



Informatica®
10.5.7

Glossário

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.

Informatica e o logotipo Informatica 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.

Sujeito aos seus direitos de descadastramento, o software transmitirá automaticamente para a Informatica nos EUA informações sobre o ambiente de rede e computação no qual o software é implantado, bem como sobre o uso de dados e as estatísticas do sistema da implantação. Essa transmissão é considerada parte dos Serviços conforme a política de privacidade da Informatica, e a Informatica usará e processará essas informações de acordo com a política de privacidade da Informatica, disponível em <https://www.informatica.com/in/privacy-policy.html>. Você pode desativar a coleção de uso na ferramenta Administrator.

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.

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 Curl com o Copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://curl.haxx.se/docs/copyright.html>. É permitido usar, copiar, modificar e distribuir este software com qualquer objetivo, com ou sem taxa, desde que a nota de direitos autorais acima e esta nota de permissão apareçam em todas as cópias. O produto inclui software copyright 2001-2005 (©) MetaStuff, Ltd. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://www.dom4j.org/license.html>.

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

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

Este produto inclui o 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>.

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

As informações contidas neste documento estão sujeitas a alteração sem aviso prévio. Se você encontrar quaisquer problemas nesta documentação, informe-os em infa_documentation@informatica.com.

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.

Partes deste software e/ou documentação estão sujeitas a copyright detido por terceiros. Os avisos de terceiros necessários são incluídos no produto.

Data da Publicação: 2025-02-26

Conteúdo

Prefácio.....	4
Recursos da Informatica.	4
Informatica Network.	4
Base de Dados de Conhecimento da Informatica.	4
Documentação da Informatica.	4
Matrizes de Disponibilidade de Produto da Informatica.	5
Informatica Velocity.	5
Informatica Marketplace.	5
Suporte Global a Clientes da Informatica.	5
Apêndice A: Glossário.....	6

Prefácio

Consulte o *Glossário Informatica®* para saber mais sobre a terminologia dos produtos Informatica que você pode usar com o Informatica Administrator, o Informatica Analyst e o Informatica Developer.

Recursos da Informatica

A Informatica oferece uma variedade de recursos de produtos através da Rede da Informatica e outros portais on-line. Use os recursos para obter o máximo de seus produtos e soluções da Informatica e para aprender com outros usuários da Informatica e especialistas no assunto.

Informatica Network

A Informatica Network é a porta de entrada para muitos recursos, incluindo a Base de Dados de Conhecimento da Informatica e o Suporte Global a Clientes da Informatica. Para acessar a Informatica Network, visite <https://network.informatica.com>.

Como membro da Informatica Network, você tem as seguintes opções:

- Pesquisar por recursos do produto na Base de Dados de Conhecimento.
- Visualizar informações sobre disponibilidade de produtos.
- Criar e 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 encontrar recursos de produtos, como artigos de instruções, práticas recomendadas, tutoriais em vídeo e respostas a perguntas frequentes.

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

Documentação da Informatica

Use o Portal de Documentação da Informatica para explorar uma extensa biblioteca de documentação para versões de produtos atuais e recentes. Para explorar o Portal de Documentação, visite <https://docs.informatica.com>.

Em caso de dúvidas, comentários ou ideias sobre a documentação do produto, entre em contato com a equipe da Documentação da Informatica em infa_documentation@informatica.com.

Matrizes de Disponibilidade de Produto da Informatica

As Matrizes de Disponibilidade de Produto (PAMs) indicam as versões dos sistemas operacionais, os bancos de dados e tipos de fontes e destinos de dados com os quais uma versão de produto é compatível. Veja as PAMs da Informatica 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 e baseada em experiências reais de centenas de projetos de gerenciamento de dados. O Informatica Velocity representa o conhecimento coletivo dos consultores da Informatica que trabalham com organizações em todo o mundo para planejar, desenvolver, implantar e manter soluções de gerenciamento de dados bem-sucedidas.

Encontre os recursos do Informatica Velocity em <http://velocity.informatica.com>. Se você tiver dúvidas, comentários ou ideias sobre o Informatica Velocity, entre em contato com os Serviços Profissionais da Informatica em ips@informatica.com.

Informatica Marketplace

O Informatica Marketplace é um fórum onde você pode encontrar soluções que ampliam e aprimoram suas implementações da Informatica. Aproveite as centenas de soluções dos desenvolvedores e parceiros da Informatica no Marketplace para melhorar sua produtividade e agilizar o tempo de implementação em seus projetos. Encontre o Informatica Marketplace em <https://marketplace.informatica.com>.

Suporte Global a Clientes da Informatica

Você pode entrar em contato com um Centro de Suporte Global por telefone ou por meio da 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: <https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

Para encontrar recursos de suporte on-line na Rede da Informatica, visite <https://network.informatica.com> e selecione a opção eSupport.

APÊNDICE A

Glossário

ambiente de tempo de execução

O ambiente que você configura para executar um mapeamento ou um perfil. O ambiente de tempo de execução pode ser nativo ou o Hive.

ambiente de validação

O ambiente que você configura para validar um mapeamento ou um perfil. Você valida um mapeamento ou um perfil para garantir que ele possa ser executado em um ambiente de tempo de execução. O ambiente de validação pode ser Hive e/ou nativo.

