Guia do usuário

AWS Entity Resolution

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

AWS Entity Resolution: Guia do usuário

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

As marcas comerciais e imagens comerciais da Amazon não podem ser usadas no contexto de nenhum produto ou serviço que não seja da Amazon, nem de qualquer maneira que possa gerar confusão entre os clientes ou que deprecie ou desprestigie a Amazon. Todas as outras marcas comerciais que não pertencem à Amazon pertencem a seus respectivos proprietários, que podem ou não ser afiliados, patrocinados pela Amazon ou ter conexão com ela.

Table of Contents

O que AWS Entity Resolutioné 1
Você é um AWS Entity Resolution usuário iniciante? 1
Características do AWS Entity Resolution 2
Serviços relacionados 5
Acessando AWS Entity Resolution 6
Preços para AWS Entity Resolution 6
Configuração 7
Inscrevendo-se para AWS 7
Criando um usuário administrador 7
Criação de uma função do IAM para um usuário do console 8
Criação de uma função de trabalho no fluxo de trabalho 10
Prepare tabelas de dados de entrada 17
Preparando dados de entrada primários 17
Etapa 1: Preparar tabelas de dados primárias 17
Etapa 2: Salve sua tabela de dados de entrada em um formato de dados compatível 19
Etapa 3: Carregue sua tabela de dados de entrada para o Amazon S3 19
Etapa 4: criar uma AWS Glue tabela 20
Etapa 4: criar uma tabela particionada AWS Glue 22
Preparando dados de entrada de terceiros 24
Etapa 1: Assine um serviço de provedor em AWS Data Exchange 25
Etapa 2: Preparar tabelas de dados de terceiros 26
Etapa 3: Salve sua tabela de dados de entrada em um formato de dados compatível 31
Etapa 4: Carregue sua tabela de dados de entrada para o Amazon S3 32
Etapa 5: criar uma AWS Glue tabela 32
Mapeamento de esquemas 34
Criação de um mapeamento de esquema 35
Clonando um mapeamento de esquema 48
Editando um mapeamento de esquema 48
Excluindo um mapeamento de esquema 49
namespace de ID 51
Fonte do namespace de ID 52
Criação de uma fonte de namespace de ID (com base em regras) 52
Criação de uma fonte de namespace de ID (serviços do provedor) 56
Destino do namespace de ID 59

Criação de um destino de namespace de ID (método baseado em regras) 59
Criação de um destino de namespace de ID (método de serviços do provedor) 62
Editando um namespace de ID 63
Excluindo um namespace de ID 64
Adicionar ou atualizar uma política de recursos para um namespace de ID 64
Fluxo de trabalho correspondente 66
Criação de um fluxo de trabalho de correspondência baseado em regras 68
Tipo de regra avançada 69
Tipo de regra simples 85
Criação de um fluxo de trabalho de correspondência baseado em aprendizado de máquina 95
Criação de um fluxo de trabalho de correspondência baseado em serviços do provedor 100
Criando um fluxo de trabalho correspondente com LiveRamp 101
Criando um fluxo de trabalho correspondente com TransUnion 109
Criando um fluxo de trabalho correspondente com o UID 2.0 115
Editando um fluxo de trabalho correspondente 120
Excluindo um fluxo de trabalho correspondente 121
Modificando ou gerando um Match ID 121
Procurando um Match ID 126
Excluindo registros de um fluxo de trabalho de correspondência baseado em regras ou em ML 129
Solução de problemas 130
Recebi um arquivo de erro depois de executar um fluxo de trabalho correspondente 130
Fluxo de trabalho de mapeamento de ID 132
Fluxo de trabalho de mapeamento de ID para um Conta da AWS 133
Pré-requisitos 134
Criação de um fluxo de trabalho de mapeamento de ID (baseado em regras) 135
Criação de um fluxo de trabalho de mapeamento de ID (serviços do provedor) 141
Fluxo de trabalho de mapeamento de ID em dois Contas da AWS 147
Pré-requisitos 148
Criação de um fluxo de trabalho de mapeamento de ID (baseado em regras) 149
Criação de um fluxo de trabalho de mapeamento de ID (serviços do provedor) 155
Executar um fluxo de trabalho de mapeamento de ID 161
Executando um fluxo de trabalho de mapeamento de ID com um novo destino de saída 162
Editando um fluxo de trabalho de mapeamento de ID 165
Excluindo um fluxo de trabalho de mapeamento de ID 165

Adicionar ou atualizar uma política de recursos para um fluxo de trabalho de mapeamento de ID 166
Integração com provedores 167
Requisitos 167
Listar um serviço de provedor em AWS Data Exchange 168
Identifique seus atributos 169
Solicite a AWS Entity Resolution especificação OpenAPI 169
Usando a especificação OpenAPI 170
Integração de processamento em lote 170
Integração de processamento síncrono 173
Testando a integração de um provedor 174
Segurança 183
Proteção de dados 183
Criptografia de dados em repouso para AWS Entity Resolution 185
Gerenciamento de chaves 186
AWS PrivateLink 196
Gerenciamento de identidade e acesso 198
Público 199
Autenticar com identidades 199
Gerenciar o acesso usando políticas 203
Como AWS Entity Resolution funciona com o IAM 206
Exemplos de políticas baseadas em identidade 213
AWS políticas gerenciadas 216
Solução de problemas 218
Validação de conformidade 220
AWS Entity Resolution melhores práticas de conformidade 222
Resiliência 222
Monitoramento 224
CloudTrail troncos 224
AWS Entity Resolution informações em CloudTrail 225
Entendendo as entradas do arquivo de AWS Entity Resolution log 226
CloudWatch Registros 226
Configurando a entrega do log 227
Desativando o registro (console) 234
Lendo os registros 235
AWS CloudFormation recursos 238

Resolução e AWS CloudFormation modelos de entidades da AWS 238
Saiba mais sobre AWS CloudFormation 240
Cotas 241
Histórico do documento 249
Glossário 255
Nome do recurso da Amazon (ARN) 255
Tipo de atributo 255
Processamento automático 255
AWS KMS key ARN 255
Texto não criptografado 255
Nível de confiança (ConfidenceLevel) 256
Descriptografia 256
Criptografia 256
Group name 256
Hash 256
Protocolo de hash () HashingProtocol 256
Método de mapeamento de ID 257
Fluxo de trabalho de mapeamento de ID 257
namespace de ID 257
Campo de entrada 258
ARN da fonte de entrada (ARN) InputSource 258
Correspondência baseada em aprendizado de máquina 258
Processamento manual 258
Many-to-Many combinando 258
ID da partida (MatchID) 259
Tecla de correspondência (MatchKey) 259
Nome da chave de correspondência 260
Regra de partida (MatchRule) 260
Correspondência 260
Fluxo de trabalho correspondente 260
Descrição do fluxo de trabalho correspondente 260
Nome do fluxo de trabalho correspondente 261
Metadados de fluxo de trabalho correspondentes 261
Normalização () ApplyNormalization 261
Name 262
E-mail 262

Telefone 263
Endereço 263
Hashado 266
ID de origem 266
Normalização (ApplyNormalization) — somente com base em ML 267
Name 267
E-mail 267
Telefone 267
One-to-One combinando 268
Saída 268
Saídas 3Path 269
OutputSourceConfig 269
Correspondência baseada em serviços de provedores 269
Correspondência baseada em regras 269
Schema 270
Descrição do esquema 270
Nome do esquema 270
Mapeamento de esquemas 270
ARN de mapeamento de esquema 271
ID exclusivo 271
..... cclxxii

O que AWS Entity Resolutioné

AWS Entity Resolution é um serviço que ajuda você a combinar, vincular e aprimorar registros relacionados armazenados em vários aplicativos, canais e armazenamentos de dados. Você pode começar a usar fluxos de trabalho de resolução de entidades que são flexíveis, escaláveis e podem se conectar aos seus aplicativos e provedores de serviços de dados existentes.

AWS Entity Resolution oferece técnicas avançadas de correspondência, como correspondência baseada em regras, correspondência baseada em aprendizado de máquina (correspondência de ML) e correspondência liderada por provedores de serviços de dados. Essas técnicas podem ajudá-lo a vincular e aprimorar com mais precisão os registros relacionados de informações de clientes, códigos de produtos ou códigos de dados comerciais.

Você pode usar AWS Entity Resolution para criar uma visão unificada das interações com os clientes vinculando eventos recentes (como cliques em anúncios, abandono de carrinho e compras) a sinais pseudonimizados de seus provedores de serviços de dados em um ID de entidade exclusivo. Você também pode acompanhar melhor os produtos que usam códigos diferentes (por exemplo, SKU, UPC) em suas lojas. Você pode usar AWS Entity Resolution para controlar a precisão da correspondência e proteger melhor a segurança dos dados, minimizando a movimentação dos dados.

Tópicos

- Você é um AWS Entity Resolution usuário iniciante?
- Características do AWS Entity Resolution
- Serviços relacionados
- Acessando AWS Entity Resolution
- Preços para AWS Entity Resolution

Você é um AWS Entity Resolution usuário iniciante?

Se você é usuário iniciante do AWS Entity Resolution, recomendamos que comece lendo as seguintes seções:

- Características do AWS Entity Resolution
- Acessando AWS Entity Resolution

- Configurar AWS Entity Resolution

Características do AWS Entity Resolution

AWS Entity Resolution inclui os seguintes recursos:

- Preparação de dados flexível e personalizável

  AWS Entity Resolution lê seus dados AWS Glue para usar como entradas para processamento de partidas. Você pode especificar no máximo 20 entradas de dados. AWS Entity Resolution processa cada linha da tabela de entrada de dados como um registro, com uma entidade exclusiva servindo como chave primária. AWS Entity Resolution pode operar em conjuntos de dados criptografados. Primeiro, defina o mapeamento do esquema AWS Entity Resolution para entender quais campos de entrada você deseja usar no fluxo de trabalho correspondente. Você pode trazer seu próprio esquema de dados, ou blueprint, a partir de uma entrada de AWS Glue dados existente. Ou você pode criar seu esquema personalizado usando uma interface de usuário interativa ou um editor JSON. Por padrão, AWS Entity Resolution também normaliza as entradas de dados antes da correspondência para melhorar o processamento da correspondência, como remover caracteres especiais e espaços extras e formatar texto em minúsculas. Se a entrada de dados já estiver normalizada, você poderá desativar a normalização. Também fornecemos uma GitHub biblioteca, que você pode usar para personalizar ainda mais o processo de normalização de dados de acordo com suas necessidades.

- Fluxos de trabalho de correspondência de entidades configuráveis

  Um fluxo de trabalho de correspondência de entidades é uma sequência de etapas que você configura para saber AWS Entity Resolution como combinar sua entrada de dados e onde gravar a saída de dados consolidada. Você pode configurar um ou mais fluxos de trabalho correspondentes para comparar diferentes entradas de dados e usar diferentes técnicas de correspondência, como correspondência baseada em regras, correspondência de aprendizado de máquina ou correspondência liderada por provedor de serviços de dados sem resolução de entidades ou experiência em ML. Você também pode visualizar o status do trabalho dos fluxos de trabalho e métricas correspondentes existentes, como número do recurso, número de registros processados e número de correspondências encontradas.

  - Ready-to-use correspondência baseada em regras

    Essa técnica de correspondência inclui um conjunto de ready-to-use regras no AWS Management Console ou AWS Command Line Interface (AWS CLI). Você pode usar essas

regras para encontrar registros relacionados com base em seus campos de entrada. Você também pode personalizar as regras adicionando ou removendo campos de entrada para cada regra, excluindo regras, reorganizando a prioridade da regra e criando novas regras. Você também pode redefinir as regras para retorná-las às configurações originais. A saída de dados em seu bucket do Amazon Simple Storage Service (Amazon S3) tem grupos de correspondência AWS Entity Resolution que são gerados usando a técnica de correspondência baseada em regras. Cada grupo de correspondência tem o número da regra usado para gerar a correspondência associada a ele para ajudar você a entender a correspondência. Por exemplo, o número da regra pode demonstrar a precisão de cada grupo de correspondência, de forma que a regra um seja mais precisa do que a regra dois.

- Correspondência pré-configurada baseada em aprendizado de máquina (correspondência de ML)

  Essa técnica de correspondência inclui um modelo de ML pré-configurado para encontrar correspondências em todas as suas entradas de dados, especialmente nos registros baseados no consumidor. O modelo usa todos os campos de entrada associados aos tipos de dados de nome, endereço de e-mail, número de telefone, endereço e data de nascimento. O modelo gera grupos de correspondência de registros relacionados com uma pontuação de confiança em cada grupo, explicando a qualidade da correspondência em relação a outros grupos de correspondência. O modelo considera os campos de entrada ausentes e analisa todo o registro em conjunto para representar uma entidade. A saída de dados em seu bucket do Amazon S3 tem grupos de correspondência que são AWS Entity Resolution gerados usando a correspondência de ML. É aqui que cada grupo de correspondência tem uma pontuação de confiança associada de 0,0—1,0, o que indica a precisão da partida.

- Combinando registros com provedores de serviços de dados

  Com isso, AWS Entity Resolution você pode combinar, vincular e aprimorar seus registros com os principais fornecedores de serviços de dados e conjuntos de dados licenciados para expandir sua capacidade de entender, alcançar e atender seus clientes. Por exemplo, você pode acrescentar atributos aos seus dados para aprimorar seus registros ou pode melhorar a interoperabilidade dos sistemas e plataformas com os quais trabalha para atingir suas metas de negócios. Você pode usar esse fluxo de trabalho correspondente com alguns cliques, eliminando a necessidade de criar e manter integrações proprietárias complexas. Você deve ter um contrato de licença com esses provedores de serviços de dados para aproveitar essa técnica de correspondência.

- Processamento manual em massa e processamento incremental automático

Você pode usar o processamento de dados para ajudar a converter suas entradas ou entradas de dados em uma tabela de saída de dados consolidada com registros semelhantes que tenham uma ID de correspondência comum gerada usando configurações de fluxo de trabalho de correspondência de entidades. Usando a API AWS Management Console e/ou o AWS CLI, você pode executar o processamento manual em massa sob demanda, com base em seu pipeline de dados de extração, transformação e carregamento (ETL) existente, que reprocessa todos os dados para quaisquer novas correspondências e atualizações para correspondências existentes. Além disso, para cenários de correspondência baseados em regras, você pode iniciar o processamento incremental automático para que, assim que novos dados estejam disponíveis em seu bucket do Amazon S3, o serviço leia esses novos registros e os compare com os registros existentes. Isso mantém suas correspondências atualizadas com quaisquer alterações nos dados do Amazon S3.

- Pesquisa quase em tempo real

  Pesquisar qualquer campo de entidade por meio da operação da AWS Entity Resolution GetMatchId API ajuda você a recuperar de forma síncrona um ID de correspondência existente. Você pode ligar AWS Entity Resolution com atributos de informações de identificação pessoal (PII) adquiridos por meio de diferentes fontes e canais. AWS Entity Resolution faz o hash desses atributos para proteção de dados e recupera a ID de correspondência correspondente para vincular e combinar o cliente. Por exemplo, você pode obter uma inscrição na web com um nome, e-mail e endereço para correspondência associados. Use a operação de AWS Entity Resolution GetMatchId API para descobrir se esse cliente ou entidade já existe nos resultados correspondentes armazenados em seu bucket do S3, junto com o ID de correspondência da entidade correspondente associado a ele. Depois de obter a ID de correspondência da entidade, você pode encontrar as informações transacionais associadas a ela em seus aplicativos de origem, como seus sistemas de gerenciamento de relacionamento com o cliente (CRM) ou plataforma de dados do cliente (CDP).

- Proteção de dados e regionalização por design

  AWS Entity Resolution oferece um recurso de criptografia padrão que pode ajudá-lo a proteger seus dados e fornece uma chave de criptografia para cada entrada de dados no serviço. Por exemplo, AWS Entity Resolution oferece a flexibilidade de trazer dados criptografados e com hash do lado do servidor para executar fluxos de trabalho de correspondência baseados em regras. AWS Entity Resolution oferece suporte à regionalização, o que significa que seus fluxos de trabalho correspondentes são executados para processar seus dados da mesma forma Região da AWS de onde você está usando o serviço. Você também pode criptografar e fazer o hash da saída de dados no Amazon S3 antes de usar seus dados resolvidos em outros aplicativos.

- Transcodificação multipartidária

  AWS Entity Resolution ajuda você a definir suas fontes de dados e as configurações correspondentes entre várias partes que desejam usar uma colaboração de dados, como em AWS Clean Rooms.

Serviços relacionados

Os itens a seguir Serviços da AWS estão relacionados a AWS Entity Resolution:

- Amazon S3

  Armazene os dados que você traz para AWS Entity Resolution o Amazon S3.

  Para obter mais informações, consulte O que é o Amazon S3? no Guia do usuário do Amazon Simple Storage Service.

- AWS Glue

  Crie AWS Glue tabelas a partir de seus dados no Amazon S3 para uso em. AWS Entity Resolution

  Para obter mais informações, consulte O que é AWS Glue? no Guia do AWS Glue desenvolvedor.

- AWS CloudTrail

  Use AWS Entity Resolution com CloudTrail registros para aprimorar sua análise da AWS service (Serviço da AWS) atividade.

  Para obter mais informações, consulte Registrando chamadas de AWS Entity Resolution API usando AWS CloudTrail.

- AWS CloudFormation

  Crie os seguintes recursos em AWS CloudFormation: AWS::EntityResolution::MatchingWorkflow, AWS::EntityResolution::SchemaMapping, AWS::EntityResolution:IdMappingWorkflow, AWS::EntityResolution::IdNamespace e AWS::EntityResolution::PolicyStatement

  Para obter mais informações, consulte Crie recursos de resolução de entidades da AWS com AWS CloudFormation.

Acessando AWS Entity Resolution

Você pode acessar AWS Entity Resolution por meio das seguintes opções:

- Diretamente pelo AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
- Programaticamente por meio da API. AWS Entity Resolution Para obter mais informações, consulte a Referência da API do AWS Entity Resolution .
  - Se você planeja chamar a AWS Entity Resolution API no AWS Lambda Runtime, crie seu próprio pacote de implantação e inclua a versão desejada da biblioteca do AWS SDK. Para obter mais informações, consulte os exemplos a seguir no Guia do AWS Lambda desenvolvedor:
    - Implemente funções Java Lambda com arquivamentos de arquivos.zip ou JAR
    - Trabalhando com arquivos de arquivos.zip para funções Python Lambda

Preços para AWS Entity Resolution

Para obter informações sobre a definição de preço, consulte Definição de preço do AWS Entity Resolution.

Configurar AWS Entity Resolution

Antes de usar AWS Entity Resolution pela primeira vez, inscreva-se AWS e crie um usuário administrador para criar funções.

Inscrevendo-se para AWS

Se você já tem um Conta da AWS, pule esta etapa.

Se você não tiver um Conta da AWS, conclua as etapas a seguir para criar um.

Para se inscrever em um Conta da AWS

1. Abra a https://portal.aws.amazon.com/billing/inscrição.
2. Siga as instruções online.

   Parte do procedimento de inscrição envolve receber uma chamada telefônica ou uma mensagem de texto e inserir um código de verificação pelo teclado do telefone.

   Quando você se inscreve em um Conta da AWS, um Usuário raiz da conta da AWSé criado. O usuário-raiz tem acesso a todos os Serviços da AWS e recursos na conta. Como prática recomendada de segurança, atribua o acesso administrativo a um usuário e use somente o usuário-raiz para executar tarefas que exigem acesso de usuário-raiz.

Criando um usuário administrador

Para criar um usuário administrador, selecione uma das opções a seguir.

Selecionar uma forma de gerenciar o administrador	Para	Por	Você também pode
Centro de Identidade do IAM (Recomendado)	Usar credenciais de curto prazo para acessar a AWS. Isso está de acordo com as práticas recomendadas de segurança. Para obter informações sobre as práticas recomendadas, consulte Práticas recomendadas de segurança no IAM no Guia do usuário do IAM.	Seguindo as instruções em Conceitos básicos no Guia do usuário do AWS IAM Identity Center .	Configure o acesso programático configurando o AWS CLI para uso AWS IAM Identity Center no Guia do AWS Command Line Interface usuário.
No IAM (Não recomendado)	Usar credenciais de longo prazo para acessar a AWS.	Seguindo as instruções em Criar um acesso de emergência para um usuário do IAM no Guia do usuário do IAM.	Configurar o acesso programático, com base em Gerenciar chaves de acesso para usuários do IAM no Guia do usuário do IAM.

Criação de uma função do IAM para um usuário do console

Conclua o procedimento a seguir se estiver usando o AWS Entity Resolution console.

Para criar um perfil do IAM

1. Faça login no console do IAM (https://console.aws.amazon.com/iam/) com sua conta de administrador.
2. Em Gerenciamento de acesso, escolha Perfis.

   Você pode usar Funções para criar credenciais de curto prazo, o que é recomendado para aumentar a segurança. Você também pode escolher Usuários para criar credenciais de longo prazo.
3. Selecione Criar perfil.
4. No assistente de criação de função, em Tipo de entidade confiável, escolha Conta da AWS.
5. Mantenha a opção Esta conta selecionada e, em seguida, escolha Avançar.
6. Em Adicionar permissões, escolha Criar política.

   Uma nova guia será aberta.

   a. Selecione a guia JSON e, em seguida, adicione políticas de acordo com as habilidades concedidas ao usuário do console. AWS Entity Resolution oferece as seguintes políticas gerenciadas com base em casos de uso comuns:

      - AWS política gerenciada: AWSEntity ResolutionConsoleFullAccess
      - AWS política gerenciada: AWSEntity ResolutionConsoleReadOnlyAccess

   b. Escolha Próximo: Etiquetas, adicionar tags (opcional) e escolha Próximo: Revisão.

   c. Em Política de revisão, insira um Nome e uma Descrição e revise o Resumo.

   d. Escolha Criar política.

      Você criou uma política para um membro da colaboração.

   e. Volte para a guia original e, em Adicionar permissões, insira o nome da política que você acabou de criar. (Você pode precisar recarregar a página.)

   f. Marque a caixa de seleção ao lado do nome da política que você criou e escolha Avançar.
7. Na página Nome, revisar e criar, insira um nome de perfil e uma descrição.

   a. Revise Selecionar entidades confiáveis, insira o Conta da AWS para o nome da pessoa ou pessoas que assumirão a função (se necessário).

   b. Revise as permissões em Adicionar permissões e edite, se necessário.

   c. Revise as tags e adicione tags, se necessário.

d. Selecione Criar perfil.

Criação de uma função de trabalho de fluxo de trabalho para AWS Entity Resolution

AWS Entity Resolution usa uma função de trabalho de fluxo de trabalho para executar um fluxo de trabalho. Você pode criar esse perfil usando o console se você tiver as permissões necessárias do IAM. Se você não tiver `CreateRole` permissões, peça ao administrador que crie a função.

Para criar uma função de trabalho de fluxo de trabalho para AWS Entity Resolution

1. Faça login no console do IAM em https://console.aws.amazon.com/iam/com sua conta de administrador.
2. Em Gerenciamento de acesso, escolha Perfis.

   Você pode usar Funções para criar credenciais de curto prazo, o que é recomendado para aumentar a segurança. Você também pode escolher Usuários para criar credenciais de longo prazo.
3. Selecione Criar perfil.
4. No assistente Criar perfil, para Tipo de entidade confiável, escolha Política de confiança personalizada.
5. Copie e cole a seguinte política de confiança personalizada no editor JSON.

   JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "entityresolution.amazonaws.com"
                ]
            },
            "Action": "sts:AssumeRole"
        }
```

```
    ]
}
```

6. Escolha Próximo.
7. Em Adicionar permissões, escolha Criar política.

   Uma nova guia é exibida.

   a. Copie e cole a política a seguir no editor JSON.

> **Note**
>
> O exemplo de política a seguir oferece suporte às permissões necessárias para ler os recursos de dados correspondentes, como Amazon S3 e. AWS Glue No entanto, talvez seja necessário modificar essa política dependendo de como você configurou suas fontes de dados.
> Seus AWS Glue recursos e os recursos subjacentes do Amazon S3 devem estar no mesmo Região da AWS que. AWS Entity Resolution
> Você não precisa conceder AWS KMS permissões se suas fontes de dados não estiverem criptografadas ou descriptografadas.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject",
                "s3:ListBucket",
                "s3:GetBucketLocation"
            ],
            "Resource": [
                "arn:aws:s3:::{{input-buckets}}",
                "arn:aws:s3:::{{input-buckets}}/*"
            ],
            "Condition": {
                "StringEquals": {
```

```
                    "s3:ResourceAccount": [
                        "{{accountId}}"
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:PutObject",
                "s3:ListBucket",
                "s3:GetBucketLocation"
            ],
            "Resource": [
                "arn:aws:s3:::{{output-bucket}}",
                "arn:aws:s3:::{{output-bucket}}/*"
            ],
            "Condition": {
                "StringEquals": {
                    "s3:ResourceAccount": [
                        "{{accountId}}"
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "glue:GetDatabase",
                "glue:GetTable",
                "glue:GetPartition",
                "glue:GetPartitions",
                "glue:GetSchema",
                "glue:GetSchemaVersion",
                "glue:BatchGetPartition"
            ],
            "Resource": [
                "arn:aws:glue:us-east-1:{{accountId}}:database/{{input-
databases}}",
                "arn:aws:glue:us-east-1:{{accountId}}:table/{{input-
database}}/{{input-tables}}",
                "arn:aws:glue:us-east-1:{{accountId}}:catalog"
            ]
        }
```

```
    ]
}
```

Substitua cada *{{user input placeholder}}* por suas próprias informações.

aws-region	Região da AWS de seus recursos. Seus AWS Glue recursos, recursos e AWS KMS recursos subjacentes do Amazon S3 devem estar no mesmo Região da AWS que. `AWS Entity Resolution`
accountId	Sua Conta da AWS identidade.
input-buckets	Buckets do Amazon S3 que contêm os objetos de dados subjacentes de AWS Glue onde `AWS Entity Resolution` serão lidos.
output-buckets	Buckets do Amazon S3 onde `AWS Entity Resolution` gerarão os dados de saída.
input-databases	AWS Glue bancos de dados de onde `AWS Entity Resolution` lerá.

b. (Opcional) Se o bucket de entrada do Amazon S3 for criptografado usando a chave KMS do cliente, adicione o seguinte:

```
{
    "Effect": "Allow",
    "Action": [
        "kms:Decrypt"
    ],
    "Resource": [
        "arn:aws:kms:{{aws-region}}:{{accountId}}:key/{{inputKeys}}"
    ]
}
```

Substitua cada *{{user input placeholder}}* por suas próprias informações.

aws-region	Região da AWS de seus recursos. Seus AWS Glue recursos, recursos e AWS KMS recursos subjacentes do Amazon S3 devem estar no mesmo Região da AWS que. `AWS Entity Resolution`
accountId	Sua Conta da AWS identidade.
inputKeys	Entrada gerenciada de chaves AWS Key Management Service. Se suas fontes de entrada forem criptografadas, `AWS Entity Resolution` deverá descripto grafar seus dados usando sua chave.

c. (Opcional) Se os dados que estão sendo gravados no bucket de saída do Amazon S3 precisarem ser criptografados, adicione o seguinte:

```
        {
            "Effect": "Allow",
            "Action": [
                "kms:GenerateDataKey",
                "kms:Encrypt"
            ],
            "Resource": [
                "arn:aws:kms:{{aws-region}}:{{accountId}}:key/{{outputKeys}}"
            ]
        }
```

Substitua cada *{{user input placeholder}}* por suas próprias informações.

aws-region	Região da AWS de seus recursos. Seus AWS Glue recursos, recursos e AWS KMS recursos subjacentes do Amazon S3 devem estar no mesmo Região da AWS que. `AWS Entity Resolution`
accountId	Sua Conta da AWS identidade.

outputKeys	Entrada gerenciada de chaves AWS Key Management Service. Se você precisar que suas fontes de saída sejam criptografadas, `AWS Entity Resolution` deverá criptografar os dados de saída usando sua chave.

d. (Opcional) Se você tiver uma assinatura com um serviço de provedor por meio AWS Data Exchange de e quiser usar uma função existente para um fluxo de trabalho baseado em serviços de provedor, adicione o seguinte:

```
        {
            "Effect": "Allow",
            "Sid": "DataExchangePermissions",
            "Action": "dataexchange:SendApiAsset",
            "Resource": [
                "arn:aws:dataexchange:{{aws-region}}::data-sets/{{datasetId}}/
revisions/{{revisionId}}/assets/{{assetId}}"
            ]
        }
```

Substitua cada *{{user input placeholder}}* por suas próprias informações.

aws-region	O Região da AWS local onde o recurso do provedor é concedido. Você pode encontrar esse valor no ARN do ativo no AWS Data Exchange console. Por exemplo: `arn:aws:dataexchange:us-east-2::data-sets/111122223333/revisions/339ffc64444examplef3bc15cf0b2346b/assets/546468b8dexamplea37bfc73b8f79fefa`
datasetId	O ID do conjunto de dados, encontrado no AWS Data Exchange console.

revisionId	A revisão do conjunto de dados, encontrada no AWS Data Exchange console.
assetId	O ID do ativo, encontrado no AWS Data Exchange console.

8. Volte para a guia original e, em Adicionar permissões, insira o nome da política que você acabou de criar. (Você pode precisar recarregar a página.)
9. Marque a caixa de seleção ao lado do nome da política que você criou e escolha Avançar.
10. Para Nome, revisar e criar, insira um nome de perfil e uma descrição.

