address_Validation_Hybrid	Valida a capacidade de entrega de endereços brasileiros. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_BRA_Address_Validation_Hybrid_w_Geocoding	Valida a capacidade de entrega de endereços brasileiros e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_BRA_Address_Validation_Multiline	Valida a capacidade de entrega de endereços brasileiros. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_BRA_Address_Validation_Multiline_w_Geocoding	Valida a capacidade de entrega de endereços brasileiros e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.

## Regras de Composição do Brasil

Use as regras de composição no acelerador do Brasil para adicionar um conjunto de regras a um mapeamento como um único objeto. Uma regra de composição é uma regra que faz uso da lógica de outras regras do acelerador.

As regras de composição se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```

A seguinte tabela descreve a regra de composição no acelerador do Brasil:

Nome	Descrição
rule_BRA_Contact_Data	Analisa, padroniza e valida os dados de contato do Brasil, como endereços, números de telefone e números do Cadastro de Pessoas Físicas (CPF).

A regra `rule_BRA_Contact_Data` lê mapplets de várias pastas no repositório. A regra também inclui uma transformação não reutilizável.

A seguinte tabela lista os nomes e as localizações de repositório das regras e da transformação em `rule_BRA_Contact_Data`:

Regra	Localização
Case_Converter	Transformação não reutilizável
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_BRA_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_BRA_Company_Suffix_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_BRA_Personal_CPF_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_BRA_Personal_Name_Parse_Validate	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_BRA_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_BRA_Phone_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_BRA_Prenome_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_BRA_Salutation_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing

## Regras de Limpeza de Dados de Contato do Brasil

Use as regras de limpeza de dados de contato para analisar, padronizar e validar dados sobre contatos comerciais e indivíduos.

As regras de limpeza de dados de contato se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados de contato no acelerador do Brasil:

Nome	Descrição
rule_BRA_Gender_Assignment	Atribui o sexo de acordo com o nome. A regra retorna "M" para nomes masculinos, "F" para nomes feminino e "U" quando o sexo é desconhecido. Por exemplo, a regra atribui a "João Coelho" o sexo "M" de masculino.
rule_BRA_Given_Name_Standard	Gera determinados nomes com base em nomes brasileiros.
rule_BRA_Personal_CPF_Validation	Valida a verificação de dígitos para os números de Cadastro de Pessoas Físicas (CPF).
rule_BRA_Personal_Name_Parse_Validate	Analisa os valores de nome de pessoa em portas diferentes. A regra cria portas para valores como título, nome, nome do meio e sobrenome. A regra também indica se o nome pode ser um nome de empresa e valida a ortografia do nome. A saída da regra inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.
rule_BRA_Personal_PIS_PASEP_Validation	Valida os números de previdência social do Brasil.
rule_BRA_Personal_Voter_Registration_Validation	Valide os dígitos de seleção nos números de título de eleitor do Brasil.
rule_BRA_Phone_Number_Parse	Analisa um número de telefone brasileiro em uma cadeia. A regra analisa o primeiro número de telefone dos dados, lendo da esquerda para a direita. A regra retorna um número de telefone e também retorna uma string que contém o texto de entrada com o número de telefone removido.
rule_BRA_Phone_Number_Standardization	Padroniza os números de telefone brasileiros. A regra retorna o número de telefone nos seguintes formatos: - Padrão - nn nnnn nnnn - Hífens - nn-nnnn-nnnn - Sem Espaços - nnnnnnnnnn
rule_BRA_Phone_Validation	Valida o código de área e o tamanho dos números de telefone brasileiros. Essa regra retorna códigos que indicam se o código de área e o comprimento de um número de telefone são válidos.
rule_BRA_Prenome_Assignment	Gera um título de acordo com o sexo. Você pode alterar a variável de expressão female_prenome de "Sra." para "Srta.".
rule_BRA_Salutation_Assignment	Gera saudações formais e casuais com base em nomes e tokens de nome. Por exemplo, quando os dados de entrada contêm "Sr. João Coelho", a regra gera a saudação formal "Prezado Sr. Coelho" e a saudação casual "Prezado João,". É possível alterar o prefixo e a pontuação editando as variáveis na transformação de Expressão dq_Generate_Salutation.

## Dependências de Regras de Limpeza de Dados de Contato Principais

O acelerador do Brasil depende das regras de limpeza de dados de contato do acelerador Principal a seguir:

- rule\_Email\_Parse\_Into\_Mailbox\_Domain
- rule\_Email\_Validation

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados do Contato Principal” na página 22.](#)

# Regras de Limpeza de Dados Corporativos do Brasil

Use as regras de limpeza de dados corporativos no acelerador do Brasil para padronizar e validar os dados corporativos.

As regras de limpeza de dados corporativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados corporativos no acelerador do Brasil:

Nome	Descrição
rule_BRA_Company_CNPJ_Validation	Valida os números do Cadastro Nacional da Pessoa Jurídica (CNPJ). Os números do CNPJ identificam as empresas brasileiras.
rule_BRA_Company_Suffix_Standardization	Padroniza os sufixos de empresas brasileiras.

# Regras de Limpeza de Dados Gerais do Brasil

Use as regras de limpeza de dados gerais para identificar os tipos de informações contidas nos campos de entrada.

As regras de limpeza de dados gerais se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados gerais no acelerador do Brasil:

Nome	Descrição
rule_BRA_NER_Field_Identification	Identifica o tipo de informações contidas em um campo de entrada. A regra pode identificar nomes, IDs Pessoais, nomes de empresa, datas e dados de endereço brasileiros. A regra retorna um rótulo que descreve o tipo de dados de entrada. A regra usa dados de referência e técnicas de correspondência probabilística para identificar os tipos de informações.

## Dependências de Regras de Limpeza de Dados Gerais Principais

O acelerador do Brasil depende das regras de limpeza de dados gerais a seguir do acelerador Principal:

- rule\_Assign\_DQ\_GeocodingStatus\_Description
- rule\_Assign\_DQ\_Mailability\_Score\_Description
- rule\_Assign\_DQ\_Match\_Code\_Descriptions
- rule\_Remove\_Extra\_Spaces
- rule\_Remove\_Non\_Numbers
- rule\_Remove\_Punctuation
- rule\_Remove\_Punctuation\_and\_Space
- rule\_Replace\_Limited\_Punct\_with\_Space
- rule\_TitleCase

- rule\_UpperCase

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados Gerais Principais” na página 23.](#)

## Regras de Eliminação de Duplicação e Correspondência do Brasil

Use as regras de correspondência e de eliminação de duplicação para medir os níveis de semelhança entre os registros em conjuntos de dados.

As regras de correspondência e de eliminação de duplicação se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

A tabela a seguir descreve as regras de correspondência e de eliminação de duplicação no acelerador do Brasil:

Nome	Descrição
mplt_BRA_Company_Name_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de empresa e em dados de endereço brasileiros. O mapplet usa uma combinação de caracteres dos valores de nome de empresa e dos valores de código postal para gerar chaves de grupo.
mplt_BRA_Famillyname_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas em dados brasileiros com base nos nomes de família e nos endereços. O mapplet usa uma combinação de caracteres dos valores de sobrenome e de código postal para gerar chaves de grupo.
mplt_BRA_Firstname_and_CPF_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base no nome e no número do Cadastro de Pessoas Físicas (CPF). O mapplet gera chaves de grupo com base em números de CPF.
mplt_BRA_IMO_Company_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados brasileiros com base em nomes e endereços de empresas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_BRA_IMO_Famillyname_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas em dados brasileiros com base nos nomes de família e nos endereços. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_BRA_IMO_Individual_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados brasileiros com base em nomes e endereços de pessoas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_BRA_IMO_Personal_Name_and_Data_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados brasileiros com base em nomes de pessoas e em dados pessoais. Os campos da coluna de dados pessoais devem conter um único tipo de dados, como o número de telefone, o e-mail ou o número do Cadastro de Pessoas Físicas. O mapplet gera chaves de grupo com base em dados pessoais.

Nome	Descrição
mplt_BRA_Individual_Name_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas e em dados de endereços brasileiros. O mapplet usa uma combinação de caracteres dos valores de sobrenome e de código postal para gerar chaves de grupo.
mplt_BRA_Individual_Name_and_CPF_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base nos nomes de pessoas brasileiras e nos números do Cadastro de Pessoas Físicas (CPF). O mapplet gera chaves de grupo com base em números de CPF.
mplt_BRA_Individual_Name_and_Date_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas brasileiras e em dados de data. O mapplet gera chaves de grupo com base em dados de data.
mplt_BRA_Individual_Name_and_Email_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas brasileiras e em endereços de e-mail. O mapplet gera chaves de grupo com base nos dados de endereço de e-mail.
mplt_BRA_Individual_Name_and_Phone_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas brasileiras e em números de telefone. O mapplet gera chaves de grupo geradas de números de telefone.
mplt_Company_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base no nome de empresa. O mapplet gera os códigos Soundex dos valores de nome de empresa e usa os códigos Soundex como chaves de grupo.
rule_BRA_Company_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes de empresa e em dados de endereços brasileiros.
rule_BRA_Familyname_and_Address_MatchScore	Gera uma pontuação de correspondência com base em sobrenomes e em dados de endereços brasileiros.
rule_BRA_Firstname_and_CPF_MatchScore	Gera uma pontuação de correspondência com base no nome e no número do Cadastro de Pessoas Físicas (CPF).
rule_BRA_Individual_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos dados de endereços brasileiros.
rule_BRA_Individual_Name_and_CPF_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos dados de endereços brasileiros.
rule_BRA_Individual_Name_and_Phone_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos números de telefone.
rule_Company_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de empresa.
rule_Individual_Name_and_Date_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nas datas.
rule_Individual_Name_and_Email_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos endereços de e-mail.

# Mapeamentos Demonstrativos do Brasil

Os mapeamentos demonstrativos do acelerador do Brasil usam várias regras para demonstrar processos de qualidade de dados.

Os mapeamentos demonstrativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules_Demo\BRA_Accelerator
```

O acelerador contém os seguintes mapeamentos demonstrativos:

## **m\_BRA\_customer\_data\_demo**

Analisa, padroniza e valida dados brasileiros.

## **m\_BRA\_customer\_matching\_demo**

Analisa e padroniza os dados de identidade do Brasil e realiza a análise de correspondência de identidade nos dados.

O mapeamento analisa as seguintes combinações de dados e gera clusters de correspondência para cada combinação:

- Dados de nome e endereço da pessoa
- Nome e número de telefone da pessoa

## CAPÍTULO 6

# Acelerador dos Serviços Financeiros

Este capítulo inclui os seguintes tópicos:

- [Visão Geral do Acelerador dos Serviços Financeiros, 72](#)
- [Regras de Limpeza de Dados de Contato dos Serviços Financeiros, 72](#)
- [Regras de Limpeza de Dados Financeiros dos Serviços Financeiros, 73](#)
- [Regras de Limpeza de Dados Gerais dos Serviços Financeiros, 76](#)
- [Regras de Eliminação de Duplicação e Correspondência dos Serviços Financeiros, 76](#)

## Visão Geral do Acelerador dos Serviços Financeiros

Use o acelerador dos Serviços Financeiros para verificar e aprimorar os dados de organizações no setor de serviços financeiros.

O acelerador dos Serviços Financeiros inclui regras que executam os seguintes processos de qualidade de dados:

- Limpeza de dados de contato
- Limpeza de dados financeiros
- Limpeza de dados gerais
- Correspondência e eliminação de duplicação

O acelerador depende das regras de limpeza de dados instaladas pelo acelerador Principal.

## Regras de Limpeza de Dados de Contato dos Serviços Financeiros

Use as regras de limpeza de dados de contato para padronizar dados de contato sobre contatos comerciais e indivíduos.

As regras de limpeza de dados de contato se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
```



A tabela a seguir descreve a regra de limpeza de dados de contato no acelerador dos Serviços Financeiros:

Nome	Descrição
rule_USA_Given_Name_Standard	Gera determinados nomes com base em nomes dos Estados Unidos. Por exemplo, a regra padroniza o nome "Bob" como o nome "Robert".

## Regras de Limpeza de Dados Financeiros dos Serviços Financeiros

Use as regras de limpeza de dados financeiros para analisar, padronizar e validar dados financeiros.

As regras de limpeza de dados financeiros se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Financial_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados financeiros no acelerador dos Serviços Financeiros:

Nome	Descrição
rule_Account_Status_Validation	Valida o status da conta. A regra requer dados de referência do status da conta.
rule_Accrual_Period_Validation	Valida que a data inicial é anterior à data final.
rule_Age_For_Account_Validation	Valida a idade do cliente para o tipo de conta. A regra usa a tabela de referência age_per_account_inf. Você deve atualizar a tabela de referência com os seus próprios dados.
rule_Beta_Coefficient_Validation	Verifica se a cadeia do coeficiente Beta é um número. A regra indica se a cadeia é um número positivo, negativo, um zero ou se não é um número.
rule_BIC_SWIFT_Code_Validation	Valida um código de Bank Identifier Code (BIC) ou Society for Worldwide Interbank Financial Telecommunication (SWIFT) pela validação de reconhecimento de padrão e de código de país.
rule_CAN_Transit_Number_Validation	Usa transações de fundos eletrônicas e em papel para validar o formato de um número de trânsito canadense.
rule_Credit_Card_Expiry_Check	Valida a data de validade de um cartão de crédito. A regra compara a data de validade do cartão de crédito com a data do sistema e identifica as datas expiradas. A regra aceita uma cadeia de sete caracteres no formato MM/YYYY.
rule_Credit_Card_Security_Code_Validation	Verifica se o código de segurança do cartão de crédito é um número inteiro que contém três ou quatro dígitos.
rule_Currency_Code_Country_Validation	Verifica se o código da moeda é válido para o código de países ISO de três caracteres.
rule_Currency_Code_Validation	Valida o código da moeda. A regra retorna "Válido" ou "Inválido".

Nome	Descrição
rule_CUSIP_Validation	Valida o formato e o tamanho do valor do dígito de verificação. A regra retorna um status que descreve a validade do valor do dígito de verificação e uma mensagem que explica o status.
rule_Delta_Validation	Verifica se o valor delta é positivo, negativo ou zero.
rule_Dividend_Yield_Validation	Verifica se a cadeia de resultado de dividendo é um número maior ou igual a zero. A regra retorna se a cadeia é um número positivo, negativo, um zero ou não é um número.
rule_EAD_Drawn_Balance_Validation	Verifica se a quantidade listada no risco de inadimplência (EAD) não é menor que o saldo utilizado. A regra segue as diretrizes para cálculo de EAD da Autoridade de Serviços Financeiros do Reino Unido.
rule_EAD_Validation	Verifica se a cadeia de risco à inadimplência (EAD) é um número. A regra retorna se a cadeia é um número positivo, negativo, um zero ou não é um número.
rule_EPS_Validation	Verifica se a entrada é um número maior ou igual a zero.
rule_Ex_Dividend_Date_Validation	Valida se a data do ex-dividendo e a data do registro são datas válidas, e se a data do ex-dividendo é anterior à data do registro. A regra identifica datas com uma diferença de mais de 15 dias como não válidas. A regra retorna a diferença em dias entre a data do registro e a data do ex-dividendo.
rule_Gamma_Validation	Verifica se a cadeia Gamma é um número. A regra retorna se a cadeia é um número positivo, negativo, um zero ou não é um número.
rule_GBR_Bank_Account_Parse	Analisa cadeias numéricas de oito dígitos como números de contas bancárias do Reino Unido.
rule_GBR_Bank_Account_Validation	Valida números de contas bancárias do Reino Unido. A regra retorna os códigos que indicam se a entrada é numérica e se tem o número correto de dígitos.
rule_GBR_Bank_Sort_Code_Parse	Analisa cadeias numéricas de seis dígitos como códigos de classificação bancária do Reino Unido. A regra analisa as cadeias de números nos seguintes formatos: <ul style="list-style-type: none"> <li>- Números consecutivos (999999)</li> <li>- Números delimitados por um traço (99-99-99)</li> </ul>
rule_GBR_Bank_Sort_Code_Standardize	Padroniza um código de classificação bancária do Reino Unido com o formato "NN-NN-NN".
rule_GBR_Bank_Sort_Code_Validation	Valida o formato e o tamanho dos códigos de classificação bancária do Reino Unido que são padronizados com o formato delimitado por traço (99-99-99). A regra retorna uma porta de Status que descreve a validade do código de classificação e uma porta de Nota de Validação que explica o status. Se o prefixo do código de classificação corresponder a uma atribuição conhecida de um banco do Reino Unido, a porta de Nota de Validação incluirá o nome do banco.
rule_Interest_Rate_Within_Range	Verifica se o valor decimal de taxas de juros está dentro do intervalo especificado. O intervalo é definido pelas duas portas de variável na transformação de Expressão. A regra retorna "True" ou "False".
rule_ISIN_Code_Validation	Valida que o valor de entrada é um International Securities Identification Number (ISIN). A regra verifica a estrutura do valor e verifica o dígito de verificação.