Ambiente Hadoop

Um ambiente que você pode configurar para executar um mapeamento ou um perfil em um Cluster Hadoop. Você deve configurar o Hadoop como o ambiente de validação e de tempo de execução.

ambiente não nativo

Um cluster distribuído fora do domínio Informatica, como o Hadoop ou o Databricks, em que o Serviço de Integração de Dados pode enviar por push o processamento em tempo de execução. Se o Serviço de Integração de Dados não enviar o processamento para um ambiente não nativo, ele processará o trabalho no ambiente nativo do domínio Informatica.

ambiente nativo

O ambiente padrão no domínio Informatica que executa um mapeamento, um fluxo de trabalho ou um perfil. O Serviço de Integração de Dados realiza o processo de extração, transformação e carregamento de dados. Como alternativa, o Serviço de Integração de Dados pode enviar o processamento para um ambiente não nativo, que é um cluster distribuído fora do domínio Informatica, como o Hadoop ou o Databricks.

aplicativo

Um objeto que você pode implantar em um Serviço de Integração de Dados. Um objeto implantável que pode conter objetos de dados, mapeamentos, serviços de dados SQL, serviços da Web e fluxos de trabalho.

aplicativo completo

Um aplicativo que você deve reimplantar no Serviço de Integração de Dados após a edição de objetos de aplicativo.

aplicativo de tempo de execução

Um aplicativo que é implantado em um Serviço de Integração de Dados e contém instâncias de tempo de execução de objetos de aplicativo.

Aplicativo em tempo de design

Um aplicativo que você edita na Developer tool. O aplicativo contém instâncias de tempo de design de objetos de aplicativo.

aplicativo incremental

Um aplicativo que você pode atualizar implantando um patch de aplicativo no Serviço de Integração de Dados.

arquivo de esquema XSD

Um arquivo *.xsd contendo um esquema XML que define os elementos, os atributos e a estrutura de documentos XML.

audit

Um conjunto de regras e condições que validam os dados na origem e no destino de um mapeamento. As regras de auditoria calculam um valor agregado para uma coluna de dados na origem ou no destino. As condições de auditoria comparam várias regras de auditoria ou regras de auditoria e valores constantes.

banco de dados virtual

Um serviço de dados SQL que você pode consultar. Ele contém objetos virtuais e fornece uma exibição uniforme dos dados a partir de origens de dados distintas e heterogêneas.

biblioteca de definições de tipo

Um objeto no repositório do Modelo que armazena as definições de tipo de dados complexo para um mapeamento ou mapplet.

big data

Um conjunto de dados que é tão grande e complexo que não pode ser processado por meio de ferramentas de gerenciamento de banco de dados padrão.

cache de conjunto de resultados

Um cache que contém os resultados de cada consulta de serviço de dados SQL ou solicitação de serviço da Web. Com o cache de conjunto de resultados, o Serviço de Integração de Dados retorna resultados em cache quando os usuários executam consultas idênticas. O cache de conjunto de resultados reduz o tempo de execução para consultas idênticas.

chave candidata

Uma coluna ou um conjunto de colunas que identifica exclusivamente cada linha de origem em uma tabela de banco de dados.

chave documentada

Uma chave primária declarada no banco de dados de origem.

chave inferida

Uma coluna ou um conjunto de colunas que a ferramenta Analyst ou Desenvolvedor infere como uma chave candidata com base nos dados da coluna.

cluster Hadoop

Um cluster de máquinas que está configurado para executar serviços e aplicativos Hadoop. Um cluster Hadoop típico inclui um nó mestre e vários nós de trabalho. O nó mestre executa os daemons mestre JobTracker e NameNode. Um nó do funcionário executa os daemons DataNode e TaskTracker. Em pequenos clusters, o nó mestre também pode executar os daemons de funcionário.

coluna dependente

Em uma dependência funcional, a coluna contendo valores que são determinados por uma coluna determinante.

coluna determinante

Em uma dependência funcional, um conjunto de colunas que determina o valor da coluna dependente. Se o determinante tem zero colunas, o dependente é uma constante.

componente de inicialização

O componente executável iniciado em primeiro lugar pelo Data Transformation quando ele executa uma Transformação de Processador de Dados.

CompressionCodec

Interface de compactação Hadoop. Um codec é a implementação de um algoritmo de compactação/descompactação. No Hadoop, um codec é representado por uma implementação da interface CompressionCodec.

configuração de tipo

Um conjunto de propriedades de porta complexa que especificam o tipo de dados dos elementos de tipo de dados complexo ou o esquema dos dados.

contêiner

Uma alocação de memória e recursos de CPU em um nó com a função de computação. Um serviço de aplicativo usa o contêiner para realizar remotamente cálculos no nó. Por exemplo, um Serviço de Integração de Dados que é executado em uma grade pode executar remotamente um mapeamento em um contêiner de dados de um nó com a função de cálculo.

conteúdo de e-mail dinâmico

Conteúdo de e-mail definido em um parâmetro ou variável de fluxo de trabalho.

correspondência direta

Em uma pesquisa global, uma correspondência direta é um ativo que corresponde a toda a consulta de pesquisa. Na pesquisa de descoberta, uma correspondência direta é uma correspondência com alguns ou todos os metadados do ativo que correspondem à consulta de pesquisa.

correspondência indireta

Uma correspondência nos resultados de pesquisas de descoberta que está vinculada ao ativo que corresponde diretamente a uma parte da consulta de pesquisa ou a toda ela.

curadoria

O processo de validação e gerenciamento dos metadados descobertos em uma fonte de dados de forma que os metadados estejam prontos para uso e preparação de relatórios.

dados hierárquicos

Um conjunto de dados que está relacionado hierarquicamente. O relacionamento hierárquico é representado como um esquema. As transformações do Informatica usam tipos de dados complexos para representar dados hierárquicos.

dados virtuais

As informações obtidas quando você consulta tabelas virtuais ou executa procedimentos armazenados em um serviço de dados SQL.