> **Note**
>
> O nome da função deve corresponder ao padrão nas `passRole` permissões concedidas ao membro que pode passar o `workflow job role` para criar um fluxo de trabalho correspondente.
> Por exemplo, se você estiver usando a política `AWSEntityResolutionConsoleFullAccess` gerenciada, lembre-se de incluir `entityresolution` no nome da sua função.

    a. Revise Selecionar entidades confiáveis e edite, se necessário.
    b. Revise as permissões em Adicionar permissões e edite, se necessário.
    c. Revise as tags e adicione tags, se necessário.
    d. Selecione Criar perfil.

A função de trabalho do fluxo de trabalho para AWS Entity Resolution foi criada.

Prepare tabelas de dados de entrada

Em AWS Entity Resolution, cada uma de suas tabelas de dados de entrada contém registros de origem. Esses registros contêm identificadores de consumidores, como nome, sobrenome, endereço de e-mail ou número de telefone. Esses registros de origem podem ser combinados com outros registros de origem fornecidos na mesma tabela de dados ou em outras tabelas de dados de entrada. Cada registro deve ter uma ID de registro exclusiva (ID exclusivo) e você deve defini-la como uma chave primária ao criar um mapeamento de esquema dentro AWS Entity Resolution dela.

Cada tabela de dados de entrada está disponível como uma AWS Glue tabela apoiada pelo Amazon S3. Você pode usar seus dados primários que já estão no Amazon S3 ou importar tabelas de dados de outros provedores de SaaS terceirizados para o Amazon S3. Depois de fazer o upload dos dados para o Amazon S3, você pode usar um AWS Glue rastreador para criar uma tabela de dados no. AWS Glue Data Catalog Em seguida, você pode usar a tabela de dados como entrada para AWS Entity Resolution.

As seções a seguir descrevem como preparar dados primários e dados de terceiros.

Tópicos

- Preparando dados de entrada primários
- Preparando dados de entrada de terceiros

Preparando dados de entrada primários

As etapas a seguir descrevem como preparar dados primários para uso em um fluxo de trabalho de correspondência baseado em regras, fluxo de trabalho de correspondência baseadoem aprendizado de máquina ou fluxo de trabalho de mapeamento de ID.

Etapa 1: Preparar tabelas de dados primárias

Cada tipo de fluxo de trabalho correspondente tem um conjunto diferente de recomendações e diretrizes para ajudar a garantir o sucesso.

Para preparar tabelas de dados primárias, consulte a tabela a seguir:

Diretrizes de tabelas de dados primárias

Tipo de fluxo de trabalho	Obrigatório

Fluxo de trabalho de correspondência baseado em regras com o tipo de regra avançada	• É necessário um ID exclusivo.
	• O ID exclusivo não excede 38 caracteres.
	• (Opcional) Uma coluna DELETE que especifica de quais registros remover AWS Entity Resolution após o término do processam ento do fluxo de trabalho. O valor padrão é *false* se a coluna existir sem nenhum valor. Os registros com a coluna DELETE definida como *true* serão excluídos. Os registros com a coluna DELETE definida como *false* ou vazia serão processados por AWS Entity Resolution.
	
	O esquema deve ter uma coluna DELETE com tipo `String` `matchKey` e sem e. `groupName`
	
	Note
	O Look up match ID (`GetMatchID`) não é suportado porque o tipo de regra avançada para a cadência de processamento manual não armazena nenhum dado ingerido.
	
	No exemplo a seguir, `S1` será ingerido e `S2` excluído.
	
	Example
	
	`sourceID, name, lastName, DELETE`
	`S1, name, lastname, false`
	`S2, name2, lastname2, true`

fluxo de trabalho de correspondência baseado em regras com tipo de regra simples	• É necessário um ID exclusivo. • O ID exclusivo não excede 38 caracteres.

Tipo de fluxo de trabalho	Obrigatório

fluxo de trabalho de correspondência baseado em aprendizado de máquina	• É necessário um ID exclusivo.
	• O conjunto de dados contém um dos seguintes tipos:
	  • **`Full Name`**
	  • **`Full Address`**
	  • **`Full phone`**
	  • **`Email address`**
	  • **`Date`**— com um nome de chave Match de data de nascimento

Fluxo de trabalho de mapeamento de ID	• É necessário um ID exclusivo. • O ID exclusivo não excede 257 caracteres.

Etapa 2: Salve sua tabela de dados de entrada em um formato de dados compatível

Se você já salvou seus dados de entrada primários em um formato de dados compatível, você pode pular esta etapa.

Para serem usados AWS Entity Resolution, os dados de entrada devem estar em um formato AWS Entity Resolution compatível.

AWS Entity Resolution suporta os seguintes formatos de dados:

- valor separado por vírgula (CSV)
- Parquet

Etapa 3: Carregue sua tabela de dados de entrada para o Amazon S3

Se você já tem sua tabela de dados primários no Amazon S3, você pode pular esta etapa.

> **Note**
>
> Os dados de entrada devem ser armazenados no Amazon Simple Storage Service (Amazon S3) no Conta da AWS mesmo local Região da AWS e no qual você deseja executar o fluxo de trabalho correspondente.

Para carregar sua tabela de dados de entrada para o Amazon S3

1. Faça login no AWS Management Console e abra o console do Amazon S3 em. https://console.aws.amazon.com/s3/
2. Escolha Buckets e, em seguida, escolha um bucket para armazenar sua tabela de dados.
3. Escolha Upload e siga as instruções.
4. Escolha a guia Objetos para visualizar o prefixo do onde seus dados são armazenados. Anote o nome da pasta.

   Você pode selecionar a pasta para visualizar a tabela de dados.

Etapa 4: criar uma AWS Glue tabela

> **Note**
>
> Se você precisar de AWS Glue tabelas particionadas, vá para. Etapa 4: criar uma tabela particionada AWS Glue

Os dados de entrada no Amazon S3 devem ser catalogados AWS Glue e representados como uma tabela. AWS Glue Para obter mais informações sobre como criar uma AWS Glue tabela com o Amazon S3 como entrada, consulte Como trabalhar com rastreadores no AWS Glue console no Guia do desenvolvedor.AWS Glue

Nesta etapa, você configura um rastreador AWS Glue que rastreia todos os arquivos em seu bucket do S3 e cria uma tabela. AWS Glue

> **Note**
>
> AWS Entity Resolution atualmente não oferece suporte a locais do Amazon S3 registrados com. AWS Lake Formation

Para criar uma AWS Glue tabela

1. Faça login no AWS Management Console e abra o AWS Glue console em https://console.aws.amazon.com/glue/.
2. Na barra de navegação, selecione Crawlers.
3. Selecione seu bucket do S3 na lista e escolha Criar rastreador.
4. Na página Definir propriedades do rastreador, insira uma Descrição opcional do nome do rastreador e escolha Avançar.
5. Continue na página Adicionar crawler, especificando os detalhes.
6. Na página Escolher uma função do IAM, escolha Escolher um perfil do IAM existente e, em seguida, escolha Avançar.

   Você também pode escolher Criar um perfil do IAM ou fazer com que seu administrador crie o perfil do IAM, se necessário.
7. Em Criar uma programação para esse crawler, mantenha a Frequência padrão (Executar sob demanda) e escolha Avançar.
8. Em Configurar a saída do rastreador, insira o AWS Glue banco de dados e escolha Avançar.
9. Revise todos os detalhes e escolha Concluir.
10. Na página Crawlers, marque a caixa de seleção ao lado do bucket S3 e escolha Executar crawler.
11. Depois que o rastreador terminar de ser executado, na barra de AWS Glue navegação, escolha Bancos de dados e, em seguida, escolha o nome do banco de dados.
12. Na página Banco de dados, escolha Tabelas em {nome do seu banco de dados}.
    a. Visualize as tabelas no AWS Glue banco de dados.
    b. Para visualizar o esquema de uma tabela, selecione uma tabela específica.
    c. Anote o nome do AWS Glue banco de dados e o nome AWS Glue da tabela.

Agora você está pronto para criar um mapeamento de esquema. Para obter mais informações, consulte Criação de um mapeamento de esquema.

Etapa 4: criar uma tabela particionada AWS Glue

> **Note**
>
> O recurso de AWS Glue particionamento em só AWS Entity Resolution é compatível com fluxos de trabalho de mapeamento de ID. Esse recurso AWS Glue de particionamento permite que você escolha partições específicas para processamento. AWS Entity Resolution Se você não precisar de AWS Glue tabelas particionadas, pule esta etapa.

Uma AWS Glue tabela particionada reflete automaticamente as novas partições na AWS Glue tabela quando você adiciona novas pastas à estrutura de dados (como uma nova pasta de dia em menos de um mês).

Ao criar uma AWS Glue tabela particionada em AWS Entity Resolution, você pode especificar quais partições deseja processar em um fluxo de trabalho de mapeamento de ID. Então, toda vez que você executa o fluxo de trabalho de mapeamento de ID, somente os dados nessas partições são processados, em vez de processar todos os dados na AWS Glue tabela inteira. Esse recurso permite um processamento de dados mais preciso, eficiente e econômico AWS Entity Resolution, oferecendo maior controle e flexibilidade no gerenciamento de suas tarefas de resolução de entidades.

Você pode criar uma AWS Glue tabela particionada para a conta de origem em um fluxo de trabalho de mapeamento de ID.

Primeiro, você deve catalogar os dados de entrada no Amazon S3 AWS Glue e representá-los como uma AWS Glue tabela. Para obter mais informações sobre como criar uma AWS Glue tabela com o Amazon S3 como entrada, consulte Como trabalhar com rastreadores no AWS Glue console no Guia do desenvolvedor.AWS Glue

Nesta etapa, você configura um rastreador AWS Glue que rastreia todos os arquivos em seu bucket do S3 e, em seguida, cria uma tabela particionada. AWS Glue

> **Note**
>
> AWS Entity Resolution atualmente não oferece suporte a locais do Amazon S3 registrados com. AWS Lake Formation

Para criar uma tabela particionada AWS Glue

1. Faça login no AWS Management Console e abra o AWS Glue console em https://console.aws.amazon.com/glue/.
2. Na barra de navegação, selecione Crawlers.
3. Selecione seu bucket do S3 na lista e escolha Criar rastreador.
4. Na página Definir propriedades do rastreador, insira um Nome do rastreador, uma Descrição opcional e escolha Avançar.
5. Continue na página Adicionar crawler, especificando os detalhes.
6. Na página Escolher uma função do IAM, escolha Escolher um perfil do IAM existente e, em seguida, escolha Avançar.

   Você também pode escolher Criar um perfil do IAM ou fazer com que seu administrador crie o perfil do IAM, se necessário.
7. Em Criar uma programação para esse crawler, mantenha a Frequência padrão (Executar sob demanda) e escolha Avançar.
8. Em Configurar a saída do rastreador, insira o AWS Glue banco de dados e escolha Avançar.
9. Revise todos os detalhes e escolha Concluir.
10. Na página Crawlers, marque a caixa de seleção ao lado do bucket S3 e escolha Executar crawler.
11. Depois que o rastreador terminar de ser executado, na barra de AWS Glue navegação, escolha Bancos de dados e, em seguida, escolha o nome do banco de dados.
12. Na página Banco de dados, em Tabelas, escolha a tabela a ser particionada.
13. Na visão geral da tabela, selecione o menu suspenso Ações e escolha Editar tabela.
    a. Em Propriedades da tabela, escolha Adicionar.
    b. Para a nova chave, insira**`aerPushDownPredicateString`**.
    c. Para o novo Valor, insira**`'<PartitionKey>=<PartitionValue'`**.
    d. Anote o nome do AWS Glue banco de dados e o nome AWS Glue da tabela.

Agora está tudo pronto para:

- Crie um mapeamento de esquema e, em seguida, crie um fluxo de trabalho de mapeamento de ID para um Conta da AWS.

- Crie uma fonte de namespace de ID, crie um destino de namespace de ID e, em seguida, crie um fluxo de trabalho de mapeamento de ID entre duas. Contas da AWS

Preparando dados de entrada de terceiros

Os serviços de dados de terceiros fornecem identificadores que podem ser combinados com seus identificadores conhecidos.

AWS Entity Resolution atualmente oferece suporte aos seguintes serviços de provedores de dados terceirizados:

Serviços de provedores de dados

Nome da empresa	Disponível Regiões da AWS	Identificador
LiveRamp	Leste dos EUA (Norte da Virgínia) (us-east-1), Leste dos EUA (Ohio) (us-east-2) e Oeste dos EUA (Oregon) (us-west-2)	ID da rampa
TransUnion	Leste dos EUA (Norte da Virgínia) (us-east-1), Leste dos EUA (Ohio) (us-east-2) e Oeste dos EUA (Oregon) (us-west-2)	TransUnion Indivíduo e doméstico IDs
ID unificada 2.0	Leste dos EUA (Norte da Virgínia) (us-east-1), Leste dos EUA (Ohio) (us-east-2) e Oeste dos EUA (Oregon) (us-west-2)	UID bruto 2

As etapas a seguir descrevem como preparar dados de terceiros para usar um fluxo de trabalho de correspondência baseado no serviço do provedor ou um fluxo de trabalho de mapeamento de ID baseado no serviço do provedor.

Tópicos

- Etapa 1: Assine um serviço de provedor em AWS Data Exchange
- Etapa 2: Preparar tabelas de dados de terceiros
- Etapa 3: Salve sua tabela de dados de entrada em um formato de dados compatível
- Etapa 4: Carregue sua tabela de dados de entrada para o Amazon S3
- Etapa 5: criar uma AWS Glue tabela

Etapa 1: Assine um serviço de provedor em AWS Data Exchange

Se você tiver uma assinatura com um serviço de provedor por meio de AWS Data Exchange, poderá executar um fluxo de trabalho correspondente com um dos seguintes serviços de provedor para combinar seus identificadores conhecidos com seu provedor preferido. Seus dados serão combinados com um conjunto de entradas definido pelo seu provedor preferido.

Para assinar um serviço de provedor em AWS Data Exchange

1. Veja a lista de provedores em AWS Data Exchange. As seguintes listas de fornecedores estão disponíveis:
   - LiveRamp
     - LiveRampResolução de identidade
     - LiveRampTranscodificação
   - TransUnion
     - TruAudience Resolução e enriquecimento de identidade
   - ID unificada 2.0
     - Resolução de identidade unificada de ID 2.0
2. Conclua uma das etapas a seguir, dependendo do tipo de oferta.
   - Oferta privada — Se você já tem um relacionamento com um fornecedor, siga o procedimento de produtos e ofertas privadas no Guia AWS Data Exchange do usuário para aceitar uma oferta privada em AWS Data Exchange.
   - Traga sua própria assinatura — Se você já tem uma assinatura de dados existente com um provedor, siga o procedimento de ofertas Traga sua própria assinatura (BYOS) no Guia do AWS Data Exchange usuário para aceitar uma oferta de BYOS em. AWS Data Exchange

3. Depois de assinar um serviço de provedor em AWS Data Exchange, você pode criar um fluxo de trabalho correspondente ou um fluxo de trabalho de mapeamento de ID com esse serviço de provedor.

Para obter mais informações sobre como acessar um produto do provedor que contém APIs, consulte Acessando um produto de API no Guia do AWS Data Exchange usuário.

Etapa 2: Preparar tabelas de dados de terceiros

Cada serviço terceirizado tem um conjunto diferente de recomendações e diretrizes para ajudar a garantir um fluxo de trabalho de correspondência bem-sucedido.

Para preparar tabelas de dados de terceiros, consulte a tabela a seguir:

Diretrizes de serviços para provedores de dados

Serviço do provedor	É necessário um ID exclusivo?	Ações
LiveRamp	Sim	Garanta o seguinte: • O ID exclusivo pode ser seu próprio identificador pseudônimo ou um ID de linha. • O formato e a normalização do arquivo de entrada de dados estão alinhados com as LiveRamp diretrizes. Para obter mais informações sobre as diretrizes de formatação do arquivo de entrada para o fluxo de trabalho correspon dente, consulte Executar resolução de identidade por meio do ADX na LiveRamp documentação. Para obter mais informações sobre as diretrizes de formatação do arquivo de entrada para o fluxo de trabalho de mapeamento de ID, consulte Executar

Serviço do provedor	É necessário um ID exclusivo?	Ações
		transcodificação por meio do ADX na documentação. LiveRamp

Serviço do provedor	É necessário um ID exclusivo?	Ações

TransUnion	Sim	Verifique se o seguinte é uma coluna `string` de tipo na exibição de entrada:
		• É necessário um ID exclusivo e pode ser um ID de CRM, um ID de contato, um ID de usuário ou qualquer ID exclusivo.
		• **`Name`**
		  • **`First Name`**podem ser maiúsculas ou minúsculas, apelidos são suportados, mas títulos e sufixos devem ser excluídos .
		  • **`Last Name`**podem ser minúsculas ou maiúsculas, as iniciais médias devem ser excluídas.
		• **`Address`**
		  • **`Street address1`**e **`Street address1`** é combinado em uma única **`Full address`** linha, se presente.
		  • **`City`**é separado do**`Full address`**.
		  • **`Zip`**(ou**`zip plus4`**), sem nenhum caractere especial, como espaços, hífens ou espaços em branco. Use nulos se não houver dados.
		  • **`State`**é especificado como um código de 2 letras em maiúsculas.
		• • **`Phone`**
		  • **`Phone number`**deve ter 10 dígitos, sem caracteres especiais, como espaços ou hífens.

Serviço do provedor	É necessário um ID exclusivo?	Ações

		• **`Email addresses`** é texto simples ou cadeias de caracteres minúsculas SHA256 com hash.
		• **`Date of Birth`**está no `yyy-mm-dd` formato y.
		• **`Digital identifiers`** (Dispositivo IDs) pode incluir IDs com hífens (dispositivo bruto de 36 caracteres IDsMAIDs/IFAs) e sem hífens (dispositivo com hash de 32 e 40 caracteres/). IDs MAIDs IFAs
		  • **`IPV4`**é um endereço IP de 32 bits expresso em notação decimal com pontos. Por exemplo: `192.0.2.1`
		  • **`IPV6`**é um endereço IP de 128 bits expresso em notação hexadecimal, separado por dois pontos. Por exemplo: `2001:db8:0000:0000:0000:0000:0000:0001`
		  • **`MAID`**(ID de publicidade móvel) é uma sequência alfanumérica exclusiva atribuída a um dispositivo móvel para fins publicitários. Uma MAID geralmente tem 36 caracteres. Por exemplo: `a1b2c3d4-5678-90ab-cdef-EXAMPLE11111`

Serviço do provedor	É necessário um ID exclusivo?	Ações
ID unificada 2.0	Sim	Garanta o seguinte: • O ID exclusivo não pode ser um hash. • Um **`Phone number`** ou **`Email addresses`** é usado no esquema, não ambos. • UID2 suporta e-mail e número de telefone para UID2 geração. No entanto, se os dois valores estiverem presentes no mapeament o do esquema, o fluxo de trabalho duplicará cada registro na saída. Um registro usa o e-mail para UID2 geração e o segundo registro usa o número de telefone. Se seus dados incluírem uma combinação de e-mails e números de telefone e você não quiser essa duplicação de registros na saída, a melhor abordagem é criar um fluxo de trabalho separado para cada um, com mapeamentos de esquema separados. Nesse cenário, siga as etapas duas vezes: crie um fluxo de trabalho para e-mails e outro separado para números de telefone. **Note** Um e-mail ou número de telefone específico, em qualquer momento específico, resulta no mesmo UID2 valor bruto, independentemente de quem fez a solicitação. UID2s Os crus são criados pela adição de sais de baldes de sal que são girados aproximadamente uma

Serviço do provedor	É necessário um ID exclusivo?	Ações
		vez por ano, fazendo com que o cru também seja girado UID2 com ele. Diferentes baldes de sal giram em épocas diferentes ao longo do ano. AWS Entity Resolution atualment e não acompanha a rotação de baldes de sal e crus UID2s, por isso é recomendável que você regenere o cru diariamente. UID2s Para obter mais informações, consulte Com que frequência as atualizações increment ais devem UID2s ser atualizadas? na documentação do UID 2.0.

Etapa 3: Salve sua tabela de dados de entrada em um formato de dados compatível

Se você já salvou seus dados de entrada de terceiros em um formato de dados compatível, você pode pular esta etapa.

Para serem usados AWS Entity Resolution, os dados de entrada devem estar em um formato AWS Entity Resolution compatível.

AWS Entity Resolution suporta os seguintes formatos de dados:

- valor separado por vírgula (CSV)

> **Note**
>
> LiveRamp só oferece suporte a arquivos CSV.

- Parquet

Etapa 4: Carregue sua tabela de dados de entrada para o Amazon S3

Se você já tem sua tabela de dados de terceiros no Amazon S3, você pode pular esta etapa.

> **Note**
>
> Os dados de entrada devem ser armazenados no Amazon Simple Storage Service (Amazon S3) no Conta da AWS mesmo local Região da AWS e no qual você deseja executar o fluxo de trabalho correspondente.

Para carregar sua tabela de dados de entrada para o Amazon S3

1. Faça login no AWS Management Console e abra o console do Amazon S3 em. https://console.aws.amazon.com/s3/
2. Escolha Buckets e, em seguida, escolha um bucket para armazenar sua tabela de dados.
3. Escolha Upload e siga as instruções.
4. Escolha a guia Objetos para visualizar o prefixo do onde seus dados são armazenados. Anote o nome da pasta.

   Você pode selecionar a pasta para visualizar a tabela de dados.

Etapa 5: criar uma AWS Glue tabela

Os dados de entrada no Amazon S3 devem ser catalogados AWS Glue e representados como uma tabela. AWS Glue Para obter mais informações sobre como criar uma AWS Glue tabela com o Amazon S3 como entrada, consulte Como trabalhar com rastreadores no AWS Glue console no Guia do desenvolvedor.AWS Glue

> **Note**
>
> AWS Entity Resolution não oferece suporte a tabelas particionadas.

Nesta etapa, você configura um rastreador AWS Glue que rastreia todos os arquivos em seu bucket do S3 e cria uma tabela. AWS Glue

> **Note**
>
> AWS Entity Resolution atualmente não oferece suporte a locais do Amazon S3 registrados com. AWS Lake Formation

Para criar uma AWS Glue tabela

1. Faça login no AWS Management Console e abra o AWS Glue console em https://console.aws.amazon.com/glue/.
2. Na barra de navegação, selecione Crawlers.
3. Selecione o bucket do S3 na lista e escolha Adicionar crawler.
4. Na página Adicionar crawler, insira um nome do crawler e escolha Avançar.
5. Continue na página Adicionar crawler, especificando os detalhes.
6. Na página Escolher uma função do IAM, escolha Escolher um perfil do IAM existente e, em seguida, escolha Avançar.

   Você também pode escolher Criar um perfil do IAM ou fazer com que seu administrador crie o perfil do IAM, se necessário.
7. Em Criar uma programação para esse crawler, mantenha a Frequência padrão (Executar sob demanda) e escolha Avançar.
8. Em Configurar a saída do rastreador, insira o AWS Glue banco de dados e escolha Avançar.
9. Revise os detalhes e depois escolha Concluir.
10. Na página Crawlers, marque a caixa de seleção ao lado do bucket S3 e escolha Executar crawler.
11. Depois que o rastreador terminar de ser executado, na barra de AWS Glue navegação, escolha Bancos de dados e, em seguida, escolha o nome do banco de dados.
12. Na página Banco de dados, escolha Tabelas em {nome do seu banco de dados}.

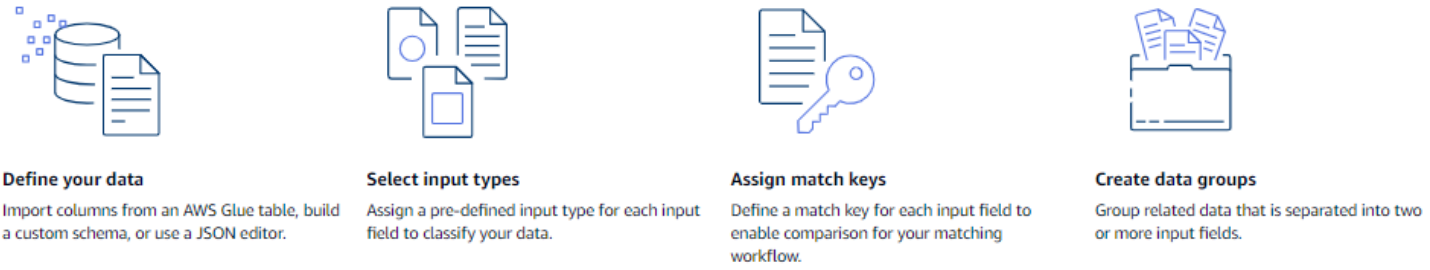
    a. Visualize as tabelas no AWS Glue banco de dados.
    b. Para visualizar o esquema de uma tabela, selecione uma tabela específica.
    c. Anote o nome do AWS Glue banco de dados e o nome AWS Glue da tabela.

Agora você está pronto para criar um mapeamento de esquema. Para obter mais informações, consulte Criação de um mapeamento de esquema.

Definir dados de entrada usando mapeamento de esquema

Um mapeamento de esquema define os dados de entrada que você deseja resolver. Ele também fornece metadados sobre os dados de entrada, como os tipos de atributos das colunas (campos de entrada) e com quais colunas corresponder.

Ao criar um mapeamento de esquema, primeiro você define seus campos de entrada e tipos de atributos e, em seguida, define suas chaves de correspondência e dados relacionados ao grupo. O diagrama a seguir resume como criar um mapeamento de esquema.



Antes de criar um mapeamento de esquema, você deve primeiro configurar AWS Entity Resolution e preparar suas tabelas de dados. Para obter mais informações, consulte Configurar AWS Entity Resolution e Prepare tabelas de dados de entrada.

Depois de criar um mapeamento de esquema, você pode fazer o seguinte:

- Crie um fluxo de trabalho correspondente para encontrar correspondências entre diferentes entradas de dados.
- Crie uma fonte de namespace de ID que você possa usar em um fluxo de trabalho de mapeamento de ID para traduzir dados de uma fonte para um destino.
- Crie um fluxo de trabalho de mapeamento de ID dentro do mesmo Conta da AWS usando seu mapeamento de esquema como fonte.

Tópicos

- Criação de um mapeamento de esquema
- Clonando um mapeamento de esquema
- Editando um mapeamento de esquema
- Excluindo um mapeamento de esquema

Criação de um mapeamento de esquema

Esse procedimento descreve o processo de criação de um mapeamento de esquema usando o AWS Entity Resolution console.

Há três maneiras de criar um mapeamento de esquema:

- Importar dados de entrada existentes usando a AWS Glue opção Importar de — Use esse método de criação para definir campos de entrada começando com colunas pré-preenchidas de uma AWS Glue tabela usando um fluxo guiado.
- Definindo manualmente os dados de entrada usando a opção Criar esquema personalizado — Use esse método de criação para definir manualmente os campos de entrada usando um fluxo guiado.
- Criar manualmente usando a opção Usar editor JSON — Use um editor JSON para criar, usar uma amostra ou importar manualmente os dados de entrada existentes.

> **Note**
>
> Os campos ID exclusivo e Entrada não estão disponíveis com essa opção.

Import from AWS Glue

Para criar mapeamento de esquema importando dados de entrada existentes do AWS Glue

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Preparação de dados, escolha Mapeamentos do esquema.
3. Na página Mapeamentos do esquema, no canto superior direito, escolha Criar mapeamento do esquema.
4. Para a Etapa 1: Especificar detalhes do esquema, faça o seguinte:
    a. Em Nome e método de criação, insira um nome de mapeamento do esquema e uma Descrição opcional.
    b. Em Método de criação, escolha Importar de AWS Glue.

c. Escolha o AWS Glue banco de dados na lista suspensa e, em seguida, escolha a AWS Glue tabela na lista suspensa.

   Para criar uma nova tabela, acesse o AWS Glue console https://console.aws.amazon.com/glue/. Para obter mais informações, consulte AWS Glue as tabelas no Guia AWS Glue do usuário.

d. Para ID exclusivo, especifique a coluna que faz referência distinta a cada linha de seus dados.

   Example

   Por exemplo: **`Primary_key`**, **`Row_ID`** ou **`Record_ID`**.

   > **Note**
   >
   > A coluna ID exclusiva é obrigatória. O ID exclusivo deve ser um identificador exclusivo em uma única tabela. No entanto, em tabelas diferentes, o ID exclusivo pode ter valores duplicados. Se a ID exclusiva não for especificada, não for exclusiva na mesma fonte ou se sobrepor em termos de nomes de atributos nas fontes, AWS Entity Resolution rejeitará o registro quando o fluxo de trabalho correspondente for executado. Se você estiver usando esse mapeamento de esquema em um fluxo de trabalho de correspondência baseado em regras, a ID exclusiva não deverá exceder 38 caracteres.

e. Em Campos de entrada, escolha as colunas que você deseja usar para correspondência e para passagem opcional.

   Você pode escolher um máximo de 34 colunas no total para combinar e passar.

   i. Em Correspondência, escolha as colunas que você deseja usar como campos de entrada para correspondência.

      Você pode escolher um máximo de 24 colunas no total para correspondência.

   ii. Selecione Adicionar colunas para passar se quiser especificar as colunas que não são usadas para correspondência.

   iii. (Opcional) Em Passar, escolha as colunas a serem incluídas como colunas de passagem.

f. (Opcional) Se você quiser ativar Tags para o recurso, escolha Adicionar nova tag e, em seguida, insira o par Chave e Valor.

g. Escolha Próximo.