Nome	Descrição
rule_Loan_to_Value_Ratio	Calcula o empréstimo conforme o índice de valor, que é o valor do empréstimo dividido pelo valor da propriedade.
rule_Loss_Given_Default_Validation	Verifica se a cadeia é numérica e se é um valor positivo, negativo ou zero.
rule_Market_Cap_Validation	Verifica se a entrada é um número maior ou igual a zero.
rule_Maturity_Date_Validation	Valida se a data de vencimento é maior que a data do sistema.
rule_Positive_Close_Price_Value_Validation	Verifica se a entrada é um número maior que zero.
rule_Positive_Coupon_Percent_Validation	Verifica se a entrada é um número maior que zero.
rule_Positive_Last_Price_Value_Validation	Verifica se a entrada é um número maior que zero.
rule_Positive_Open_Price_Validation	Verifica se a entrada é um número maior que zero.
rule_Positive_Volume_Validation	Verifica se a entrada é um número maior que zero.
rule_Price_Earnings_Ratio_Validation	Verifica se o índice de preço-para-receita é um número positivo no intervalo de 0 a 100.
rule_Probability_of_Default_Validation	Verifica se há probabilidade do valor padrão ser numérico e indica se ele é um valor positivo, negativo ou zero. Caso positivo, a regra retornará mensagens de status para os valores nos seguintes intervalos: - <= 0,1 - > 0,1 e <= 0,5 - > .5 e <= 1 - > 1
rule_Rating_Code_Validation	Verifica se uma classificação pertence à escala de classificação da Standard & Poor's, da Moody's ou a uma lista definida pelo usuário.
rule_Rating_Date_Validation	Verifica se a data de classificação é um ano posterior à data do sistema.
rule_Risk_Weighted_Asset_Validation	Verifica se um ativo ponderado pelo risco é um número positivo.
rule_SEDOL_Validation	Valida um código de Stock Exchange Daily Official List (SEDOL) verificando o dígito de formato e de verificação.
rule_Stock_Exchange_Validation	Valida a maioria das bolsas de valores mundiais pelo nome e símbolo.
rule_USA_Routing_Number_Validation	Valida um número de roteamento formatado padrão de reconhecimento de caracteres em tinta magnética (MICR). Valida o Federal Reserve Bank Associado, a estrutura da entrada e o cálculo da soma de verificação.
rule_Volatility_Validation	Verifica se o valor de volatilidade é um número maior ou igual a zero.

# Regras de Limpeza de Dados Gerais dos Serviços Financeiros

Use as regras de limpeza de dados gerais para identificar os tipos de informações contidas nos campos de entrada.

As regras de limpeza de dados gerais se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\General _Data_Cleansing
```

A tabela a seguir descreve a regra de limpeza de dados gerais no acelerador dos Serviços Financeiros:

Nome	Descrição
rule_Postive_Number_Validation	Verifica se o número é um número maior do que zero.

## Dependências de Regras de Limpeza de Dados Gerais Principais

O acelerador dos Serviços Financeiros depende das seguintes regras de limpeza de dados gerais do acelerador Principal:

- rule\_Remove\_Punctuation
- rule\_Remove\_Punctuation\_and\_Space
- rule\_Remove\_Space
- rule\_UpperCase

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados Gerais Principais” na página 23](#).

# Regras de Eliminação de Duplicação e Correspondência dos Serviços Financeiros

Use as regras de correspondência e de eliminação de duplicação para gerar pontuações de correspondência e identificar registros duplicados.

As regras de correspondência e de eliminação de duplicação se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

A tabela a seguir descreve as regras de correspondência e de eliminação de duplicação no acelerador dos Serviços Financeiros:

Nome	Descrição
mplt_Company_Name_and_Address_Match	Identifica linhas duplicadas com base no nome da empresa e nos dados de endereço americanos. O mapplet usa uma combinação de caracteres dos valores de nome de empresa e dos valores de código postal para gerar chaves de grupo.
mplt_Company_Name_Match	Identifica linhas duplicadas com base no nome da empresa. O mapplet gera os códigos Soundex dos valores de nome de empresa e usa os códigos Soundex como chaves de grupo.
mplt_Familyname_and_Address_Match	Identifica linhas duplicadas com base no sobrenome e nos dados de endereço americanos. O mapplet usa uma combinação de caracteres dos valores de sobrenome e dos valores de código postal para gerar chaves de grupo.
mplt_Individual_Name_and_Address_Match	Identifica linhas duplicadas com base em nomes de pessoas e nos dados de endereço americanos. O mapplet gera códigos NYSIIS dos valores de sobrenome e usa os códigos NYSIIS como chaves de grupo.
mplt_Individual_Name_and_Date_Match	Identifica linhas duplicadas com base em nomes de pessoas e dados de data. O mapplet gera chaves de grupo provenientes dos dados de data.
mplt_Individual_Name_and_Email_Match	Identifica linhas duplicadas com base em nomes de pessoas e endereços de e-mail. O mapplet gera chaves provenientes dos dados de endereço de e-mail.
mplt_Individual_Name_and_Phone_Match	Identifica linhas duplicadas com base em nomes de pessoas e em números de telefone. O mapplet gera chaves de grupo provenientes de números de telefone.
mplt_Individual_Name_Match	Identifica linhas duplicadas com base nos nomes da pessoa. O mapplet gera códigos NYSIIS dos valores de sobrenome e usa os códigos NYSIIS como chaves de grupo.
rule_Company_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes de empresa e em endereços americanos.
rule_Company_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de empresa.
rule_Familyname_and_Address_MatchScore	Gera uma pontuação de correspondência com base em sobrenomes e em endereços americanos.
rule_Individual_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes de pessoas e em endereços americanos.
rule_Individual_Name_and_Date_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nas datas.
rule_Individual_Name_and_Email_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos endereços de e-mail.
rule_Individual_Name_and_Phone_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos números de telefone.
rule_Individual_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas.

## CAPÍTULO 7

# Acelerador da França

Este capítulo inclui os seguintes tópicos:

- [Visão Geral do Acelerador da França, 78](#)
- [Regras de Limpeza de Dados de Endereço da França, 78](#)
- [Regras de Composição da França, 79](#)
- [Regras de Limpeza de Dados de Contato da França, 80](#)
- [Regras de Limpeza de Dados Corporativos da França, 82](#)
- [Regras de Limpeza de Dados Gerais da França, 83](#)
- [Regras de Eliminação de Duplicação e Correspondência da França, 83](#)
- [Mapeamentos Demonstrativos da França, 85](#)

## Visão Geral do Acelerador da França

Use as regras no acelerador da França para verificar e aprimorar os dados de organizações da França.

O acelerador da França inclui regras que executam os seguintes processos de qualidade de dados:

- Limpeza de dados de endereço
- Limpeza de dados de contato
- Limpeza de dados corporativos
- Limpeza de dados gerais
- Correspondência e eliminação de duplicação

O acelerador depende das regras de limpeza de dados instaladas pelo acelerador Principal.

## Regras de Limpeza de Dados de Endereço da França

Use as regras de limpeza de dados de endereço para analisar, padronizar e validar dados de endereço.

As regras de limpeza de dados de endereço se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```

A seguinte tabela descreve as regras de limpeza de dados de endereço no acelerador da França:

Nome	Descrição
rule_FRA_Address_Parse_Hybrid	Analisa os endereços franceses não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_FRA_Address_Parse_Multiline	Analisa os endereços franceses não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_FRA_Address_Validation_Distinct_w_Geocoding	Valida a capacidade de entrega de endereços franceses e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_FRA_Address_Validation_Distinct	Valida a capacidade de entrega de endereços franceses. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_FRA_Address_Validation_Hybrid_w_Geocoding	Valida a capacidade de entrega de endereços franceses e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_FRA_Address_Validation_Hybrid	Valida a capacidade de entrega de endereços franceses. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_FRA_Address_Validation_Multiline_w_Geocoding	Valida a capacidade de entrega de endereços franceses e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_FRA_Address_Validation_Multiline	Valida a capacidade de entrega de endereços franceses. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.

## Regras de Composição da França

Use as regras de composição no acelerador da França para adicionar um conjunto de regras a um mapeamento como um único objeto. Uma regra de composição é uma regra que faz uso da lógica de outras regras do acelerador.

As regras de composição no acelerador da França são instaladas na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```

## Regra de Composição para Dados de Contato da França

A tabela a seguir descreve a regra de composição para dados de contato no acelerador da França:

Nome	Descrição
rule_FRA_Contact_Data	Analisa, padroniza e valida dados de contato da França, como endereços e números de telefone.

A regra `rule_FRA_Contact_Data` lê mapplets de várias pastas no repositório. A regra também inclui uma transformação não reutilizável.

A seguinte tabela lista os nomes e as localizações de repositório das regras e da transformação em `rule_FRA_Contact_Data`:

Regra	Localização
Case_Converter	Transformação não reutilizável
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_FRA_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_FRA_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_FRA_Gender_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_FRA_Multi_Person_Name_Parse	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_FRA_Phone_Number_Standardize	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_FRA_Phone_Number_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_FRA_Prenome_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_FRA_Salutation_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

## Regras de Limpeza de Dados de Contato da França

Use as regras de limpeza de dados de contato para analisar, padronizar e validar dados sobre contatos comerciais e indivíduos.

As regras de limpeza de dados de contato se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
```



A tabela a seguir descreve as regras de limpeza de dados de contato no acelerador da França:

Nome	Descrição
rule_FRA_Gender_Assignment	Atribui o sexo de acordo com os nomes. A regra retorna "M" para nomes masculinos, "F" para nomes feminino e "U" quando o sexo é desconhecido. Por exemplo, a regra atribui a "Jean Leclerc" o sexo "M" de masculino.
rule_FRA_Given_Name_Standard	Gera determinados nomes com base em nomes franceses.
rule_FRA_INSEE_Parse	Analisa o número do Institut National de la Statistique et des Études Économiques (INSEE) francês em uma cadeia.
rule_FRA_INSEE_Standardization	Padroniza o número do INSEE francês como uma cadeia de 13 dígitos ou 13 dígitos seguidos de um espaço e da chave de controle de dois dígitos.
rule_FRA_INSEE_Validation	Valida o número do INSEE com base no sexo, na data e nos valores do Code Officiel Géographique (COG).
rule_FRA_Multi_Person_Name_Parse	<p>Analisa os valores de nome de pessoa em portas diferentes. A regra cria portas para valores como título, nome, nome do meio e sobrenome.</p> <p>A saída da regra inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p> <p>Quando os dados de nome identificam mais de uma pessoa, a regra cria uma porta de saída para cada nome completo. Por exemplo, a regra pode ler o nome "Jean et Marianne Leclerc" e criar portas de saída para "Jean Leclerc" e "Marianne Leclerc".</p>
rule_FRA_Personal_Name_Parse_FML	<p>Analisa os valores em um nome de pessoa em portas diferentes.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Nome, nome do meio e sobrenome</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p>
rule_FRA_Personal_Name_Parse_LFM	<p>Analisa os valores em um nome de pessoa em portas diferentes.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Sobrenome, nome e nome do meio</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p>
rule_FRA_Phone_Number_Parse	<p>Analisa um número de telefone francês em uma cadeia. A regra analisa o primeiro número de telefone dos dados, lendo da direita para a esquerda.</p> <p>A regra reconhece números de telefone que usam zeros à esquerda, códigos de discagem internacional ou extensões que começam com o símbolo de jogo da velha. A regra processa os seguintes símbolos de pontuação: o sinal de adição, os parênteses e o símbolo de jogo da velha. Antes de executar a regra, remova todas as outras pontuações, incluindo os espaços duplos.</p> <p>A regra retorna um número de telefone e também retorna uma string que contém o texto de entrada com o número de telefone removido.</p>

Nome	Descrição
rule_FRA_Phone_Number_Standardize	Padroniza os números de telefone franceses para formatos de discagem internacional e local. A regra reconhece números de telefone que usam zeros à esquerda, códigos de discagem internacional ou extensões que começam com o símbolo de jogo da velha.
rule_FRA_Phone_Number_Validation	Valida o código de área e o tamanho dos números de telefone franceses. A regra retorna a região do número de telefone, assim como códigos que indicam se o código de área e o tamanho de um número de telefone são válidos.
rule_FRA_Prenome_Assignment	Gera um título de acordo com o sexo.
rule_FRA_Salutation_Assignment	Gera saudações formais e casuais com base em nomes e tokens de nome. Por exemplo, quando os dados de entrada contêm "M. Jean Leclerc", a regra gera a saudação formal "Monsieur Leclerc" e a saudação casual "Cher Jean,". É possível alterar o prefixo e a pontuação editando as variáveis na transformação de Expressão dq_Generate_Salutation.

### Dependências de Regras de Limpeza de Dados de Contato Principais

O acelerador da França depende das regras de limpeza de dados de contato a seguir do acelerador Principal:

- rule\_Email\_Validation

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados do Contato Principal” na página 22.](#)

## Regras de Limpeza de Dados Corporativos da França

Use as regras de limpeza de dados corporativos para padronizar os dados corporativos.

As regras de limpeza de dados corporativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados corporativos de acelerador da França:

Nome	Descrição
rule_FRA_Company_Name_Standardization	Padroniza os nomes de empresa para os valores da tabela de referência.
rule_FRA_SIRET_Number_Parse	Analisa o número do système d'identification du répertoire des établissements (SIRET) francês em uma cadeia.
rule_FRA_SIRET_Number_Standardize	Padroniza um número de 14 dígitos para o formato NNN NNN NNN NNNN sem considerar o espaçamento ou a pontuação da cadeia. Não há padronização para cadeias com menos de 14 dígitos.
rule_FRA_SIRET_Number_Validation	Valida um número do système d'identification du répertoire des établissements (SIRET). A regra presume que o número esteja no formato padrão depois da aplicação da regra rule_FRA_SIRET_Number_Standardization.

# Regras de Limpeza de Dados Gerais da França

Use as regras de limpeza de dados gerais para identificar os tipos de informações contidas nos campos de entrada.

As regras de limpeza de dados gerais se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados gerais no acelerador da França:

Nome	Descrição
rule_FRA_NER_Field_Identification	Identifica o tipo de informações contidas em um campo de entrada. A regra pode identificar nomes, IDs Pessoais, nomes de empresa, datas e dados de endereço franceses. A regra retorna um rótulo que descreve o tipo de dados de entrada. A regra usa dados de referência para identificar os tipos de informações. A regra usa técnicas de correspondência probabilística para identificar os tipos de informações.