DataNode

Um nó HDFS que armazena dados no Sistema de Arquivos Hadoop. Um cluster HDFS pode ter mais de um DataNode, com dados replicados entre eles.

Data Transformation Manager (DTM) de execução

O componente de computação do Serviço de Integração de Dados que extrai, transforma e carrega dados para concluir uma tarefa de transformação de dados.

definição de tipo de dados aninhado

Uma definição de tipo de dados complexo que referencia outras definições de tipo de dados complexo.

definição de tipo de dados complexo

Uma representação reutilizável do esquema dos dados que você referencia em uma porta de estrutura ou em uma porta complexa que contém elementos do tipo struct. Uma ou mais portas complexas podem usar a definição de tipo de dados complexo.

definição de tipo de dados recursivo

Uma definição de tipo de dados aninhado na qual uma das definições de tipo de dados complexo em qualquer nível é a mesma que a de um pai.

dependência funcional

O relacionamento entre um conjunto de colunas em uma determinada tabela, na qual a coluna determinante determina funcionalmente a coluna dependente.

descoberta de chave externa

O processo que localiza colunas em uma fonte de dados que correspondem às colunas de chave primária na fonte de dados pai.

descoberta de chave primária

O processo para identificar uma coluna ou combinação de colunas que identificam exclusivamente uma linha em uma fonte de dados.

descoberta de dependência funcional

O processo que localiza relacionamentos de dependência funcional entre as colunas de uma fonte de dados.

Descoberta de Estrutura Inteligente

Um serviço no Informatica Intelligent Cloud Services que usa algoritmos de machine learning para identificar automaticamente os tipos de informações em um arquivo de entrada. A Descoberta de Estrutura Inteligente cria uma estrutura inteligente, um modelo do padrão e tipos de campos de dados descobertos no arquivo. O modelo inclui elementos simples e elementos complexos, dependendo da organização dos dados no arquivo.

descoberta do domínio de dados

O processo que identifica todos os domínios de dados associados a uma coluna com base nos valores ou no nome da coluna.

descoberta empresarial

O processo que localiza estatísticas de perfil de coluna, domínios de dados, chaves primárias e chaves externas em um grande número de fontes de dados distribuídas por várias conexões ou esquemas.

desenvolvimento baseado em equipe

A colaboração de membros da equipe em um projeto de desenvolvimento. A colaboração inclui funcionalidades como a criação de versão por meio de check-out e check-in de objetos do repositório.

destinatário

Um usuário ou grupo no domínio Informatica que recebe uma notificação durante um fluxo de trabalho.

destinatário dinâmico

Um destinatário de notificação definido em um parâmetro ou variável de fluxo de trabalho.

destino dinâmico

Um arquivo simples ou destino relacional para um mapeamento que pode sofrer alterações em tempo de execução. Transformações de Gravação podem definir colunas de destino em tempo de execução com base no fluxo de mapeamento ou de um destino associado. Transformações de Gravação também podem soltar e substituir a tabela de destino em tempo de execução.

Distribuição da Cloudera Incluindo o Apache Hadoop (CDH)

Versão da Cloudera da estrutura de software Hadoop com código-fonte aberto.

documento de origem

Um documento que é a entrada de uma transformação de Processador de Dados.

documento de saída

Um documento que é o resultado de uma transformação de Processador de Dados.

domínio de dados

Um objeto de repositório do Modelo predefinido ou definido pelo usuário que representa o significado funcional de uma coluna com base nos dados da coluna ou no nome da coluna. Exemplos: número de Previdência Social, número do cartão de crédito e ID de e-mail.

endereço de e-mail dinâmico

Um endereço de e-mail definido em um parâmetro ou variável de fluxo de trabalho.

entrada de tarefa

Dados que são transmitidos para uma tarefa a partir de parâmetros e variáveis de fluxo de trabalho. A tarefa usa os dados de entrada para concluir uma unidade de trabalho.

especificação de ingestão em massa

Uma configuração que determina como os dados de origem de um banco de dados relacional são ingeridos em um local específico em um data lake ou em um cluster Hadoop.

esquema

Uma definição da estrutura de dados. As portas complexas do tipo de dados struct usam definições de tipo de dados complexo para representar o esquema.

esquema virtual

Um esquema em um banco de dados virtual que define a estrutura do banco de dados.

esquema XML

Uma definição dos elementos, dos atributos e da estrutura usados em documentos XML. O esquema está em conformidade com o padrão de Esquema XML do World Wide Web Consortium e é armazenado como um arquivo *.xsd.

estrutura inteligente

Um modelo do padrão e tipos de campos de dados que a Descoberta de Estrutura Inteligente identifica em um arquivo de entrada. O modelo inclui elementos simples e elementos complexos, dependendo da organização dos dados no arquivo.

evento

Um objeto de fluxo de trabalho que inicia ou termina o fluxo de trabalho. Um evento representa algo que acontece quando o fluxo de trabalho é executado. O editor exibe eventos como círculos.

evento de Processador de Dados

Uma ocorrência durante a execução de uma transformação de Processador de Dados.