5. Para a Etapa 2: Mapear campos de entrada, defina os campos de entrada que você deseja usar para correspondência e para passagem opcional.

   a. Para campos de entrada para correspondência, para cada campo de entrada,

      - Especifique o tipo de atributo para classificar os dados.
      - Especifique o nome da chave de correspondência para permitir a comparação do campo de entrada com seu fluxo de trabalho correspondente. Por padrão, determinados nomes de chaves de correspondência são automaticamente associados a tipos de atributos específicos.
      - Marque a caixa de seleção Com hash se o valor da coluna desse campo de entrada estiver com hash ou deixe a caixa de seleção em branco se o valor for texto não criptografado.

> **Note**
>
> Se você estiver criando um mapeamento de esquema para usar com a técnica de correspondência baseada em serviços do LiveRamp provedor, poderá:
>
> - Especifique o tipo de atributo para o ID do provedor como LiveRamp ID.
> - Especifique o tipo de atributo para o campo de nome como vários campos (como nome, sobrenome) ou em um campo.
> - Especifique o tipo de atributo para o campo de endereço residencial como vários campos (como endereço 1, endereço 2) ou em um campo (endereço completo).
>
>   Se corresponder a um endereço, é necessário um CEP (CEP).
> - Se você incluir e-mail (endereço de e-mail) ou telefone (número de telefone) com um nome, esses campos podem corresponder ao endereço da rua.

> **Note**
>
> Se você estiver criando um mapeamento de esquema para usar com a técnica de correspondência baseada em serviços do TransUnion provedor, poderá especificar qualquer um dos seguintes tipos de atributos:
>
> - Nome completo, primeiro nome, sobrenome
> - Endereço completo, endereço 1, cidade, estado, país, código postal
> - Número de telefone
> - Endereço de e-mail
> - Data
> - Identificadores digitais: IPV4, IPV6, ou MAID

> **Note**
>
> Se você estiver criando um mapeamento de esquema para usar com o fluxo de trabalho de correspondência baseado em aprendizado de máquina, seu conjunto de dados deverá conter pelo menos um dos seguintes tipos de atributos:
>
> - Nome completo
> - Endereço completo
> - Telefone completo
> - Endereço de e-mail
> - Data com uma chave de correspondência (nome da data de nascimento)
>
> Não especifique o tipo de atributo para nenhum desses atributos como uma string personalizada.

b. (Opcional) Para campos de entrada a serem transmitidos, adicione os campos de entrada que não serão correspondidos e o status de hash correspondente.

   O status de hash indica se o valor da coluna desse campo de entrada é criptografado ou não criptografado.

c. Escolha Próximo.

6. Para a Etapa 3: Agrupar dados, você pode agrupar os campos de entrada Nome, Endereço e Número de telefone se eles tiverem sido separados em vários campos.

   Essa etapa concatena os campos de entrada relacionados em um campo, o que permite compará-los como um campo em um fluxo de trabalho correspondente.

   Se você não tiver nenhum dado mapeado para os campos de entrada Nome, Endereço ou Número de telefone, essa seção ficará em branco.

   Você também pode adicionar mais grupos se tiver mais tipos de dados.

   a. Se você quiser agrupar os dados de entrada do Nome:

      Em Nome completo, escolha dois ou mais campos de entrada que você deseja agrupar.

      O nome do grupo e a chave de correspondência são automaticamente associados ao tipo de dados.

      Você pode atualizar o nome do grupo e a tecla de correspondência com uma chave de correspondência personalizada que pode conter até 255 caracteres, incluindo letras, números, sublinhados (_) ou hífens (-).

      Escolha Adicionar grupo para adicionar outro grupo.

      > **Note**
      >
      > A normalização só é suportada para o nome completo.
      > Se você quiser normalizar os subtipos de nome completo, atribua os seguintes subtipos ao grupo Nome completo: Nome, segundo nome e sobrenome.

   b. Se você quiser agrupar os dados de entrada de endereço:

      Em Endereço completo, escolha dois ou mais campos de campos de entrada que você deseja agrupar.

      O nome do grupo e a chave de correspondência são automaticamente associados ao tipo de dados.

Você pode atualizar o nome do grupo e a tecla de correspondência com uma chave de correspondência personalizada que pode conter até 255 caracteres, incluindo letras, números, sublinhados (_) ou hífens (-).

Escolha Adicionar grupo para adicionar outro grupo.

> **Note**
>
> A normalização só é suportada para endereço completo.
> Se você quiser normalizar os subtipos de endereço completo, atribua os seguintes subtipos ao grupo Endereço completo: Endereço 1, Endereço 2: nome do endereço 3, nome da cidade, estado, país e código postal.

c. Se você quiser agrupar os dados de entrada do telefone:

Para Telefone completo, escolha dois ou mais campos de campos de entrada que você deseja agrupar.

O nome do grupo e a chave de correspondência são automaticamente associados ao tipo de dados.

Você pode atualizar o nome do grupo e a tecla de correspondência com uma chave de correspondência personalizada que pode conter até 255 caracteres, incluindo letras, números, sublinhados (_) ou hífens (-).

Escolha Adicionar grupo para adicionar outro grupo.

> **Note**
>
> A normalização só é suportada para o telefone completo.
> Se você quiser normalizar os subtipos de telefone completo, atribua os seguintes subtipos ao grupo de telefone completo: Número de telefone e Código do país do telefone.

d. Escolha Próximo.

7. Para a Etapa 4: revisar e criar, faça o seguinte:

a. Revise as seleções feitas nas etapas anteriores e edite, se necessário.

b. Escolha Criar mapeamento de esquema.

> **Note**
>
> Você não pode modificar um mapeamento de esquema depois de associá-lo a um fluxo de trabalho. Você pode clonar um mapeamento de esquema se quiser usar uma configuração existente para criar um novo mapeamento de esquema.

Depois de criar o mapeamento do esquema, você estará pronto para criar um fluxo de trabalho correspondente ou criar um namespace de ID.

Build custom schema

Para criar um mapeamento de esquema usando a opção Criar esquema personalizado

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Preparação de dados, escolha Mapeamentos do esquema.
3. Na página Mapeamentos do esquema, no canto superior direito, escolha Criar mapeamento do esquema.
4. Para a Etapa 1: Especificar detalhes do esquema, faça o seguinte:

   a. Em nome e método de criação, insira um nome de mapeamento do esquema e uma Descrição opcional.

   b. Em Método de criação, escolha Criar esquema personalizado.

   c. Em ID exclusiva, insira uma ID exclusiva para identificar cada linha de seus dados.

      Example

      Por exemplo: **`Primary_key`**, **`Row_ID`** ou **`Record_ID`**.

> **Note**
>
> A coluna ID exclusiva é obrigatória. O ID exclusivo deve ser um identificador exclusivo em uma única tabela. No entanto, em tabelas diferentes, o ID exclusivo pode ter valores duplicados. Se a ID exclusiva não for especificada, não for

exclusiva na mesma fonte ou se sobrepor em termos de nomes de atributos nas fontes, AWS Entity Resolution rejeitará o registro quando o fluxo de trabalho correspondente for executado. Se você estiver usando esse mapeamento de esquema em um fluxo de trabalho de correspondência baseado em regras, a ID exclusiva não deverá exceder 38 caracteres.

d. (Opcional) Se você quiser ativar Tags para o recurso, escolha Adicionar nova tag e, em seguida, insira o par Chave e Valor.

e. Escolha Próximo.

5. Para a Etapa 2: Mapear campos de entrada, defina os campos de entrada que você deseja usar para correspondência e para passagem opcional.

   Você pode definir um máximo de 34 colunas no total para correspondência e passagem.

   a. Em Campos de entrada para correspondência, insira um campo de entrada.

   b. Selecione o tipo de atributo para classificar os dados.

> **Note**
>
> Se você estiver criando um mapeamento de esquema para usar com a técnica de correspondência baseada no serviço do LiveRamp provedor, poderá especificar o tipo de atributo providerID como ID. LiveRamp Se você quiser incluir dados de PII na saída, deverá especificar o tipo de atributo como Cadeia de caracteres personalizada.

> **Note**
>
> Se você estiver criando um mapeamento de esquema para usar com a técnica de correspondência baseada em serviços do TransUnion provedor, poderá especificar qualquer um dos seguintes tipos de atributos:
>
> - Nome completo, primeiro nome, sobrenome
> - Endereço completo, endereço 1, cidade, estado, país, código postal
> - Número de telefone
> - Endereço de e-mail
> - Data

- Identificadores digitais: IPV4, IPV6, ou MAID

> **Note**
>
> Se você estiver criando um mapeamento de esquema para usar com o fluxo de trabalho de correspondência baseado em aprendizado de máquina, seu conjunto de dados deverá conter pelo menos um dos seguintes tipos de atributos:
>
> - Nome completo
> - Endereço completo
> - Telefone completo
> - Endereço de e-mail
> - Data com uma chave de correspondência (nome da data de nascimento)
>
> Não especifique o tipo de atributo para nenhum desses atributos como uma string personalizada.

c. Selecione o nome da chave de correspondência para permitir a comparação do campo de entrada com seu fluxo de trabalho correspondente.

   Por padrão, determinados nomes de chaves de correspondência são automaticamente associados a tipos de atributos específicos.

d. Marque a caixa de seleção Com hash se o valor da coluna desse campo de entrada estiver com hash ou deixe a caixa de seleção em branco se o valor for texto não criptografado.

e. Escolha Adicionar campo de entrada para adicionar mais campos de entrada.

   Você pode adicionar no máximo 24 campos de entrada no total para correspondência.

f. (Opcional) Para os campos de entrada a serem transmitidos, adicione os campos de entrada que não serão correspondidos e o status de hash correspondente.

g. Escolha Próximo.

6. Para a Etapa 3: Agrupar dados, você pode agrupar os campos de entrada Nome, Endereço e Número de telefone se eles tiverem sido separados em vários campos.

Essa etapa concatena os campos de entrada relacionados em um campo, o que permite compará-los como um campo em um fluxo de trabalho correspondente.

Se você não tiver nenhum dado mapeado nos campos de entrada Nome, Endereço e Número de telefone, essa seção ficará em branco.

Você também pode adicionar mais grupos se tiver mais tipos de dados.

a. Se você quiser agrupar os dados de entrada do Nome:

   Em Nome completo, escolha dois ou mais campos de entrada que você deseja agrupar.

   O nome do grupo e a chave de correspondência são automaticamente associados ao tipo de dados.

   Você pode atualizar o nome do grupo e a tecla de correspondência com uma chave de correspondência personalizada que pode conter até 255 caracteres, incluindo letras, números, sublinhados (_) ou hífens (-).

   Escolha Adicionar grupo para adicionar outro grupo.

   > **Note**
   >
   > A normalização só é suportada para o nome completo.
   > Se você quiser normalizar os subtipos de nome completo, atribua os seguintes subtipos ao grupo Nome completo: Nome, segundo nome e sobrenome.

b. Se você quiser agrupar os dados de entrada de endereço:

   Em Endereço completo, escolha dois ou mais campos de campos de entrada que você deseja agrupar.

   O nome do grupo e a chave de correspondência são automaticamente associados ao tipo de dados.

   Você pode atualizar o nome do grupo e a tecla de correspondência com uma chave de correspondência personalizada que pode conter até 255 caracteres, incluindo letras, números, sublinhados (_) ou hífens (-).

   Escolha Adicionar grupo para adicionar outro grupo.

> **Note**
>
> A normalização só é suportada para endereço completo.
> Se você quiser normalizar os subtipos de endereço completo, atribua os seguintes subtipos ao grupo Endereço completo: Endereço 1, Endereço 2: nome do endereço 3, nome da cidade, estado, país e código postal.

c. Se você quiser agrupar os dados de entrada do telefone:

Para Telefone completo, escolha dois ou mais campos de campos de entrada que você deseja agrupar.

O nome do grupo e a chave de correspondência são automaticamente associados ao tipo de dados.

Você pode atualizar o nome do grupo e a tecla de correspondência com uma chave de correspondência personalizada que pode conter até 255 caracteres, incluindo letras, números, sublinhados (_) ou hífens (-).

Escolha Adicionar grupo para adicionar outro grupo.

> **Note**
>
> A normalização só é suportada para o telefone completo.
> Se você quiser normalizar os subtipos de telefone completo, atribua os seguintes subtipos ao grupo de telefone completo: Número de telefone e Código do país do telefone.

d. Escolha Próximo.

7. Para a Etapa 4: revisar e criar, faça o seguinte:

a. Revise as seleções feitas nas etapas anteriores e edite, se necessário.

b. Escolha Criar mapeamento de esquema.

> **Note**
>
> Você não pode modificar um mapeamento de esquema depois de associá-lo a um fluxo de trabalho. Você pode clonar um mapeamento de esquema se quiser usar uma configuração existente para criar um novo mapeamento de esquema.

Depois de criar o mapeamento do esquema, você estará pronto para criar um fluxo de trabalho correspondente ou criar um namespace de ID.

Use JSON editor

Para criar um mapeamento de esquema usando o editor JSON

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Preparação de dados, escolha Mapeamentos do esquema.
3. Na página Mapeamentos do esquema, no canto superior direito, escolha Criar mapeamento do esquema.
4. Para a Etapa 1: Especificar detalhes do esquema, faça o seguinte:
    a. Em nome e método de criação, insira um nome de mapeamento do esquema e uma Descrição opcional.
    b. Em Método de criação, escolha Usar editor JSON.
    c. (Opcional) Se você quiser ativar Tags para o recurso, escolha Adicionar nova tag e, em seguida, insira o par Chave e Valor.
    d. Escolha Próximo.
5. Para a Etapa 2: Especifique o mapeamento:
    a. Comece a criar o esquema no editor JSON ou escolha uma das seguintes opções com base em sua meta:

Seu objetivo	Opção recomendada
Comece a criar seu mapeamento de esquema	Insira uma amostra de JSON e edite as informações conforme necessário.
Use um arquivo JSON existente	Importar do arquivo

> **Note**
>
> A normalização só é suportada para os seguintes tipos: `NAMEADDRESS`,`PHONE`, e. `EMAIL_ADRESS`
> Se você quiser normalizar os `NAME` subtipos, atribua os seguintes subtipos ao `NAME` groupName:,, e `NAME_FIRST NAME_MIDDLE NAME_LAST`
> Se você quiser normalizar os `ADDRESS` subtipos, atribua os seguintes subtipos ao `ADDRESS` GroupName:,,,,, e. `ADDRESS_STREET1 ADDRESS_STREET2 ADDRESS_STREET3 ADDRESS_CITY ADDRESS_STATE ADDRESS_COUNTRY ADDRESS_POSTALCODE`
> Se você quiser normalizar os **`PHONE`** subtipos, atribua os seguintes subtipos ao groupName: e. **`PHONE`** `PHONE_NUMBER PHONE_COUNTRYCODE`

b. Escolha Próximo.

6. Para a Etapa 3: Revise e crie:

   a. Revise as seleções feitas nas etapas anteriores e edite, se necessário.

   b. Escolha Criar mapeamento de esquema.

> **Note**
>
> Você não pode modificar um mapeamento de esquema depois de associá-lo a um fluxo de trabalho. Você pode clonar um mapeamento de esquema se quiser usar uma configuração existente para criar um novo mapeamento de esquema.

Depois de criar o mapeamento do esquema, você estará pronto para criar um fluxo de trabalho correspondente ou criar um namespace de ID.

Clonando um mapeamento de esquema

Você pode clonar um mapeamento de esquema se quiser usar uma configuração existente para criar um novo mapeamento de esquema.

Para clonar um mapeamento de esquema:

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Preparação de dados, escolha Mapeamentos do esquema.
3. Escolha o mapeamento do esquema.
4. Escolha Clonar.
5. Na página Especificar detalhes do esquema, faça as alterações necessárias e escolha Avançar.
6. Na página Escolher técnica de correspondência, faça as alterações necessárias e escolha Avançar.
7. Na página Campos de entrada do mapa, faça as alterações necessárias e escolha Avançar.
8. Na página Dados do grupo, faça as alterações necessárias e escolha Avançar.
9. Na página Revisar e salvar, faça as alterações necessárias e escolha Clonar mapeamento do esquema.

Editando um mapeamento de esquema

Você só pode editar um mapeamento de esquema antes de associá-lo a um fluxo de trabalho. Depois de associar um mapeamento de esquema a um fluxo de trabalho, você não pode editá-lo. Você pode clonar um mapeamento de esquema se quiser usar uma configuração existente para criar um novo mapeamento de esquema.

Para editar um mapeamento de esquema:

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Preparação de dados, escolha Mapeamentos do esquema.
3. Escolha o mapeamento do esquema.

4. Escolha Editar.
5. Na página Especificar detalhes do esquema, faça as alterações necessárias e escolha Avançar.
6. Na página Escolher técnica de correspondência, faça as alterações necessárias e escolha Avançar.
7. Na página Campos de entrada do mapa, faça as alterações necessárias e escolha Avançar.
8. Na página Dados do grupo, faça as alterações necessárias e escolha Avançar.

> **Note**
>
> A normalização só é suportada para o nome completo, endereço completo, telefone completo e endereço de e-mail.
> Se você quiser normalizar os subtipos de nome completo, atribua os seguintes subtipos ao grupo Nome completo: Nome, segundo nome e sobrenome.
> Se você quiser normalizar os subtipos de endereço completo, atribua os seguintes subtipos ao grupo Endereço completo: Endereço 1, Endereço 2: nome do endereço 3, nome da cidade, estado, país e código postal.
> Se você quiser normalizar os subtipos de telefone completo, atribua os seguintes subtipos ao grupo de telefone completo: Número de telefone e Código do país do telefone.

9. Na página Revisar e salvar, faça as alterações necessárias e escolha Editar mapeamento do esquema.

Excluindo um mapeamento de esquema

Você não pode excluir um mapeamento de esquema quando ele está associado a um fluxo de trabalho correspondente. Primeiro, você deve remover o mapeamento do esquema de todos os fluxos de trabalho correspondentes associados antes de excluí-lo.

Para excluir um mapeamento de esquema:

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Preparação de dados, escolha Mapeamentos do esquema.
3. Escolha o mapeamento do esquema.

4. Escolha Excluir.
5. Confirme a exclusão e escolha Excluir.

Defina os dados de entrada usando um namespace de ID

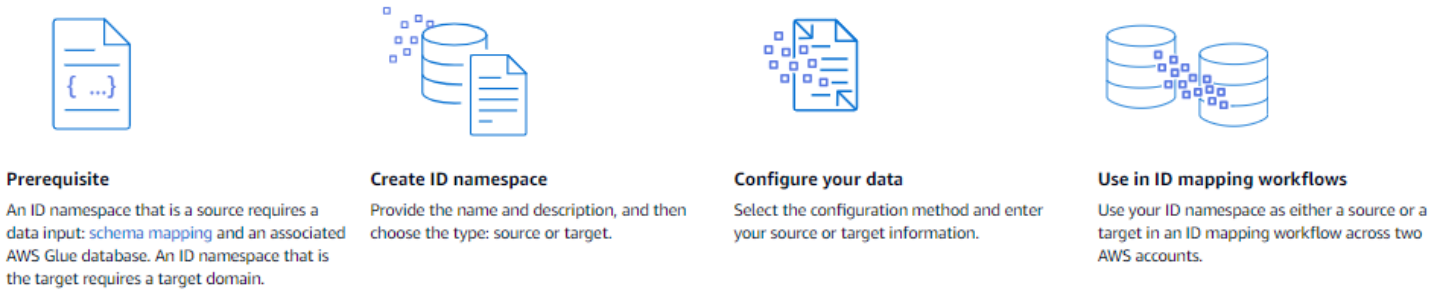
Um namespace de ID é um invólucro em torno da tabela de dados de entrada. Você usa um namespace de ID para fornecer metadados explicando seus dados de entrada e técnicas de correspondência e como usá-los em um fluxo de trabalho de mapeamento de ID.

Há dois tipos de namespace de ID: Origem e Destino.

- A Fonte contém configurações para os dados de origem que são AWS Entity Resolution processados em um fluxo de trabalho de mapeamento de ID.
- O Target contém uma configuração dos dados de destino para os quais todas as fontes resolvem.

Você pode definir os dados de entrada que deseja resolver em dois Contas da AWS em um fluxo de trabalho de mapeamento de ID. Um participante cria uma fonte de namespace de ID e outro cria um destino de namespace de ID. Depois que os participantes criarem a origem e o destino, você poderá executar um fluxo de trabalho de mapeamento de ID para traduzir os dados da origem para o destino.

O diagrama a seguir resume como criar um namespace de ID para usar em um fluxo de trabalho de mapeamento de ID.



As seções a seguir descrevem como criar uma origem de namespace de ID e um destino de namespace de ID.

Tópicos

- Fonte do namespace de ID
- Destino do namespace de ID
- Editando um namespace de ID
- Excluindo um namespace de ID

- Adicionar ou atualizar uma política de recursos para um namespace de ID

Fonte do namespace de ID

A fonte do namespace de ID é a fonte dos dados em um fluxo de trabalho de mapeamento de ID.

Antes de criar uma fonte de namespace de ID, você deve primeiro criar um mapeamento de esquema ou um fluxo de trabalho correspondente, dependendo do seu caso de uso. Para obter mais informações, consulte Criação de um mapeamento de esquema e Combine os dados de entrada usando um fluxo de trabalho correspondente.

Depois de criar uma fonte de namespace de ID, você pode usá-la junto com um destino de namespace de ID em um fluxo de trabalho de mapeamento de ID. Para obter mais informações, consulte Mapeie dados de entrada usando um fluxo de trabalho de mapeamento de ID.

Há duas maneiras de criar uma fonte de namespace de ID no AWS Entity Resolution console: o método baseado em regras ou o método de serviços do provedor.

Tópicos

- Criação de uma fonte de namespace de ID (com base em regras)
- Criação de uma fonte de namespace de ID (serviços do provedor)

Criação de uma fonte de namespace de ID (com base em regras)

Este tópico descreve o processo de criação de uma fonte de namespace de ID usando o método baseado em regras. Esse método usa regras de correspondência para traduzir dados primários de uma fonte para um destino em um fluxo de trabalho de mapeamento de ID.

> **Note**
>
> Se os dados de entrada forem a fonte, eles deverão ter um mapeamento de esquema e um AWS Glue banco de dados associado.

Para criar uma fonte de namespace de ID (com base em regras)

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.

2. No painel de navegação esquerdo, em Preparação de dados, escolha Namespaces de ID.
3. Na página de namespaces de ID, no canto superior direito, escolha Criar namespace de ID.
4. Para obter detalhes, faça o seguinte:
   a. Em Nome do namespace ID, insira um nome exclusivo.
   b. (Opcional) Em Descrição, insira uma descrição opcional.
   c. Para o tipo de namespace de ID, escolha Origem.
5. Para o método de namespace ID, escolha Baseado em regras.
6. Em Entrada de dados, escolha o tipo de entrada que você deseja usar e, em seguida, execute as ações recomendadas.

Tipo de entrada	Ações recomendadas
Um mapeamento de esquema existente	1. Escolha Mapeamento do esquema. 2. Escolha o AWS Glue banco de dados, a AWS Glue tabela e o mapeamento do esquema na lista suspensa. Você pode adicionar até 20 entradas de dados.
Um fluxo de trabalho de correspondência existente	1. Escolha o fluxo de trabalho de correspon dência. 2. Escolha a conta associada ao namespace de ID: Sua Conta da AWS ou Outra. Conta da AWS 3. Dependendo do tipo de conta, selecione o nome do fluxo de trabalho correspon dente ou insira o ARN do fluxo de trabalho correspondente.

7. Para parâmetros de regra, faça o seguinte.
   a. Especifique os controles da regra escolhendo uma das opções a seguir com base em sua meta.

Seu objetivo	Opção recomendada
Permitir regras da origem e do destino	Sem preferência
Escolha se uma fonte, um destino ou ambos podem fornecer regras em um fluxo de trabalho de mapeamento de ID	Regras limitadas

Os controles de regras devem ser compatíveis entre a origem e o destino para serem usados em um fluxo de trabalho de mapeamento de ID. Por exemplo, se um namespace de ID de origem limitar as regras ao destino, mas o namespace de ID de destino limitar as regras à origem, isso vai gerar um erro.

b. Especifique as regras de correspondência escolhendo uma das seguintes opções com base no seu tipo de entrada de dados.

Tipo de entrada de dados	Ação recomendada
Mapeamento de esquemas	Escolha Adicionar outra regra para adicionar uma regra correspondente. Você pode aplicar até 25 regras de correspondência para definir seus critérios de correspondência.
Fluxo de trabalho correspondente	Escolha Usar regras do fluxo de trabalho correspondente ou Fornecer novas regras para definir suas regras de correspon dência.

8. Para parâmetros de comparação e correspondência, faça o seguinte.

   a. Especifique o tipo de comparação escolhendo uma das opções a seguir com base em sua meta.

Seu objetivo	Opção recomendada
Permita que qualquer tipo de comparação seja usado ao criar o fluxo de trabalho de mapeamento de ID.	Sem preferência
Encontre qualquer combinação de correspondências nos dados armazenados em vários campos de entrada, independe ntemente de os dados estarem no mesmo campo de entrada ou em um campo de entrada diferente.	Vários campos de entrada
Limite a comparação em um único campo de entrada, quando dados semelhant es armazenados em vários campos de entrada não devem ser correspondidos.	Campo de entrada único

b. Especifique o tipo de correspondência de registros escolhendo uma das seguintes opções com base em sua meta.

Seu objetivo	Opção recomendada
Permita que qualquer tipo de comparação seja usado ao criar o fluxo de trabalho de mapeamento de ID.	Sem preferência
Limite o tipo de correspondência de registro para armazenar somente um registro correspondente na origem para cada registro correspondente no destino ao criar o fluxo de trabalho de mapeament o de ID.	Correspondência limitada de registros and Uma fonte para um alvo

Seu objetivo	Opção recomendada
Limite o tipo de correspondência de registro para armazenar todos os registros correspondentes na origem para cada registro correspondente no destino ao criar o fluxo de trabalho de mapeamento de ID.	Correspondência limitada de registros and Várias fontes para um único alvo

> **Note**
>
> Você deve especificar limitações compatíveis para os namespaces de ID de origem e destino. Por exemplo, se um namespace de ID de origem limitar as regras ao destino, mas o namespace de ID de destino limitar as regras à origem, isso vai gerar um erro.

9. Especifique as permissões de acesso ao serviço escolhendo um nome de função de serviço existente na lista suspensa.
10. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e insira o par de chave e valor.
11. Selecione Criar namespace.

A fonte do namespace de ID é criada. Agora você está pronto para criar um destino de namespace de ID.

Criação de uma fonte de namespace de ID (serviços do provedor)

Este tópico descreve o processo de criação de uma fonte de namespace de ID usando o método Provider services. Esse método usa um serviço de provedor chamado LiveRamp. LiveRamp converte dados codificados de terceiros de uma fonte para um destino durante um fluxo de trabalho de mapeamento de ID.

> **Note**
>
> Se os dados de entrada forem a fonte, eles deverão ter um mapeamento de esquema e um AWS Glue banco de dados associado.

Para criar uma fonte de namespace de ID (serviços do provedor)

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Preparação de dados, escolha Namespaces de ID.
3. Na página de namespaces de ID, no canto superior direito, escolha Criar namespace de ID.
4. Para obter detalhes, faça o seguinte:
   a. Em Nome do namespace ID, insira um nome exclusivo.
   b. (Opcional) Em Descrição, insira uma descrição opcional.
   c. Para o tipo de namespace de ID, escolha Origem.
5. Para o método de namespace ID, escolha Provider services.

> **Note**
>
> AWS Entity Resolution atualmente oferece o serviço de LiveRamp provedor como um método de namespace de ID. Se você tiver uma assinatura LiveRamp, o status aparecerá como Assinado. Para obter mais informações sobre como assinar LiveRamp, consulteEtapa 1: Assine um serviço de provedor em AWS Data Exchange.

6. Em Entrada de dados, escolha o AWS Glue banco de dados, a AWS Glue tabela e o mapeamento do esquema na lista suspensa.

   Você pode adicionar até 20 entradas de dados.
7. Para especificar as permissões de acesso ao serviço, escolha uma opção e execute a ação recomendada.

Opção	Ação recomendada
Criar e usar um novo perfil de serviço	• AWS Entity Resolution cria uma função de serviço com a política necessária para essa tabela. • O nome padrão da função de serviço éentityresolution-id-mapping-workflow-<timestamp> .

Opção	Ação recomendada
	• Você deve ter permissões para criar perfis e anexar políticas. • Se seus dados de entrada estiverem criptografados, escolha a opção Esses dados são criptografados por uma chave KMS. Em seguida, insira uma AWS KMS chave usada para descriptografar sua entrada de dados.
Use um perfil de serviço existente	1. Escolha um nome do perfil de serviço existente na lista suspensa. A lista de perfis é exibida se você tiver permissões para listar funções. Se você não tiver permissões para listar perfis, insira o nome do recurso da Amazon (ARN) do perfil que você deseja usar. Se não houver perfis de serviço existente s, a opção de Usar um perfil de serviço existente não estará disponível. 2. Para visualizar o perfil de serviço, selecione o link externo Visualizar no IAM. Por padrão, AWS Entity Resolution não tenta atualizar a política de função existente para adicionar as permissões necessárias.

8. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e insira o par de chave e valor.
9. Selecione Criar namespace.

A fonte do namespace de ID é criada. Agora você está pronto para criar um destino de namespace de ID.