O acelerador da França depende das regras de limpeza de dados gerais a seguir do acelerador Principal:

- rule\_Assign\_DQ\_GeocodingStatus\_Description
- rule\_Assign\_DQ\_Mailability\_Score\_Description
- rule\_Assign\_DQ\_Match\_Code\_Description
- rule\_Luhn\_Algorithm
- rule\_Remove\_Extra\_Spaces
- rule\_Remove\_Parentheses
- rule\_Remove\_Punctuation
- rule\_Remove\_Punctuation\_and\_Space
- rule\_Replace\_Limited\_Punct\_with\_Space
- rule\_UpperCase

Para obter mais informações sobre essas regras, consulte ["Regras de Limpeza de Dados Gerais Principais" na página 23](#).

## Regras de Eliminação de Duplicação e Correspondência da França

Use as regras de correspondência e de eliminação de duplicação para gerar pontuações de correspondência e identificar registros duplicados.

As regras de correspondência e de eliminação de duplicação no acelerador da França são instaladas na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

A tabela a seguir descreve as regras de correspondência e de eliminação de duplicação no acelerador da França:

Nome	Descrição
mplt_Company_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base no nome de empresa. O mapplet gera os códigos Soundex dos valores de nome de empresa e usa os códigos Soundex como chaves de grupo.
mplt_FRA_Company_Name_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes e endereços de empresas. O mapplet usa uma combinação de caracteres dos valores de nome de empresa e dos valores de código postal para gerar chaves de grupo.
mplt_FRA_Famillyname_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes e endereços de família. O mapplet usa uma combinação de caracteres dos valores de sobrenome e de código postal para gerar chaves de grupo.
mplt_FRA_Firstname_and_INSEE_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base no número do French Institut National de la Statistique et des Études Économiques (INSEE). O mapplet gera chaves de grupo dos dados de número do INSEE.
mplt_FRA_Firstname_Surname_DOB_and_Postcode_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de nomes pessoais, datas de nascimento e códigos postais. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_FRA_IMO_Company_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados franceses com base em nomes e endereços de empresas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_FRA_IMO_Famillyname_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados franceses com base em nomes e endereços de família. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_FRA_IMO_Individual_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados franceses com base em nomes e endereços de pessoas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_FRA_IMO_Personal_Name_and_Data_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados franceses com base em nomes de pessoas e em dados pessoais. Os campos na coluna de dados pessoais devem conter um único tipo de dados, como número de telefone, e-mail ou número do Institut National de la Statistique et des Études Économiques (INSEE). O mapplet gera chaves de grupo provenientes de dados pessoais.
mplt_FRA_Individual_Name_and_Date_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas francesas e em dados de data. O mapplet gera chaves de grupo com base em datas.
mplt_FRA_Individual_Name_and_Email_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas francesas e em endereços de e-mail. O mapplet gera chaves de grupo com base em dados de endereço de e-mail.
mplt_FRA_Individual_Name_and_INSEE_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas francesas e nos números do INSEE. O mapplet gera chaves de grupo provenientes de dados de número do INSEE.

Nome	Descrição
mplt_FRA_Individual_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas francesas. O mapplet gera códigos NYSIIS dos valores de sobrenome e usa os códigos NYSIIS como chaves de grupo.
rule_Company_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de empresa.
rule_FRA_Company_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes de empresa e em endereços franceses.
rule_FRA_Familyname_and_Address_MatchScore	Gera uma pontuação de correspondência com base em sobrenomes e em endereços franceses.
rule_FRA_Firstname_and_INSEE_MatchScore	Gera uma pontuação de correspondência com base em nomes e em todos os dados da coluna de dados pessoais, como número de telefone, e-mail ou número do INSEE.
rule_FRA_Firstname_Surname_DOB_and_Postcode_MatchScore	Gera uma pontuação de correspondência com base em sobrenomes, datas de nascimento e códigos postais.
rule_FRA_Individual_Name_and_INSEE_MatchScore	Gera uma pontuação de correspondência com base em nomes de pessoas e em números do INSEE.
rule_Individual_Name_and_Date_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nas datas.
rule_Individual_Name_and_Email_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos endereços de e-mail.
rule_Individual_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas.

## Mapeamentos Demonstrativos da França

Os mapeamentos demonstrativos no acelerador da França usam várias regras para demonstrar os processos de qualidade de dados.

Os mapeamentos demonstrativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules_Demo\FRA_Accelerator
```

O acelerador contém os seguintes mapeamentos demonstrativos:

### **m\_FRA\_customer\_data\_demo**

Analisa, padroniza e valida dados franceses.

### **m\_FRA\_customer\_matching\_demo**

Analisa e padroniza os dados de identidade da Austrália e da Nova Zelândia e realiza a análise de correspondência de identidade nos dados.

O mapeamento analisa as seguintes combinações de dados e gera clusters de correspondência para cada combinação:

- Dados de nome e endereço da pessoa
- Nome e número de telefone da pessoa

## CAPÍTULO 8

# Acelerador da Alemanha

Este capítulo inclui os seguintes tópicos:

- [Visão Geral do Acelerador da Alemanha, 87](#)
- [Regras de Limpeza de Dados de Endereço da Alemanha, 87](#)
- [Regras de Composição da Alemanha, 88](#)
- [Regras de Limpeza de Dados de Contato da Alemanha, 89](#)
- [Regras de Limpeza de Dados Corporativos da Alemanha, 91](#)
- [Regras de Limpeza de Dados Gerais da Alemanha, 91](#)
- [Regras de Eliminação de Duplicação e Correspondência da Alemanha, 92](#)
- [Mapeamentos Demonstrativos da Alemanha, 95](#)

## Visão Geral do Acelerador da Alemanha

Use as regras no acelerador da Alemanha para verificar e aprimorar os dados de organizações da Alemanha.

O acelerador da Alemanha inclui regras que executam os seguintes processos de qualidade de dados:

- Limpeza de dados de endereço
- Limpeza de dados de contato
- Limpeza de dados corporativos
- Limpeza de dados gerais
- Correspondência e eliminação de duplicação

O acelerador depende das regras de limpeza de dados instaladas pelo acelerador Principal.

## Regras de Limpeza de Dados de Endereço da Alemanha

Use as regras de limpeza de dados de endereço para analisar, padronizar e validar dados de endereço.

As regras de limpeza de dados de endereço se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados de endereço no acelerador da Alemanha:

Nome	Descrição
rule_DEU_Address_Parse_Hybrid	Analisa os endereços alemães não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_DEU_Address_Parse_Multiline	Analisa os endereços alemães não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_DEU_Address_Validation_Distinct_w_Geocoding	Valida a capacidade de entrega de endereços alemães e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_DEU_Address_Validation_Distinct	Valida a capacidade de entrega de endereços alemães. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_DEU_Address_Validation_Hybrid_w_Geocoding	Valida a capacidade de entrega de endereços alemães e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_DEU_Address_Validation_Hybrid	Valida a capacidade de entrega de endereços alemães. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_DEU_Address_Validation_Multiline_w_Geocoding	Valida a capacidade de entrega de endereços alemães e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_DEU_Address_Validation_Multiline	Valida a capacidade de entrega de endereços alemães. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.

## Regras de Composição da Alemanha

Use as regras de composição no acelerador da Alemanha para adicionar um conjunto de regras a um mapeamento como um único objeto. Uma regra de composição é uma regra que faz uso da lógica de outras regras do acelerador.

As regras de composição se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```



## Regra de Composição para Dados de Contato da Alemanha

A tabela a seguir descreve a regra de composição para dados de contato no acelerador da Alemanha:

Nome	Descrição
rule_DEU_Contact_Data	Analisa, padroniza e valida dados de contato da Alemanha, como endereços e números de telefone.

A regra `rule_DEU_Contact_Data` lê mapplets de várias pastas no repositório. A regra também inclui uma transformação não reutilizável.

A seguinte tabela lista os nomes e as localizações de repositório das regras e da transformação em `rule_DEU_Contact_Data`:

Regra	Localização
Case_Converter	Transformação não reutilizável
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_DEU_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_DEU_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_DEU_Gender_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_DEU_Multi_Person_Name_Parse	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_DEU_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_DEU_Phone_Number_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_DEU_Prenome_Assignment	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
Atribuição de rule_DEU_Salutation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing

## Regras de Limpeza de Dados de Contato da Alemanha

Use as regras de limpeza de dados de contato para analisar, padronizar e validar dados sobre contatos comerciais e indivíduos.

As regras de limpeza de dados de contato se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Contact_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados de contato no acelerador da Alemanha:

Nome	Descrição
rule_DEU_Gender_Assignment	Atribui o sexo de acordo com os nomes. A regra retorna "M" para nomes masculinos, "F" para nomes feminino e "U" quando o sexo é desconhecido. Por exemplo, a regra atribui a "Hans Adler" o sexo "M" de masculino.
rule_DEU_Given_Name_Standard	Gera determinados nomes com base em nomes alemães.
rule_DEU_Multi_Person_Name_Parse	<p>Analisa os valores de nome de pessoa em portas diferentes. A regra cria portas para valores como título, nome, nome do meio e sobrenome.</p> <p>A saída da regra inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p> <p>Quando os dados de nome identificam mais de uma pessoa, a regra cria uma porta de saída para cada nome completo. Por exemplo, a regra pode ler o nome "Hans und Maria Adler" e criar portas de saída para "Aans Adler" e "Maria Adler".</p>
rule_DEU_Personal_Name_Parsing_FML	<p>Analisa os valores em um nome de pessoa em portas diferentes.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Nome, nome do meio e sobrenome</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p>
rule_DEU_Personal_Name_Parsing_LFM	<p>Analisa os valores em um nome de pessoa em portas diferentes.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Sobrenome, nome e nome do meio</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p>
rule_DEU_Phone_Number_Parse	<p>Analisa um número de telefone alemão em uma cadeia. A regra analisa o primeiro número de telefone dos dados, lendo da direita para a esquerda.</p> <p>A regra reconhece números de telefone que usam zeros à esquerda, códigos de discagem internacional ou extensões que começam com o símbolo de jogo da velha. A regra processa os seguintes símbolos de pontuação: o sinal de adição, os parênteses e o símbolo de jogo da velha. Antes de executar a regra, remova todas as outras pontuações, incluindo os espaços duplos.</p> <p>A regra retorna um número de telefone e também retorna uma string que contém o texto de entrada com o número de telefone removido.</p>
rule_DEU_Phone_Number_Standardization	Padroniza os números de telefone alemães para os formatos de discagem internacional e local. A regra reconhece números de telefone que usam zeros à esquerda, códigos de discagem internacional ou extensões que começam com o símbolo de jogo da velha.
rule_DEU_Phone_Number_Validation	Valida o código de área e o tamanho dos números de telefone alemães. A regra retorna a região do número de telefone, assim como códigos que indicam se o código de área e o tamanho de um número de telefone são válidos.

Nome	Descrição
rule_DEU_Prenome_Assignment	Gera um título de acordo com o sexo.
rule_DEU_Salutation_Assignment	Gera saudações formais e casuais com base em nomes e tokens de nome. Por exemplo, quando os dados de entrada contêm "Herr Hans Adler", a regra gera a saudação formal "Sehr geehrter Herr Adler" e a saudação casual "Lieber Hans,". É possível alterar o prefixo e a pontuação editando as variáveis na transformação de Expressão dq_Generate_Salutation.

### Dependências de Regras de Limpeza de Dados de Contato Principais

O acelerador da Alemanha depende das regras de limpeza de dados de contato a seguir do acelerador Principal:

- rule\_Email\_Validation

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados do Contato Principal” na página 22.](#)

## Regras de Limpeza de Dados Corporativos da Alemanha

Use as regras de limpeza de dados corporativos para padronizar os dados corporativos.

As regras de limpeza de dados corporativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados corporativos no acelerador da Alemanha:

Nome	Descrição
rule_DEU_Company_Name_Standardization	Usa tabelas de referência para padronizar os nomes de empresa.

## Regras de Limpeza de Dados Gerais da Alemanha

Use as regras de limpeza de dados gerais para identificar os tipos de informações contidas nos campos de entrada.

As regras de limpeza de dados gerais se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados gerais no acelerador da Alemanha:

Nome	Descrição
rule_DEU_NER_Field_Identification	Identifica o tipo de informações contidas em um campo de entrada. A regra pode identificar nomes, IDs Pessoais, nomes de empresa, datas e dados de endereço alemães. A regra retorna um rótulo que descreve o tipo de dados de entrada. A regra usa dados de referência para identificar os tipos de informações. A regra usa técnicas de correspondência probabilística para identificar os tipos de informações.

### Dependências de Regras de Limpeza de Dados Gerais Principais

O acelerador da Alemanha depende das regras de limpeza de dados gerais a seguir do acelerador Principal:

- rule\_Assign\_DQ\_GeocodingStatus\_Description
- rule\_Assign\_DQ\_Mailability\_Score\_Description
- rule\_Assign\_DQ\_Match\_Code\_Descriptions
- rule\_Remove\_Extra\_Spaces
- rule\_Remove\_Hyphen
- rule\_Remove\_Leading\_Zero
- rule\_Remove\_Parentheses
- rule\_Remove\_Period\_Parentheses
- rule\_Remove\_Punctuation
- rule\_Remove\_Punctuation\_and\_Space
- rule\_Remove\_Space
- rule\_Replace\_Limited\_Punct\_with\_Space
- rule\_UpperCase

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados Gerais Principais” na página 23](#).

## Regras de Eliminação de Duplicação e Correspondência da Alemanha

Use as regras de correspondência e de eliminação de duplicação para gerar pontuações de correspondência e identificar registros duplicados.

As regras de correspondência e de eliminação de duplicação se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

A tabela a seguir descreve as regras de correspondência e de eliminação de duplicação no acelerador da Alemanha:

Nome	Descrição
mplt_Company_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base nos nomes de empresa. O mapplet gera os códigos Soundex dos valores de nome de empresa e usa os códigos Soundex como chaves de grupo.
mplt_DEU_Company_Name_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados alemães com base em dados de nomes e endereços de empresas. O mapplet usa uma combinação de caracteres dos valores de nome de empresa e dos valores de código postal para gerar chaves de grupo.
mplt_DEU_Famillyname_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados alemães com base em dados de sobrenomes e endereços de empresas. O mapplet usa uma combinação de caracteres dos valores de sobrenome e dos valores de código postal para gerar chaves de grupo.
mplt_DEU_Firstname_3CharsSurname_DOB_and_Postcode_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas em dados alemães com base nos nomes pessoais, nos três primeiros caracteres do nome de família, na data de nascimento e nos códigos postais. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_DEU_Firstname_and_PID_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados alemães com base em nomes pessoais e em IDs pessoais agrupadas. O mapplet gera chaves de grupo com base em dados de ID pessoais.
mplt_DEU_Firstname_Surname_2ElementsDOB_and_Postcode_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas em dados alemães com base nos nomes pessoais, nos dois elementos da data de nascimento e nos códigos postais. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_DEU_Firstname_Surname_DOB_and_Postcode_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas em dados alemães com base nos nomes pessoais, na data de nascimento e nos códigos postais. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_DEU_IMO_Company_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados alemães com base em nomes e endereços de empresas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_DEU_IMO_Famillyname_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados alemães com base em sobrenomes e em endereços. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_DEU_IMO_Individual_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados alemães com base em nomes e endereços de pessoas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_DEU_IMO_Personal_Name_and_Data_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados alemães com base em nomes de pessoas e em dados pessoais. Os campos da coluna de dados pessoais devem conter um único tipo de dados, como o número de telefone, o e-mail ou a ID pessoal. O mapplet gera chaves de grupo com base em dados pessoais.
mplt_DEU_Individual_Name_and_Date_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas e em dados de data agrupados por data. O mapplet gera chaves de grupo com base em dados de data.

Nome	Descrição
mplt_DEU_Individual_Name_and_Email_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados alemães com base em nomes de pessoas e em endereços de e-mail. O mapplet gera chaves de grupo com base em dados de endereço de e-mail.
mplt_DEU_Individual_Name_and_Phone_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados alemães com base em nomes e números de telefone de pessoas. O mapplet gera chaves de grupo dos dados de número de telefone.
mplt_DEU_Individual_Name_and_PID_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados alemães com base em nomes de pessoas e em IDs pessoais. O mapplet gera chaves de grupo com base em dados de ID pessoais.
mplt_DEU_Individual_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas em dados alemães com base nos nomes de pessoas. O mapplet gera códigos NYSIIS dos valores de sobrenome e usa o código NYSIIS como chaves de grupo.
rule_Company_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de empresa.
rule_DEU_Company_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes e endereços de empresas.
rule_DEU_Familyname_and_Address_MatchScore	Gera uma pontuação de correspondência com base em sobrenomes e em endereços.
rule_DEU_Firstname_3CharsSurname_DOB_and_Postcode_MatchScore	Gera uma pontuação de correspondência com base nos nomes, nos primeiros três caracteres do sobrenome, na data de nascimento e nos códigos postais.
rule_DEU_Firstname_and_PID_MatchScore	Gera uma pontuação de correspondência com base nos nomes e em todos os dados da coluna de dados pessoais, como número de telefone, e-mail ou ID pessoal.
rule_DEU_Firstname_Surname_2ElementsDOB_and_Postcode_MatchScore	Gera uma pontuação de correspondência com base nos nomes pessoais, na data de nascimento e nos códigos postais. <b>Nota:</b> Presume-se que o formato de entrada da data de nascimento é DD/MM/YYYY.
rule_DEU_Firstname_Surname_DOB_and_Postcode_MatchScore	Gera uma pontuação de correspondência com base nos sobrenomes, na data de nascimento e nos códigos postais.
rule_DEU_Individual_Name_and_Phone_MatchScore	Gera uma pontuação de correspondência com base em nomes de pessoas e em números de telefone.
rule_Familyname_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes e endereços de família.
rule_Individual_Name_and_Date_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nas datas.
rule_Individual_Name_and_Email_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos endereços de e-mail.

Nome	Descrição
rule_Individual_Name_and_SSN_MatchScore	Gera uma pontuação de correspondência com base em nomes e em todos os dados da coluna de dados pessoais, como número de telefone, e-mail ou número do SSN.
rule_Individual_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas.

## Mapeamentos Demonstrativos da Alemanha

Os mapeamentos demonstrativos no acelerador da Alemanha usam várias regras para demonstrar processos de qualidade.

Os mapeamentos demonstrativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules_Demo\DEU_Accelerator
```