Exceção

Uma exceção é um padrão, um valor ou uma frequência de uma coluna nos resultados do perfil que não estão dentro de um intervalo de valores esperado.

executor Blaze

Um componente do DTM que pode simplificar e converter um mapeamento em um plano de execução do Blaze executado em um cluster Hadoop.

executor do Hive

Um componente do DTM que pode simplificar e converter um mapeamento ou um perfil em um plano de execução do Hive executado em um cluster Hadoop.

exemplo de documento de origem

Uma amostra dos documentos que são processados pela transformação de Processador de Dados.

exibição virtual de dados

Um banco de dados virtual definido por um serviço de dados SQL que você pode consultar como se fosse um banco de dados físico.

expressão de predicado

Uma expressão que filtra os dados em um mapeamento. Uma expressão de predicado retorna "true" ou "false".

Ferramenta Informatica Monitoring

A ferramenta Informatica Monitoring (a ferramenta Monitoring) é um aplicativo que fornece um link direto para a guia Monitorar da ferramenta Administrator. A guia Monitorar mostra propriedades, estatísticas de tempo de execução e relatórios de tempo de execução sobre os objetos de integração que são executados em um Serviço de Integração de Dados.

fluxo de sequência

Um conector entre objetos de fluxo de trabalho que especifica a ordem em que o Serviço de Integração de Dados executa os objetos. O editor exibe fluxos de sequência como setas.

fluxo de sequência condicional

Um fluxo de sequência que inclui uma expressão que o Serviço de Integração de Dados avalia como "true" ou "false". Se a expressão for avaliada como "true", o Serviço de Integração de Dados executará o próximo objeto do fluxo de trabalho. Se a expressão for avaliada como "false", o Serviço de Integração de Dados não executará o próximo objeto do fluxo de trabalho.

fluxo de sequência padrão

O fluxo de sequência de saída a partir de um gateway Exclusivo que é sempre avaliado como "true". Quando todos os outros fluxos de sequência condicionais são avaliados como "false", o Serviço de Integração de Dados executa o objeto conectado ao fluxo de sequência de saída padrão.

fluxo de trabalho

Uma representação gráfica de um conjunto de eventos, tarefas e decisões que definem um processo de negócios. Use a ferramenta Developer para adicionar objetos a um fluxo de trabalho e conectar objetos a fluxos de sequência. O Serviço de Integração de Dados usa as instruções configuradas no fluxo de trabalho para executar os objetos.

função complexa

Um tipo de função predefinida na qual a entrada ou o tipo de retorno é do tipo de dados complexo.

função de nó

A finalidade de um nó. Um nó com a função de serviço pode executar os serviços de aplicativo. Um nó com a função de cálculo pode executar cálculos solicitados por serviços de aplicativo remotos. Um nó com as duas funções pode executar os serviços de aplicativo e realizar localmente cálculos para esses serviços.

função de usuário

Coleta de privilégios atribuídos a um usuário ou grupo. Você pode atribuir funções a usuários e grupos para o domínio e para alguns serviços de aplicativo nesse domínio.

gateway

Um objeto de fluxo de trabalho que divide e mescla caminhos no fluxo de trabalho com base na maneira como o Serviço de Integração de Dados avalia expressões em fluxos de sequência condicionais. O editor exibe gateways como losangos.

Gateway exclusivo

Um gateway que representa uma decisão tomada em um fluxo de trabalho. Quando um gateway Exclusivo divide o fluxo de trabalho, o Serviço de Integração de Dados toma uma decisão para usar uma das ramificações de saída. Quando um gateway Exclusivo mescla o fluxo de trabalho, o Serviço de Integração de Dados aguarda a conclusão de uma ramificação de entrada antes de disparar a ramificação de saída.

glossário do domínio de dados

Um contêiner para todos os domínios de dados e grupos de domínios de dados na ferramenta Analyst ou na ferramenta Developer.

grupo de métricas

Um grupo de métricas definidas pelo usuário.

grupo do domínio de dados

Uma coleção de domínios de dados sob uma categoria de domínios de dados específica.

Hive

Uma infraestrutura de data warehouse construída com base no Hadoop. O Hive oferece suporte a uma linguagem semelhante à SQL, denominada HiveQL, para sumarização, consulta e análise de dados.

ID da instância de fluxo de trabalho

Número que define exclusivamente uma instância de fluxo de trabalho executada.

implantar

Para tornar os objetos em um aplicativo acessíveis para os usuários finais. Dependendo dos tipos de objetos no aplicativo, os usuários finais podem executar consultas com base nesses objetos, acessar serviços da Web, executar mapeamentos ou executar fluxos de trabalho.

Informatica Administrator

O Informatica Administrator (a ferramenta Administrator) é um aplicativo que consolida as tarefas administrativas para objetos de domínio, como serviços, nós, licenças e grades. Gerencie o domínio e a segurança do domínio por meio da ferramenta Administrador.

Informatica Developer

O Informatica Developer (a ferramenta Developer) é um aplicativo usado para projetar soluções de integração de dados. O repositório do Modelo armazena os objetos criados na ferramenta Developer.