Destino do namespace de ID

O destino do namespace de ID é o destino dos dados em um fluxo de trabalho de mapeamento de ID. Todas as fontes são direcionadas para o alvo.

Antes de criar um destino de namespace de ID, você deve primeiro criar um fluxo de trabalho correspondente ou ter uma assinatura de um provedor service (LiveRamp), dependendo do seu caso de uso. Para obter mais informações, consulte Combine os dados de entrada usando um fluxo de trabalho correspondente e Etapa 1: Assine um serviço de provedor em AWS Data Exchange.

Depois de criar um destino de namespace de ID, você pode usá-lo junto com uma fonte de namespace de ID em um fluxo de trabalho de mapeamento de ID. Para obter mais informações, consulte Mapeie dados de entrada usando um fluxo de trabalho de mapeamento de ID.

Há duas maneiras de criar um destino de namespace de ID no AWS Entity Resolution console: o método baseado em regras ou o método de serviços do provedor.

Tópicos

- Criação de um destino de namespace de ID (método baseado em regras)
- Criação de um destino de namespace de ID (método de serviços do provedor)

Criação de um destino de namespace de ID (método baseado em regras)

Este tópico descreve o processo de criação de um destino de namespace de ID usando o método baseado em regras. Esse método usa regras de correspondência para traduzir dados primários de uma fonte para um destino durante um fluxo de trabalho de mapeamento de ID.

Para criar um destino de namespace de ID (baseado em regras)

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Preparação de dados, escolha Namespaces de ID.
3. Na página de namespaces de ID, no canto superior direito, escolha Criar namespace de ID.
4. Para obter detalhes, faça o seguinte:

a. Em Nome do namespace ID, insira um nome exclusivo.

b. (Opcional) Em Descrição, insira uma descrição opcional.

c. Para o tipo de namespace de ID, escolha Target.

5. Para o método de namespace ID, escolha Baseado em regras.

6. Para entrada de dados, em Fluxo de trabalho correspondente, faça o seguinte.

a. Escolha a conta associada ao namespace de ID: Sua Conta da AWS ou Outra. Conta da AWS

b. Dependendo do tipo de conta, selecione o nome do fluxo de trabalho correspondente ou insira o ARN do fluxo de trabalho correspondente.

7. Para parâmetros de regra, faça o seguinte.

a. Especifique os controles da regra escolhendo uma das opções a seguir com base em sua meta.

Seu objetivo	Opção recomendada
Permitir regras da origem e do destino	Sem preferência
Escolha se uma fonte, um destino ou ambos podem fornecer regras em um fluxo de trabalho de mapeamento de ID	Regras limitadas

Os controles de regras devem ser compatíveis entre a origem e o destino para serem usados em um fluxo de trabalho de mapeamento de ID. Por exemplo, se um namespace de ID de origem limitar as regras ao destino, mas o namespace de ID de destino limitar as regras à origem, isso vai gerar um erro.

b. Para regras de correspondência, adiciona AWS Entity Resolution automaticamente as regras do fluxo de trabalho correspondente.

8. Para parâmetros de comparação e correspondência, faça o seguinte.

a. Especifique o tipo de comparação escolhendo uma das opções a seguir com base em sua meta.

Seu objetivo	Opção recomendada
Permita que qualquer tipo de comparação seja usado ao criar o fluxo de trabalho de mapeamento de ID.	Sem preferência
Encontre qualquer combinação de correspondências nos dados armazenados em vários campos de entrada, independe ntemente de os dados estarem no mesmo campo de entrada ou em um campo de entrada diferente.	Vários campos de entrada
Limite a comparação em um único campo de entrada, quando dados semelhant es armazenados em vários campos de entrada não devem ser correspondidos.	Campo de entrada único

b. Especifique o tipo de correspondência de registros escolhendo uma das seguintes opções com base em sua meta.

Seu objetivo	Opção recomendada
Permita que qualquer tipo de comparação seja usado ao criar o fluxo de trabalho de mapeamento de ID.	Sem preferência
Limite o tipo de correspondência de registro para armazenar somente um registro correspondente na origem para cada registro correspondente no destino ao criar o fluxo de trabalho de mapeament o de ID.	Correspondência limitada de registros and Uma fonte para um alvo

Seu objetivo	Opção recomendada
Limite o tipo de correspondência de registro para armazenar todos os registros correspondentes na origem para cada registro correspondente no destino ao criar o fluxo de trabalho de mapeamento de ID.	Correspondência limitada de registros and Várias fontes para um único alvo

> **Note**
>
> Você deve especificar limitações compatíveis para os namespaces de ID de origem e destino. Por exemplo, se um namespace de ID de origem limitar as regras ao destino, mas o namespace de ID de destino limitar as regras à origem, isso vai gerar um erro.

9. Especifique as permissões de acesso ao serviço escolhendo um nome de função de serviço existente na lista suspensa.
10. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e insira o par de chave e valor.
11. Selecione Criar namespace.

O destino do namespace ID é criado. Depois de criar os namespaces de ID (origem e destino) necessários para um fluxo de trabalho de mapeamento de ID, você está pronto para criar um fluxo de trabalho de mapeamento de ID.

Criação de um destino de namespace de ID (método de serviços do provedor)

Este tópico descreve o processo de criação de um destino de namespace de ID usando o método Provider services. Esse método usa um serviço de provedor chamado LiveRamp. LiveRamp converte dados codificados de terceiros de uma fonte para um destino durante um fluxo de trabalho de mapeamento de ID.

Para criar um destino de namespace de ID (serviços do provedor)

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Preparação de dados, escolha Namespaces de ID.
3. Na página de namespaces de ID, no canto superior direito, escolha Criar namespace de ID.
4. Para obter detalhes, faça o seguinte:
   a. Em Nome do namespace ID, insira um nome exclusivo.
   b. (Opcional) Em Descrição, insira uma descrição opcional.
   c. Para o tipo de namespace de ID, escolha Target.
5. Para o método de namespace ID, escolha Provider services.

> **Note**
>
> AWS Entity Resolution atualmente oferece o serviço de LiveRamp provedor como um método de namespace de ID.
> Se você tiver uma assinatura LiveRamp, o status aparecerá como Assinado.
> Para obter mais informações sobre como assinar LiveRamp, consulteEtapa 1: Assine um serviço de provedor em AWS Data Exchange.

6. Em Domínio de destino, insira o identificador de domínio LiveRamp do cliente destinado à transcodificação que LiveRamp fornece.
7. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e insira o par de chave e valor.
8. Selecione Criar namespace.

O destino do namespace ID é criado. Depois de criar os namespaces de ID (origem e destino) necessários para um fluxo de trabalho de mapeamento de ID, você está pronto para criar o fluxo de trabalho de mapeamento de ID.

Editando um namespace de ID

Você só pode editar um namespace de ID antes de associá-lo a um fluxo de trabalho de mapeamento de ID. Depois de associar um namespace de ID a um fluxo de trabalho de mapeamento de ID, você não pode editá-lo.

Para editar um namespace de ID:

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Preparação de dados, escolha Namespaces de ID.
3. Escolha o namespace de ID.
4. Escolha Editar.
5. Na página Editar namespace ID, faça as alterações necessárias e escolha Salvar.

Excluindo um namespace de ID

Você não pode excluir um namespace de ID quando ele está associado a um fluxo de trabalho de mapeamento de ID. Primeiro, você deve remover o mapeamento do esquema de todos os fluxos de trabalho de mapeamento de ID associados antes de excluí-lo.

Para excluir um namespace de ID:

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Preparação de dados, escolha Namespaces de ID.
3. Escolha o namespace de ID.
4. Escolha Excluir.
5. Confirme a exclusão e escolha Excluir.

Adicionar ou atualizar uma política de recursos para um namespace de ID

Uma política de recursos permite que o criador do recurso de mapeamento de ID acesse seu recurso de namespace de ID.

Para adicionar ou atualizar uma política de recursos

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha namespaces de ID.

3. Escolha o namespace de ID.
4. Na página de detalhes do namespace ID, escolha a guia Permissões.
5. Na seção Política de recursos, escolha Editar.
6. Adicione ou atualize a política no editor JSON.
7. Escolha Salvar alterações.

Combine os dados de entrada usando um fluxo de trabalho correspondente

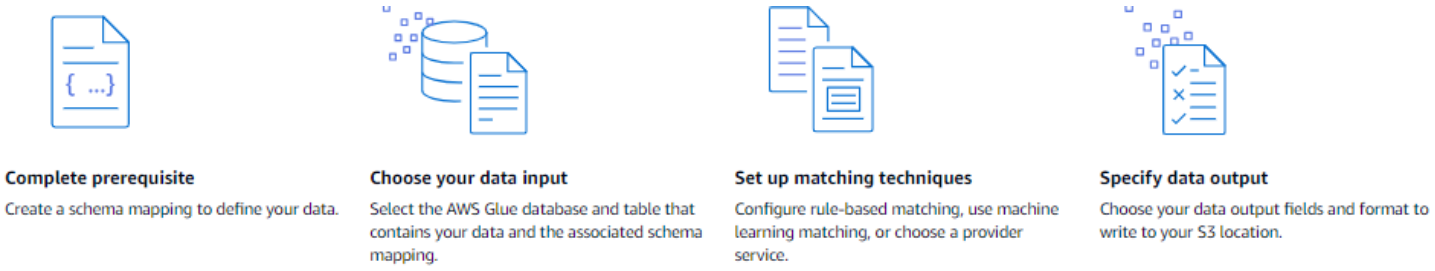
Um fluxo de trabalho de correspondência é um trabalho de processamento de dados que combina e compara dados de diferentes fontes de entrada e determina quais deles correspondem com base em diferentes técnicas de correspondência. Ele produz uma tabela de saída de dados.

Ao criar um fluxo de trabalho correspondente, primeiro você especifica suas entradas de dados, etapas de normalização e, em seguida, escolhe as técnicas de correspondência e a saída de dados desejadas. AWS Entity Resolution lê seus dados do local ou locais especificados e encontra uma correspondência entre dois ou mais registros em seus dados. Em seguida, ele atribui uma ID de correspondência aos registros no conjunto de dados correspondente. AWS Entity Resolution em seguida, grava os arquivos de saída de dados em um local que você escolher. Você pode usar AWS Entity Resolution para fazer o hash dos dados de saída, se desejar, ajudando você a manter o controle sobre seus dados.

Um fluxo de trabalho correspondente pode ter várias execuções e os resultados (acertos ou erros) são gravados em uma pasta com o `jobId` como nome.

A saída de dados contém um arquivo para correspondências bem-sucedidas e um arquivo para erros. A saída de dados pode conter vários campos. Os resultados bem-sucedidos são gravados em uma `success` pasta que contém vários arquivos, e cada arquivo contém um subconjunto dos registros bem-sucedidos. Da mesma forma, os erros são gravados em uma `error` pasta com vários campos, cada um contendo um subconjunto dos registros de erro. Para obter mais informações sobre a solução de problemas de erros, consulteSolução de problemas de fluxos de trabalho correspondentes.

O diagrama a seguir resume como criar um fluxo de trabalho correspondente.



Antes de criar um fluxo de trabalho correspondente, você deve primeiro criar um mapeamento de esquema. Para obter mais informações, consulte Criação de um mapeamento de esquema.

Há três maneiras de criar um fluxo de trabalho correspondente, com base em técnicas de correspondência: baseado em regras, baseado em aprendizadode máquina ou baseado em serviços do provedor.

Depois de criar e executar um fluxo de trabalho correspondente, você pode fazer o seguinte:

- Visualize os resultados no local do S3 que você especificou. Os fluxos de trabalho correspondentes são gerados IDs após a indexação dos dados.
- Use a saída da correspondência baseada em regras ou da correspondência de aprendizado de máquina (ML) como uma entrada para a correspondência baseada em serviços do provedor ou vice-versa para atender às suas necessidades comerciais.

Por exemplo, para economizar nos custos de assinatura do provedor, você pode primeiro executar a correspondência baseada em regras para encontrar correspondências em seus dados. Em seguida, você pode enviar um subconjunto de registros incomparáveis para a correspondência baseada em serviços do provedor.

Tópicos

- Criação de um fluxo de trabalho de correspondência baseado em regras
- Criação de um fluxo de trabalho de correspondência baseado em aprendizado de máquina
- Criação de um fluxo de trabalho de correspondência baseado em serviços do provedor
- Editando um fluxo de trabalho correspondente
- Excluindo um fluxo de trabalho correspondente
- Modificando ou gerando uma ID de correspondência para um fluxo de trabalho de correspondência baseado em regras
- Procurando um ID de correspondência para um fluxo de trabalho de correspondência baseado em regras
- Excluindo registros de um fluxo de trabalho de correspondência baseado em regras ou em ML
- Solução de problemas de fluxos de trabalho correspondentes

Criação de um fluxo de trabalho de correspondência baseado em regras

A correspondência baseada em regras é um conjunto hierárquico de regras de correspondência em cascata, sugerido por AWS Entity Resolution, com base nos dados que você insere e é totalmente configurável por você. O fluxo de trabalho de correspondência baseado em regras permite comparar texto não criptografado ou dados com hash para encontrar correspondências exatas com base nos critérios que você personaliza.

Quando AWS Entity Resolution encontra uma correspondência entre dois ou mais registros em seus dados, ele atribui:

- Um ID de correspondência para os registros no conjunto de dados correspondente
- A regra de correspondência que gerou a partida.

Ao criar um fluxo de trabalho de correspondência baseado em regras no AWS Entity Resolution, você deve escolher um tipo de regra simples ou avançado. O tipo de regra determina a complexidade das condições de regra que você pode criar. Você não pode alterar o tipo de regra depois de criar o fluxo de trabalho.

Você pode usar a tabela a seguir para comparar os dois tipos de regras e determinar qual delas se adequa ao seu caso de uso.

Gráfico de comparação de tipos de regras

Caso de uso	Tipo de regra avançada	Tipo de regra simples
Mapeamentos de esquema mapeados com tipos de entrada one-to-one	Sim	Não
Mapeamento de esquema com várias colunas de dados mapeadas para os mesmos tipos de entrada	Não	Sim
Suporta correspondência exata e difusa	Sim	Não (somente correspondência exata)

Caso de uso	Tipo de regra avançada	Tipo de regra simples
Suporta operadores AND, OR e parênteses	Sim	Não (somente operador AND)
Suporta fluxos de trabalho em lote	Sim	Sim
Oferece suporte a fluxos de trabalho incrementais	Sim	Sim
Suporta fluxos de trabalho em tempo real	Não	Sim
Suporta fluxos de trabalho de mapeamento de ID	Não	Sim

Depois de determinar qual tipo de regra você deseja usar, use os tópicos a seguir para criar um fluxo de trabalho de correspondência baseado em regras com o tipo de regra Avançado ou Simples.

Tópicos

- Criação de um fluxo de trabalho de correspondência baseado em regras com o tipo de regra Avançado
- Criação de um fluxo de trabalho de correspondência baseado em regras com o tipo de regra simples

Criação de um fluxo de trabalho de correspondência baseado em regras com o tipo de regra Avançado

O procedimento a seguir demonstra como criar um fluxo de trabalho de correspondência baseado em regras com o tipo de regra Avançado usando o AWS Entity Resolution console ou a API. `CreateMatchingWorkflow`

Console

Para criar um fluxo de trabalho de correspondência baseado em regras com o tipo de regra Avançado usando o console

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Correspondência.
3. Na página Fluxos de trabalho correspondentes, no canto superior direito, escolha Criar fluxo de trabalho correspondente.
4. Para a Etapa 1: Especificar os detalhes correspondentes do fluxo de trabalho, faça o seguinte:

   a. Insira um nome de fluxo de trabalho correspondente e uma Descrição opcional.

   b. Em Entrada de dados, escolha um AWS Glue banco de dados na lista suspensa, selecione a AWS Glue tabela e, em seguida, o mapeamento do esquema correspondente.

      Você pode adicionar até 19 entradas de dados.

      > **Note**
      >
      > Para usar regras avançadas, seus mapeamentos de esquema devem atender aos seguintes requisitos:
      >
      > 1. Cada campo de entrada deve ser mapeado para uma chave de correspondência exclusiva, a menos que os campos estejam agrupados.
      > 2. Se os campos de entrada estiverem agrupados, eles poderão compartilhar a mesma chave de correspondência.
      >
      >   Por exemplo, o mapeamento de esquema a seguir seria válido para regras avançadas:
      >
      >    ```
      >    firstName: { matchKey: 'name', groupName: 'name' }
      >
      >    lastName: { matchKey: 'name', groupName: 'name' }
      >    ```

Nesse caso, os `lastName` campos `firstName` e são agrupados e compartilham a mesma chave de correspondência de nome, o que é permitido.

Revise seus mapeamentos de esquema e atualize-os para seguir essa regra de one-to-one correspondência, a menos que os campos estejam agrupados corretamente, para usar as regras avançadas.

3. Se sua tabela de dados tiver uma coluna DELETE, o tipo do mapeamento do esquema deve ser `String` e você não pode ter um `matchKey` e. `groupName`

c. A opção Normalizar dados é selecionada por padrão, para que as entradas de dados sejam normalizadas antes da correspondência. Se você não quiser normalizar dados, desmarque a opção Normalizar dados.

> **Note**
>
> A normalização só é suportada nos seguintes cenários em Criar mapeamento de esquema:
>
> - Se os seguintes subtipos de nome estiverem agrupados: Nome, segundo nome, sobrenome.
> - Se os seguintes subtipos de endereço estiverem agrupados: Endereço 1, Endereço 2, Endereço 3, Cidade, Estado, País, Código postal.
> - Se os seguintes subtipos de telefone estiverem agrupados: Número de telefone, Código do país do telefone.

d. Para especificar as permissões de acesso ao serviço, escolha uma opção e execute a ação recomendada.

Opção	Ação recomendada
Criar e usar um novo perfil de serviço	• AWS Entity Resolution cria uma função de serviço com a política necessária para essa tabela. • O nome do perfil de serviço padrão é `entityresolution-matching-workflow-<timestamp>` . • Você deve ter permissões para criar perfis e anexar políticas. • Se seus dados de entrada estiverem criptografados, você poderá escolher a opção Esses dados são criptografados com uma chave KMS e, em seguida, inserir uma AWS KMS chave que será usada para descriptografar sua entrada de dados.

Opção	Ação recomendada
Use um perfil de serviço existente	1. Escolha um nome do perfil de serviço existente na lista suspensa. A lista de perfis é exibida se você tiver permissões para listar funções. Se você não tiver permissões para listar perfis, insira o nome do recurso da Amazon (ARN) do perfil que você deseja usar. Se não houver perfis de serviço existentes, a opção de Usar um perfil de serviço existente não estará disponível. 2. Para visualizar o perfil de serviço, selecione o link externo Visualizar no IAM. Por padrão, AWS Entity Resolution não tenta atualizar a política de função existente para adicionar as permissõe s necessárias.

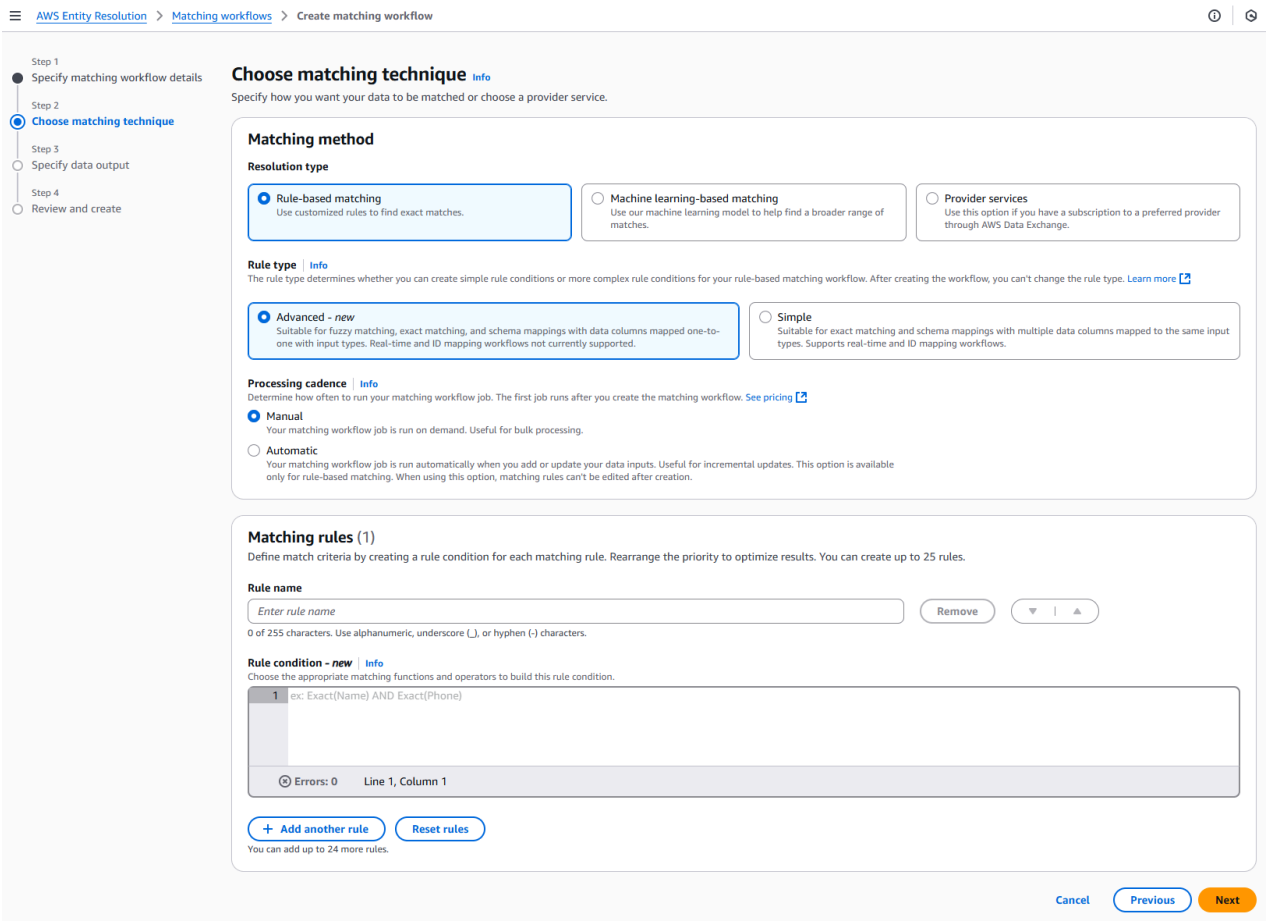
e. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e, em seguida, insira o par de chave e valor.

f. Escolha Próximo.

5. Para a Etapa 2: Escolha a técnica de correspondência:

a. Em Método de correspondência, escolha Correspondência baseada em regras.

b. Em Tipo de regra, escolha Avançado.



c. Em Cadência de processamento, selecione uma das opções a seguir.

- Escolha Manual para executar um fluxo de trabalho sob demanda para uma atualização em massa
- Escolha Automático para executar um fluxo de trabalho assim que novos dados estiverem em seu bucket do S3

> **Note**
>
> Se você escolher Automático, certifique-se de ter EventBridge as notificações da Amazon ativadas para seu bucket do S3. Para obter instruções sobre como habilitar a Amazon EventBridge usando o console do S3, consulte Habilitando a Amazon EventBridge no Guia do usuário do Amazon S3.

d. Em Regras de correspondência, insira um nome de regra e, em seguida, crie a condição da regra escolhendo as funções e operadores de correspondência apropriados na lista suspensa com base em sua meta.

Você pode criar até 25 regras.

Você deve combinar uma função de correspondência difusa (Cosine, Levenshtein ou Soundex) com uma função de correspondência exata (Exact,) usando o operador AND. ExactManyToMany

Você pode usar a tabela a seguir para ajudar a decidir que tipo de função ou operador deseja usar, dependendo da sua meta.

Seu objetivo	Função ou operador recomendados	Modificador opcional recomendado	Prós
Combine cadeias de caracteres idênticas em dados precisos, mas não corresponda a valores vazios.	Exato	EmptyValues=Processo	
Combine cadeias de caracteres idênticas em dados precisos e ignore valores vazios.	Exato (***matchKey***)	EmptyValues=Ignorar	
Combine vários registros nas teclas de partida. Adequado para emparelhamentos flexíveis. Limite: 15 teclas de partida	ExactManyToMany(*matchKey*, *matchKey*, ...)	n/a	

Seu objetivo	Função ou operador recomendados	Modificador opcional recomendado	Prós
Meça a semelhança entre as representações numéricas dos dados, mas não faça a correspondência em valores vazios. Adequado para texto, números ou uma combinação de ambos.	Cosseno	EmptyValues=Processo	Simples e eficiente. Funciona bem com texto longo quando combinado com a ponderação TF-IDF. Bom para correspondência exata baseada em palavras.
Meça a semelhança entre representações numéricas de dados e ignore valores vazios.	Cosseno (*matchKey*,*threshold*,...)	EmptyValues=Ignorar	Lida bem com erros de digitação, erros ortográficos e transposições. Eficaz em uma ampla variedade de tipos de PII. Bom para sequências curtas (por exemplo, nomes ou números de telefone).
Conte o número mínimo de alterações necessárias para transformar uma palavra em outra, mas não corresponda aos valores vazios. Adequado para texto com pequenas diferenças na ortografia.	Levenshtein	EmptyValues=Processo	

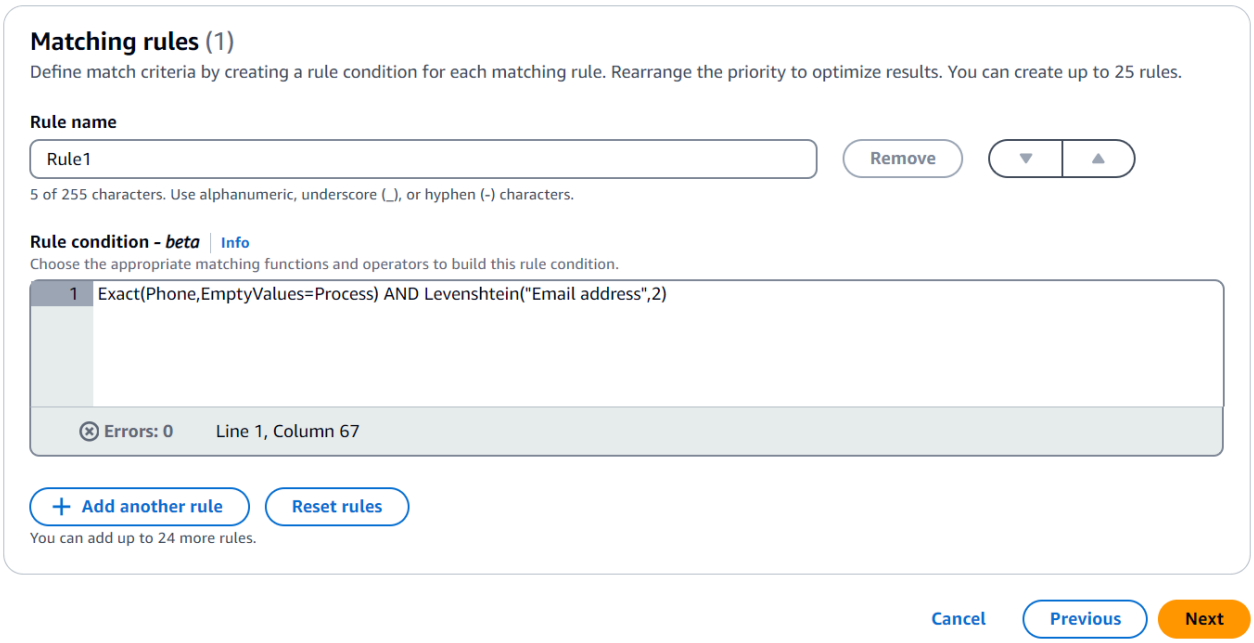
Seu objetivo	Função ou operador recomendados	Modificador opcional recomendado	Prós
Conte o número mínimo de alterações necessárias para transformar uma palavra em outra e ignore valores vazios.	Levenshtein (,,...) ***matchKey threshold***	EmptyValues=Ignorar	
Compare e combine cadeias de texto com base em quão parecidas elas soam, mas não coincidem em valores vazios. Adequado para texto com variações na ortografia ou pronúncia.	Soundex	EmptyValues=Processo	Eficaz para correspondência fonética, identificando palavras com sons semelhantes. Rápido e computacionalmente barato. Bom para combinar nomes com pronúncias semelhantes, mas com grafias diferentes.
Compare e combine cadeias de texto com base na semelhança entre elas e ignore valores vazios.	Som (1) ***matchKey***	EmptyValues=Ignorar	
Combine funções.	E	n/a	
Funções separadas.	OU	n/a	

Seu objetivo	Função ou operador recomendados	Modificador opcional recomendado	Prós
Agrupe condições para criar condições aninhadas.	(…)	n/a	

Example Condição de regra que corresponde aos números de telefone e e-mail

Veja a seguir um exemplo de uma condição de regra que corresponde a registros em números de telefone (chave de correspondência de telefone) e endereços de e-mail (chave de correspondência de endereço de e-mail):

```
Exact(Phone,EmptyValues=Process) AND Levenshtein("Email address",2)
```



A tecla Phone match usa a função de correspondência exata para combinar sequências idênticas. A tecla Phone match processa valores vazios na correspondência usando o modificador EmptyValues=Process.

A chave de correspondência de endereço de e-mail usa a função de correspondência Levenshtein para combinar dados com erros ortográficos usando o limite padrão do algoritmo de distância Levenshtein de 2. A tecla de correspondência de e-mail não usa nenhum modificador opcional.