O acelerador contém os seguintes mapeamentos demonstrativos:

### **m\_DEU\_customer\_data\_demo**

Analisa, padroniza e valida dados alemães.

### **m\_DEU\_customer\_matching\_demo**

Analisa e padroniza os dados de identidade da Alemanha e realiza a análise de correspondência de identidade nos dados.

O mapeamento analisa as seguintes combinações de dados e gera clusters de correspondência para cada combinação:

- Dados de nome e endereço da pessoa
- Nome e número de telefone da pessoa

## CAPÍTULO 9

# Acelerador de Portugal

Este capítulo inclui os seguintes tópicos:

- [Visão Geral do Acelerador de Portugal, 96](#)
- [Regras de Limpeza de Dados de Endereço de Portugal, 96](#)
- [Regras de Composição de Portugal, 97](#)
- [Regras de Limpeza de Dados de Contato de Portugal, 98](#)
- [Regras Limpeza de Dados Corporativos de Portugal, 100](#)
- [Regras de Limpeza de Dados Gerais de Portugal, 100](#)
- [Regras de Eliminação de Duplicação e Correspondência de Portugal, 101](#)
- [Mapeamentos Demonstrativos de Portugal, 103](#)

## Visão Geral do Acelerador de Portugal

Use as regras no acelerador de Portugal para verificar e aprimorar os dados de organizações de Portugal.

O acelerador de Portugal inclui regras que executam os seguintes processos de qualidade de dados:

- Limpeza de dados de endereço
- Limpeza de dados de contato
- Limpeza de dados corporativos
- Limpeza de dados gerais
- Correspondência e eliminação de duplicação

O acelerador depende das regras de limpeza de dados instaladas pelo acelerador Principal.

## Regras de Limpeza de Dados de Endereço de Portugal

Use as regras de limpeza de dados de endereço para analisar e validar dados de endereço.

As regras de limpeza de dados de endereço se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```



A tabela a seguir descreve as regras de limpeza de dados de endereço no acelerador de Portugal:

Nome	Descrição
rule_PRT_Address_Parse_Hybrid	Analisa os endereços portugueses não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_PRT_Address_Parse_Multiline	Analisa os endereços portugueses não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_PRT_Address_Validation_Distinct_w_Geocoding	Valida a capacidade de entrega de endereços portugueses e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_PRT_Address_Validation_Distinct	Valida a capacidade de entrega de endereços portugueses. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_PRT_Address_Validation_Hybrid_w_Geocoding	Valida a capacidade de entrega de endereços portugueses e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_PRT_Address_Validation_Hybrid	Valida a capacidade de entrega de endereços portugueses. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_PRT_Address_Validation_Multiline_w_Geocoding	Valida a capacidade de entrega de endereços portugueses e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_PRT_Address_Validation_Multiline	Valida a capacidade de entrega de endereços portugueses. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.

## Regras de Composição de Portugal

Use as regras de composição no acelerador de Portugal para adicionar um conjunto de regras a um mapeamento como um único objeto. Uma regra de composição é uma regra que faz uso da lógica de outras regras do acelerador.

As regras de composição se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```

## Regra de Composição para Dados de Contato de Portugal

A tabela a seguir descreve a regra de composição para dados de contato de Portugal no acelerador de Portugal:

Nome	Descrição
rule_PRT_Contact_Data	Analisa, padroniza e valida os dados de contato de Portugal, como endereços, números de telefone e números do Número de Identificação Fiscal (NIF).

A regra `rule_PRT_Contact_Data` lê mapplets de várias pastas no repositório. A regra também inclui uma transformação não reutilizável.

A seguinte tabela lista os nomes e as localizações de repositório das regras e da transformação em `rule_PRT_Contact_Data`:

Regra	Localização
Case_Converter	Transformação não reutilizável
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_PRT_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_PRT_NIF_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_NIF_Validate	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_Personal_Name_Parse_Validate	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_Phone_Number_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_Prenome_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_PRT_Salutation_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

## Regras de Limpeza de Dados de Contato de Portugal

Use as regras de limpeza de dados de contato para analisar, padronizar e validar dados sobre contatos comerciais e indivíduos.

As regras de limpeza de dados de contato se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados de contato no acelerador de Portugal:

Nome	Descrição
rule_PRT_Gender_Assignment	Atribui o sexo de acordo com os nomes. A regra retorna "M" para nomes masculinos, "F" para nomes feminino e "U" quando o sexo é desconhecido. Por exemplo, a regra atribui a "Artur Cruz" o sexo "M" de masculino.
rule_PRT_Given_Name_Standard	Gera determinados nomes com base em nomes portugueses.
rule_PRT_NIF_Parse	Analisa os números de Número de Identificação Fiscal (NIF) das cadeias. A regra retorna os números de ID e também retorna uma cadeia que contém o texto de entrada com os números de ID removidos.
rule_PRT_NIF_Standardization	Padroniza os números de Número de Identificação Fiscal (NIF) como cadeias de nove dígitos. A regra remove caracteres alfabéticos, símbolos e espaços.
rule_PRT_NIF_Validate	Valida um Número de Identificação Fiscal (NIF) com base no dígito de verificação em cada número. A regra requer que a entrada seja uma cadeia numérica de nove dígitos sem espaços.
rule_PRT_Personal_Name_Parse_Validate	Analisa os valores de nome de pessoa em portas diferentes. A regra cria portas para valores como título, nome, nome do meio e sobrenome. A regra também indica se o nome pode ser um nome de empresa e valida a ortografia do nome. A saída da regra inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.
rule_PRT_Phone_Number_Parse	Analisa um número de telefone português em uma cadeia. A regra analisa o primeiro número de telefone dos dados, lendo da direita para a esquerda. A regra retorna um número de telefone e também retorna uma string que contém o texto de entrada com o número de telefone removido.
rule_PRT_Phone_Number_Standardization	Padroniza os números de telefone portugueses para formatos de discagem internacional e local.
rule_PRT_Phone_Number_Validation	Valida o código de área e o tamanho dos números de telefone portugueses. A regra retorna a região do número de telefone, assim como códigos que indicam se o código de área e o tamanho de um número de telefone são válidos.
rule_PRT_Prefix_Assignment	Gera um título de acordo com o sexo. Você pode alterar a variável de expressão female_prefix de "Sra." para "Srta.".
rule_PRT_Salutation_Assignment	Gera saudações formais e casuais com base em nomes e tokens de nome. Por exemplo, quando os dados de entrada contêm "Sr. Artur Cruz", a regra gera a saudação formal "Prezado Sr. Cruz" e a saudação casual "Prezado Artur,". É possível alterar o prefixo e a pontuação editando as variáveis na transformação de Expressão dq_Generate_Salutation.

## Dependências de Regras de Limpeza de Dados de Contato Principais

O acelerador de Portugal depende das regras de limpeza de dados de contato a seguir do acelerador Principal:

- rule\_Email\_Validation

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados do Contato Principal” na página 22.](#)

# Regras Limpeza de Dados Corporativos de Portugal

Use as regras de limpeza de dados corporativos para analisar, padronizar e validar dados corporativos.

As regras de limpeza de dados corporativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados corporativos no acelerador de Portugal:

Nome	Descrição
rule_PRT_Company_Name_Standardization	Padroniza os nomes de empresa portugueses para os valores da tabela de referência.
rule_PRT_NIPC_Parse	Analisa um Número de Identificação de Pessoa Colectiva (NIPC). A regra retorna o NIPC e também retorna uma cadeia que contém o texto de entrada com o NIPC removido.
rule_PRT_NIPC_Standardize	Padroniza um Número de Identificação de Pessoa Colectiva (NIPC) como cadeias de nove dígitos. A regra remove caracteres alfabéticos, símbolos e espaços.
rule_PRT_NIPC_Validate	Valida um Número de Identificação Pessoa Colectiva (NIPC) com base no dígito de verificação em cada número. A regra requer que a entrada seja uma cadeia de nove dígitos.

# Regras de Limpeza de Dados Gerais de Portugal

Use as regras de limpeza de dados gerais para identificar os tipos de informações contidas nos campos de entrada.

As regras de limpeza de dados gerais se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados gerais no acelerador de Portugal:

Nome	Descrição
rule_PRT_NER_Field_Identification	Identifica o tipo de informações contidas em um campo de entrada. A regra pode identificar nomes, IDs Pessoais, nomes de empresa, datas e dados de endereço portugueses. A regra retorna um rótulo que descreve o tipo de dados de entrada. A regra usa dados de referência para identificar os tipos de informações. A regra usa técnicas de correspondência probabilística para identificar os tipos de informações.

## Dependências de Regras de Limpeza de Dados Gerais Principais

O acelerador de Portugal depende das regras de limpeza de dados gerais a seguir do acelerador Principal:

- rule\_Assign\_DQ\_GeocodingStatus\_Description
- rule\_Assign\_DQ\_ElementResultStatus\_Description
- rule\_Assign\_DQ\_Mailability\_Score\_Description
- rule\_Assign\_DQ\_Match\_Code\_Descriptions

- rule\_Parse\_First\_Word
- rule\_Remove\_Extra\_Spaces
- rule\_Remove\_Non\_Numbers
- rule\_Remove\_Punctuation
- rule\_Remove\_Punctuation\_and\_Space
- rule\_Replace\_Limited\_Punct\_with\_Space
- rule\_UpperCase

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados Gerais Principais” na página 23](#).

## Regras de Eliminação de Duplicação e Correspondência de Portugal

Use as regras de correspondência e de eliminação de duplicação para gerar pontuações de correspondência e identificar registros duplicados.

As regras de correspondência e de eliminação de duplicação se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

A tabela a seguir descreve as regras de correspondência e de eliminação de duplicação no acelerador de Portugal:

Nome	Descrição
mplt_Company_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base no nome de empresa. O mapplet gera os códigos Soundex dos valores de nome de empresa e usa os códigos Soundex como chaves de grupo.
mplt_PRT_Company_Name_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados portugueses com base em dados de nomes e endereços de empresas. O mapplet usa uma combinação de caracteres dos valores de nome de empresa e dos valores de código postal para gerar chaves de grupo.
mplt_PRT_Famillyname_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados portugueses com base em dados de sobrenomes e endereços de empresas. O mapplet usa uma combinação de caracteres dos valores de sobrenome e de código postal para gerar chaves de grupo.
mplt_PRT_Firstname_and_NIF_BI_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas nos dados portugueses com base em nomes e números de identificação pessoal, como Número de Identificação Fiscal (NIF) e Bilhete de Identidade (BI). O mapplet gera chaves de grupo com base nos dados de número de identificação pessoal.
mplt_PRT_IMO_Company_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados portugueses com base em nomes e endereços de empresas. O mapplet gera chaves de grupo com base nos dados do código postal.

Nome	Descrição
mplt_PRT_IMO_Famillyname_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados portugueses com base em nomes e endereços de família. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_PRT_IMO_Individual_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados portugueses com base em nomes e endereços de pessoas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_PRT_IMO_Personal_Name_and_Data_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados portugueses com base em nomes de pessoas e em dados pessoais. Os campos da coluna de dados pessoais devem conter um único tipo de dados, como o número de telefone, o e-mail ou o Número de Identificação Fiscal (NIF). O mapplet gera chaves de grupo provenientes de dados pessoais.
mplt_PRT_Individual_Name_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados portugueses com base em nomes de pessoas e em dados de endereço. O mapplet usa uma combinação de caracteres dos valores de sobrenome e de código postal para gerar chaves de grupo.
mplt_PRT_Individual_Name_and_Date_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados portugueses com base em nomes de pessoas e em dados de data. O mapplet gera chaves de grupo com base em dados de data.
mplt_PRT_Individual_Name_and_Email_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados portugueses com base em nomes de pessoas e em endereços de e-mail. O mapplet gera chaves de grupo com base em dados de endereço de e-mail.
mplt_PRT_Individual_Name_and_Phone_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados portugueses com base em nomes e números de telefone de pessoas. O mapplet gera chaves de grupo dos dados de número de telefone.
mplt_PRT_Individual_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas em dados portugueses com base nos nomes de pessoas. O mapplet gera códigos NYSIIS dos valores de sobrenome e usa os códigos NYSIIS como chaves de grupo.
rule_Company_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de empresa.
rule_PRT_Company_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes de empresa e em dados de endereços portugueses.
rule_PRT_Famillyname_and_Address_MatchScore	Gera uma pontuação de correspondência com base em sobrenomes e em dados de endereços portugueses.
rule_PRT_Firstname_and_NIF_BI_MatchScore	Gera uma pontuação de correspondência com base em dados do nome, Número de Identificação Fiscal (NIF) e números de Bilhete de Identidade (BI).
rule_PRT_Individual_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes de pessoas e em dados de endereços portugueses.
rule_PRT_Individual_Name_and_Date_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nas datas.
rule_PRT_Individual_Name_and_Email_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos endereços de e-mail.

Nome	Descrição
rule_PRT_Individual_Name_and_Phone_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos números de telefone.
rule_PRT_Individual_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas.

## Mapeamentos Demonstrativos de Portugal

Os mapeamentos demonstrativos no acelerador de Portugal usam várias regras para demonstrar os processos de qualidade de dados.