Informatica Mass Ingestion

O Informatica Mass Ingestion (a ferramenta de Ingestão em Massa) é um aplicativo que você pode usar para configurar, implantar, executar e monitorar especificações de ingestão em massa.

ingestão em massa

O movimento de grandes quantidades de dados entre um banco de dados relacional e um data lake ou um cluster Hadoop.

instância de fluxo de trabalho

A representação em tempo de execução de um fluxo de trabalho. Ao executar um fluxo de trabalho a partir de um aplicativo implantado, você executa uma instância desse fluxo de trabalho. Você pode executar simultaneamente várias instâncias do mesmo fluxo de trabalho.

instância do DTM

Uma representação lógica específica da transformação do Data Transformation Manager (DTM) que o Serviço de Integração de Dados cria para executar um trabalho. Com base em como você configura o Serviço de Integração de Dados, as instâncias do DTM podem ser executadas no processo do Serviço de Integração de Dados, em um processo DTM separado no nó local ou em um processo do DTM separado em um nó remoto.

JobTracker

Um serviço Hadoop que coordena tarefas de mapa e redução e que as programa para execução em TaskTrackers.

linhagem de scorecard

Um diagrama que mostra a origem dos dados, descreve o caminho e mostra como os dados fluem para uma métrica ou grupo de métricas em um scorecard. Na análise da linhagem de scorecard, as caixas ou nós representam objetos. As setas representam relacionamentos de fluxo de dados.

link em tempo de execução

Um link de grupo para grupo que usa uma diretiva e/ou um parâmetro para determinar quais portas devem ser vinculadas entre os grupos em tempo de execução.

logical Data Transformation Manager (LDTM)

Um componente de serviço do Serviço de Integração de Dados que otimiza e compila trabalhos e, em seguida, os envia para o Data Transformation Manager (DTM) de execução.

mapeamento

Um conjunto de entradas e saídas vinculadas por objetos de transformação que definem as regras para a transformação de dados.

mapeamento de grade

Um mapeamento do Informatica que o mecanismo Blaze compila e distribui em um cluster de nós.

mapeamento de gravação de objeto de dados lógicos

Um mapeamento que grava dados em destinos usando um objeto de dados lógicos como entrada. Ele contém um ou mais objetos de dados lógicos como a entrada e um objeto de dados físicos como o destino.

mapeamento de leitura de objeto de dados lógicos

Um mapeamento que fornece uma exibição dos dados por meio de um objeto de dados lógicos. Ele contém um ou mais objetos de dados físicos como origens e um objeto de dados lógicos como o mapeamento de saída.

mapeamento de objeto de dados lógicos

Um mapeamento que vincula um objeto de dados lógicos a um ou mais objetos de dados físicos. Ele pode incluir lógica de transformação.

mapeamento de operação

Um mapeamento que realiza a operação de serviço da Web para o cliente de serviço da Web. Um mapeamento de operação pode conter uma transformação de Entrada, uma transformação de Saída e várias transformações de Falha.

mapeamento de tabela virtual

Um mapeamento que contém uma tabela virtual como destino.

mapeamento dinâmico

Um mapeamento no qual você pode alterar as origens, os destinos e a lógica da transformação em tempo de execução com base em parâmetros e regras definidos. Você pode configurar mapeamentos dinâmicos para permitir alterações de metadados em origens e destinos. Você pode determinar quais portas são recebidas por uma transformação, quais portas devem ser usadas na lógica da transformação e quais links devem ser estabelecidos entre grupos de transformações.

mapplet

Um objeto reutilizável que contém um conjunto de transformações que você pode usar em vários mapeamentos ou validar como uma regra.

MapReduce

Um modelo de programação para o processamento de grandes volumes de dados em paralelo.

metastore

Um banco de dados usado pelo Hive para armazenar metadados das tabelas do Hive armazenadas no HDFS. Metastores podem ser locais, incorporados ou remotos.

métrica

Uma coluna de uma fonte de dados ou saída de uma regra que faz parte de um scorecard.

Minigráfico

Um minigráfico é um gráfico de linhas que exibe a variação em um valor nulo, exclusivo ou não exclusivo por todas as cinco execuções de perfil consecutivas mais recentes.

modelo de objeto de dados lógicos

Um modelo de dados que descreve os dados em uma organização e o relacionamento entre eles. Ele contém objetos de dados lógicos e define relacionamentos entre eles.

Módulo de Serviço de Fluxo de Trabalho

Um componente no Serviço de Integração de Dados que gerencia solicitações para executar fluxos de trabalho.

Módulo de Serviço SQL

O serviço de componente no Serviço de Integração de Dados que gerencia consultas SQL enviadas a um serviço de dados SQL a partir de uma ferramenta de cliente de terceiros.

Módulo do Serviço da Web

Um componente no Serviço de Integração de Dados que gerencia solicitações de operação de serviços da Web enviadas a um serviço da Web a partir de um cliente de serviço da Web.