O operador AND combina a função de correspondência exata e a função de correspondência Levenshtein.

Example Condição de regra usada ExactManyToMany para realizar a correspondência de teclas de correspondência

Veja a seguir um exemplo de uma condição de regra que combina registros em três campos de endereço (HomeAddresschave de BillingAddresscorrespondência, chave de ShippingAddresscorrespondência e chave de correspondência) para encontrar possíveis correspondências verificando se alguma delas tem valores idênticos.

O `ExactManyToMany` operador avalia todas as combinações possíveis dos campos de endereço especificados para identificar correspondências exatas entre dois ou mais endereços. Por exemplo, ele detectaria se os endereços `HomeAddress` correspondem a `BillingAddress` ou `ShippingAddress` ou se todos os três endereços correspondem exatamente.

```
ExactManyToMany(HomeAddress, BillingAddress, ShippingAddress)
```

Example Condição de regra que usa agrupamento

Na correspondência avançada baseada em regras com condições difusas, o sistema primeiro agrupa os registros em clusters com base nas correspondências exatas. Depois que esses clusters iniciais são formados, o sistema aplica filtros de correspondência difusa para identificar correspondências adicionais em cada cluster. Para um desempenho ideal, você deve selecionar condições de correspondência exatas com base em seus padrões de dados para criar clusters iniciais bem definidos.

Veja a seguir um exemplo de uma condição de regra que combina várias correspondências exatas com um requisito de correspondência difusa. Ele usa `AND` operadores para verificar se três campos —`FullName`, Data de nascimento (`DOB`) e `Address` — coincidem exatamente entre os registros. Também permite pequenas variações no `InternalID` campo usando uma distância de Levenshtein de. 1 A

distância de Levenshtein mede o número mínimo de edições de um único caractere necessárias para transformar uma string em outra. Uma distância de 1 significa `InternalIDs` que ela corresponderá à diferença em apenas um caractere (como um único erro de digitação, exclusão ou inserção). Essa combinação de condições ajuda a identificar registros que provavelmente representarão a mesma entidade, mesmo que haja pequenas discrepâncias no identificador.

```
Exact(FullName) AND Exact(DOB) AND Exact(Address) and
 Levenshtein(InternalID, 1)
```

e. Escolha Próximo.

6. Para a Etapa 3: Especifique a saída e o formato dos dados:

   a. Em Destino e formato de saída de dados, escolha a localização do Amazon S3 para a saída de dados e se o formato dos dados será dados normalizados ou dados originais.

   b. Em Criptografia, se você optar por Personalizar as configurações de criptografia, insira o ARN da AWS KMS chave.

   c. Visualize a saída gerada pelo sistema.

   d. Para Saída de dados, decida quais campos você deseja incluir, ocultar ou mascarar e, em seguida, execute as ações recomendadas com base em suas metas.

Seu objetivo	Ação recomendada
Incluir campos	Mantenha o estado de saída como Incluído.
Ocultar campos (excluir da saída)	Escolha o campo Saída e, em seguida, escolha Ocultar.
Campos de máscara	Escolha o campo Saída e, em seguida, escolha Saída de hash.
Redefinir as configurações anteriores	Escolha Redefinir.

   e. Escolha Próximo.

7. Para a Etapa 4: Revise e crie:

   a. Revise as seleções feitas nas etapas anteriores e edite, se necessário.

b. Escolha Criar e executar.

Uma mensagem aparece indicando que o fluxo de trabalho correspondente foi criado e que o trabalho foi iniciado.

8. Na página de detalhes do fluxo de trabalho correspondente, na guia Métricas, veja o seguinte em Métricas do último trabalho:

- O Job ID.
- O status da tarefa de fluxo de trabalho correspondente: Em fila, em andamento, concluída, com falha
- O tempo concluído para o trabalho do fluxo de trabalho.
- O número de registros processados.
- O número de registros não processados.
- A partida única IDs gerada.
- O número de registros de entrada.

Você também pode visualizar as métricas de trabalho para trabalhos de fluxo de trabalho correspondentes que foram executados anteriormente no Histórico de trabalhos.

9. Após a conclusão do trabalho de fluxo de trabalho correspondente (o status é concluído), você pode acessar a guia Saída de dados e selecionar sua localização no Amazon S3 para visualizar os resultados.
10. (Somente tipo de processamento manual) Se você criou um fluxo de trabalho correspondente baseado em regras com o tipo de processamento Manual, você pode executar o fluxo de trabalho correspondente a qualquer momento escolhendo Executar fluxo de trabalho na página de detalhes do fluxo de trabalho correspondente.
11. (Somente tipo de processamento automático) Se sua tabela de dados tiver uma coluna DELETE, então:

- Os registros definidos `true` na coluna DELETE são excluídos.
- Os registros definidos `false` na coluna DELETE são ingeridos no S3.

Para obter mais informações, consulte Etapa 1: Preparar tabelas de dados primárias.

API

Para criar um fluxo de trabalho de correspondência baseado em regras com o tipo de regra Avançado usando a API

> **Note**
>
> Por padrão, o fluxo de trabalho usa processamento padrão (em lote). Para usar o processamento incremental (automático), você deve configurá-lo explicitamente.

1. Abra um terminal ou prompt de comando para fazer a solicitação da API.
2. Crie uma solicitação POST para o seguinte endpoint:

```
/matchingworkflows
```

3. No cabeçalho da solicitação, defina o tipo de conteúdo como application/json.

> **Note**
>
> Para obter uma lista completa das linguagens de programação compatíveis, consulte a Referência AWS Entity Resolution da API.

4. Para o corpo da solicitação, forneça os seguintes parâmetros JSON necessários:

```
{
   "description": "string",
   "incrementalRunConfig": {
      "incrementalRunType": "string"
   },
   "inputSourceConfig": [
      {
         "applyNormalization": boolean,
         "inputSourceARN": "string",
         "schemaName": "string"
      }
   ],
   "outputSourceConfig": [
      {
         "applyNormalization": boolean,
         "KMSArn": "string",
```

```
         "output": [
            {
               "hashed": boolean,
               "name": "string"
            }
         ],
         "outputS3Path": "string"
      }
   ],
   "resolutionTechniques": {
      "providerProperties": {
         "intermediateSourceConfiguration": {
            "intermediateS3Path": "string"
         },
         "providerConfiguration": JSON value,
         "providerServiceArn": "string"
      },
      "resolutionType": "RULE_MATCHING",
      "ruleBasedProperties": {
         "attributeMatchingModel": "string",
         "matchPurpose": "string",
         "rules": [
            {
               "matchingKeys": [ "string" ],
               "ruleName": "string"
            }
         ]
      },
      "ruleConditionProperties": {
         "rules": [
            {
               "condition": "string",
               "ruleName": "string"
            }
         ]
      }
   },
   "roleArn": "string",
   "tags": {
      "string" : "string"
   },
   "workflowName": "string"
}
```

Em que:

- `workflowName`(obrigatório) — Deve ser exclusivo e ter entre 1—255 caracteres que correspondam ao padrão [a-zA-z_0-9-] *
- `inputSourceConfig`(obrigatório) — Lista de 1—20 configurações de fonte de entrada
- `outputSourceConfig`(obrigatório) — Exatamente uma configuração de fonte de saída
- `resolutionTechniques`(obrigatório) — Defina como “RULE_MATCHING” como o tipo de resolução para correspondência baseada em regras
- `roleArn`(obrigatório) — ARN da função do IAM para execução do fluxo de trabalho
- `ruleConditionProperties`(obrigatório) — Lista de condições da regra e o nome da regra correspondente.

Os parâmetros opcionais incluem:

- `description`— Até 255 caracteres
- `incrementalRunConfig`— Configuração incremental do tipo de execução
- `tags`— Até 200 pares de valores-chave

5. (Opcional) Para usar o processamento incremental em vez do processamento padrão (em lote), adicione o seguinte parâmetro ao corpo da solicitação:

```
"incrementalRunConfig": {
   "incrementalRunType": "AUTOMATIC"
}
```

6. Envie a solicitação .
7. Se for bem-sucedido, você receberá uma resposta com o código de status 200 e um corpo JSON contendo:

```
{
   "workflowArn": "string",
   "workflowName": "string",
   // Plus all configured workflow details
}
```

8. Se a chamada não for bem-sucedida, você poderá receber um dos seguintes erros:

- 400 — ConflictException se o nome do fluxo de trabalho já existir

- 400 — ValidationException se a entrada falhar na validação
- 402 — ExceedsLimitException se os limites da conta forem excedidos
- 403 — AccessDeniedException se você não tiver acesso suficiente
- 429 — ThrottlingException se a solicitação foi limitada
- 500 — InternalServerException se houver uma falha de serviço interno

Criação de um fluxo de trabalho de correspondência baseado em regras com o tipo de regra simples

O procedimento a seguir demonstra como criar um fluxo de trabalho de correspondência baseado em regras com o tipo de regra simples usando o AWS Entity Resolution console ou a API. `CreateMatchingWorkflow`

Console

Para criar um fluxo de trabalho de correspondência baseado em regras com o tipo de regra simples usando o console

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Correspondência.
3. Na página Fluxos de trabalho correspondentes, no canto superior direito, escolha Criar fluxo de trabalho correspondente.
4. Para a Etapa 1: Especificar os detalhes correspondentes do fluxo de trabalho, faça o seguinte:

   a. Insira um nome de fluxo de trabalho correspondente e uma Descrição opcional.

   b. Em Entrada de dados, escolha um AWS Glue banco de dados na lista suspensa, selecione a AWS Glue tabela e, em seguida, o mapeamento do esquema correspondente.

      Você pode adicionar até 19 entradas de dados.

   c. A opção Normalizar dados é selecionada por padrão, para que as entradas de dados sejam normalizadas antes da correspondência. Se você não quiser normalizar dados, desmarque a opção Normalizar dados.

> **Note**
>
> A normalização só é suportada nos seguintes cenários em Criar mapeamento de esquema:
>
> - Se os seguintes subtipos de nome estiverem agrupados: Nome, segundo nome, sobrenome.
> - Se os seguintes subtipos de endereço estiverem agrupados: Endereço 1, Endereço 2, Endereço 3, Cidade, Estado, País, Código postal.
> - Se os seguintes subtipos de telefone estiverem agrupados: Número de telefone, Código do país do telefone.

d. Para especificar as permissões de acesso ao serviço, escolha uma opção e execute a ação recomendada.

Opção	Ação recomendada
Criar e usar um novo perfil de serviço	• AWS Entity Resolution cria uma função de serviço com a política necessária para essa tabela. • O nome do perfil de serviço padrão é `entityresolution-matching-workflow-<timestamp>` . • Você deve ter permissões para criar perfis e anexar políticas. • Se seus dados de entrada estiverem criptografados, você poderá escolher a opção Esses dados são criptografados com uma chave KMS e, em seguida, inserir uma AWS KMS chave que será usada para descriptografar sua entrada de dados.

Opção	Ação recomendada
Use um perfil de serviço existente	1. Escolha um nome do perfil de serviço existente na lista suspensa. A lista de perfis é exibida se você tiver permissões para listar funções. Se você não tiver permissões para listar perfis, insira o nome do recurso da Amazon (ARN) do perfil que você deseja usar. Se não houver perfis de serviço existentes, a opção de Usar um perfil de serviço existente não estará disponível. 2. Para visualizar o perfil de serviço, selecione o link externo Visualizar no IAM. Por padrão, AWS Entity Resolution não tenta atualizar a política de função existente para adicionar as permissõe s necessárias.

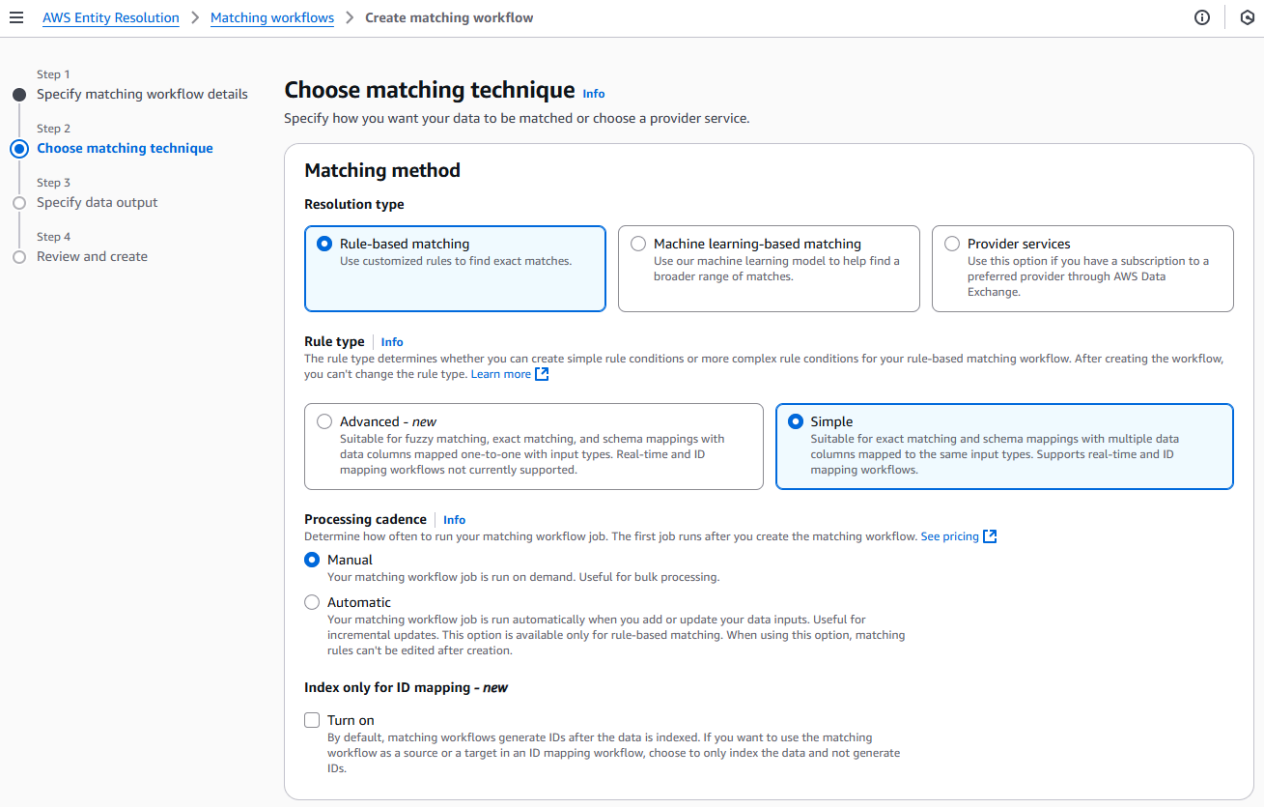
e. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e, em seguida, insira o par de chave e valor.

f. Escolha Próximo.

5. Para a Etapa 2: Escolha a técnica de correspondência:

a. Em Método de correspondência, escolha Correspondência baseada em regras.

b. Em Tipo de regra, escolha Simples.



c. Em Cadência de processamento, selecione uma das opções a seguir.

- Escolha Manual para executar um fluxo de trabalho sob demanda para uma atualização em massa
- Escolha Automático para executar um fluxo de trabalho assim que novos dados estiverem em seu bucket do S3

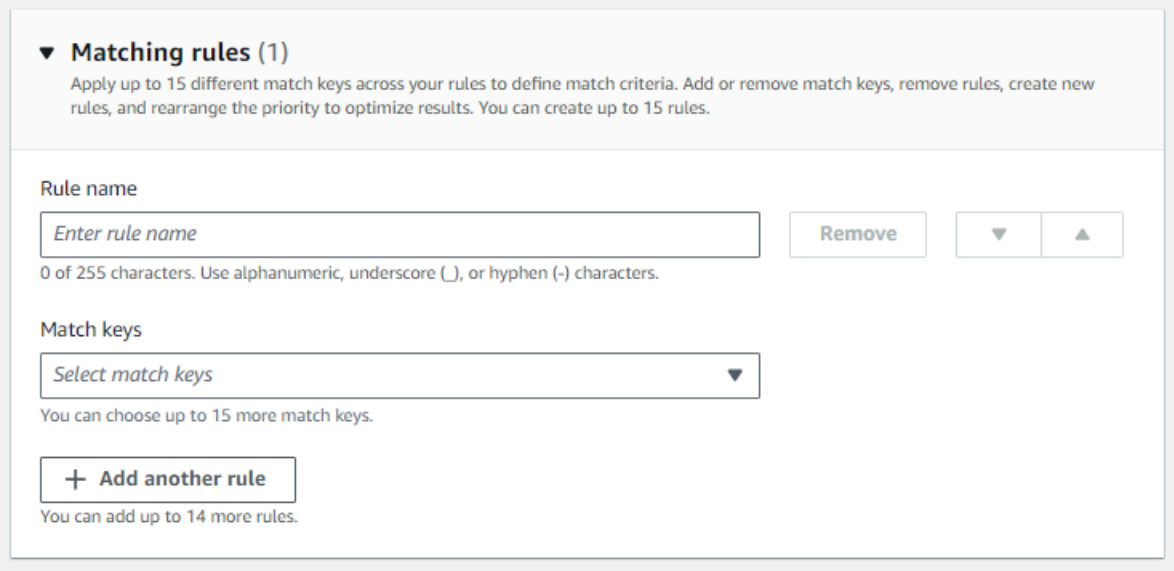
> **Note**
>
> Se você escolher Automático, certifique-se de ter EventBridge as notificações da Amazon ativadas para seu bucket do S3. Para obter instruções sobre como habilitar a Amazon EventBridge usando o console do S3, consulte Habilitando a Amazon EventBridge no Guia do usuário do Amazon S3.

d. (Opcional) Para indexar somente para mapeamento de ID, você pode optar por ativar a capacidade de indexar somente os dados e não gerar IDs.

Por padrão, o fluxo de trabalho correspondente é gerado IDs após a indexação dos dados.

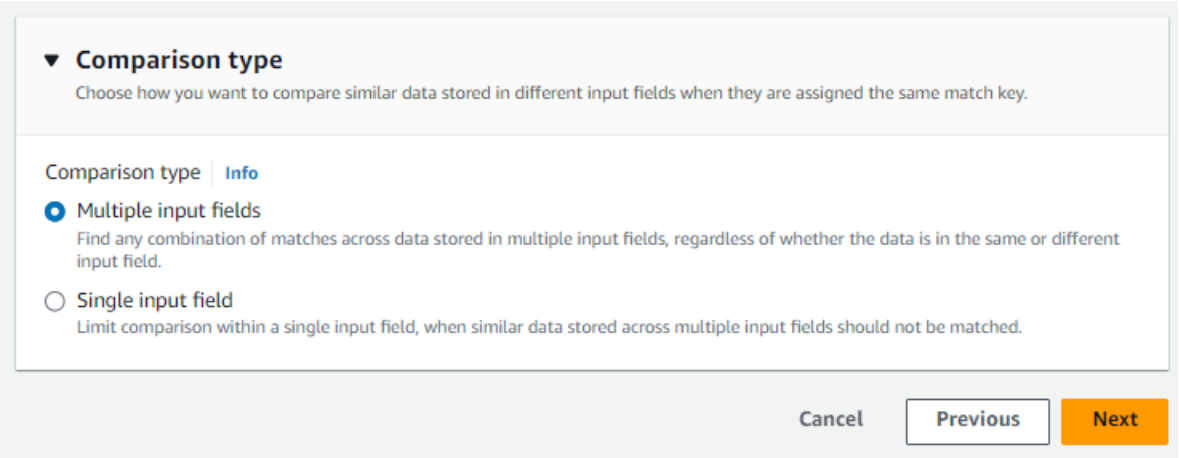
e. Em Regras de correspondência, insira um nome de regra e escolha as chaves de correspondência para essa regra.

   Você pode criar até 15 regras e aplicar até 15 chaves de correspondência diferentes em suas regras para definir critérios de correspondência.



f. Em Tipo de comparação, escolha uma das opções a seguir com base em sua meta.

Seu objetivo	Opção recomendada
Encontre qualquer combinação de correspondências nos dados armazenados em vários campos de entrada	Vários campos de entrada
Limitar a comparação a um único campo de entrada	Campo de entrada único



g. Escolha Próximo.

6. Para a Etapa 3: Especifique a saída e o formato dos dados:

   a. Em Destino e formato de saída de dados, escolha a localização do Amazon S3 para a saída de dados e se o formato dos dados será dados normalizados ou dados originais.

   b. Em Criptografia, se você optar por Personalizar as configurações de criptografia, insira o ARN da AWS KMS chave.

   c. Visualize a saída gerada pelo sistema.

   d. Para Saída de dados, decida quais campos você deseja incluir, ocultar ou mascarar e, em seguida, execute as ações recomendadas com base em suas metas.

Seu objetivo	Ação recomendada
Incluir campos	Mantenha o estado de saída como Incluído.
Ocultar campos (excluir da saída)	Escolha o campo Saída e, em seguida, escolha Ocultar.
Campos de máscara	Escolha o campo Saída e, em seguida, escolha Saída de hash.
Redefinir as configurações anteriores	Escolha Redefinir.

   e. Escolha Próximo.

7. Para a Etapa 4: Revise e crie:

a. Revise as seleções feitas nas etapas anteriores e edite, se necessário.

b. Escolha Criar e executar.

Uma mensagem aparece indicando que o fluxo de trabalho correspondente foi criado e que o trabalho foi iniciado.

8. Na página de detalhes do fluxo de trabalho correspondente, na guia Métricas, veja o seguinte em Métricas do último trabalho:

- O Job ID.
- O status da tarefa de fluxo de trabalho correspondente: Em fila, em andamento, concluída, com falha
- O tempo concluído para o trabalho do fluxo de trabalho.
- O número de registros processados.
- O número de registros não processados.
- A partida única IDs gerada.
- O número de registros de entrada.

Você também pode visualizar as métricas de trabalho para trabalhos de fluxo de trabalho correspondentes que foram executados anteriormente no Histórico de trabalhos.

9. Após a conclusão do trabalho de fluxo de trabalho correspondente (o status é concluído), você pode acessar a guia Saída de dados e selecionar sua localização no Amazon S3 para visualizar os resultados.

10. (Somente tipo de processamento manual) Se você criou um fluxo de trabalho de correspondência baseado em regras com o tipo de processamento Manual, você pode executar o fluxo de trabalho correspondente a qualquer momento escolhendo Executar fluxo de trabalho na página de detalhes do fluxo de trabalho correspondente.

API

Para criar um fluxo de trabalho de correspondência baseado em regras com o tipo de regra simples usando a API

> **Note**
>
> Por padrão, o fluxo de trabalho usa processamento padrão (em lote). Para usar o processamento incremental (automático), você deve configurá-lo explicitamente.

1. Abra um terminal ou prompt de comando para fazer a solicitação da API.
2. Crie uma solicitação POST para o seguinte endpoint:

```
/matchingworkflows
```

3. No cabeçalho da solicitação, defina o tipo de conteúdo como application/json.

> **Note**
>
> Para obter uma lista completa das linguagens de programação compatíveis, consulte a Referência AWS Entity Resolution da API.

4. Para o corpo da solicitação, forneça os seguintes parâmetros JSON necessários:

```
{
   "description": "string",
   "incrementalRunConfig": {
      "incrementalRunType": "string"
   },
   "inputSourceConfig": [
      {
         "applyNormalization": boolean,
         "inputSourceARN": "string",
         "schemaName": "string"
      }
   ],
   "outputSourceConfig": [
      {
         "applyNormalization": boolean,
         "KMSArn": "string",
```

```
         "output": [
            {
               "hashed": boolean,
               "name": "string"
            }
         ],
         "outputS3Path": "string"
      }
   ],
   "resolutionTechniques": {
      "providerProperties": {
         "intermediateSourceConfiguration": {
            "intermediateS3Path": "string"
         },
         "providerConfiguration": JSON value,
         "providerServiceArn": "string"
      },
      "resolutionType": "RULE_MATCHING",
      "ruleBasedProperties": {
         "attributeMatchingModel": "string",
         "matchPurpose": "string",
         "rules": [
            {
               "matchingKeys": [ "string" ],
               "ruleName": "string"
            }
         ]
      },
      "ruleConditionProperties": {
         "rules": [
            {
               "condition": "string",
               "ruleName": "string"
            }
         ]
      }
   },
   "roleArn": "string",
   "tags": {
      "string" : "string"
   },
   "workflowName": "string"
}
```

Em que:

- `workflowName`(obrigatório) — Deve ser exclusivo e ter entre 1—255 caracteres que correspondam ao padrão [a-zA-z_0-9-] *
- `inputSourceConfig`(obrigatório) — Lista de 1—20 configurações de fonte de entrada
- `outputSourceConfig`(obrigatório) — Exatamente uma configuração de fonte de saída
- `resolutionTechniques`(obrigatório) — Defina como “RULE_MATCHING” para correspondência baseada em regras
- `roleArn`(obrigatório) — ARN da função do IAM para execução do fluxo de trabalho
- `ruleConditionProperties`(obrigatório) — Lista de condições da regra e o nome da regra correspondente.

Os parâmetros opcionais incluem:

- `description`— Até 255 caracteres
- `incrementalRunConfig`— Configuração incremental do tipo de execução
- `tags`— Até 200 pares de valores-chave

5. (Opcional) Para usar o processamento incremental em vez do processamento padrão (em lote), adicione o seguinte parâmetro ao corpo da solicitação:

```
"incrementalRunConfig": {
   "incrementalRunType": "AUTOMATIC"
}
```

6. Envie a solicitação .
7. Se for bem-sucedido, você receberá uma resposta com o código de status 200 e um corpo JSON contendo:

```
{
   "workflowArn": "string",
   "workflowName": "string",
   // Plus all configured workflow details
}
```

8. Se a chamada não for bem-sucedida, você poderá receber um dos seguintes erros:

- 400 — ConflictException se o nome do fluxo de trabalho já existir

- 400 — ValidationException se a entrada falhar na validação
- 402 — ExceedsLimitException se os limites da conta forem excedidos
- 403 — AccessDeniedException se você não tiver acesso suficiente
- 429 — ThrottlingException se a solicitação foi limitada
- 500 — InternalServerException se houver uma falha de serviço interno

Criação de um fluxo de trabalho de correspondência baseado em aprendizado de máquina

A correspondência baseada em aprendizado de máquina é um processo predefinido que tenta combinar registros em todos os dados que você insere. O fluxo de trabalho de correspondência baseado em aprendizado de máquina permite comparar dados de texto não criptografado para encontrar uma ampla variedade de correspondências usando um modelo de aprendizado de máquina.

> **Note**
>
> O modelo de aprendizado de máquina não suporta a comparação de dados com hash.

Quando AWS Entity Resolution encontra uma correspondência entre dois ou mais registros em seus dados, ele atribui:

- Um ID de correspondência para os registros no conjunto de dados correspondente
- A porcentagem do nível de confiança da partida.

Você pode usar a saída de um fluxo de trabalho de correspondência baseado em ML como entrada para a correspondência de provedores de serviços de dados ou vice-versa para atingir suas metas específicas. Por exemplo, você pode executar uma correspondência baseada em ML para encontrar correspondências em suas fontes de dados em seus próprios registros primeiro. Se um subconjunto não corresponder, você poderá executar a correspondência baseada no serviço do provedor para encontrar correspondências adicionais.

Para criar um fluxo de trabalho de correspondência baseado em ML:

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Correspondência.
3. Na página Fluxos de trabalho correspondentes, no canto superior direito, escolha Criar fluxo de trabalho correspondente.
4. Para a Etapa 1: Especificar os detalhes correspondentes do fluxo de trabalho, faça o seguinte:

   a. Insira um nome de fluxo de trabalho correspondente e uma Descrição opcional.

   b. Em Entrada de dados, escolha um AWS Glue banco de dados na lista suspensa, selecione a AWS Glue tabela e, em seguida, o mapeamento do esquema correspondente.

      Você pode adicionar até 20 entradas de dados.

   c. A opção Normalizar dados é selecionada por padrão, para que as entradas de dados sejam normalizadas antes da correspondência. Se você não quiser normalizar dados, desmarque a opção Normalizar dados.

      A correspondência baseada em aprendizado de máquina apenas normalizaName, e. Telefone E-mail

   d. Para especificar as permissões de acesso ao serviço, escolha uma opção e execute a ação recomendada.