Os mapeamentos demonstrativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules_Demo\PRT_Accelerator
```

O acelerador contém os seguintes mapeamentos demonstrativos:

### **m\_PRT\_customer\_data\_demo**

Analisa, padroniza e valida dados portugueses.

### **m\_PRT\_customer\_matching\_demo**

Analisa e padroniza os dados de identidade de Portugal e realiza a análise de correspondência de identidade nos dados.

O mapeamento analisa as seguintes combinações de dados e gera clusters de correspondência para cada combinação:

- Dados de nome e endereço da pessoa
- Nome e número de telefone da pessoa

## CAPÍTULO 10

# Acelerador da Espanha

Este capítulo inclui os seguintes tópicos:

- [Visão Geral do Acelerador da Espanha, 104](#)
- [Regras de Limpeza de Dados de Endereço da Espanha, 104](#)
- [Regras de Limpeza de Dados de Contato da Espanha, 106](#)
- [Regras de Limpeza de Dados Corporativos da Espanha, 107](#)
- [Regras de Limpeza de Dados Gerais da Espanha, 108](#)
- [Regras de Correspondência e de Eliminação de Duplicação da Espanha, 108](#)
- [Mapeamentos Demonstrativos da Espanha, 111](#)

## Visão Geral do Acelerador da Espanha

Use as regras no acelerador da Espanha para verificar e aprimorar os dados de organizações da Espanha.

O acelerador da Espanha inclui regras que executam os seguintes processos de qualidade de dados:

- Limpeza de dados de endereço
- Limpeza de dados de contato
- Limpeza de dados corporativos
- Limpeza de dados gerais
- Correspondência e eliminação de duplicação

O acelerador depende das regras de limpeza de dados instaladas pelo acelerador Principal.

## Regras de Limpeza de Dados de Endereço da Espanha

Use as regras de limpeza de dados de endereço para analisar, padronizar e validar dados de endereço.

As regras de limpeza de dados de endereço se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```



A tabela a seguir descreve as regras de limpeza de dados de endereço no acelerador da Espanha:

Nome	Descrição
rule_ESP_Address_Parse_Hybrid	Analisa os endereços espanhóis não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_ESP_Address_Parse_Multiline	Analisa os endereços espanhóis não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_ESP_Address_Validation_Discrete_w_Geocoding	Valida a capacidade de entrega de endereços espanhóis e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_ESP_Address_Validation_Discrete	Valida a capacidade de entrega de endereços espanhóis. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_ESP_Address_Validation_Hybrid_w_Geocoding	Valida a capacidade de entrega de endereços espanhóis e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_ESP_Address_Validation_Hybrid	Valida a capacidade de entrega de endereços espanhóis. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_ESP_Address_Validation_Multiline_w_Geocoding	Valida a capacidade de entrega de endereços espanhóis e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_ESP_Address_Validation_Multiline	Valida a capacidade de entrega de endereços espanhóis. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.

# Regras de Limpeza de Dados de Contato da Espanha

Use as regras de limpeza de dados de contato para analisar, padronizar e validar dados sobre contatos comerciais e indivíduos.

As regras de limpeza de dados de contato se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados de contato no acelerador da Espanha:

Nome	Descrição
rule_ESP_CIF_Parse	Analisa o Certificado de Identificación Fiscal (CIF) da Espanha.
rule_ESP_CIF_Standardization	Padroniza o Certificado de Identificación Fiscal (CIF) da Espanha.
rule_ESP_CIF_Validation	Valida o Certificado de Identificación Fiscal (CIF) da Espanha.
rule_ESP_DNI_Parse	Analisa o Documento Nacional de Identidad (DNI) da Espanha.
rule_ESP_DNI_Standardization	Padroniza o Documento Nacional de Identidad (DNI) da Espanha.
rule_ESP_DNI_Validation	Valida o Documento Nacional de Identidad (DNI) da Espanha.
rule_ESP_Gender_Assignment	Atribui o sexo de acordo com os nomes. A regra retorna "M" para nomes masculinos, "F" para nomes feminino e "U" quando o sexo é desconhecido. Por exemplo, a regra atribui a "Juan Garcia" o sexo "M" de masculino.
rule_ESP_Given_Name_Standard	Gera determinados nomes com base em apelidos espanhóis.
rule_ESP_NIE_Parse	Analisa o Número de Identidad de Extranjero (NIE) da Espanha.
rule_ESP_NIE_Standardization	Padroniza o Número de Identidad de Extranjero (NIE) da Espanha.
rule_ESP_NIE_Validation	Valida o Número de Identidad de Extranjero (NIE) da Espanha.
rule_ESP_NIF_Parse	Analisa o Número de Identificación Fiscal (NIF) da Espanha em uma cadeia.
rule_ESP_NIF_Standardization	Padroniza o Número de Identificación Fiscal (NIF) da Espanha.
rule_ESP_NIF_Validation	Valida o Número de Identificación Fiscal (NIF) da Espanha.
rule_ESP_Personal_Name_Parse	Analisa nomes de pessoas espanholas.
rule_ESP_Phone_Number_Parse	Analisa um número de telefone espanhol em uma cadeia. A regra analisa o primeiro número de telefone dos dados, lendo da direita para a esquerda. A regra retorna um número de telefone e também retorna uma string que contém o texto de entrada com o número de telefone removido.

Nome	Descrição
rule_ESP_Phone_Number_Standardization	Padroniza os números de telefone espanhóis para formatos de discagem internacional e local. A regra reconhece números de telefone que usam zeros à esquerda, códigos de discagem internacional ou extensões que começam com o símbolo de jogo da velha.
rule_ESP_Phone_Number_Validation	Valida o código de área e o tamanho dos números de telefone espanhóis. A regra retorna a região do número de telefone, assim como códigos que indicam se o código de área e o tamanho de um número de telefone são válidos.
rule_ESP_Phone_Parse_Standardize_Validate	Analisa os números de telefone espanhóis e padroniza o formato. Valida o código de área e o tamanho dos números de telefone espanhóis.
rule_ESP_Prenome_Assignment	Gera um título de acordo com o sexo.
rule_ESP_Salutation_Assignment	Gera saudações formais e casuais com base em nomes e tokens de nome. Por exemplo, quando os dados de entrada contêm "Sr. Juan Garcia", a regra gera a saudação formal "Estimado Sr. Garcia" e a saudação casual "Querido Juan,". É possível alterar o prefixo e a pontuação editando as variáveis na transformação de Expressão dq_Generate_Salutation.

### Dependências de Regras de Limpeza de Dados de Contato Principais

O acelerador da Espanha depende das regras de limpeza de dados de contato a seguir do acelerador Principal:

- rule\_Email\_Validation

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados do Contato Principal” na página 22](#).

## Regras de Limpeza de Dados Corporativos da Espanha

Use as regras de limpeza de dados corporativos para padronizar os dados corporativos.

O acelerador da Espanha depende da regra de limpeza de dados corporativos a seguir do acelerador Principal:

- rule\_Company\_Name\_Standardization

# Regras de Limpeza de Dados Gerais da Espanha

Use as regras de limpeza de dados gerais para identificar os tipos de informações contidas nos campos de entrada.

As regras de limpeza de dados gerais se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]Rules\General_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados gerais no acelerador da Espanha:

Nome	Descrição
rule_ESP_NER_Field_Identification	Identifica o tipo de informações contidas em um campo de entrada. A regra pode identificar nomes, IDs Pessoais, nomes de empresa, datas e dados de endereço espanhóis. A regra retorna um rótulo que descreve o tipo de dados de entrada. A regra usa técnicas de correspondência probabilística para identificar os tipos de informações.

O acelerador da Espanha depende das regras de limpeza de dados gerais a seguir do acelerador Principal:

- rule\_Assign\_DQ\_ElementResultStatus\_Description
- rule\_Assign\_DQ\_GeocodingStatus\_Description
- rule\_Assign\_DQ\_Match\_Code\_Descriptions
- rule\_Remove\_Extra\_Spaces
- rule\_Remove\_Leading\_Zero
- rule\_Remove\_Limited\_Punctuation
- rule\_Remove\_Non\_Numbers
- rule\_Remove\_Punctuation\_and\_Space
- rule\_Remove\_Punctuation
- rule\_Replace\_limited\_Punct\_with\_Space
- rule\_Translate\_Diacritic\_Characters
- rule\_UpperCase

# Regras de Correspondência e de Eliminação de Duplicação da Espanha

Use as regras de correspondência e de eliminação de duplicação para gerar pontuações de correspondência e identificar registros duplicados.

As regras de correspondência e de eliminação de duplicação se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

A tabela a seguir descreve as regras de correspondência e de eliminação de duplicação no acelerador da Espanha:

Nome	Descrição
mplt_Company_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base no nome de empresa.
mplt_ESP_Company_Name_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados espanhóis com base em dados de nomes e endereços de empresas. O mapplet usa uma combinação de caracteres dos valores de nome de empresa e dos valores de código postal para gerar chaves de grupo.
mplt_ESP_Famillyname_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados espanhóis com base em dados de sobrenomes e endereços de empresas. O mapplet usa uma combinação de caracteres dos valores de sobrenome e dos valores de código postal para gerar chaves de grupo.
mplt_ESP_Firstname_and_NIF_BI_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas em dados espanhóis com base em nomes e em números de identificação pessoal, como o Número de Identificación Fiscal (NIF). O mapplet gera chaves de grupo com base nos dados de número de identificação pessoal.
mplt_ESP_IMO_Company_Name_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas em dados espanhóis com base nos nomes de empresa. O mapplet gera os códigos Soundex dos valores de nome de empresa e usa os códigos Soundex como chaves de grupo.
mplt_ESP_IMO_Company_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados espanhóis com base em nomes e endereços de empresa. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_ESP_IMO_Famillyname_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados espanhóis com base em nomes e endereços de família. O mapplet gera chaves de grupo provenientes de dados do código postal.
mplt_ESP_IMO_Individual_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados espanhóis com base em nomes e endereços de pessoas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_ESP_IMO_Personal_Name_and_Data_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados espanhóis com base em nomes de pessoas e em dados pessoais. Os campos da coluna de dados pessoais devem conter um único tipo de dados, como o número de telefone ou e-mail. O mapplet gera chaves de grupo com base em dados pessoais.
mplt_ESP_Individual_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas em dados espanhóis com base nos nomes de pessoas. O mapplet gera códigos NYSIIS dos valores de sobrenome e usa os códigos NYSIIS como chaves de grupo.

Nome	Descrição
mplt_ESP_Individual_Name_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados espanhóis com base em nomes de pessoas e em dados de endereço. O mapplet usa uma combinação de caracteres dos valores de sobrenome e dos valores de código postal para gerar chaves de grupo.
mplt_ESP_Individual_Name_and_Date_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados espanhóis com base em nomes de pessoas e em datas. O mapplet gera chaves de grupo com base em dados de data.
mplt_ESP_Individual_Name_and_Email_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados espanhóis com base em endereços de e-mail e em nomes de pessoas. O mapplet gera chaves de grupo com base em dados de endereço de e-mail.
mplt_ESP_Individual_Name_and_Phone_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados espanhóis com base em nomes e números de telefone de pessoas. O mapplet gera chaves de grupo dos dados de número de telefone.
rule_Company_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de empresa.
rule_ESP_Company_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes de empresa e em dados de endereços espanhóis.
rule_ESP_Famillyname_and_Address_MatchScore	Gera uma pontuação de correspondência com base em sobrenomes e em dados de endereços espanhóis.
rule_ESP_Firstname_and_NIF_BI_Matchscore	Gera uma pontuação de correspondência com base em nomes e nos números do Número de Identificación Fiscal (NIF).
rule_ESP_Individual_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas.
rule_ESP_Individual_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes de pessoas e em dados de endereços espanhóis.
rule_ESP_Individual_Name_and_Date_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nas datas.
rule_ESP_Individual_Name_and_Email_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos endereços de e-mail.
rule_ESP_Individual_Name_and_Phone_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos números de telefone.

# Mapeamentos Demonstrativos da Espanha

Os mapeamentos demonstrativos no acelerador da Espanha usam várias regras para demonstrar processos de qualidade de dados.

Os mapeamentos demonstrativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules_Demo\ESP_Accelerator
```

O acelerador contém os seguintes mapeamentos demonstrativos:

## **m\_ESP\_customer\_data\_demo**

Analisa, padroniza e valida dados espanhóis.

## **m\_ESP\_customer\_matching\_demo**

Analisa e padroniza os dados de identidade da Espanha e realiza a análise de correspondência de identidade nos dados.

O mapeamento analisa as seguintes combinações de dados e gera clusters de correspondência para cada combinação:

- Dados de nome e endereço da pessoa
- Nome e número de telefone da pessoa

## CAPÍTULO 11

# Acelerador do Reino Unido

Este capítulo inclui os seguintes tópicos:

- [Visão Geral do Acelerador do Reino Unido, 112](#)
- [Regras de Limpeza de Dados de Endereço do Reino Unido, 112](#)
- [Regras de Composição do Reino Unido, 114](#)
- [Regras de Limpeza de Dados de Contato do Reino Unido, 115](#)
- [Regras de limpeza de dados corporativos do Reino Unido, 118](#)
- [Regras de Limpeza de Dados Financeiros do Reino Unido, 118](#)
- [Regras de Limpeza de Dados Gerais do Reino Unido, 119](#)
- [Regras de Eliminação de Duplicação e Correspondência do Reino Unido, 119](#)
- [Mapeamentos Demonstrativos do Reino Unido, 122](#)

## Visão Geral do Acelerador do Reino Unido

Use as regras no acelerador do Reino Unido para verificar e aprimorar dados de organizações do Reino Unido.

O acelerador do Reino Unido inclui regras que executam os seguintes processos de qualidade de dados:

- Limpeza de dados de endereço
- Limpeza de dados de contato
- Limpeza de dados financeiros
- Correspondência e eliminação de duplicação

O acelerador depende das regras de limpeza de dados instaladas pelo acelerador Principal.

## Regras de Limpeza de Dados de Endereço do Reino Unido

Use as regras de limpeza de dados de endereço para analisar, padronizar e validar dados de endereço.

As regras de limpeza de dados de endereço se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```



A tabela a seguir descreve as regras de limpeza de dados de endereço no acelerador do Reino Unido:

Nome	Descrição
rule_GBR_Address_Parse_Hybrid	Analisa os endereços do Reino Unido não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_GBR_Address_Parse_Multiline	Analisa os endereços do Reino Unido não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_GBR_Address_Validation_Distinct_w_Geocoding	Valida a capacidade de entrega de endereços do Reino Unido e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_GBR_Address_Validation_Distinct	Valida a capacidade de entrega de endereços do Reino Unido. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_GBR_Address_Validation_Hybrid_w_Geocoding	Valida a capacidade de entrega de endereços do Reino Unido e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_GBR_Address_Validation_Hybrid	Valida a capacidade de entrega de endereços do Reino Unido. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_GBR_Address_Validation_Multiline_w_Geocoding	Valida a capacidade de entrega de endereços do Reino Unido e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_GBR_Address_Validation_Multiline	Valida a capacidade de entrega de endereços do Reino Unido. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_GBR_Postcode_Parse	Analisa os códigos postais do Reino Unido.

Nome	Descrição
rule_GBR_Postcode_Standardize	<p>Padroniza os códigos postais do Reino Unido. A regra exige que a entrada siga formatos predefinidos.</p> <p>A regra padroniza as entradas que correspondam aos seguintes padrões:</p> <ul style="list-style-type: none"> <li>- A9 9AA</li> <li>- A99 9AA</li> <li>- AA9 9AA</li> <li>- AA99 9AA</li> <li>- A9A 9AA</li> <li>- AA9A 9AA</li> <li>- GIR 0AA</li> </ul> <p>A letra A representa um caractere alfabético e o número 9 representa um dígito.</p>
rule_GBR_Postcode_Validate	<p>Valida os códigos postais do Reino Unido. A regra faz a correspondência de códigos postais padronizados com códigos postais válidos do Reino Unido. Se a regra não encontrar um código postal correspondente, ela verificará se o código postal segue o padrão do Reino Unido.</p>

## Regras de Composição do Reino Unido

Use as regras de composição no acelerador do Reino Unido para adicionar um conjunto de regras para um mapeamento como um único objeto. Uma regra de composição é uma regra que faz uso da lógica de outras regras do acelerador.