NameNode

Um nó no cluster Hadoop que gerencia o espaço de nome do sistema de arquivos, mantém a árvore do sistema de arquivos e os metadados para todos os arquivos e diretórios da árvore.

nó

Uma representação de um nível na hierarquia de uma mensagem de serviço da Web.

objeto de dados físicos

Uma representação física de dados que é usada para operações de leitura, pesquisa ou gravação em recursos.

objeto de dados lógicos

Um objeto que descreve uma entidade lógica em uma organização. Ele possui atributos e chaves e descreve os relacionamentos entre atributos.

objeto de dados personalizado

Um objeto de dados físicos que usa um ou mais recursos relacionais ou objetos de dados relacionais como origens. Você usa um objeto de dados personalizado para realizar tarefas, como associar dados de recursos relacionados ou filtrar linhas. O objeto de dados personalizado usa uma única conexão e instrução SQL para as tabelas de origem.

operação de gravação do objeto de dados

O objeto de repositório que contém as propriedades necessárias para executar determinadas operações em tempo de execução em destinos. Uma operação de gravação do objeto de dados está associada a um objeto de dados de destino.

operação de leitura do objeto de dados

O objeto de repositório que contém as propriedades necessárias para executar determinadas operações em tempo de execução em origens. Uma operação de leitura do objeto de dados está associada a um objeto de dados de origem.

operador complexo

Um tipo de operador para fazer referência a nomes de elemento ou acessar elementos em um tipo de dados complexo.

origem dinâmica

Um arquivo simples ou origem relacional para um mapeamento que pode sofrer alterações em tempo de execução. Transformações de Leitura e Pesquisa podem obter alterações de definição ou metadados diretamente da origem. Se você usar um parâmetro para a origem, poderá alterá-la em tempo de execução.

otimização baseada em custos

Método de otimização que reduz o tempo de execução para mapeamentos que realizam operações de associação. Com a otimização baseada em custos, o Serviço de Integração de Dados cria diferentes planos para executar um mapeamento e calcula um custo para cada plano. O Serviço de Integração de Dados executa o plano com o menor custo. O Serviço de Integração de Dados calcula o custo com base em estatísticas de banco de dados, E/S, CPU, rede e memória.

otimização com projeção antecipada

Método de otimização que reduz a quantidade de dados que se move entre transformações no mapeamento. Usando a otimização com projeção antecipada, o Serviço de Integração de Dados identifica portas não utilizadas e remove os links entre as portas em um mapeamento.

otimização com seleção antecipada

Método de otimização que reduz o número de linhas que passam pelo mapeamento. Usando a otimização com seleção antecipada, o Serviço de Integração de Dados aproxima os filtros da origem do mapeamento no pipeline.

otimização com semi-associação

Método de otimização que reduz o número de linhas extraídas da origem. Usando a otimização com semi-associação, o Serviço de Integração de Dados modifica as operações de associação em um mapeamento. O Serviço de Integração de Dados aplica o método de otimização com semi-associação a uma transformação de Associador quando um grupo de entrada maior possui linhas que não correspondem a um grupo de entrada menor na condição de associação. O Serviço de Integração de Dados lê as linhas do grupo menor, localiza as linhas correspondentes no grupo maior e realiza a operação de associação.

otimização de empilhamento

Método de otimização que envia a lógica de transformação a um banco de dados de origem ou destino. Com a otimização de empilhamento, o Serviço de Integração de Dados converte a lógica de transformação em consultas SQL e envia essas consultas SQL ao banco de dados. O banco de dados executa as consultas SQL para processar os dados.

otimização de predicado

Método de otimização que simplifica ou regrava as expressões de predicado em um mapeamento. Com a otimização de predicado, o Serviço de Integração de Dados tenta aplicar expressões de predicado com a maior antecedência possível para aumentar o desempenho do mapeamento.

parâmetro de fluxo de trabalho

Um valor constante que você define antes da execução do fluxo de trabalho. Parâmetros mantêm o mesmo valor por toda a execução do fluxo de trabalho. Você define o valor do parâmetro em um arquivo de parâmetros. Todos os parâmetros de fluxo de trabalho são definidos pelo usuário.

particionamento

O processo de dividir os dados subjacentes em subconjuntos que podem ser executados em vários segmentos de processamento. Quando os administradores permitem que o Serviço de Integração de Dados maximize o paralelismo, o serviço aumenta o número de segmentos de processamento, o que pode otimizar o desempenho de mapeamento e criação de perfil.

pasta

Um contêiner para objetos no repositório do Modelo. Use pastas para organizar objetos em um projeto e crie pastas para agrupar objetos com base em necessidades de negócios.

patch de aplicativo

Um objeto que você pode implantar em um Serviço de Integração de Dados. Ele contém os metadados que você pode usar para atualizar um aplicativo incremental implantado.

perfil

Um objeto que contém regras para descobrir padrões em dados de origem. Execute um perfil para avaliar a estrutura de dados e verificar se as colunas de dados contêm o tipo de informações que você espera.

perfil de associação

Um tipo de perfil que determina o grau de sobreposição entre um conjunto de um ou mais colunas em uma fonte de dados e um conjunto similar na mesma fonte de dados ou em outra.

perfil de coluna

Um tipo de perfil que determina as características das colunas em uma fonte de dados, como frequência de valor, porcentagens, padrões e tipos de dados.

perfil de descoberta empresarial

Um tipo de perfil usado para executar a descoberta empresarial.

perfil de objeto de dados

Um objeto de repositório que define o tipo de análise a ser executado em uma fonte de dados.

perfil do sistema operacional

Um tipo de segurança que os Serviços de Integração de Dados no UNIX ou Linux usam para isolar o ambiente do usuário de tempo de execução. O perfil do sistema operacional contém o nome de usuário do sistema operacional, variáveis do processo de serviço, variáveis de ambiente e permissões. O Serviço de Integração de Dados executa mapeamentos, fluxos de trabalhos, trabalhos de criação de perfil e scorecards com as permissões do usuário do sistema operacional e as propriedades definidas no perfil do sistema operacional.