Opção	Ação recomendada
Criar e usar um novo perfil de serviço	• AWS Entity Resolution cria uma função de serviço com a política necessária para essa tabela. • O nome do perfil de serviço padrão é `entityresolution-matching-workflow-<timestamp>` . • Você deve ter permissões para criar perfis e anexar políticas. • Se seus dados de entrada estiverem criptografados, escolha a opção Esses dados são criptografados por uma chave

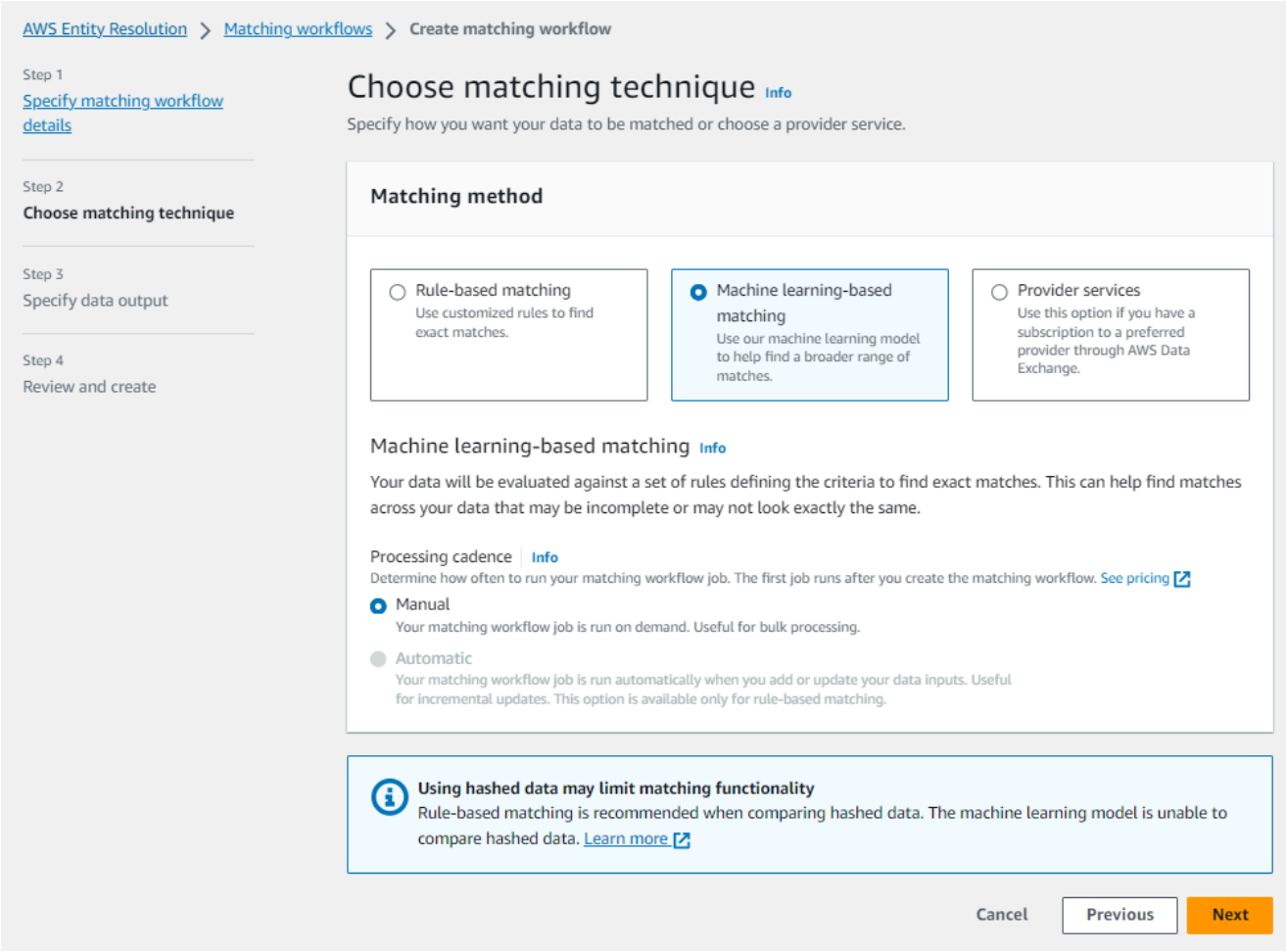
Opção	Ação recomendada
	KMS. Em seguida, insira uma AWS KMS chave usada para descriptografar sua entrada de dados.
Use um perfil de serviço existente	1. Escolha um nome do perfil de serviço existente na lista suspensa. A lista de perfis é exibida se você tiver permissões para listar funções. Se você não tiver permissões para listar perfis, insira o nome do recurso da Amazon (ARN) do perfil que você deseja usar. Se não houver perfis de serviço existentes, a opção de Usar um perfil de serviço existente não estará disponível. 2. Para visualizar o perfil de serviço, selecione o link externo Visualizar no IAM. Por padrão, AWS Entity Resolution não tenta atualizar a política de função existente para adicionar as permissões necessárias.

e. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e, em seguida, insira o par de chave e valor.

f. Escolha Próximo.

5. Para a Etapa 2: Escolha a técnica de correspondência:

   a. Em Método de correspondência, escolha Correspondência baseada em aprendizado de máquina.



b. Em Cadência de processamento, a opção Manual é selecionada.

Essa opção permite que você execute um fluxo de trabalho sob demanda para uma atualização em massa.

> **Note**
>
> O processamento automático (incremental) não é suportado para fluxos de trabalho de correspondência baseados em aprendizado de máquina.

c. Escolha Próximo.

6. Para a Etapa 3: Especifique a saída e o formato dos dados:

a. Em Destino e formato de saída de dados, escolha a localização do Amazon S3 para a saída de dados e se o formato dos dados será dados normalizados ou dados originais.

b. Em Criptografia, se você optar por Personalizar as configurações de criptografia, insira o ARN da AWS KMS chave.

c. Visualize a saída gerada pelo sistema.

d. Para Saída de dados, decida quais campos você deseja incluir, ocultar ou mascarar e, em seguida, execute as ações recomendadas com base em suas metas.

Seu objetivo	Opção recomendada
Incluir campos	Mantenha o estado de saída como Incluído.
Ocultar campos (excluir da saída)	Escolha o campo Saída e, em seguida, escolha Ocultar.
Campos de máscara	Escolha o campo Saída e, em seguida, escolha Saída de hash.
Redefinir as configurações anteriores	Escolha Redefinir.

e. Escolha Próximo.

7. Para a Etapa 4: Revise e crie:

a. Revise as seleções feitas nas etapas anteriores e edite, se necessário.

b. Escolha Criar e executar.

Uma mensagem aparece indicando que o fluxo de trabalho correspondente foi criado e que o trabalho foi iniciado.

8. Na página de detalhes do fluxo de trabalho correspondente, na guia Métricas, veja o seguinte em Métricas do último trabalho:

- O Job ID.
- O status da tarefa de fluxo de trabalho correspondente: Em fila, em andamento, concluída, com falha
- O tempo concluído para o trabalho do fluxo de trabalho.
- O número de registros processados.
- O número de registros não processados.
- A partida única IDs gerada.

- O número de registros de entrada.

Você também pode visualizar as métricas de trabalho para trabalhos de fluxo de trabalho correspondentes que foram executados anteriormente no Histórico de trabalhos.

9. Após a conclusão do trabalho de fluxo de trabalho correspondente (o status é concluído), você pode acessar a guia Saída de dados e selecionar sua localização no Amazon S3 para visualizar os resultados.
10. (Somente tipo de processamento manual) Se você criou um fluxo de trabalho de correspondência baseado em aprendizado de máquina com o tipo de processamento Manual, você pode executar o fluxo de trabalho correspondente a qualquer momento escolhendo Executar fluxo de trabalho na página de detalhes do fluxo de trabalho correspondente.

Criação de um fluxo de trabalho de correspondência baseado em serviços do provedor

A correspondência baseada no serviço do provedor permite que você combine seus identificadores conhecidos com o provedor de serviços de dados de sua preferência.

AWS Entity Resolution atualmente oferece suporte aos seguintes serviços de provedor de dados:

- LiveRamp
- TransUnion
- ID unificada 2.0

Para obter mais informações sobre os serviços de provedor suportados, consultePreparando dados de entrada de terceiros.

Você pode usar uma assinatura pública para esses provedores AWS Data Exchange ou negociar uma oferta privada diretamente com o provedor de dados. Para obter mais informações sobre como criar uma nova assinatura ou reutilizar uma assinatura existente de um serviço de provedor, consulteEtapa 1: Assine um serviço de provedor em AWS Data Exchange.

As seções a seguir descrevem como criar um fluxo de trabalho de correspondência baseado no provedor.

Tópicos

- Criando um fluxo de trabalho correspondente com LiveRamp
- Criando um fluxo de trabalho correspondente com TransUnion
- Criando um fluxo de trabalho correspondente com o UID 2.0

Criando um fluxo de trabalho correspondente com LiveRamp

Se você tiver uma assinatura do LiveRamp serviço, poderá criar um fluxo de trabalho compatível com o LiveRamp serviço para realizar a resolução de identidade.

O LiveRamp serviço fornece um identificador chamado RampID. O RampID é um dos mais usados IDs em plataformas de demanda para criar um público para uma campanha publicitária. Usando um fluxo de trabalho correspondente com LiveRamp, você pode resolver endereços de e-mail com hash para RAMPIDs.

> **Note**
>
> AWS Entity Resolution suporta atribuição de RampID baseada em PII.

Esse fluxo de trabalho requer um bucket de preparação de dados do Amazon S3 no qual você deseja que a saída do fluxo de trabalho correspondente seja gravada temporariamente. Antes de criar um fluxo de trabalho de mapeamento de ID com LiveRamp, adicione as seguintes permissões ao intervalo de preparação de dados.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::715724997226:root"

            },
            "Action": [
                "s3:PutObject",
                "s3:GetObject",
```

```
                "s3:GetObjectVersion",
                "s3:DeleteObject"
            ],
            "Resource": [
                "arn:aws:s3:::<staging-bucket>",
                "arn:aws:s3:::<staging-bucket>/*"
            ]
        },
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::715724997226:root"
            },
            "Action": [
                "s3:ListBucket",
                "s3:GetBucketLocation",
                "s3:GetBucketPolicy",
                "s3:ListBucketVersions",
                "s3:GetBucketAcl"
            ],
            "Resource": [
                "arn:aws:s3:::<staging-bucket>",
                "arn:aws:s3:::<staging-bucket>/*"
            ]
        }
    ]
}
```

Substitua cada *<user input placeholder>* por suas próprias informações.

staging-bucket	Bucket Amazon S3 que armazena temporari amente seus dados enquanto executa um fluxo de trabalho baseado em serviços do provedor.

Para criar um fluxo de trabalho correspondente com LiveRamp:

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Correspondência.

3. Na página Fluxos de trabalho correspondentes, no canto superior direito, escolha Criar fluxo de trabalho correspondente.
4. Para a Etapa 1: Especificar os detalhes correspondentes do fluxo de trabalho, faça o seguinte:

   a. Insira um nome de fluxo de trabalho correspondente e uma Descrição opcional.

   b. Em Entrada de dados, escolha um AWS Glue banco de dados na lista suspensa, selecione a AWS Glue tabela e, em seguida, selecione o mapeamento do esquema correspondente.

      Você pode adicionar até 20 entradas de dados.

   c. A opção Normalizar dados é selecionada por padrão, para que as entradas de dados sejam normalizadas antes da correspondência.

      > **Note**
      >
      > A normalização só é suportada nos seguintes cenários em Criar mapeamento de esquema:
      >
      > - Se os seguintes subtipos de nome estiverem agrupados: Nome, segundo nome, sobrenome.
      > - Se os seguintes subtipos de endereço estiverem agrupados: Endereço 1, Endereço 2: nome do endereço 3, nome da cidade, estado, país, código postal.
      > - Se os seguintes subtipos de telefone estiverem agrupados: Número de telefone, Código do país do telefone.

      Se você estiver usando o processo de resolução somente por e-mail, desmarque a opção Normalizar dados, pois somente e-mails com hash são usados para dados de entrada.

   d. Para especificar as permissões de acesso ao serviço, escolha uma opção e execute a ação recomendada.

Opção	Ação recomendada
Criar e usar um novo perfil de serviço	• AWS Entity Resolution cria uma função de serviço com a política necessária para essa tabela.

Opção	Ação recomendada
	• O nome do perfil de serviço padrão é `entityresolution-matching-workflow-<timestamp>` . • Você deve ter permissões para criar perfis e anexar políticas. • Se seus dados de entrada estiverem criptografados, escolha a opção Esses dados são criptografados por uma chave KMS. Em seguida, insira uma AWS KMS chave usada para descriptografar sua entrada de dados.
Use um perfil de serviço existente	1. Escolha um nome do perfil de serviço existente na lista suspensa. A lista de perfis é exibida se você tiver permissões para listar funções. Se você não tiver permissões para listar perfis, insira o nome do recurso da Amazon (ARN) do perfil que você deseja usar. Se não houver perfis de serviço existentes, a opção de Usar um perfil de serviço existente não estará disponível. 2. Para visualizar o perfil de serviço, selecione o link externo Visualizar no IAM. Por padrão, AWS Entity Resolution não tenta atualizar a política de função existente para adicionar as permissões necessárias.

e. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e, em seguida, insira o par de chave e valor.

f. Escolha Próximo.

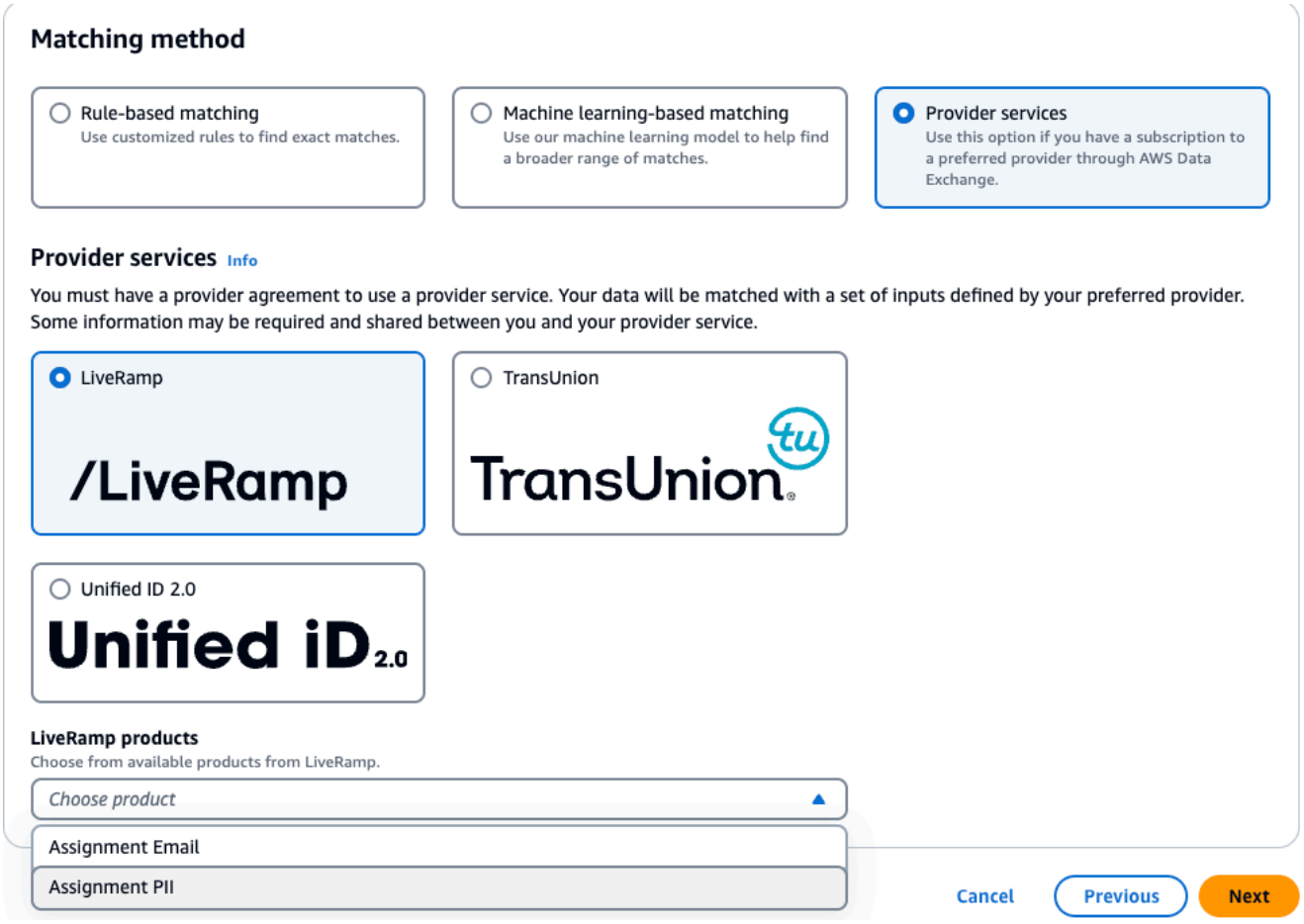
5. Para a Etapa 2: Escolha a técnica de correspondência:

a. Em Método de correspondência, escolha Serviços do provedor.

b. Para serviços do provedor, escolha LiveRamp.

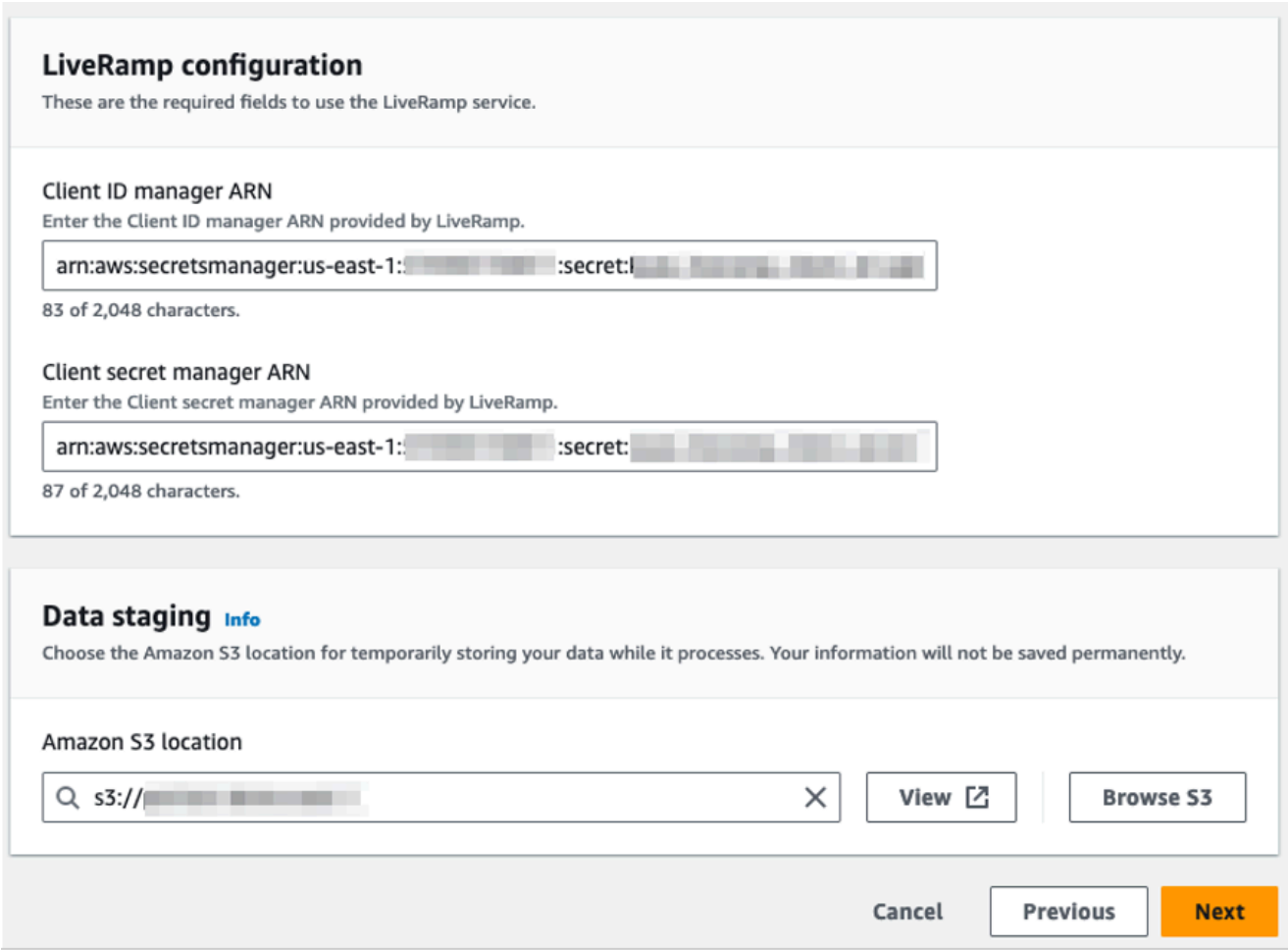
> **Note**
>
> Certifique-se de que o formato e a normalização do arquivo de entrada de dados estejam alinhados com as diretrizes do serviço do provedor.
> Para obter mais informações sobre as diretrizes de formatação do arquivo de entrada para o fluxo de trabalho correspondente, consulte Executar resolução de identidade por meio do ADX na LiveRamp documentação.

c. Para LiveRamp produtos, escolha um produto na lista suspensa.



> **Note**
>
> Se você escolher Atribuição PII, deverá fornecer pelo menos uma coluna não identificadora ao realizar a resolução da entidade. Por exemplo, GÊNERO.

d. Para LiveRamp configuração, insira um ARN do gerenciador de ID do cliente e um ARN do gerenciador secreto do cliente.



e. Para preparação de dados, escolha o local do Amazon S3 para o armazenamento temporário de seus dados enquanto eles são processados.

   Você deve ter permissão para a localização do Amazon S3 de armazenamento de dados. Para obter mais informações, consulte Criação de uma função de trabalho de fluxo de trabalho para AWS Entity Resolution.

f. Escolha Próximo.

6. Para a Etapa 3: Especifique a saída de dados:

a. Em Destino e formato de saída de dados, escolha a localização do Amazon S3 para a saída de dados e se o formato dos dados será dados normalizados ou dados originais.

b. Em Criptografia, se você optar por Personalizar as configurações de criptografia, insira o ARN da AWS KMS chave.

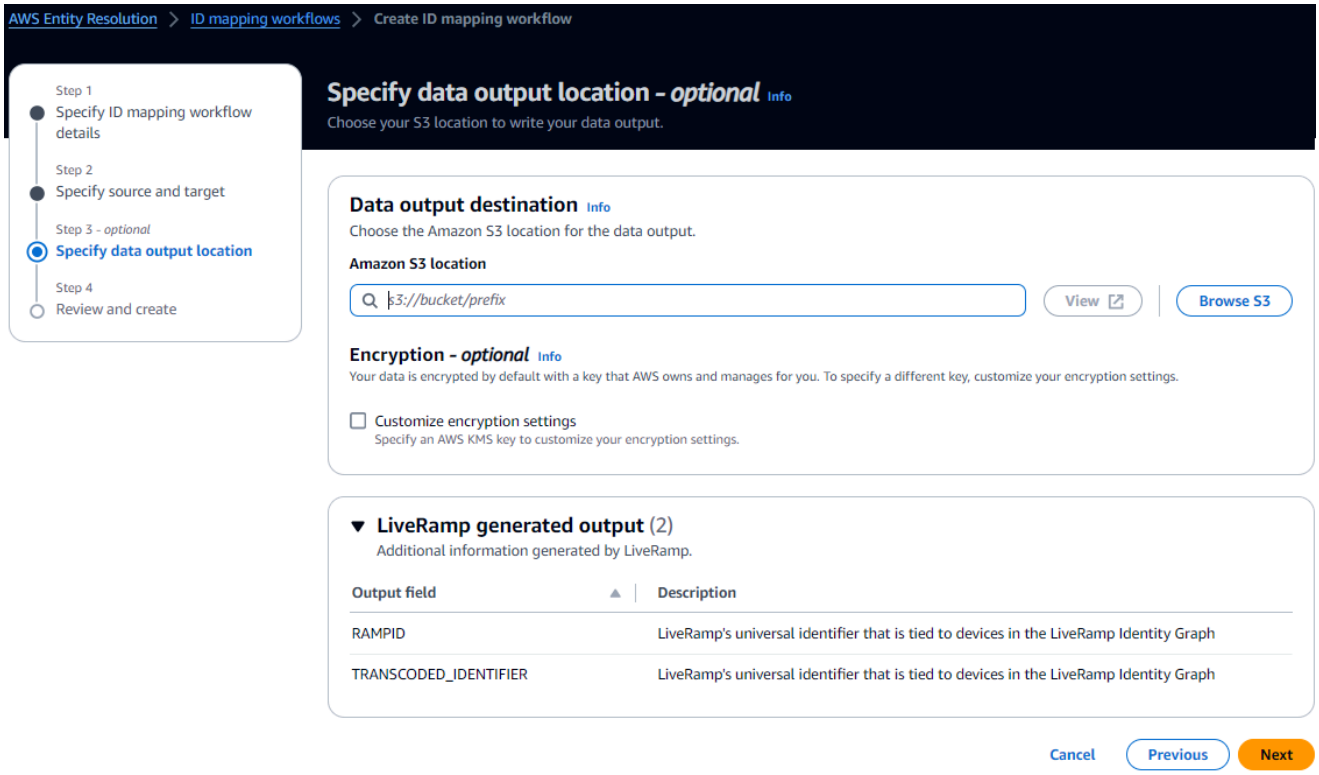
c. Visualize a saída LiveRamp gerada.

   Essas são as informações adicionais geradas pelo LiveRamp.

d. Para Saída de dados, decida quais campos você deseja incluir, ocultar ou mascarar e, em seguida, execute as ações recomendadas com base em suas metas.

> **Note**
>
> Se você tiver escolhido LiveRamp, devido aos filtros LiveRamp de privacidade que removem informações de identificação pessoal (PII), alguns campos exibirão um estado de saída indisponível.

Seu objetivo	Opção recomendada
Incluir campos	Mantenha o estado de saída como Incluído.
Ocultar campos (excluir da saída)	Escolha o campo Saída e, em seguida, escolha Ocultar.
Campos de máscara	Escolha o campo Saída e, em seguida, escolha Saída de hash.
Redefinir as configurações anteriores	Escolha Redefinir.



e. Escolha Próximo.

7. Para a Etapa 4: Revise e crie:

   a. Revise as seleções feitas nas etapas anteriores e edite, se necessário.

   b. Escolha Criar e executar.

      Uma mensagem aparece indicando que o fluxo de trabalho correspondente foi criado e que o trabalho foi iniciado.

8. Na página de detalhes do fluxo de trabalho correspondente, na guia Métricas, veja o seguinte em Métricas do último trabalho:

   - O Job ID.
   - O status da tarefa de fluxo de trabalho correspondente: Em fila, em andamento, concluída, com falha
   - O tempo concluído para o trabalho do fluxo de trabalho.
   - O número de registros processados.
   - O número de registros não processados.
   - A partida única IDs gerada.

- O número de registros de entrada.

Você também pode visualizar as métricas de trabalho para trabalhos de fluxo de trabalho correspondentes que foram executados anteriormente no Histórico de trabalhos.

9. Após a conclusão do trabalho de fluxo de trabalho correspondente (o status é concluído), você pode acessar a guia Saída de dados e selecionar sua localização no Amazon S3 para visualizar os resultados.

Criando um fluxo de trabalho correspondente com TransUnion

Se você tiver uma assinatura do TransUnion serviço, poderá melhorar a compreensão do cliente vinculando, combinando e aprimorando os registros relacionados ao cliente armazenados em canais diferentes com chaves eletrônicas TransUnion pessoais e domésticas e mais de 200 atributos de dados.

O TransUnion serviço fornece identificadores conhecidos como TransUnion Indivíduo e Domicílio IDs. TransUnion fornece atribuição de ID (também conhecida como codificação) de identificadores conhecidos, como nome, endereço, número de telefone e endereço de e-mail.

Esse fluxo de trabalho requer um bucket de preparação de dados do Amazon S3 no qual você deseja que a saída do fluxo de trabalho correspondente seja gravada temporariamente. Antes de criar um fluxo de trabalho correspondente com TransUnion, adicione as seguintes permissões ao intervalo de preparação de dados.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::381491956555:root"

            },
            "Action": [
                "s3:PutObject",
                "s3:GetObject",
```

```
                "s3:GetObjectVersion",
                "s3:DeleteObject"
            ],
            "Resource": [
                "arn:aws:s3:::<staging-bucket>",
                "arn:aws:s3:::<staging-bucket>/*"
            ]
        },
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::381491956555:root"
            },
            "Action": [
                "s3:ListBucket",
                "s3:GetBucketLocation",
                "s3:GetBucketPolicy",
                "s3:ListBucketVersions",
                "s3:GetBucketAcl"
            ],
            "Resource": [
                "arn:aws:s3:::<staging-bucket>",
                "arn:aws:s3:::<staging-bucket>/*"
            ]
        }
    ]
}
```

Substitua cada *<user input placeholder>* por suas próprias informações.

staging-bucket	Bucket Amazon S3 que armazena temporari amente seus dados enquanto executa um fluxo de trabalho baseado em serviços do provedor.

Para criar um fluxo de trabalho correspondente com TransUnion:

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Correspondência.

3. Na página Fluxos de trabalho correspondentes, no canto superior direito, escolha Criar fluxo de trabalho correspondente.
4. Para a Etapa 1: Especificar os detalhes correspondentes do fluxo de trabalho, faça o seguinte:

   a. Insira um nome de fluxo de trabalho correspondente e uma Descrição opcional.

   b. Em Entrada de dados, escolha um AWS Glue banco de dados na lista suspensa, selecione a AWS Glue tabela e, em seguida, selecione o mapeamento do esquema correspondente.

      Você pode adicionar até 20 entradas de dados.

   c. A opção Normalizar dados é selecionada por padrão, para que as entradas de dados sejam normalizadas antes da correspondência. Se você não quiser normalizar dados, desmarque a opção Normalizar dados.

      > **Note**
      >
      > A normalização só é suportada nos seguintes cenários em Criar mapeamento de esquema:
      >
      > - Se os seguintes subtipos de nome estiverem agrupados: Nome, segundo nome, sobrenome.
      > - Se os seguintes subtipos de endereço estiverem agrupados: Endereço 1, Endereço 2: nome do endereço 3, nome da cidade, estado, país, código postal.
      > - Se os seguintes subtipos de telefone estiverem agrupados: Número de telefone, Código do país do telefone.

   d. Para especificar as permissões de acesso ao serviço, escolha uma opção e execute a ação recomendada.