As regras de composição se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```

### Regra de Composição para Dados de Contato do Reino Unido

A tabela a seguir descreve a regra de composição para dados de contato no acelerador do Reino Unido:

Nome	Descrição
rule_GBR_Contact_Data	Analisa, padroniza e valida os dados de contato do Reino Unido, como endereços, números de telefone e números do National Insurance Numbers (NINO).

A regra `rule_GBR_Contact_Data` lê mapplets de várias pastas no repositório. A regra também inclui uma transformação não reutilizável.

A seguinte tabela lista os nomes e as localizações de repositório das regras e da transformação em `rule_GBR_Contact_Data`:

Nome	Localização
Case_Converter	Transformação não reutilizável
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

Nome	Localização
rule_GBR_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_GBR_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_GBR_Gender_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_GBR_Multi_Person_Name_Parse	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_GBR_NINO_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_GBR_NINO_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_GBR_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_GBR_Phone_Number_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Prenome_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Salutation_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

## Regras de Limpeza de Dados de Contato do Reino Unido

Use as regras de limpeza de dados de contato para analisar, padronizar e validar dados sobre contatos comerciais e indivíduos.

As regras de limpeza de dados de contato se encontram na seguinte localização do repositório:

[Informatica\_DQ\_Content]\Rules\Contact\_Data\_Cleansing

A tabela a seguir descreve as regras de limpeza de dados de contato no acelerador do Reino Unido:

Nome	Descrição
rule_GBR_Driver_Number_Parse	Analisa as cadeias que correspondem ao formato de números de carteira de habilitação do Reino Unido.
rule_GBR_Driver_Number_Validation	Valida os números de carteira de habilitação do Reino Unido com base nos requisitos do Government Data Standards Catalog do Reino Unido.
rule_GBR_Gender_Assignment	Atribui o sexo de acordo com os nomes. A regra retorna "M" para nomes masculinos, "F" para nomes feminino e "U" quando o sexo é desconhecido. Por exemplo, a regra atribui a "John Smith" o sexo "M" de masculino.
rule_GBR_Given_Name_Standard	Gera determinados nomes com base em apelidos do Reino Unido.

Nome	Descrição
rule_GBR_Multi_Person_Name_Parse	<p>Analisa os valores de nome de pessoa em portas diferentes. A regra cria portas para valores como título, nome, nome do meio e sobrenome.</p> <p>A saída da regra inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p> <p>Quando os dados de nome identificam mais de uma pessoa, a regra cria uma porta de saída para cada nome completo. Por exemplo, a regra pode ler o nome "John and Jane Smith" e criar portas de saída para "John Smith" e "Jane Smith".</p>
rule_GBR_NHS_Number_Parse	Analisa os números de National Health Service (NHS) de uma cadeia. A regra retorna o número do NHS e também retorna uma cadeia que contém o texto de entrada com o número do NHS removido.
rule_GBR_NHS_Number_Standardize	Padroniza os números do National Health Service (NHS) com o formato padrão (999 999 9999). A regra requer que a entrada seja uma cadeia de 10 dígitos.
rule_GBR_NHS_Number_Validate	Valida os números do National Health Service (NHS) com base no dígito de verificação em cada número. A regra requer que a entrada seja uma cadeia de 10 dígitos.
rule_GBR_NINO_Conformity_Check	Valida o padrão de um National Insurance Number (NINO) do Reino Unido. A regra não verifica se um NINO está correto ou ativo.
rule_GBR_NINO_Parse	Analisa os National Insurance Number (NINO) do Reino Unido com base nas cadeias. A regra retorna o NINO e também retorna uma cadeia que contém o texto de entrada com o NINO removido.
rule_GBR_NINO_Standardization	<p>Padroniza os National Insurance Numbers (NINO) do Reino Unido com os dois formatos mais típicos. A regra retorna os seguintes formatos, onde C representa caracteres alfabéticos e N representa numerais:</p> <ul style="list-style-type: none"> <li>- CC NN NN NN C</li> <li>- CCNNNNNNC</li> </ul> <p>A regra formata todos os caracteres alfabéticos como maiúsculas. A regra requer que a entrada esteja em conformidade com o padrão de um NINO.</p>
rule_GBR_NINO_Validation	Valida um National Insurance Number (NINO) do Reino Unido. A regra não verifica se um NINO está ativo.
rule_GBR_Passport_Number_MR_Parse	Analisa os números de passaporte do Reino Unido no formato estendido. O formato estendido é o formato legível por máquina para números de passaporte.
rule_GBR_Passport_Number_Parse	Analisa os números de passaporte do Reino Unido que usam o formato especificado pelo Government Data Standards Catalogue. A regra analisa todas as cadeias de nove dígitos.
rule_GBR_Passport_Number_Validation	Valida os números de passaporte do Reino Unido que usam o formato especificado pelo Government Data Standards Catalogue.

Nome	Descrição
rule_GBR_Personal_Name_Parsing_FML	<p>Analisa os valores em um nome de pessoa em portas diferentes.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Nome, nome do meio e sobrenome</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p>
rule_GBR_Personal_Name_Parsing_LFM	<p>Analisa os valores em um nome de pessoa em portas diferentes.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Sobrenome, nome e nome do meio</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p>
rule_GBR_Phone_Number_Parse	<p>Analisa um número de telefone do Reino Unido em uma cadeia. A regra analisa o primeiro número de telefone dos dados, lendo da direita para a esquerda.</p> <p>A regra reconhece os números de telefone que usam zeros à esquerda, o código de discagem internacional "+44" e as extensões que começam com o símbolo de jogo da velha. A regra processa os seguintes símbolos de pontuação: o sinal de adição, os parênteses e o símbolo de jogo da velha. Antes de executar a regra, remova todas as outras pontuações, incluindo os espaços duplos.</p> <p>A regra retorna um número de telefone e também retorna uma string que contém o texto de entrada com o número de telefone removido.</p>
rule_GBR_Phone_Number_Standardization	<p>Padroniza os números de telefone do Reino Unido para formatos de discagem internacional e local. A regra reconhece os números de telefone que usam zeros à esquerda, o código de discagem internacional "+44" e as extensões que começam com o símbolo de jogo da velha.</p>
rule_GBR_Phone_Number_Validation	<p>Valida o código de área e o tamanho dos números de telefone do Reino Unido. A regra retorna a região do número de telefone, assim como códigos que indicam se o código de área e o tamanho de um número de telefone são válidos.</p>
rule_Prefix_Assignment	<p>Gera um título de acordo com o sexo. Você pode alterar a variável de expressão female_prefix de "Sra." para "Srta.".</p>
rule_Salutation_Assignment	<p>Gera saudações formais e casuais com base em nomes e tokens de nome. Por exemplo, quando os dados de entrada contêm "Sr. John Smith", a regra gera a saudação formal "Prezado Sr. Smith," e a saudação casual "Prezado John,". É possível alterar o prefixo e a pontuação editando as variáveis na transformação de Expressão dq_Generate_Salutation.</p>

# Regras de limpeza de dados corporativos do Reino Unido

Use as regras de limpeza de dados corporativos para padronizar os dados corporativos.

As regras de limpeza de dados corporativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
```

A seguinte tabela descreve as regras de limpeza de dados corporativos de acelerador do Reino Unido:

Nome	Descrição
rule_GBR_Company_Name_Standardization	Padroniza um nome da empresa e fornece o acrônimo para o nome quando possível.

# Regras de Limpeza de Dados Financeiros do Reino Unido

Use as regras de limpeza de dados financeiros para analisar, padronizar e validar dados financeiros.

As regras de limpeza de dados financeiros se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Financial_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados financeiros no acelerador do Reino Unido:

Nome	Descrição
rule_GBR_Bank_Account_Parse	Analisa cadeias numéricas de oito dígitos como números de contas bancárias do Reino Unido.
rule_GBR_Bank_Account_Validation	Valida números de contas bancárias do Reino Unido. A regra retorna os códigos que indicam se a entrada é numérica e se tem o número correto de dígitos.
rule_GBR_Bank_Sort_Code_Parse	Analisa cadeias numéricas de seis dígitos como códigos de classificação bancária do Reino Unido. A regra analisa as cadeias de números nos seguintes formatos: <ul style="list-style-type: none"><li>- Números consecutivos (999999)</li><li>- Números delimitados por um traço (99-99-99)</li></ul>
rule_GBR_Bank_Sort_Code_Validation	Valida o formato e o tamanho dos códigos de classificação bancária do Reino Unido que são padronizados com o formato delimitado por traço (99-99-99). A regra retorna uma porta de Status que descreve a validade do código de classificação e uma porta de Nota de Validação que explica o status. Se o prefixo do código de classificação corresponder a uma atribuição conhecida de um banco do Reino Unido, a porta de Nota de Validação incluirá o nome do banco.
rule_GBR_Bank_Sort_Code_Standardize	Padroniza um código de classificação bancária do Reino Unido com o formato "NN-NN-NN".

# Regras de Limpeza de Dados Gerais do Reino Unido

Use as regras de limpeza de dados gerais para identificar os tipos de informações contidas nos campos de entrada.

As regras de limpeza de dados gerais se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados gerais no acelerador do Reino Unido:

Nome	Descrição
rule_GBR_NER_Field_Identification	Identifica o tipo de informações contidas em um campo de entrada. A regra pode identificar nomes, IDs Pessoais, nomes de empresa, datas e dados de endereço do Reino Unido. A regra retorna um rótulo que descreve o tipo de dados de entrada. A regra usa dados de referência para identificar os tipos de informações. A regra usa técnicas de correspondência probabilística para identificar os tipos de informações.

O acelerador do Reino Unido depende das seguintes regras de limpeza de dados gerais do acelerador Principal:

- rule\_Assign\_DQ\_GeocodingStatus\_Description
- rule\_Assign\_DQ\_Mailability\_Score\_Description
- rule\_Assign\_DQ\_Match\_Code\_Descriptions
- rule\_Remove\_Extra\_Spaces
- rule\_Remove\_Leading\_Zero
- rule\_Remove\_Period\_Parentheses
- rule\_Remove\_Punctuation
- rule\_Remove\_Punctuation\_and\_Space
- rule\_Remove\_Space
- rule\_Replace\_Limited\_Punct\_with\_Space
- rule\_UpperCase

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados Gerais Principais” na página 23](#).

## Regras de Eliminação de Duplicação e Correspondência do Reino Unido

Use as regras de correspondência e de eliminação de duplicação para medir os níveis de semelhança entre os registros em conjuntos de dados.

As regras de correspondência e de eliminação de duplicação se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

A tabela a seguir descreve as regras de correspondência e de eliminação de duplicação no acelerador do Reino Unido:

Nome	Descrição
mplt_GBR_Company_Name_Postcode_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados do Reino Unido com base em nomes de empresa e em códigos postais. O mapplet gera chaves de grupo com base no código postal.
mplt_GBR_Famlyname_and_NINO_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados do Reino Unido com base em sobrenomes e em National Insurance Number (NINO). O mapplet gera chaves de grupo com base em dados do NINO.
mplt_GBR_Famlyname_and_Postcode_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados do Reino Unido com base em sobrenomes e em códigos postais do Reino Unido. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_GBR_Firstname_3CharsSurname_DOB_and_Postcode_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados do Reino Unido com base nos seguintes dados: <ul style="list-style-type: none"> <li>- Nome</li> <li>- Os três primeiros caracteres do sobrenome</li> <li>- Data de nascimento</li> <li>- código postal</li> </ul> O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_GBR_Firstname_Surname_2ElementsDOB_and_Postcode_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados do Reino Unido com base nos seguintes dados: <ul style="list-style-type: none"> <li>- Nomes de pessoa</li> <li>- Quaisquer dois elementos da data de nascimento, como mês e ano</li> <li>- Código postal do Reino Unido</li> </ul> O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_GBR_Firstname_Surname_DOB_and_Postcode_Match	Usa estratégias de correspondência de campo para identificar as linhas com base nas seguintes dados: <ul style="list-style-type: none"> <li>- Nomes de pessoa</li> <li>- Data de nascimento</li> <li>- código postal</li> </ul> O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_GBR_IMO_Company_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados do Reino Unido com base em nomes e endereços de empresas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_GBR_IMO_Famlyname_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados do Reino Unido com base em nomes e endereços de família. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_GBR_IMO_Individual_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados do Reino Unido com base em nomes e endereços de pessoas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_GBR_IMO_Personal_Name_and_Data_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados do Reino Unido com base em nomes de pessoas e em dados pessoais. Os campos da coluna de dados pessoais devem conter um único tipo de dados, como o número de telefone, o e-mail ou o National Insurance Number. O mapplet gera chaves de grupo com base em dados pessoais.



Nome	Descrição
mplt_GBR_Individual_Name_and_Date_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados do Reino Unido com base em nomes de pessoas e em dados de data. O mapplet gera chaves de grupo com base em dados de data.
mplt_GBR_Individual_Name_and_Email_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados do Reino Unido com base em nomes de pessoas e em dados de endereços de e-mail. O mapplet gera chaves de grupo com base em dados de endereço de e-mail.
mplt_GBR_Individual_Name_and_NINO_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados do Reino Unido com base em nomes de pessoas e em National Insurance Numbers (NINO). O mapplet gera chaves de grupo com base em dados do NINO.
mplt_GBR_Individual_Name_and_Phone_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados do Reino Unido com base em nomes e números de telefone de pessoas. O mapplet gera chaves de grupo dos dados de número de telefone.
mplt_GBR_Individual_Name_and_Postcode_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados do Reino Unido com base em nomes de pessoas e em dados de código postal. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_GBR_Individual_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas em dados do Reino Unido com base nos nomes de pessoas. O mapplet gera códigos NYSIS dos valores de sobrenome e usa os códigos NYSIS como chaves de grupo.
rule_GBR_Familyname_and_NINO_MatchScore	Gera uma pontuação de correspondência com base em sobrenomes e nos National Identification Numbers (NINO) do Reino Unido.
rule_GBR_Familyname_and_Postcode_MatchScore	Gera uma pontuação de correspondência com base em sobrenomes e em códigos postais do Reino Unido.
rule_GBR_Firstname_3CharsSurname_DOB_and_Postcode_MatchScore	Gera uma pontuação de correspondência com base nas seguintes informações: <ul style="list-style-type: none"> <li>- Nome</li> <li>- Os três primeiros caracteres do sobrenome</li> <li>- Data de nascimento</li> <li>- Código postal</li> </ul>
rule_GBR_Firstname_Surname_2ElementsDOB_and_Postcode_MatchScore	Gera uma pontuação de correspondência com base nas seguintes informações: <ul style="list-style-type: none"> <li>- Nomes de pessoa</li> <li>- Quaisquer dois elementos da data de nascimento, como mês e ano</li> <li>- Código postal do Reino Unido</li> </ul>
rule_GBR_Firstname_Surname_DOB_and_Postcode_MatchScore	Gera uma pontuação de correspondência com base nos nomes pessoais, na data de nascimento e nos códigos postais.
rule_GBR_Individual_Name_and_NINO_MatchScore	Gera uma pontuação de correspondência com base em nomes de pessoas e nos National Insurance Numbers (NINO) do Reino Unido.
rule_GBR_Individual_Name_and_Phone_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos números de telefone.
rule_GBR_Individual_Name_and_Postcode_MatchScore	Gera uma pontuação de correspondência com base em nomes de pessoas e em códigos postais do Reino Unido.

Nome	Descrição
rule_Individual_Name_and_Date_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nas datas.
rule_Individual_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas.
rule_GBR_Company_Name_Postcode_MatchScore	Gera uma pontuação de correspondência com base em nomes de empresa e em códigos postais do Reino Unido.
rule_Individual_Name_and_Email_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos endereços de e-mail.

## Mapeamentos Demonstrativos do Reino Unido

Os mapeamentos demonstrativos no Acelerador do Reino Unido usam várias regras para demonstrar os processos de qualidade de dados.

Os mapeamentos demonstrativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules_Demo\GBR_Accelerator
```

O acelerador do Reino Unido contém os seguintes mapeamentos demonstrativos:

### **m\_GBR\_customer\_data\_demo**

Analisa, padroniza e valida dados de clientes do Reino Unido.