pesquisa de descoberta

Um tipo de pesquisa na ferramenta Analyst que identifica ativos com base em correspondências diretas à consulta de pesquisa, bem como relacionamentos com outros objetos que correspondem à consulta de pesquisa.

pipeline

Uma origem e todas as transformações e destinos que recebem dados dessa origem. Cada mapeamento contém um ou mais pipelines.

plano de execução do Hive

Uma série de tarefas do Hive geradas pelo executor do Hive depois de processar um mapeamento ou um perfil. Um plano de execução do Hive também pode ser conhecido como um fluxo de trabalho do Hive.

ponderação métrica

Um número inteiro maior que ou igual a 0 atribuído a uma métrica. Uma ponderação métrica define a contribuição da métrica para a pontuação do grupo de métricas.

ponto de partição

Um limite entre fases em um pipeline de mapeamento. Quando o particionamento está habilitado, o Serviço de Integração de Dados pode redistribuir linhas de dados em pontos de partição.

pontuação do grupo de métricas

A média ponderada computada de todas as pontuações de métrica no grupo de métricas.

pontuação métrica

A porcentagem de valores válidos em uma métrica.

porta complexa

Um tipo de porta que é atribuído a um tipo de dados complexo, como uma array, struct ou map para transmitir dados hierárquicos.

porta complexa aninhada

Uma porta complexa que contém uma definição de tipo de dados complexo aninhado.

porta dinâmica

Uma porta que pode receber uma ou mais colunas de uma transformação upstream e criar uma porta gerada para cada coluna.

porta gerada

Uma porta dentro de uma porta dinâmica que representa uma única coluna. A Developer tool gera portas com base em uma ou mais regras de entrada.

porta variável com estado

Uma porta variável que faz referência a valores de linhas anteriores.

pré-processador

Um processador de documento usado para realizar uma modificação geral de um documento de origem, antes da transformação principal.

procedimento armazenado virtual

Um conjunto de instruções de procedimentos ou fluxos de dados em um serviço de dados SQL.

processador de documentos

Um componente que opera em um documento como um todo, normalmente realizando conversões preliminares antes da análise.

processo do DTM

Um processo do sistema operacional que o Serviço de Integração de Dados inicia para executar instâncias do DTM. Com base em como você configura o Serviço de Integração de Dados, o serviço pode executar cada instância do DTM em um processo separado do DTM no nó local ou em um nó remoto.

projeto

O contêiner de nível superior para armazenar objetos criados no Informatica Analyst e no Informatica Developer. Crie projetos com base em requisitos ou metas de negócios. Os projetos aparecem no Informatica Analyst e no Informatica Developer.

regra de criação de perfil

A lógica comercial reutilizável que define condições aplicadas a dados de origem quando você executa um perfil. Use regras para validar os dados de um perfil e medir o progresso da qualidade dos dados. Você pode criar uma regra no Informatica Analyst ou no Informatica Developer.

regra de dados

A lógica comercial reutilizável que identifica uma coluna por seus valores como pertencente a um determinado domínio de dados.

regra de entrada

Uma regra que determina quais portas geradas devem ser criadas dentro de uma porta dinâmica.

regra de nome da coluna

A lógica comercial reutilizável que identifica uma coluna por seu nome como pertencente a um determinado domínio de dados.

saída de tarefa

Dados que são transmitidos de uma tarefa para variáveis de fluxo de trabalho. Ao configurar uma tarefa, você especifica os valores de saída de tarefa que deseja atribuir a variáveis de fluxo de trabalho. O Serviço de Integração de Dados copia os valores de saída de tarefa para variáveis de fluxo de trabalho quando a tarefa é concluída. O Serviço de Integração de Dados pode acessar esses valores a partir das variáveis de fluxo de trabalho ao avaliar expressões em fluxos de sequência condicionais e ao executar objetos adicionais no fluxo de trabalho.

scorecard

Uma representação gráfica de valores válidos para uma coluna de origem ou saída de uma regra nos resultados do perfil. Use scorecards para medir o progresso da qualidade dos dados.

scripts do Hive

Script na linguagem de consulta Hive que contém consultas e comandos do Hive para executar o mapeamento.

segmento de grade

Parte de um mapeamento de grade que está contido em uma tarefa de grade.

serviço de aplicativo

Um serviço que executa em um ou mais nós no domínio Informatica. Você cria e gerencia serviços de aplicativo no Informatica Administrator ou através do comando `infacmd` do programa. Os serviços de aplicativo incluem serviços que podem ter várias instâncias no domínio e serviços de sistema que podem ter uma única instância no domínio. Configure cada serviço de aplicativo com base em seus requisitos de ambiente.

serviço de dados

Um conjunto de operações reutilizáveis que você pode executar para acessar e transformar dados. Um serviço de dados fornece um modelo de dados unificado que você pode acessar por meio de um serviço da Web ou usar como base para executar uma consulta SQL.

Serviço de dados SQL

Um banco de dados virtual que pode ser consultado. Ele contém objetos virtuais e fornece uma exibição uniforme dos dados a partir de origens de dados distintas e heterogêneas.

Serviço de Ingestão em Massa

Um serviço de aplicativo que gerencia as especificações de ingestão em massa que você cria na ferramenta de Ingestão em Massa. O Serviço de Ingestão em Massa valida as especificações, as especificações de agendamentos para execução em um Serviço de Integração de Dados e monitora as estatísticas de ingestão.