Opção	Ação recomendada
Criar e usar um novo perfil de serviço	• AWS Entity Resolution cria uma função de serviço com a política necessária para essa tabela. • O nome do perfil de serviço padrão é `entityresolution-matching-workflow-<timestamp>` .

Opção	Ação recomendada
	• Você deve ter permissões para criar perfis e anexar políticas. • Se seus dados de entrada estiverem criptografados, escolha a opção Esses dados são criptografados por uma chave KMS. Em seguida, insira uma AWS KMS chave usada para descriptografar sua entrada de dados.
Use um perfil de serviço existente	1. Escolha um nome do perfil de serviço existente na lista suspensa. A lista de perfis é exibida se você tiver permissões para listar funções. Se você não tiver permissões para listar perfis, insira o nome do recurso da Amazon (ARN) do perfil que você deseja usar. Se não houver perfis de serviço existentes, a opção de Usar um perfil de serviço existente não estará disponível. 2. Para visualizar o perfil de serviço, selecione o link externo Visualizar no IAM. Por padrão, AWS Entity Resolution não tenta atualizar a política de função existente para adicionar as permissões necessárias.

e. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e, em seguida, insira o par de chave e valor.

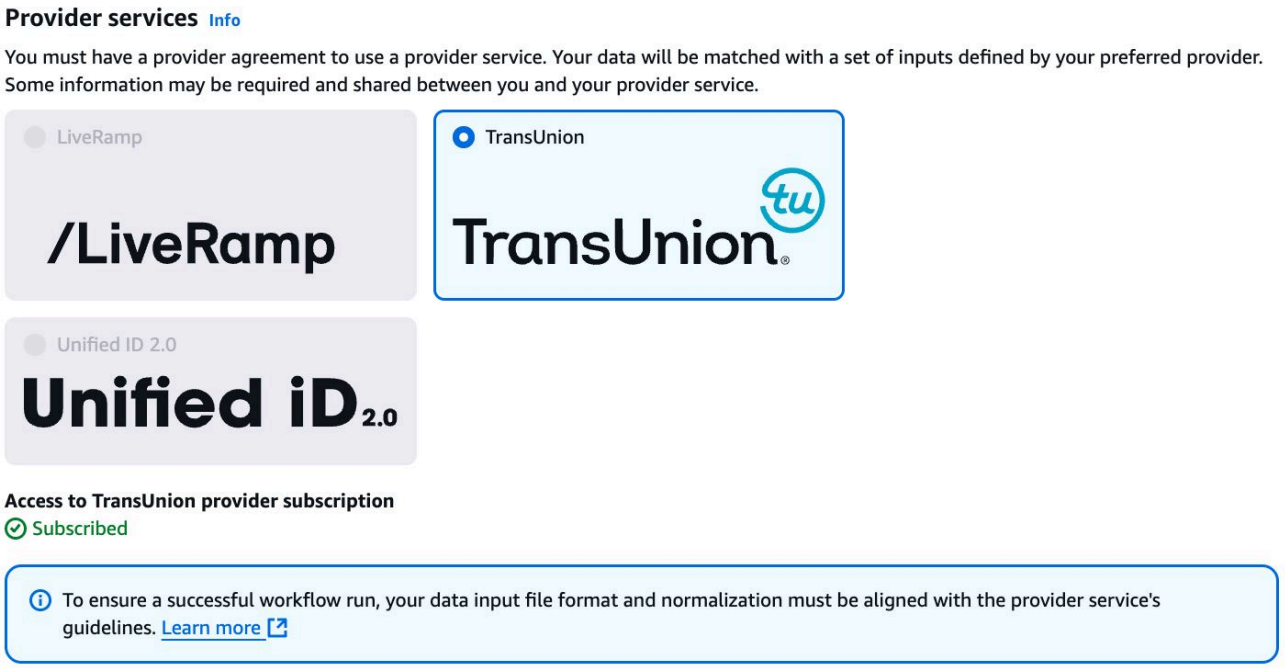
f. Escolha Próximo.

5. Para a Etapa 2: Escolha a técnica de correspondência:

a. Em Método de correspondência, escolha Serviços do provedor.

b. Para serviços do provedor, escolha TransUnion.

> **Note**
>
> Certifique-se de que o formato e a normalização do arquivo de entrada de dados estejam alinhados com as diretrizes do serviço do provedor.



c. Para preparação de dados, escolha o local do Amazon S3 para o armazenamento temporário de seus dados enquanto eles são processados.

   Você deve ter permissão para a localização do Amazon S3 de armazenamento de dados. Para obter mais informações, consulte the section called “Criação de uma função de trabalho no fluxo de trabalho”.

6. Escolha Próximo.

7. Para a Etapa 3: Especifique a saída de dados:

   a. Em Destino e formato de saída de dados, escolha a localização do Amazon S3 para a saída de dados e se o formato dos dados será dados normalizados ou dados originais.

b. Em Criptografia, se você optar por Personalizar as configurações de criptografia, insira o ARN da AWS KMS chave.

c. Visualize a saída TransUnion gerada.

   Essas são as informações adicionais geradas pelo TransUnion.

d. Para Saída de dados, decida quais campos você deseja incluir, ocultar ou mascarar e, em seguida, execute as ações recomendadas com base em suas metas.

Seu objetivo	Opção recomendada
Incluir campos	Mantenha o estado de saída como Incluído.
Ocultar campos (excluir da saída)	Escolha o campo Saída e, em seguida, escolha Ocultar.
Campos de máscara	Escolha o campo Saída e, em seguida, escolha Saída de hash.
Redefinir as configurações anteriores	Escolha Redefinir.

e. Para a saída gerada pelo sistema, visualize todos os campos incluídos.

f. Escolha Próximo.

8. Para a Etapa 4: Revise e crie:

   a. Revise as seleções feitas nas etapas anteriores e edite, se necessário.

   b. Escolha Criar e executar.

      Uma mensagem aparece indicando que o fluxo de trabalho correspondente foi criado e que o trabalho foi iniciado.

9. Na página de detalhes do fluxo de trabalho correspondente, na guia Métricas, veja o seguinte em Métricas do último trabalho:

   - O Job ID.
   - O status da tarefa de fluxo de trabalho correspondente: Em fila, em andamento, concluída, com falha
   - O tempo concluído para o trabalho do fluxo de trabalho.

- O número de registros processados.
- O número de registros não processados.
- A partida única IDs gerada.
- O número de registros de entrada.

Você também pode visualizar as métricas de trabalho para trabalhos de fluxo de trabalho correspondentes que foram executados anteriormente no Histórico de trabalhos.

10. Após a conclusão do trabalho de fluxo de trabalho correspondente (o status é concluído), você pode acessar a guia Saída de dados e selecionar sua localização no Amazon S3 para visualizar os resultados.

Criando um fluxo de trabalho correspondente com o UID 2.0

Se você tiver uma assinatura do serviço Unified ID 2.0, poderá ativar campanhas publicitárias com identidade determinística e confiar na interoperabilidade com muitos participantes UID2 habilitados em todo o ecossistema de publicidade. Para obter mais informações, consulte Visão geral do Unified ID 2.0.

O serviço Unified ID 2.0 fornece UID 2 bruto, que é usado para criar campanhas publicitárias na plataforma The Trade Desk. O UID 2.0 é gerado usando uma estrutura de código aberto.

Em um fluxo de trabalho, você pode usar um **`Email Address`** ou **`Phone number`** para UID2 geração bruta, mas não ambos. Se ambos estiverem presentes no mapeamento do esquema, o fluxo de trabalho escolherá o **`Email Address`** e o **`Phone number`** será um campo de passagem. Para oferecer suporte a ambos, crie um novo mapeamento de esquema onde **`Phone number`** está mapeado, mas **`Email Address`** não está mapeado. Em seguida, crie um segundo fluxo de trabalho usando esse novo mapeamento de esquema.

> **Note**
>
> UID2s Os crus são criados pela adição de sais de baldes de sal que são girados aproximadamente uma vez por ano, fazendo com que o cru também seja girado UID2 com ele. Portanto, é recomendável que você atualize o bruto UID2s diariamente. Para obter mais informações, consulte https://unifiedid.com/docs/how-often-should-uidgetting-started/gs-faqs#2 -incremental-updates. s-be-refreshed-for

Para criar um fluxo de trabalho correspondente com o UID 2.0:

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Correspondência.
3. Na página Fluxos de trabalho correspondentes, no canto superior direito, escolha Criar fluxo de trabalho correspondente.
4. Para a Etapa 1: Especificar os detalhes correspondentes do fluxo de trabalho, faça o seguinte:

   a. Insira um nome de fluxo de trabalho correspondente e uma Descrição opcional.

   b. Em Entrada de dados, escolha um AWS Glue banco de dados na lista suspensa, selecione a AWS Glue tabela e, em seguida, selecione o mapeamento do esquema correspondente.

      Você pode adicionar até 20 entradas de dados.

   c. Deixe a opção Normalizar dados selecionada, para que as entradas de dados (**`Email Address`**ou**`Phone number`**) sejam normalizadas antes da correspondência.

      Para obter mais informações sobre **`Email Address`** normalização, consulte Normalização de endereço de e-mail na documentação do UID 2.0.

      Para obter mais informações sobre **`Phone number`** normalização, consulte Normalização do número de telefone na documentação do UID 2.0.

   d. Para especificar as permissões de acesso ao serviço, escolha uma opção e execute a ação recomendada.

Opção	Ação recomendada
Criar e usar um novo perfil de serviço	• AWS Entity Resolution cria uma função de serviço com a política necessária para essa tabela. • O nome do perfil de serviço padrão é `entityresolution-matching-workflow-<timestamp>` . • Você deve ter permissões para criar perfis e anexar políticas.

Opção	Ação recomendada
	• Se seus dados de entrada estiverem criptografados, escolha a opção Esses dados são criptografados por uma chave KMS. Em seguida, insira uma AWS KMS chave usada para descriptografar sua entrada de dados.
Use um perfil de serviço existente	1. Escolha um nome do perfil de serviço existente na lista suspensa. A lista de perfis é exibida se você tiver permissões para listar funções. Se você não tiver permissões para listar perfis, insira o nome do recurso da Amazon (ARN) do perfil que você deseja usar. Se não houver perfis de serviço existentes, a opção de Usar um perfil de serviço existente não estará disponível. 2. Para visualizar o perfil de serviço, selecione o link externo Visualizar no IAM. Por padrão, AWS Entity Resolution não tenta atualizar a política de função existente para adicionar as permissões necessárias.

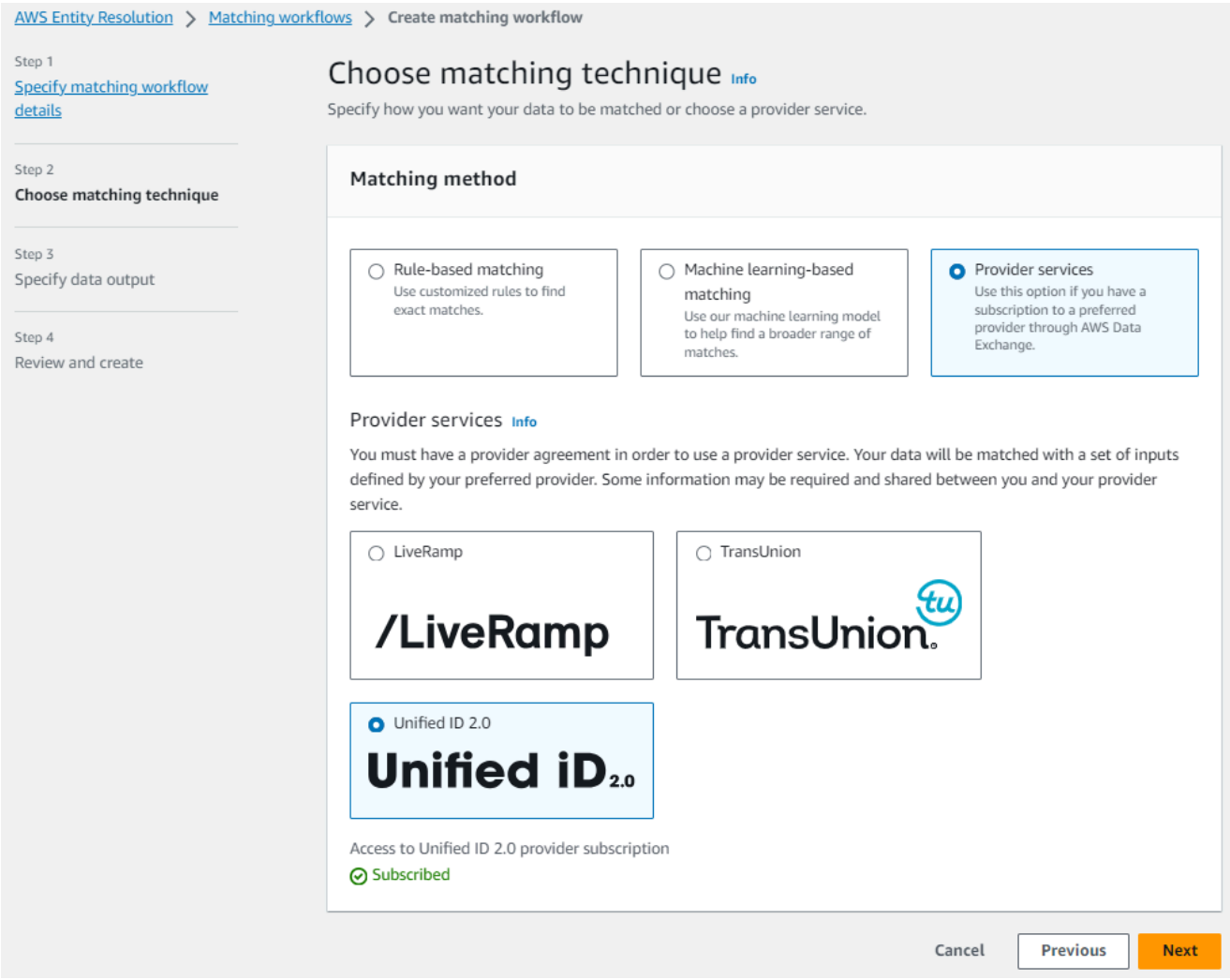
e. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e, em seguida, insira o par de chave e valor.

f. Escolha Próximo.

5. Para a Etapa 2: Escolha a técnica de correspondência:

   a. Em Método de correspondência, escolha Serviços do provedor.

b. Para serviços de provedor, escolha Unified ID 2.0.



c. Escolha Próximo.

6. Para a Etapa 3: Especifique a saída de dados:

a. Em Destino e formato de saída de dados, escolha a localização do Amazon S3 para a saída de dados e se o formato dos dados será dados normalizados ou dados originais.

b. Em Criptografia, se você optar por Personalizar as configurações de criptografia, insira o ARN da AWS KMS chave.

c. Veja a saída gerada pelo Unified ID 2.0.

Esta é uma lista de todas as informações adicionais geradas pelo UID 2.0

d. Para Saída de dados, decida quais campos você deseja incluir, ocultar ou mascarar e, em seguida, execute as ações recomendadas com base em suas metas.

Seu objetivo	Opção recomendada
Incluir campos	Mantenha o estado de saída como Incluído.
Ocultar campos (excluir da saída)	Escolha o campo Saída e, em seguida, escolha Ocultar.
Campos de máscara	Escolha o campo Saída e, em seguida, escolha Saída de hash.
Redefinir as configurações anteriores	Escolha Redefinir.

e. Para a saída gerada pelo sistema, visualize todos os campos incluídos.

f. Escolha Próximo.

7. Para a Etapa 4: Revise e crie:

   a. Revise as seleções feitas nas etapas anteriores e edite, se necessário.

   b. Escolha Criar e executar.

      Uma mensagem aparece indicando que o fluxo de trabalho correspondente foi criado e que o trabalho foi iniciado.

8. Na página de detalhes do fluxo de trabalho correspondente, na guia Métricas, veja o seguinte em Métricas do último trabalho:

   - O Job ID.
   - O status da tarefa de fluxo de trabalho correspondente: Em fila, em andamento, concluída, com falha
   - O tempo concluído para o trabalho do fluxo de trabalho.
   - O número de registros processados.
   - O número de registros não processados.
   - A partida única IDs gerada.
   - O número de registros de entrada.

Você também pode visualizar as métricas de trabalho para trabalhos de fluxo de trabalho correspondentes que foram executados anteriormente no Histórico de trabalhos.

9. Após a conclusão do trabalho de fluxo de trabalho correspondente (o status é concluído), você pode acessar a guia Saída de dados e selecionar sua localização no Amazon S3 para visualizar os resultados.

Editando um fluxo de trabalho correspondente

A edição do fluxo de trabalho correspondente permite que você mantenha seus processos de resolução de entidades up-to-date e responda às mudanças nos requisitos de sua organização ao longo do tempo. Talvez você queira ajustar os critérios, as técnicas ou as saídas de dados correspondentes para melhorar a precisão e a eficiência do processo de resolução da entidade. Se você identificar problemas ou erros nos resultados do fluxo de trabalho atual, editá-lo poderá ajudá-lo a diagnosticar e resolver esses problemas.

Para editar um fluxo de trabalho correspondente:

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Correspondência.
3. Escolha o fluxo de trabalho correspondente.
4. Na página de detalhes do fluxo de trabalho correspondente, no canto superior direito, escolha Editar fluxo de trabalho.
5. Na página Especificar detalhes do fluxo de trabalho correspondente, faça as alterações necessárias e escolha Avançar.
6. Na página Escolher técnica de correspondência, faça as alterações necessárias e escolha Avançar.

> **Important**
>
> Você pode alterar a cadência de processamento de Manual para Automática, mas depois de alterá-la para Automática, não é possível alterá-la novamente para Manual. Se a cadência de processamento já estiver definida como Automática, você não poderá alterá-la para Manual.

7. Na página Especificar saída de dados, faça as alterações necessárias e escolha Avançar.
8. Na página Revisar e salvar, faça as alterações necessárias e escolha Salvar.

Excluindo um fluxo de trabalho correspondente

Se um fluxo de trabalho correspondente não estiver mais sendo usado ou se tornar obsoleto, excluí-lo pode ajudar a manter seu espaço de trabalho organizado e organizado. Se você desenvolveu um fluxo de trabalho novo e aprimorado que substitui um antigo, excluir o fluxo de trabalho antigo pode ajudar a garantir que você esteja usando apenas a maioria up-to-date dos processos.

Para excluir um fluxo de trabalho correspondente:

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Correspondência.
3. Escolha o fluxo de trabalho correspondente.
4. Na página de detalhes do fluxo de trabalho correspondente, no canto superior direito, escolha Excluir.
5. Confirme a exclusão e escolha Excluir.

Modificando ou gerando uma ID de correspondência para um fluxo de trabalho de correspondência baseado em regras

Uma ID de correspondência é o identificador gerado AWS Entity Resolution e aplicado a cada conjunto de registros correspondente após a execução de um fluxo de trabalho correspondente. Isso faz parte dos metadados de fluxo de trabalho correspondentes incluídos na saída.

Quando precisar atualizar os registros de um cliente existente ou adicionar um novo cliente ao seu conjunto de dados, você pode usar o AWS Entity Resolution console ou a `GenerateMatchID` API. A modificação de uma ID de correspondência existente ajuda a manter a consistência ao atualizar as informações do cliente, enquanto a geração de uma nova ID de correspondência é necessária ao adicionar clientes não identificados anteriormente ao seu sistema.

> **Note**
>
> Cobranças adicionais se aplicam, independentemente de você usar o console ou a API. O tipo de processamento escolhido afeta a precisão e o tempo de resposta da operação.

> **Important**
>
> Se você revogar AWS Entity Resolution as permissões do seu bucket do S3 enquanto um trabalho estiver em andamento, ainda AWS Entity Resolution processará e cobrará pela saída dos resultados para o S3, mas não poderá entregá-los ao seu bucket. Para evitar esse problema, verifique se ele AWS Entity Resolution tem as permissões corretas para gravar em seu bucket do S3 antes de iniciar um trabalho. Se as permissões forem revogadas durante o processamento, AWS Entity Resolution tentará reenviar os resultados por até 30 dias após a conclusão do trabalho depois de restaurar as permissões corretas do bucket.

O procedimento a seguir orienta você no processo de pesquisar ou gerar um ID de correspondência, selecionar um tipo de processamento e visualizar os resultados.

Console

Para modificar ou gerar um Match ID usando o console

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Correspondência.
3. Escolha o fluxo de trabalho de correspondência baseado em regras que foi processado (o status do trabalho é Concluído).
4. Na página de detalhes do fluxo de trabalho correspondente, escolha a IDs guia Corresponder.
5. Escolha Modificar ou gerar ID de correspondência.

> **Note**
>
> A opção Modificar ou gerar ID de correspondência só está disponível para fluxos de trabalho correspondentes que usam a cadência de processamento automático. Se

> você tiver selecionado a cadência de processamento manual, essa opção aparecerá inativa. Para usar essa opção, edite seu fluxo de trabalho para usar a cadência de processamento automático. Para obter mais informações sobre a edição de fluxos de trabalho, consulteEditando um fluxo de trabalho correspondente.

6. Selecione a AWS Glue tabela na lista suspensa.

   Se houver somente uma AWS Glue tabela no fluxo de trabalho, ela será selecionada por padrão.

7. Escolha o tipo de processamento.

   - Consistente — Você pode pesquisar um ID de correspondência existente ou gerar e salvar um novo ID de correspondência imediatamente. Essa opção tem a maior precisão e o tempo de resposta mais lento.
   - Plano de fundo (mostrado `EVENTUAL` na API) — Você pode pesquisar um ID de correspondência existente ou gerar um novo ID de correspondência imediatamente. O registro atualizado é salvo em segundo plano. Essa opção tem uma resposta inicial rápida, com resultados completos disponíveis posteriormente no S3.
   - Geração rápida de ID (mostrada `EVENTUAL_NO_LOOKUP` na API) — Você pode criar uma nova ID de correspondência sem procurar uma existente. O registro atualizado é salvo em segundo plano. Essa opção tem a resposta mais rápida. É recomendado somente para registros exclusivos.

8. Para atributos de registro,

   a. Insira o valor da ID exclusiva.

   b. Insira um valor para cada chave de correspondência que corresponderá aos registros existentes com base nas regras configuradas em seu fluxo de trabalho.

9. Escolha Encontrar ID de correspondência e salve o registro.

   Uma mensagem de sucesso é exibida, informando que a ID de correspondência foi encontrada ou uma nova ID de correspondência foi gerada e o registro foi salvo.

10. Veja a ID de correspondência correspondente e a regra associada que foi salva no fluxo de trabalho correspondente na mensagem de sucesso.

11. (Opcional) Para copiar a ID de correspondência, escolha Copiar.

API

Para modificar ou gerar um Match ID usando a API

> **Note**
>
> Para chamar essa API com sucesso, você deve primeiro executar com êxito um fluxo de trabalho de correspondência baseado em regras usando a StartMatchingJob API.
> Para obter uma lista completa das linguagens de programação suportadas, consulte a seção Consulte também do GenerateMatchID.

1. Abra um terminal ou prompt de comando para fazer a solicitação da API.
2. Crie uma solicitação POST para o seguinte endpoint:

   ```
   /matchingworkflows/workflowName/generateMatches
   ```

3. No cabeçalho da solicitação, defina o tipo de conteúdo como application/json.
4. No URI da solicitação, especifique seu`workflowName`.

   O `workflowName` imperativo:

   - Ter entre 1 e 255 caracteres
   - Combine o padrão [A-zA-z_0-9-] *
5. Para o corpo da solicitação, forneça o seguinte JSON:

   ```
   {
      "processingType": "string",
      "records": [
         {
            "inputSourceARN": "string",
            "recordAttributeMap": {
               "string" : "string"
            },
            "uniqueId": "string"
         }
      ]
   }
   ```

   Em que:

- `processingType`(opcional) - O padrão é. `CONSISTENT` Escolha um desses valores:
  - `CONSISTENT`- Para maior precisão com tempo de resposta mais lento
  - `EVENTUAL`- Para uma resposta inicial mais rápida com processamento em segundo plano
  - `EVENTUAL_NO_LOOKUP`- Para uma resposta mais rápida quando se sabe que os registros são exclusivos
- `records`(obrigatório) - Matriz contendo exatamente um objeto de registro

6. Envie a solicitação .

   Se for bem-sucedido, você receberá uma resposta com o código de status 200 e um corpo JSON contendo:

```
{
   "failedRecords": [
      {
         "errorMessage": "string",
         "inputSourceARN": "string",
         "uniqueId": "string"
      }
   ],
   "matchGroups": [
      {
         "matchId": "string",
         "matchRule": "string",
         "records": [
            {
               "inputSourceARN": "string",
               "recordId": "string"
            }
         ]
      }
   ]
}
```

   Se a chamada não for bem-sucedida, você poderá receber um dos seguintes erros:

   - 403 - AccessDeniedException se você não tiver acesso suficiente
   - 404 - ResourceNotFoundException se o recurso não puder ser encontrado

- 429 - ThrottlingException se a solicitação foi limitada
- 400 - ValidationException se a entrada falhar na validação
- 500 - InternalServerException se houver uma falha de serviço interno

Procurando um ID de correspondência para um fluxo de trabalho de correspondência baseado em regras

Depois de concluir um fluxo de trabalho de correspondência baseado em regras, você pode recuperar a ID de correspondência e a regra associada para cada registro processado. Essas informações ajudam você a entender como os registros foram combinados e quais regras foram aplicadas. O procedimento a seguir demonstra como acessar esses dados usando o AWS Entity Resolution console ou a `GetMatchID` API.

Console

Para pesquisar um Match ID usando o console

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Correspondência.
3. Escolha o fluxo de trabalho de correspondência baseado em regras que foi processado (o status do trabalho é Concluído).
4. Na página de detalhes do fluxo de trabalho correspondente, escolha a IDs guia Corresponder.
5. Escolha Pesquisar ID de correspondência.

> **Note**
>
> A opção Pesquisar ID de correspondência só está disponível para fluxos de trabalho correspondentes que usam a cadência de processamento automático. Se você tiver selecionado a cadência de processamento manual, essa opção aparecerá inativa. Para usar essa opção, edite seu fluxo de trabalho para usar a cadência de processamento automático. Para obter mais informações sobre a edição de fluxos de trabalho, consulteEditando um fluxo de trabalho correspondente.

6. Execute um destes procedimentos:

Se…	Então…
Há somente um mapeamento de esquema associado a esse fluxo de trabalho.	Visualize o mapeamento do esquema selecionado por padrão.
Há mais de um mapeamento de esquema associado a esse fluxo de trabalho.	Escolha o mapeamento do esquema na lista suspensa.

7. Em Atributos de registro, insira o valor de uma chave de correspondência existente para pesquisar cada registro existente.

> **Tip**
>
> Insira o máximo de valores possível para ajudar a encontrar o Match ID.

8. A opção Normalizar dados é selecionada por padrão, para que as entradas de dados sejam normalizadas antes da correspondência. Se você não quiser normalizar dados, desmarque a opção Normalizar dados.
9. Se você quiser ver as regras de correspondência, expanda a opção Exibir regras de correspondência.
10. Escolha Look up.

    Uma mensagem de sucesso aparece informando que o Match ID foi encontrado.
11. Veja o ID de correspondência correspondente e a regra associada que foi encontrada.

API

Para pesquisar um Match ID usando a API

> **Note**
>
> Para chamar essa API com sucesso, você deve primeiro executar com êxito um fluxo de trabalho de correspondência baseado em regras usando a StartMatchingJob API.
> Para obter uma lista completa das linguagens de programação compatíveis, consulte a seção Consulte também da API de GetMatch ID.

1. Abra um terminal ou prompt de comando para fazer a solicitação da API.
2. Crie uma solicitação POST para o seguinte endpoint:

   ```
   /matchingworkflows/workflowName/matches
   ```

3. No cabeçalho da solicitação, defina o tipo de conteúdo como application/json.
4. No URI da solicitação, especifique seu`workflowName`.

   O `workflowName` imperativo:

   - Ter entre 1 e 255 caracteres
   - Combine o padrão [A-zA-z_0-9-] *

5. Para o corpo da solicitação, forneça o seguinte JSON:

   ```
   {
      "applyNormalization": boolean,
      "record": {
         "string" : "string"
      }
   }
   ```

   Em que:

   `applyNormalization`(opcional) - Defina como `true` para normalizar os atributos definidos no esquema

   `record`(obrigatório) - O registro para o qual buscar o Match ID

6. Envie a solicitação .

   Se for bem-sucedido, você receberá uma resposta com o código de status 200 e um corpo JSON contendo:

   ```
   {
      "matchId": "string",
      "matchRule": "string"
   }
   ```

   O `matchId` é o identificador exclusivo desse grupo de registros correspondentes e `matchRule` indica em qual regra o registro correspondeu.

Se a chamada não for bem-sucedida, você poderá receber um dos seguintes erros:

- 403 - AccessDeniedException se você não tiver acesso suficiente
- 404 - ResourceNotFoundException se o recurso não puder ser encontrado
- 429 - ThrottlingException se a solicitação foi limitada
- 400 - ValidationException se a entrada falhar na validação
- 500 - InternalServerException se houver uma falha de serviço interno

Excluindo registros de um fluxo de trabalho de correspondência baseado em regras ou em ML

Se precisar estar em conformidade com os regulamentos de gerenciamento de dados, você pode excluir os registros de um fluxo de trabalho de correspondência baseado em regras ou baseado em ML.