### **m\_GBR\_customer\_matching\_demo**

Analisa e padroniza os dados de identidade do Reino Unido e realiza a análise de correspondência de identidade nos dados.

O mapeamento analisa as seguintes combinações de dados e gera clusters de correspondência para cada combinação:

- Dados de nome e endereço da pessoa
- Nome e número de telefone da pessoa

## CAPÍTULO 12

# Acelerador dos EUA/Canadá

Este capítulo inclui os seguintes tópicos:

- [Visão Geral do Acelerador dos EUA/Canadá, 123](#)
- [Regras de Limpeza de Dados de Endereço dos EUA/Canadá, 123](#)
- [Regras de Composição dos EUA/Canadá, 126](#)
- [Regras de Limpeza de Dados de Contato dos EUA/Canadá, 128](#)
- [Regras de Limpeza de Dados Corporativos dos EUA/Canadá, 133](#)
- [Regras de Limpeza de Dados Gerais dos EUA/Canadá, 133](#)
- [Regras de Eliminação de Duplicação e Correspondência dos EUA/Canadá, 134](#)
- [Mapeamentos Demonstrativos dos EUA/Canadá, 137](#)

## Visão Geral do Acelerador dos EUA/Canadá

Use as regras no acelerador dos EUA/Canadá para verificar e aprimorar os dados das organizações nos Estados Unidos e no Canadá.

O acelerador dos EUA/Canadá inclui regras que executam os seguintes processos de qualidade de dados:

- Limpeza de dados de endereço
- Limpeza de dados de contato
- Limpeza de dados gerais
- Correspondência e eliminação de duplicação

O acelerador depende das regras de limpeza de dados instaladas pelo acelerador Principal.

## Regras de Limpeza de Dados de Endereço dos EUA/Canadá

Use as regras de limpeza de dados de endereço para analisar, padronizar e validar dados de endereço.

As regras de limpeza de dados de endereço se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados de endereço no acelerador dos EUA/Canadá:

Nome	Descrição
rule_CAN_Address_Certification_Hybrid	<p>Valida a entrega de endereços canadenses para os padrões SERP Programa de Reconhecimento e Avaliação de Software que o Canada Post mantém para endereços canadenses. A regra corrige erros nos endereços de entrada quando possível e inclui uma porta de status que identifica os endereços que atendem ao padrão SERP. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.</p> <p><b>Nota:</b> Quando você certifica um conjunto de registros de endereço para o padrão SERP, deve enviar um relatório de certificação para o Canada Post. A transformação de Validador de Endereço inclui campos de propriedade que podem ser preenchidos com informações para o relatório. Salve ou imprima o relatório e inclua-o com os detalhes de endereço que você enviar para o Canada Post.</p>
rule_CAN_Address_Parse_Hybrid	<p>Analisa os endereços canadenses não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.</p>
rule_CAN_Address_Parse_Multiline	<p>Analisa os endereços canadenses não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.</p>
rule_CAN_Address_Validation_Distinct_w_Geocoding	<p>Valida a capacidade de entrega de endereços canadenses e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.</p>
rule_CAN_Address_Validation_Distinct	<p>Valida a capacidade de entrega de endereços canadenses. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.</p>
rule_CAN_Address_Validation_Hybrid_w_Geocoding	<p>Valida a capacidade de entrega de endereços canadenses e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.</p>
rule_CAN_Address_Validation_Hybrid	<p>Valida a capacidade de entrega de endereços canadenses. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.</p>
rule_CAN_Address_Validation_Multiline_w_Geocoding	<p>Valida a capacidade de entrega de endereços canadenses e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.</p>
rule_CAN_Address_Validation_Multiline	<p>Valida a capacidade de entrega de endereços canadenses. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.</p>

Nome	Descrição
rule_CAN_Postcode_Validation	Valida códigos postais canadenses. A regra retorna "Válido" ou "Inválido".
rule_CAN_Province_Validation	Valida nomes de províncias canadenses. A regra retorna "Válido" ou "Inválido".
rule_USA_Address_Certification_Hybrid	<p>Valida a entrega dos endereços dos Estados Unidos para os padrões CASS (Sistema de Suporte de Precisão de Codificação) que o serviço postal dos Estados Unidos mantém para os endereços. A regra corrige erros nos endereços de entrada quando possível e inclui uma porta de status que identifica os endereços que atendem ao padrão CASS. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.</p> <p><b>Nota:</b> Quando você certifica um conjunto de registros de endereço para o padrão CASS, deve enviar um relatório de certificação para o USPS. A transformação de Validador de Endereço inclui campos de propriedade que podem ser preenchidos com informações para o relatório. Salve ou imprima o relatório e inclua-o com os detalhes de endereço que você envia ao USPS.</p>
rule_USA_Address_Parse_Hybrid	Analisa os endereços dos Estados Unidos não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_USA_Address_Parse_Multiline	Analisa os endereços dos Estados Unidos não estruturados em elementos de endereço. A regra não valida os endereços. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_USA_Address_Validation_Distinct_w_Geocoding	Valida a capacidade de entrega de endereços dos Estados Unidos e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_USA_Address_Validation_Distinct	Valida a capacidade de entrega de endereços dos Estados Unidos. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Distinta na transformação do Validador de Endereço.
rule_USA_Address_Validation_Hybrid_w_Geocoding	Valida a capacidade de entrega de registros dos endereços dos Estados Unidos e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_USA_Address_Validation_Hybrid	Valida a capacidade de entrega de registros dos endereços dos Estados Unidos. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada Híbrida na transformação do Validador de Endereço.
rule_USA_Address_Validation_Multiline_w_Geocoding	Valida a capacidade de entrega de endereços dos Estados Unidos e adiciona coordenadas de latitude e longitude a cada endereço de saída. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.

Nome	Descrição
rule_USA_Address_Validation_Multiline	Valida a capacidade de entrega de endereços dos EUA. A regra corrige os erros nos endereços de entrada quando possível. Use as regras quando você puder conectar os campos de endereço de entrada às portas de entrada de Várias Linhas na transformação do Validador de Endereço.
rule_USA_County_Validation	Valida nomes de municípios dos Estados Unidos. A regra compara os dados de entrada com os nomes de município em todos os estados. A regra retorna "Válido" ou "Inválido".
rule_USA_State_Validation	Valida os nomes de estado dos Estados Unidos. A regra retorna "Válido" ou "Inválido".
rule_USA_ZIPCode_Validation	Valida os códigos Zone Improvement Plan (ZIP) de cinco dígitos dos Estados Unidos. A regra retorna "Válido" ou "Inválido".

## Regras de Composição dos EUA/Canadá

Use as regras de composição no acelerador dos EUA/Canadá para adicionar um conjunto de regras para um mapeamento como um único objeto. Uma regra de composição é uma regra que faz uso da lógica de outras regras do acelerador.

As regras de composição se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Composite_Rules
```

A seguinte tabela descreve as regras de composição no acelerador dos EUA/Canadá:

Nome	Descrição
rule_CAN_Contact_Data	Analisa, padroniza e valida os dados de contato do Canadá, como endereços, números de telefone e Social Insurance Numbers (SIN).
rule_USA_Contact_Data	Analisa, padroniza e valida os dados de contato dos Estados Unidos, como endereços, números de telefone e Social Security Numbers (SSN).

### Regra de composição para dados de contato do Canadá

A regra `rule_CAN_Contact_Data` lê mapplets de várias pastas no repositório. A regra também inclui uma transformação não reutilizável.

A seguinte tabela lista os nomes e as localizações de repositório das regras e da transformação em `rule_CAN_Contact_Data`:

Regra	Localização
Case_Converter	Transformação não reutilizável
rule_Assign_DQ_Mailability_Score_Description	Informatica_DQ_Content\Rules\General_Data_Cleansing

Regra	Localização
rule_CAN_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_CAN_Gender_Assignment	Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_CAN_Multi_Person_Name_Parse	Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_CAN_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_CAN_Phone_Number_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_CAN_SIN_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_CAN_SIN_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Prenome_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Salutation_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

### Regra de composição para dados de contato dos Estados Unidos

A regra `rule_USA_Contact_Data` lê mapplets de várias pastas no repositório. A regra também inclui uma transformação não reutilizável.

A seguinte tabela lista os nomes e as localizações de repositório das regras e da transformação em `rule_USA_Contact_Data`:

Regra	Localização
Case_Converter	Transformação não reutilizável
rule_Assign_DQ_Mailability_Score_Description	[Informatica_DQ_Content]\Rules\General_Data_Cleansing
rule_Company_Name_Standardization	[Informatica_DQ_Content]\Rules\Corporate_Data_Cleansing
rule_Email_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Prenome_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_Salutation_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_USA_Address_Validation_Hybrid	[Informatica_DQ_Content]\Rules\Address_Data_Cleansing
rule_USA_Gender_Assignment	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_USA_Multi_Person_Name_Parse	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_USA_Phone_Number_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_USA_Phone_Number_Validation	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

Regra	Localização
rule_USA_SSN_Standardization	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing
rule_USA_SSN_Validation_post_June2011	[Informatica_DQ_Content]\Rules>Contact_Data_Cleansing

## Regras de Limpeza de Dados de Contato dos EUA/Canadá

Use as regras de limpeza de dados de contato para analisar, padronizar e validar dados sobre contatos comerciais e indivíduos.

As regras de limpeza de dados de contato se encontram na seguinte localização do repositório:

[Informatica\_DQ\_Content]\Rules>Contact\_Data\_Cleansing

A tabela a seguir descreve as regras de limpeza de dados de contato no acelerador dos EUA/Canadá:

Nome	Descrição
rule_CAN_Gender_Assignment	Atribui o sexo de acordo com os nomes. A regra retorna "M" para nomes masculinos, "F" para nomes feminino e "U" quando o sexo é desconhecido. Por exemplo, a regra atribui a "John Smith" o sexo "M" de masculino.
rule_CAN_Given_Name_Standard	Gera determinados nomes com base em nomes canadenses. Por exemplo, a regra padroniza o nome "Bob" como o nome "Robert".
rule_CAN_Multi_Person_Name_Parse	<p>Analisa os valores de nome de pessoa em portas diferentes. A regra cria portas para valores como título, nome, nome do meio e sobrenome.</p> <p>A saída da regra inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p> <p>Quando os dados de nome identificam mais de uma pessoa, a regra cria uma porta de saída para cada nome completo. Por exemplo, a regra pode ler o nome "John and Jane Smith" e criar portas de saída para "John Smith" e "Jane Smith".</p>
rule_CAN_Personal_Name_Parse_and_Standardize_FML	<p>Analisa os valores em um nome de pessoa em portas diferentes. A regra também padroniza os valores de nome.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Nome, nome do meio e sobrenome</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p>



Nome	Descrição
rule_CAN_Personal_Name_Parse_and_Standardize_LFM	<p>Analisa os valores em um nome de pessoa em portas diferentes. A regra também padroniza os valores de nome.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Sobrenome, nome e nome do meio</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p>
rule_CAN_Personal_Name_Parse_and_Standardize_FML	<p>Analisa os valores em um nome de pessoa em portas diferentes.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Sobrenome, nome e nome do meio</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p> <p><b>Nota:</b> A regra não padroniza os valores de nome. Para padronizar e analisar os valores de nomes canadenses na sequência definida pela regra, selecione rule_CAN_Personal_Name_Parse_and_Standardize_FML.</p>
rule_CAN_Personal_Name_Parse_and_Standardize_LFM	<p>Analisa os valores em um nome de pessoa em portas diferentes.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Sobrenome, nome e nome do meio</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p> <p><b>Nota:</b> A regra não padroniza os valores de nome. Para padronizar e analisar os valores de nomes canadenses na sequência definida pela regra, selecione rule_CAN_Personal_Name_Parse_and_Standardize_LFM.</p>
rule_CAN_Phone_Number_Parse	<p>Analisa um número de telefone canadense em uma cadeia. A regra analisa o primeiro número de telefone dos dados, lendo da direita para a esquerda. A regra retorna um número de telefone e também retorna uma string que contém o texto de entrada com o número de telefone removido.</p>
rule_CAN_Phone_Number_Standardization	<p>Padroniza os números de telefone canadenses. A regra retorna o número de telefone nos seguintes formatos:</p> <ul style="list-style-type: none"> <li>- Padrão - (nnn) nnn-nnnn</li> <li>- Hífens - nnn-nnn-nnnn</li> <li>- Sem Espaços - nnnnnnnnnn</li> </ul>
rule_CAN_Phone_Number_Validation	<p>Valida o código de área e o tamanho dos números de telefone canadenses. A regra retorna códigos que indicam o tipo e a validade do número de telefone. Tipos de descrição de categorias, como "toll-free".</p>
rule_CAN_Phone_Parse_Standardize_Validate	<p>Analisa um número de telefone de uma string de texto e verifica se o código de área é válido para o Canadá.</p> <p>Se o código de área for válido, a regra retornará o número de telefone em três formatos padrão. A regra também retorna um valor de status para indicar se os dados estão em conformidade com o formato padrão de um número de telefone do Canadá.</p>

Nome	Descrição
rule_CAN_Phone_w_Extension_Parse	Analisa um número de uma string de texto quando o número está em conformidade com o formato padrão de um número de telefone do Canadá. A regra inclui todos os dados de extensão telefônica quando analisa o número de telefone.
rule_CAN_SIN_Parse	Analisa um Social Insurance Number (SIN) canadense em uma cadeia. A regra retorna o SIN e também retorna uma cadeia que contém o texto de entrada com o SIN removido.
rule_CAN_SIN_Standardization	Padroniza os Social Insurance Numbers (SIN) canadenses. A regra pode ter os seguintes formatos de saída: <ul style="list-style-type: none"> <li>- Sem Pontuação - nnnnnnnnn</li> <li>- Espaço - nnn nnn nnn</li> <li>- Traço - nnn-nnn-nnn</li> </ul> Para alterar o formato, edite a variável de expressão SIN_Format na transformação de Expressão dq_Format_SIN. O padrão é "No_Punctuation".
rule_CAN_SIN_Validation	Valida os Social Insurance Numbers (SIN) canadenses. A regra usa o algoritmo de Luhn para verificar se um SIN é válido ou não. A regra retorna "Válido" ou "Inválido".
rule_Prenome_Assignment	Gera um título de acordo com o sexo. Você pode alterar a variável de expressão female_prenome de "Sra." para "Srta.".
rule_Salutation_Assignment	Gera saudações formais e casuais com base em nomes e tokens de nome. Por exemplo, quando os dados de entrada contêm "Sr. John Smith", a regra gera a saudação formal "Prezado Sr. Smith," e a saudação casual "Prezado John,". É possível alterar o prefixo e a pontuação editando as variáveis na transformação de Expressão dq_Generate_Salutation.
rule_USA_Gender_Assignment	Atribui o sexo de acordo com o nome. A regra retorna "M" para nomes masculinos, "F" para nomes feminino e "U" quando o sexo é desconhecido. Por exemplo, a regra atribui a "John Smith" o sexo "M" de masculino.
rule_USA_Given_Name_Standard	Gera determinados nomes com base em nomes dos Estados Unidos. Por exemplo, a regra padroniza o nome "Bob" como o nome "Robert".
rule_USA_Multi_Person_Name_Parse	Analisa os valores de nome de pessoa em portas diferentes. A regra cria portas para valores como título, nome, nome do meio e sobrenome. A saída da regra inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade. Quando os dados de nome identificam mais de uma pessoa, a regra cria uma porta de saída para cada nome completo. Por exemplo, a regra pode ler o nome "John and Jane Smith" e criar portas de saída para "John Smith" e "Jane Smith".
rule_USA_Personal_Name_Parse_and_Standardize_FML	Analisa os valores em um nome de pessoa em portas diferentes. A regra também padroniza os valores de nome. A regra cria as portas na seguinte sequência: <ul style="list-style-type: none"> <li>- Nome, nome do meio e sobrenome</li> </ul> A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.

Nome	Descrição
rule_USA_Personal_Name_Parse_and_Standardize_LFM	<p>Analisa os valores em um nome de pessoa em portas diferentes. A regra também padroniza os valores de nome.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Sobrenome, nome e nome do meio</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p>
rule_USA_Personal_Name_Parse_Validation	<p>Valida a atribuição de gênero para um nome. A regra calcula as probabilidades de que um valor de dados seja um nome masculino ou um nome feminino. Se o gênero for desconhecido, a regra usará os cálculos de probabilidade para atribuir um gênero ao nome.</p>
rule_USA_Personal_Name_Parsing_FML	<p>Analisa os valores em um nome de pessoa em portas diferentes.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Nome, nome do meio e sobrenome</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p> <p><b>Nota:</b> A regra não padroniza os valores de nome. Para padronizar e analisar os valores de nomes estadunidenses na sequência definida pela regra, selecione rule_USA_Personal_Name_Parse_and_Standardize_FML .</p>
rule_USA_Personal_Name_Parsing_LFM	<p>Analisa os valores em um nome de pessoa em portas diferentes.</p> <p>A regra cria as portas na seguinte sequência:</p> <ul style="list-style-type: none"> <li>- Sobrenome, nome e nome do meio</li> </ul> <p>A saída da regra também inclui uma porta que contém o nome completo da pessoa no registro. Você pode usar a porta do nome completo como uma entrada para uma transformação de Correspondência em um mapeamento de análise de correspondência de identidade.</p> <p><b>Nota:</b> A regra não padroniza os valores de nome. Para padronizar e analisar os valores de nomes estadunidenses na sequência definida pela regra, selecione rule_USA_Personal_Name_Parse_and_Standardize_LFM.</p>
rule_USA_Phone_Number_Parse	<p>Analisa um número de telefone americano em uma cadeia. A regra analisa o primeiro número de telefone dos dados, lendo da direita para a esquerda. A regra retorna um número de telefone e também retorna uma string que contém o texto de entrada com o número de telefone removido.</p>
rule_USA_Phone_Number_Standardization	<p>Padroniza os números de telefone americanos. A regra retorna o número de telefone nos seguintes formatos:</p> <ul style="list-style-type: none"> <li>- Padrão - (nnn) nnn-nnnn</li> <li>- Hífens - nnn-nnn-nnnn</li> <li>- Sem Espaços - nnnnnnnnnn</li> </ul>
rule_USA_Phone_Number_Validation	<p>Valida o código de área e o tamanho dos números de telefone americanos. Essa regra retorna códigos que indicam se o código de área e o comprimento de um número de telefone são válidos.</p>

Nome	Descrição
rule_USA_Phone_Parse_Standardize_Validate	<p>Analisa um número de telefone de uma string de texto e verifica se o código de área é válido para os Estados Unidos.</p> <p>Se o código de área for válido, a regra retornará o número de telefone em três formatos padrão. A regra também retorna um valor de status para indicar se os dados estão em conformidade com o formato padrão de um número de telefone dos Estados Unidos.</p>
rule_USA_Phone_w_Extension_Parse	<p>Analisa um número de uma string de texto quando o número está em conformidade com o formato padrão de um número de telefone dos Estados Unidos. A regra inclui todos os dados de extensão telefônica quando analisa o número de telefone.</p>
rule_USA_SSN_Parse	<p>Analisa os Social Security Numbers (SSN) dos Estados Unidos.</p>
rule_USA_SSN_Parse_Standardize_and_Validate	<p>Analisa, padroniza e valida os Social Security Numbers dos Estados Unidos de uma string maior de texto. A regra pode analisar números que incluam ou omitam traços.</p> <p>Por padrão, a regra grava os Social Security Numbers sem nenhuma pontuação. Para alterar o formato de padronização, abra a transformação dq_SSN_Format na regra e atualize a expressão na porta SSN_Format.</p>
rule_USA_SSN_Standardization	<p>Padroniza os Social Security Numbers (SSN) dos Estados Unidos. A regra pode ter os seguintes formatos de saída:</p> <ul style="list-style-type: none"> <li>- Sem Pontuação - nnnnnnnnn</li> <li>- Espaço - nnn nnn nnn</li> <li>- Traço - nnn-xxx-xxxx</li> </ul> <p>Para alterar o formato, edite a variável de expressão SSN_format na transformação de Expressão dq_SSN_Format. O padrão é "No_Punctuation".</p>
rule_USA_SSN_Validation	<p>Valida os Social Security Numbers (SSN) dos Estados Unidos. A regra valida cada SSN com relação ao tamanho, aos valores numéricos e aos valores mínimos e máximos nas seções de Área, Grupo e Número de Série.</p> <p>A seção Área abrange os primeiros três dígitos do SSN e a seção Grupo abrange o quarto e o quinto dígitos. A seção Número de Série abrange os quatro dígitos finais.</p> <p>Se o SSN tiver sido emitido antes de junho 2011, a regra também verificará se o valor da área e o valor do Grupo são uma combinação válida. A regra não verifica se o SSN é um número emitido. A regra retorna "Válido" ou "Inválido".</p>
rule_USA_SSN_Validation_post_June2011	<p>Valida os Social Security Numbers (SSN) dos Estados Unidos. A regra valida cada SSN com relação ao tamanho, aos valores numéricos e aos valores mínimos e máximos nas seções de Área, Grupo e Número de Série.</p> <p>A seção Área abrange os primeiros três dígitos do SSN e a seção Grupo abrange o quarto e o quinto dígitos. A seção Número de Série abrange os quatro dígitos finais.</p> <p>A regra não verifica se o valor de Área e o valor de Grupo são uma combinação válida. A regra não verifica se o SSN é um número emitido. A regra retorna "Válido" ou "Inválido".</p>

## Dependências de Regras de Limpeza de Dados de Contato Principais

O acelerador dos EUA/Canadá depende das regras de limpeza de dados de contato a seguir do acelerador Principal:

- rule\_Email\_Validation

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados do Contato Principal” na página 22.](#)

## Regras de Limpeza de Dados Corporativos dos EUA/Canadá

Use as regras de limpeza de dados corporativos no acelerador dos EUA/Canadá para analisar, padronizar e validar dados corporativos.

A seguinte tabela descreve as regras de limpeza de dados corporativos no acelerador dos EUA/Canadá:

Nome	Descrição
rule_NAICS_Code_Validation	Valida os códigos NAICS (North American Industry Classification System).
rule_USA_SIC_Code_Validation	Valida os códigos SIC (Classificação Industrial Padrão).

O acelerador dos EUA/Canadá depende da seguinte regra de limpeza de dados corporativos do acelerador Principal:

- rule\_Company\_Name\_Standardization

## Regras de Limpeza de Dados Gerais dos EUA/Canadá

Use as regras de limpeza de dados gerais para identificar os tipos de informações contidas nos campos de entrada.

As regras de limpeza de dados gerais se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\General_Data_Cleansing
```