Serviço de Integração de Dados

Um serviço de aplicativo que realiza trabalhos de integração de dados para o Informatica Analyst, o Informatica Developer e clientes externos. Trabalhos de integração de dados incluem visualizar dados e executar mapeamentos, perfis, serviços de dados SQL, serviços da Web e fluxos de trabalho.

Serviço de Repositório do Modelo

Um serviço de aplicativo no domínio Informatica que executa e gerencia o repositório do Modelo. Esse repositório armazena os metadados criados pelos produtos Informatica em um banco de dados relacional para ativar a colaboração entre os produtos.

serviço de sistema

Um serviço de aplicativo que pode ter uma única instância no domínio. Quando você cria o domínio, os serviços de sistema são criados para você. É possível ativar, desativar e configurar serviços de sistema.

Serviço do Gerenciador de Recursos

Um serviço do sistema que gerencia recursos de computação no domínio e despacha trabalhos para obter desempenho e escalabilidade ideais. O Serviço do Gerenciador de Recursos coleta informações sobre nós com a função de cálculo. O serviço corresponde os requisitos de trabalho à disponibilidade de recursos para identificar o melhor nó de cálculo para executar o trabalho. O Serviço do Gerenciador de Recursos se comunica com nós de cálculo em uma grade do Serviço de Integração de Dados. Ative o Serviço do Gerenciador de Recursos ao configurar uma grade do Serviço de Integração de Dados para executar trabalhos em processos remotos separados.

Sistema de Arquivos Distribuídos Hadoop (HDFS)

Um sistema de armazenamento de arquivos distribuídos usado por aplicativos Hadoop.

tabela virtual

Uma tabela em um banco de dados virtual.

tarefa

Um objeto de fluxo de trabalho que executa uma única unidade de trabalho no fluxo de trabalho, como executar um mapeamento, enviar um e-mail ou executar um comando shell. Uma tarefa representa algo que é realizado durante o fluxo de trabalho. O editor exibe tarefas como quadrados.

tarefa de comando

A tarefa de pré-processamento ou pós-processamento de dados locais para um fluxo de trabalho do mecanismo Blaze.

tarefa de grade

Uma solicitação de trabalho de processamento paralelo. Quando o mapeamento é executado no ambiente Hadoop, o executor do mecanismo Blaze envia a solicitação ao Gerenciador de Grade. Quando o mapeamento é executado no ambiente nativo, e o Serviço de Integração de Dados é executado no modo remoto, o processo do Serviço de Integração de Dados envia a solicitação ao Gerenciador de Serviços no nó de computação mestre.

tarefa do Hive

Uma tarefa no plano de execução do Hive. Um plano de execução do Hive contém várias tarefas do Hive. Uma tarefa do Hive contém um script do Hive.

tasklet

Uma partição de um segmento de grade que é executada em um DTM separado.

TaskTracker

Um nó no cluster Hadoop que executa tarefas como de mapa ou redução. TaskTrackers enviam relatórios de progresso ao JobTracker.

tipo de dados aninhado

Um tipo de dados complexo que contém pelo menos um elemento de um tipo de dados complexo. Por exemplo, um tipo de dados struct que contém um elemento do tipo array.

tipo de dados complexo

Um tipo de dados que permite representar valores múltiplos de dados em uma única posição de coluna. Os valores de dados são chamados de elementos.

tipo de dados primitivo

Um tipo de dados que permite representar um único valor de dados em uma única posição de coluna.

trabalho MapReduce

Uma unidade de trabalho que consiste nos dados de entrada, no programa MapReduce e em informações de configuração. O Hadoop executa o trabalho MapReduce dividindo-os em tarefas de mapa e de redução.

transformação

Um objeto de repositório em um mapeamento que gera, modifica ou passa dados. Cada transformação executa uma função diferente.

transformação de serviço da Web

Uma transformação que processa solicitações ou respostas de serviços da Web. Exemplos de transformações de serviços da Web incluem uma transformação de Entrada, uma transformação de Saída, uma transformação de Falha e a transformação de Consumidor de Serviço da Web.

variável de fluxo de trabalho

Um valor que podem mudar durante uma execução de fluxo de trabalho. Use variáveis de fluxo de trabalho para referenciar valores e registrar informações de tempo de execução. É possível usar variáveis de fluxo de trabalho do sistema ou definidas pelo usuário.

variável de fluxo de trabalho definida pelo usuário

Uma variável de fluxo de trabalho que captura a saída de tarefa ou os critérios que você especifica. Depois de criar uma variável de fluxo de trabalho definida pelo usuário, você configura o fluxo de trabalho para atribuir um valor de tempo de execução a essa variável.

variável de fluxo de trabalho do sistema

Uma variável de fluxo de trabalho que retorna informações de tempo de execução do sistema, como o ID da instância de fluxo de trabalho, o usuário que iniciou o fluxo de trabalho ou a hora de início do fluxo de trabalho.

warehouse de criação de perfil

Um banco de dados relacional que armazena informações de criação de perfil, como resultados de perfil e scorecard.

XMap

Um objeto de transformação de Processador de Dados que mapeia um documento de entrada XML para outro documento XML.

XPath

Uma linguagem de consulta usada para selecionar nós em um documento XML e realizar computações.