A tabela a seguir descreve as regras de limpeza de dados gerais no acelerador dos EUA/Canadá:

Nome	Descrição
rule_CAN_Field_Identification	Identifica o tipo de informação contida em um campo de entrada. A regra pode identificar nomes, IDs pessoais, nomes de empresa, datas e dados de endereço canadenses. A regra retorna um rótulo que descreve o tipo de dados de entrada. A regra usa dados de referência para identificar os tipos de informações.
rule_CAN_NER_Field_Identification	Identifica o tipo de informação contida em um campo de entrada. A regra pode identificar nomes, IDs pessoais, nomes de empresa, datas e dados de endereço canadenses. A regra retorna um rótulo que descreve o tipo de dados de entrada. A regra usa dados de referência para identificar os tipos de informações. A regra usa técnicas de correspondência probabilística para identificar os tipos de informações.

Nome	Descrição
rule_USA_Field_Identification	Identifica o tipo de informação contida em um campo de entrada. A regra pode identificar nomes, IDs pessoais, nomes de empresa, datas e dados de endereço dos Estados Unidos. A regra retorna um rótulo que descreve o tipo de dados de entrada. A regra usa dados de referência para identificar os tipos de informações.
rule_Field_North_American_Data	Identifica os seguintes tipos de campos: nome, cargo, empresa, endereço, cidade, estado ou província, CEP, país, identificação pessoal, e-mail, telefone, cartão de crédito e data.  A regra gera uma pontuação que indica o grau de confiança na identificação do campo. As pontuações mais altas indicam níveis de confiança maiores.  Se a regra não puder atribuir um tipo de campo, a regra gravará os dados na porta Out_Undetermined.
rule_USA_NER_Field_Identification	Identifica o tipo de informação contida em um campo de entrada. A regra pode identificar nomes, IDs pessoais, nomes de empresa, datas e dados de endereço dos Estados Unidos. A regra retorna um rótulo que descreve o tipo de dados de entrada. A regra usa dados de referência para identificar os tipos de informações. A regra usa técnicas de correspondência probabilística para identificar os tipos de informações.

## Dependências de Regras de Limpeza de Dados Gerais Principais

O acelerador dos EUA/Canadá depende das regras de limpeza de dados gerais do acelerador Principal a seguir:

- rule\_Assign\_DQ\_GeocodinStatus\_Description
- rule\_Assign\_DQ\_Mailability\_Score\_Description
- rule\_Assign\_DQ\_Match\_Code\_Descriptions
- rule\_Date\_Validation
- rule\_Remove\_Extra\_Spaces
- rule\_Remove\_Punctuation
- rule\_Replace\_Limited\_Punct\_with\_Space
- rule\_UpperCase

Para obter mais informações sobre essas regras, consulte [“Regras de Limpeza de Dados Gerais Principais” na página 23](#).

## Regras de Eliminação de Duplicação e Correspondência dos EUA/Canadá

Use as regras de correspondência e de eliminação de duplicação para medir os níveis de semelhança entre os registros em conjuntos de dados.

As regras de correspondência e de eliminação de duplicação se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules\Matching_Deduplication
```

A tabela a seguir descreve as regras de correspondência e de eliminação de duplicação no acelerador dos EUA/Canadá:

Nome	Descrição
mplt_CAN_IMO_Company_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados canadenses com base em nomes e endereços de empresas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_CAN_IMO_Famillyname_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas em dados canadenses com base em nomes e endereços de família. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_CAN_IMO_Individual_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados canadenses com base em nomes e endereços de pessoas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_CAN_IMO_Personal_Name_and_Data_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados canadenses com base em nomes de pessoas e em dados pessoais. Os campos da coluna de dados pessoais devem conter um único tipo de dados, como o número de telefone, o e-mail ou o Social Insurance Number. O mapplet gera chaves de grupo com base em dados pessoais.
mplt_Company_Name_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de empresa e em dados de endereço. O mapplet usa uma combinação de caracteres dos valores de nome de empresa e dos valores de código postal para gerar chaves de grupo.
mplt_Company_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base nos nomes de empresa. O mapplet gera os códigos Soundex dos valores de nome de empresa e usa os códigos Soundex como chaves de grupo.
mplt_Famillyname_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em sobrenomes e em dados de endereço. O mapplet usa uma combinação de caracteres dos valores de sobrenome e dos valores de código postal para gerar chaves de grupo.
mplt_Firstname_and_SSN_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes e nos números de Social Security dos Estados Unidos. O mapplet gera chaves de grupo dos dados de número de Social Security.
mplt_Individual_Name_and_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas e em dados de endereços americanos. O mapplet usa uma combinação de caracteres dos valores de sobrenome e dos valores de código postal para gerar chaves de grupo.
mplt_Individual_Name_and_Date_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas e em dados de data. O mapplet gera chaves de grupo com base em dados de data.
mplt_Individual_Name_and_Email_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas e em endereços de e-mail. O mapplet gera chaves de grupo com base em dados de endereço de e-mail.
mplt_Individual_Name_and_Phone_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas e em números de telefone. O mapplet gera chaves de grupo dos dados de número de telefone.

Nome	Descrição
mplt_Individual_Name_and_SSN_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base em nomes de pessoas e nos números de Social Security dos Estados Unidos. O mapplet gera chaves provenientes dos dados de número de Social Security.
mplt_Individual_Name_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas com base nos nomes de pessoas. O mapplet gera códigos NYSIIS dos valores de sobrenome e usa os códigos NYSIIS como chaves de grupo.
mplt_USA_Address_Match	Usa estratégias de correspondência de campo para identificar linhas duplicadas de dados americanos com base em dados de endereços americanos. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_USA_IMO_Company_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados americanos com base em nomes e endereços de empresas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_USA_IMO_Famillyname_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados americanos com base em nomes e endereços de família. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_USA_IMO_Individual_Name_and_Address_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados americanos com base em nomes e em endereços de pessoas. O mapplet gera chaves de grupo com base nos dados do código postal.
mplt_USA_IMO_Personal_Name_and_Data_Match	Usa estratégias de correspondência de identidade para identificar linhas duplicadas de dados americanos com base em nomes de pessoas e em dados pessoais. Os campos da coluna de dados pessoais devem conter um único tipo de dados, como o número de telefone, o e-mail ou o número de Social Security. O mapplet gera chaves de grupo com base em dados pessoais.
rule_Company_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes de empresa e em dados de endereços americanos.
rule_Company_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de empresa.
rule_Famillyname_and_Address_MatchScore	Gera uma pontuação de correspondência com base em sobrenomes e em dados de endereços americanos.
rule_Firstname_and_SSN_MatchScore	Gera uma pontuação de correspondência com base em nomes e em dados de endereços americanos.
rule_Individual_Name_and_Address_MatchScore	Gera uma pontuação de correspondência com base em nomes de pessoas e em dados de endereços americanos.
rule_Individual_Name_and_Date_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nas datas.
rule_Individual_Name_and_Email_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos endereços de e-mail.
rule_Individual_Name_and_Phone_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas e nos números de telefone.



Nome	Descrição
rule_Individual_Name_and_SSN_MatchScore	Gera uma pontuação de correspondência com base em nomes de pessoas, números de Social Security e dados de identificação.
rule_Individual_Name_MatchScore	Gera uma pontuação de correspondência com base nos nomes de pessoas.
rule_USA_Address_MatchScore	Gera uma pontuação de correspondência com base em dados de endereços americanos.

## Mapeamentos Demonstrativos dos EUA/Canadá

Os mapeamentos demonstrativos do acelerador dos EUA/Canadá usam várias regras para demonstrar processos de qualidade de dados.

Os mapeamentos demonstrativos se encontram na seguinte localização do repositório:

```
[Informatica_DQ_Content]\Rules_Demo\US_Canada_Accelerator
```

O acelerador dos EUA/Canadá inclui os seguintes mapeamentos demonstrativos:

### **m\_customer\_data\_US\_demo**

Analisa, padroniza e valida dados dos Estados Unidos e do Canadá.

### **m\_customer\_matching\_US\_demo**

Analisa e padroniza os dados de identidade dos EUA/Canadá e realiza a análise de correspondência de identidade nos dados.

O mapeamento analisa as seguintes combinações de dados e gera clusters de correspondência para cada combinação:

- Dados de nome e endereço da pessoa
- Nome e número de telefone da pessoa