Para excluir registros de um fluxo de trabalho de correspondência baseado em regras ou em ML

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Correspondência.
3. Escolha o fluxo de trabalho de correspondência baseado em regras ou baseado em ML.
4. Na página de detalhes do fluxo de trabalho correspondente, escolha Excluir exclusivo na IDs lista suspensa Ações.
5. Insira o ID exclusivo que você deseja excluir na IDs seção Exclusivo.

   Você pode inserir até 10 exclusivos IDs.
6. Especifique a fonte de entrada da qual excluir o exclusivo IDs.

   Se houver somente uma fonte de entrada para o fluxo de trabalho, a fonte de entrada será listada por padrão.

   Se você especificar apenas uma fonte de entrada, a exclusiva IDs em outras fontes de entrada não será afetada.
7. Escolha Excluir exclusivo IDs.

Solução de problemas de fluxos de trabalho correspondentes

Use as informações a seguir para ajudá-lo a diagnosticar e corrigir problemas comuns que você pode encontrar ao executar fluxos de trabalho correspondentes.

Recebi um arquivo de erro depois de executar um fluxo de trabalho correspondente

Causa comum

Um fluxo de trabalho correspondente pode ter várias execuções e os resultados (acertos ou erros) são gravados em uma pasta com o `jobId` como nome.

Os resultados bem-sucedidos de um fluxo de trabalho correspondente são gravados em uma `success` pasta que contém vários arquivos, e cada arquivo contém um subconjunto dos registros bem-sucedidos.

Os erros de um fluxo de trabalho correspondente são gravados em uma `error` pasta com vários campos, cada um contendo um subconjunto dos registros de erro.

O arquivo de erro pode ser criado pelos seguintes motivos:

- O ID exclusivo é:
  - nulo
  - ausente em uma linha de dados
  - ausente em um registro na tabela de dados
  - repetido em outra linha de dados na tabela de dados
  - não especificado
  - não é exclusivo na mesma fonte
  - não é exclusivo em várias fontes
  - sobreposições entre fontes
  - excede 38 caracteres (somente fluxo de trabalho de correspondência baseado em regras)
- Um dos campos no mapeamento do esquema inclui um nome reservado:
  - EmailAddress
  - InputSourceARN
  - MatchRule

- ID da partida
- HashingProtocol
- ConfidenceLevel
- Origem

> **Note**
>
> Se o registro no arquivo de erro for criado devido aos motivos listados anteriormente, você será cobrado, pois isso incorrerá no custo de processamento do serviço. Se o registro no arquivo de erro for causado por um erro interno do servidor, você não será cobrado.

Resolução

Para resolver esse problema

1. Verifique se a ID exclusiva é válida.

   Se a ID exclusiva não for válida, atualize a ID exclusiva em sua tabela de dados, salve a nova tabela de dados, crie um novo mapeamento de esquema e execute o fluxo de trabalho correspondente novamente.
2. Verifique se um dos campos no mapeamento do esquema inclui um nome reservado.

   Se um dos campos incluir um nome reservado, crie um novo mapeamento de esquema com um novo nome e execute o fluxo de trabalho correspondente novamente.

Mapeie dados de entrada usando um fluxo de trabalho de mapeamento de ID

Fluxo de trabalho de mapeamento de ID é um trabalho de processamento de dados que associa dados de uma fonte de dados de entrada a um destino de dados de entrada com base no método de mapeamento de ID especificado. Ele produz uma tabela de mapeamento de ID.

Um fluxo de trabalho de mapeamento de ID requer uma fonte de dados de entrada e um destino de dados de entrada. Sua fonte e destino de entrada de dados dependem do tipo de mapeamento de ID que você deseja realizar. Há duas maneiras de realizar o mapeamento de ID: com base em regras ou serviços de provedor:

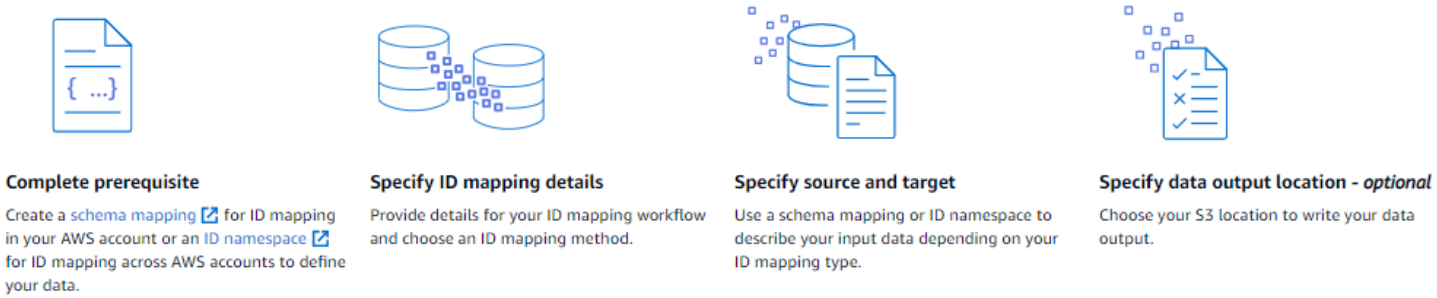
- Mapeamento de ID baseado em regras: você usa regras de correspondência para converter dados primários de uma origem em um destino.
- Mapeamento de ID de serviços do provedor — Você usa o serviço do LiveRamp provedor para traduzir dados de terceiros de uma fonte para um destino.

> **Note**
>
> O fluxo de trabalho de mapeamento de ID dos serviços do provedor AWS Entity Resolution está atualmente integrado ao LiveRamp. Se você tiver uma assinatura do LiveRamp serviço, poderá criar um fluxo de trabalho de mapeamento de ID LiveRamp para realizar a transcodificação. Com a LiveRamp transcodificação, você pode traduzir um conjunto de rampas de origem IDs em qualquer rampID de destino. Ao usar o RampID como um token para representar seus clientes, você pode evitar o compartilhamento de dados do cliente diretamente com plataformas de publicidade.
> Para obter mais informações, consulte Executar tradução por meio do ADX no site da LiveRamp documentação.

Você pode realizar o mapeamento de ID entre dois conjuntos de dados em qualquer um dos seguintes cenários:

- Dentro do seu Conta da AWS
- Em dois diferentes Contas da AWS

O diagrama a seguir resume como configurar um fluxo de trabalho de mapeamento de ID.



Tópicos

- Fluxo de trabalho de mapeamento de ID para um Conta da AWS
- Fluxo de trabalho de mapeamento de ID em dois Contas da AWS
- Executar um fluxo de trabalho de mapeamento de ID
- Executando um fluxo de trabalho de mapeamento de ID com um novo destino de saída
- Editando um fluxo de trabalho de mapeamento de ID
- Excluindo um fluxo de trabalho de mapeamento de ID
- Adicionar ou atualizar uma política de recursos para um fluxo de trabalho de mapeamento de ID

Fluxo de trabalho de mapeamento de ID para um Conta da AWS

Um fluxo de trabalho de mapeamento de ID para um Conta da AWS permite que você realize o mapeamento de ID entre dois conjuntos de dados por conta própria Conta da AWS.

Antes de criar um fluxo de trabalho de mapeamento de ID por conta própria Conta da AWS, você deve primeiro preencher os pré-requisitos.

Depois de criar e executar um fluxo de trabalho de mapeamento de ID, você pode visualizar a saída (a tabela de mapeamento de ID) e usá-la para análise.

Os tópicos a seguir orientam você por um conjunto de etapas para criar um fluxo de trabalho de mapeamento de ID no mesmo Conta da AWS.

Tópicos

- Pré-requisitos
- Criação de um fluxo de trabalho de mapeamento de ID (baseado em regras)
- Criação de um fluxo de trabalho de mapeamento de ID (serviços do provedor)

Pré-requisitos

Antes de criar um fluxo de trabalho de mapeamento de ID para um Conta da AWS usando o método de mapeamento de ID baseado em regras ou o método de mapeamento de ID do Provider Services, você deve primeiro fazer o seguinte:

- Conclua as tarefas em Configurar a resolução de entidades da AWS.
- Conclua as tarefas emPrepare tabelas de dados de entrada, dependendo do tipo de dados de entrada que você está usando.
- Crie um mapeamento de esquema ou crie um fluxo de trabalho correspondente.
- (Somente mapeamento de ID de serviços do provedor) Antes de criar um fluxo de trabalho de mapeamento de ID com LiveRamp, você deve escolher um bucket de preparação de dados do Amazon Simple Storage Service (Amazon S3) no qual deseja gravar temporariamente a saída do fluxo de trabalho de mapeamento de ID.

  Se você estiver usando o serviço do LiveRamp provedor para traduzir dados de terceiros, adicione a seguinte política de permissões, que permite acessar o intervalo de preparação de dados.

  JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::715724997226:root"

            },
            "Action": [
                "s3:PutObject",
                "s3:GetObject",
                "s3:GetObjectVersion",
                "s3:DeleteObject"
            ],
            "Resource": [
                "arn:aws:s3:::<staging-bucket>",
                "arn:aws:s3:::<staging-bucket>/*"
            ]
        },
        {
```

```
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::715724997226:root"
            },
            "Action": [
                "s3:ListBucket",
                "s3:GetBucketLocation",
                "s3:GetBucketPolicy",
                "s3:ListBucketVersions",
                "s3:GetBucketAcl"
            ],
            "Resource": [
                "arn:aws:s3:::<staging-bucket>",
                "arn:aws:s3:::<staging-bucket>/*"
            ]
        }
    ]
}
```

Na política de permissões anterior, substitua cada uma *<user input placeholder>* por suas próprias informações.

staging-bucket	O bucket do Amazon S3 que armazena temporariamente seus dados enquanto executa um fluxo de trabalho baseado em serviços do provedor.

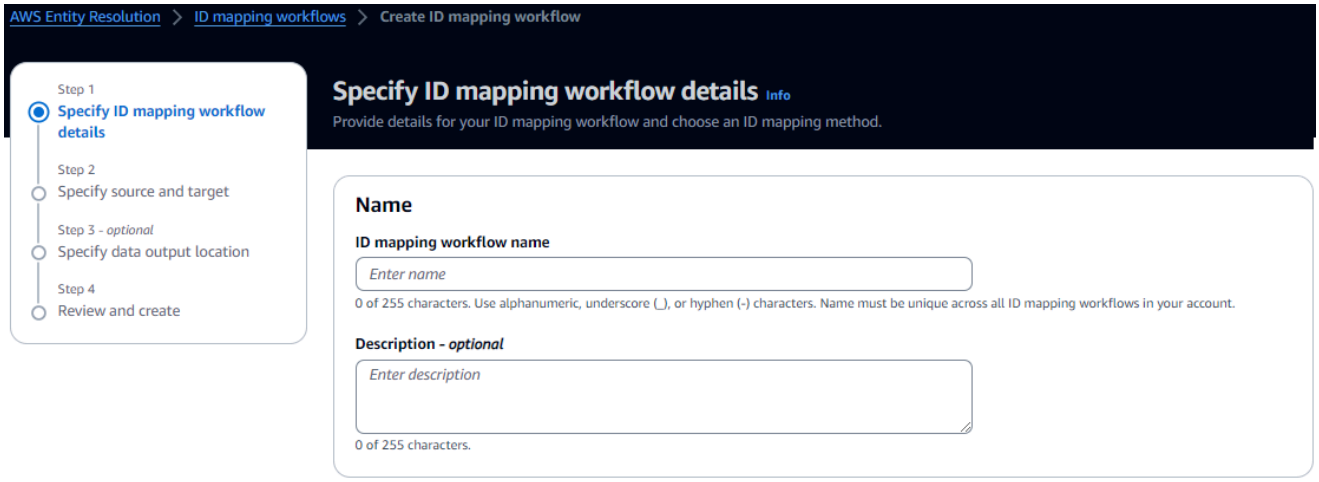
Criação de um fluxo de trabalho de mapeamento de ID (baseado em regras)

Este tópico descreve o processo de criação de um fluxo de trabalho de mapeamento de ID para um Conta da AWS que usa regras de correspondência para traduzir dados primários de uma fonte para um destino.

Para criar um fluxo de trabalho de mapeamento de ID baseado em regras para um Conta da AWS

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Mapeamento de ID.

3. Na página Fluxos de trabalho de mapeamento de ID, no canto superior direito, escolha Criar fluxo de trabalho de mapeamento de ID.
4. Para a Etapa 1: Especificar detalhes do fluxo de trabalho de mapeamento de ID, faça o seguinte.
   a. Insira um nome de fluxo de trabalho de mapeamento de ID e uma Descrição opcional.



   b. Para o método de mapeamento de ID, escolha Baseado em regras.
   c. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e, em seguida, insira o par de chave e valor.
   d. Escolha Próximo.
5. Para a Etapa 2: Especificar a origem e o destino, faça o seguinte.
   a. Em Origem, escolha o cenário que se aplica a você e, em seguida, execute a ação recomendada.

Cenário	Ação recomendada
Use seu próprio mapeamento AWS Glue de banco de dados, AWS Glue tabelas e esquemas no fluxo de trabalho de mapeamento de ID.	1. Escolha Mapeamento do esquema. 2. Selecione um AWS Gluebanco de dados na lista suspensa, selecione a AWS Glue tabela e, em seguida, selecione o mapeamento de esquema correspondente.

Cenário	Ação recomendada
	Você pode adicionar até 19 entradas de dados.
Use um fluxo de trabalho correspondente existente que aponte para os dados de registro que você deseja usar no fluxo de trabalho de mapeamento de ID.	1. Escolha Fluxo de trabalho correspon dente. 2. Selecione um fluxo de trabalho de correspondência existente na lista suspensa.

b. Para Target, selecione um fluxo de trabalho de correspondência existente na lista suspensa.

c. Para parâmetros de regra, faça o seguinte.

i. Especifique os controles de regra escolhendo uma das opções a seguir com base no seu tipo de fonte.

Tipo de origem	Ação recomendada
Fluxo de trabalho correspondente	Especifique os controles de regra escolhendo se uma origem, um destino ou ambos podem fornecer regras em um fluxo de trabalho de mapeamento de ID. Os controles de regras devem ser compatíveis entre a origem e o destino para serem usados em um fluxo de trabalho de mapeamento de ID. Por exemplo, se um namespace de ID de origem limitar as regras ao destino, mas o namespace de ID de destino limitar as regras à origem, isso vai gerar um erro.
Mapeamento de esquemas	Pule esta etapa.

ii. Para parâmetros de comparação e correspondência, o tipo de comparação é automaticamente definido como Vários campos de entrada.

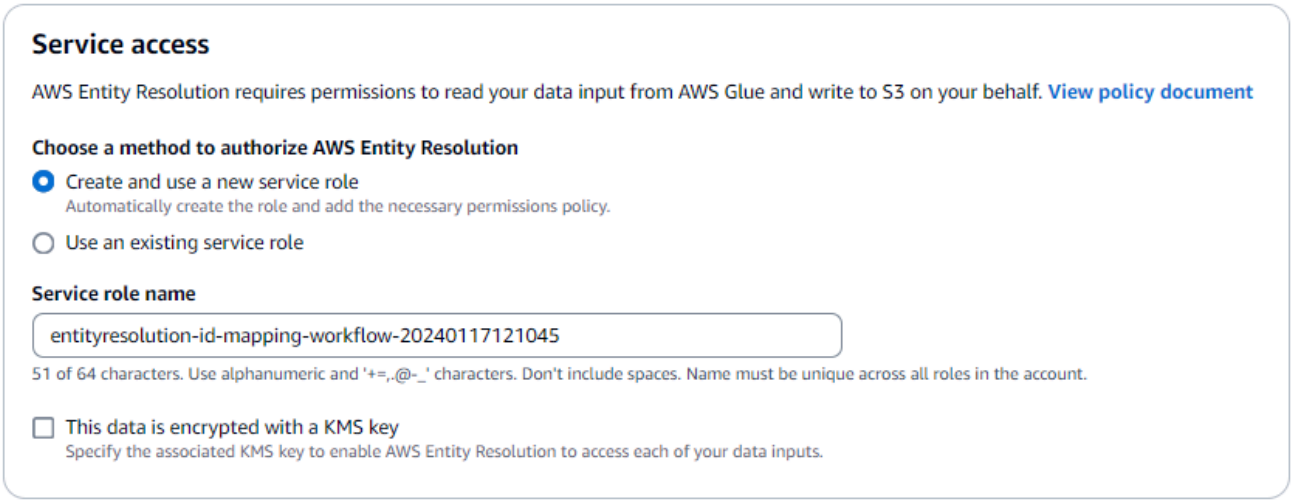
Isso ocorre porque os dois participantes haviam selecionado essa opção anteriormente.

d. Especifique o tipo de correspondência de registros escolhendo uma das seguintes opções com base em sua meta.

Seu objetivo	Opção recomendada
Limite o tipo de correspondência de registro para armazenar somente um registro correspondente na origem para cada registro correspondente no destino ao criar o fluxo de trabalho de mapeament o de ID.	Uma fonte para um alvo
Limite o tipo de correspondência de registro para armazenar todos os registros correspondentes na origem para cada registro correspondente no destino ao criar o fluxo de trabalho de mapeamento de ID.	Muitas fontes para um alvo

> **Note**
>
> Você deve especificar limitações compatíveis para os namespaces de ID de origem e destino.

e. Para especificar as permissões de acesso ao serviço, escolha uma opção e execute a ação recomendada.



Opção	Ação recomendada
Criar e usar um novo perfil de serviço	• AWS Entity Resolution cria uma função de serviço com a política necessária para essa tabela. • O nome do perfil de serviço padrão é `entityresolution-id-mapping-workflow-<timestamp>` . • Você deve ter permissões para criar perfis e anexar políticas. • Se seus dados de entrada estiverem criptografados, escolha a opção Esses dados são criptografados por uma chave KMS. Em seguida, insira uma AWS KMS chave usada para descriptografar sua entrada de dados.

Opção	Ação recomendada
Use um perfil de serviço existente	1. Escolha um nome do perfil de serviço existente na lista suspensa. A lista de perfis é exibida se você tiver permissões para listar funções. Se você não tiver permissões para listar perfis, insira o nome do recurso da Amazon (ARN) do perfil que você deseja usar. Se não houver perfis de serviço existentes, a opção de Usar um perfil de serviço existente não estará disponível. 2. Para visualizar o perfil de serviço, selecione o link externo Visualizar no IAM. Por padrão, AWS Entity Resolution não tenta atualizar a política de função existente para adicionar as permissões necessárias.

6. Escolha Próximo.
7. Para a Etapa 3: Especifique o local de saída de dados — opcional, faça o seguinte.
   a. Para Destino de saída de dados, faça o seguinte:
      i. Escolha a localização do Amazon S3 para a saída de dados.
      ii. Em Criptografia, se você optar por Personalizar as configurações de criptografia, insira o ARN da AWS KMS chave ou escolha Criar uma AWS KMS chave.
   b. Escolha Próximo.
8. Para a Etapa 4: revisar e criar, faça o seguinte.
   a. Revise as seleções feitas nas etapas anteriores e edite-as, se necessário.
   b. Escolha Criar.

Uma mensagem aparece indicando que o fluxo de trabalho de mapeamento de ID foi criado.

Depois de criar o fluxo de trabalho de mapeamento de ID, você estará pronto para executar um fluxo de trabalho de mapeamento de ID.

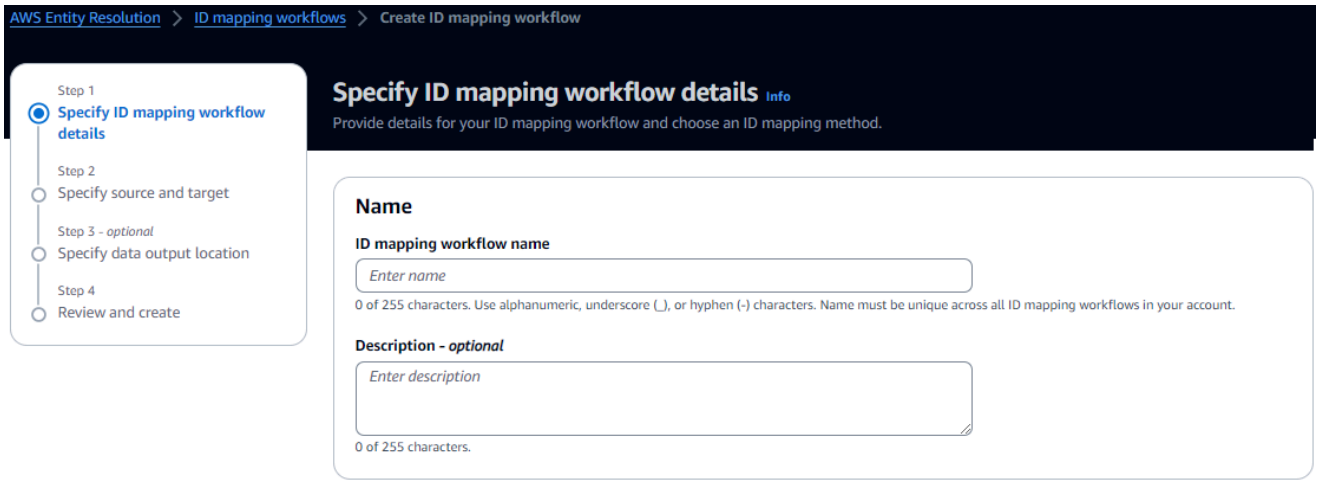
Criação de um fluxo de trabalho de mapeamento de ID (serviços do provedor)

Este tópico descreve o processo de criação de um fluxo de trabalho de mapeamento de ID para alguém Conta da AWS usando um serviço de provedor chamado LiveRamp. LiveRamp traduz um conjunto de rampa de origem IDs para outro conjunto usando rampa mantida ou derivada. IDs

Para criar um fluxo de trabalho de mapeamento de ID baseado em serviços de provedores para um Conta da AWS

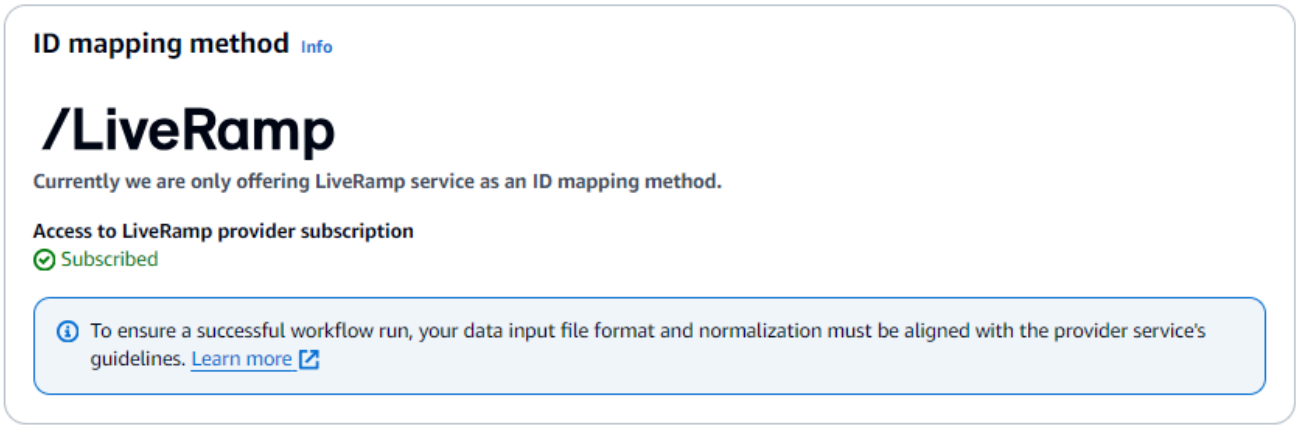
1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Mapeamento de ID.
3. Na página Fluxos de trabalho de mapeamento de ID, no canto superior direito, escolha Criar fluxo de trabalho de mapeamento de ID.
4. Para a Etapa 1: Especificar detalhes do fluxo de trabalho de mapeamento de ID, faça o seguinte.

    a. Insira um nome de fluxo de trabalho de mapeamento de ID e uma Descrição opcional.



    b. Para o método de mapeamento de ID, escolha Provider services.

AWS Entity Resolution atualmente oferece o serviço do LiveRamp provedor como um método de mapeamento de ID. Se você tiver uma assinatura LiveRamp, o status aparecerá como Assinado. Para obter mais informações sobre como assinar LiveRamp, consulteEtapa 1: Assine um serviço de provedor em AWS Data Exchange.



> **Note**
>
> Certifique-se de que o formato do arquivo de entrada de dados esteja alinhado com as diretrizes do serviço do provedor. Para obter mais informações sobre as diretrizes de formatação LiveRamp do arquivo de entrada, consulte Executar tradução por meio do ADX no site da LiveRamp documentação.

c. Para LiveRamp configuração, insira os seguintes valores que LiveRamp fornecem:

- Gerenciador de ID de cliente ARN
- Gerenciador secreto do cliente ARN

d. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e, em seguida, insira o par de chave e valor.

e. Escolha Próximo.

5. Para a Etapa 2: Especificar a origem e o destino, faça o seguinte.

a. Em Origem, escolha o cenário que se aplica a você e, em seguida, execute a ação recomendada.

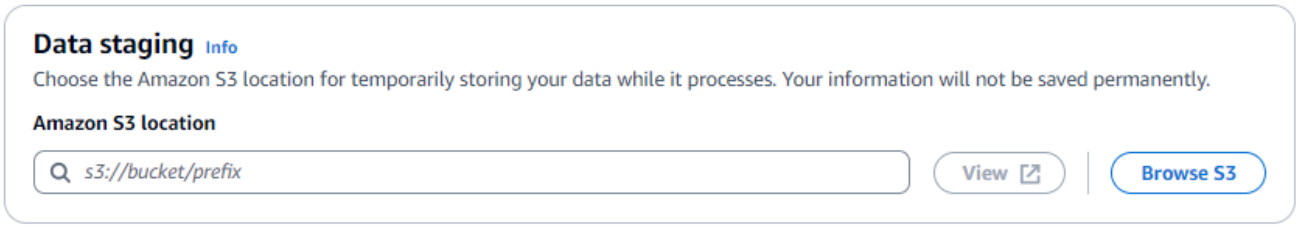
Cenário	Ação recomendada
Use seu próprio mapeamento AWS Glue de banco de dados, AWS Glue tabe las e esquemas no fluxo de trabalho de mapeamento de ID.	1. Escolha Mapeamento do esquema. 2. Selecione um AWS Gluebanco de dados na lista suspensa, selecione a AWS Glue tabela e, em seguida, selecione o mapeamento de esquema correspondente. Você pode adicionar até 19 entradas de dados.
Use um fluxo de trabalho correspondente existente que aponte para os dados de registro que você deseja usar no fluxo de trabalho de mapeamento de ID.	1. Escolha Fluxo de trabalho correspon dente. 2. Selecione um fluxo de trabalho de correspondência existente na lista suspensa.

b. Para o Target, execute uma das ações a seguir com base no método de mapeamento de ID escolhido.

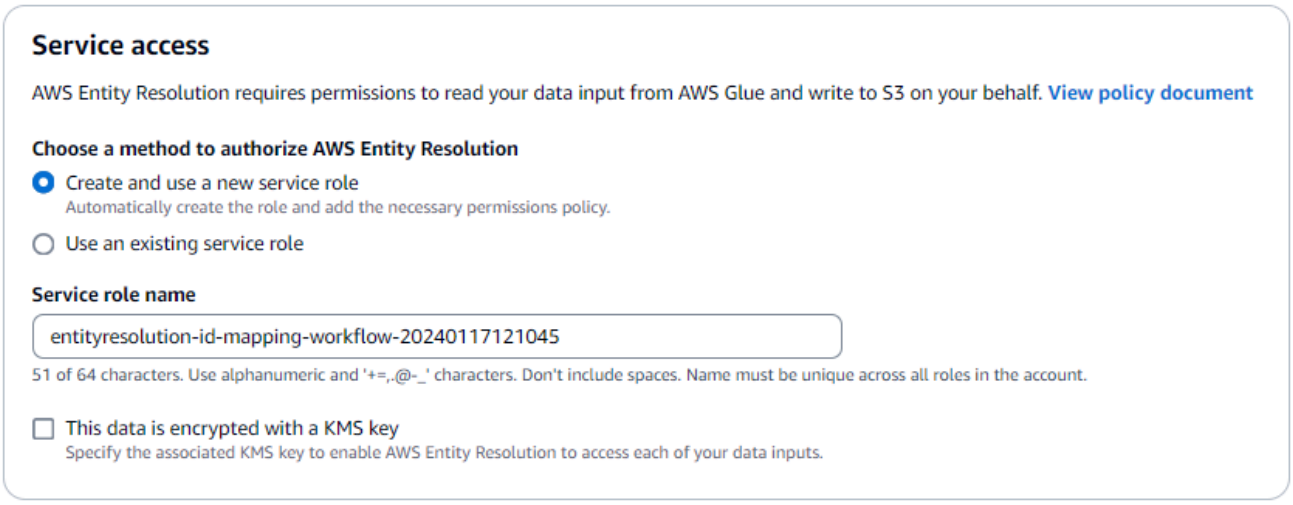
Método de mapeamento de ID	Ação recomendada
Baseado em regras	Selecione um fluxo de trabalho de correspondência existente na lista suspensa.

Método de mapeamento de ID	Ação recomendada
Serviços do provedor	Insira o identificador de domínio LiveRamp do cliente destinado à transcodificação LiveRamp fornecido no domínio de destino. 

c. Para preparação de dados, escolha o local do Amazon S3 em que você deseja gravar temporariamente a saída do fluxo de trabalho de mapeamento de ID.



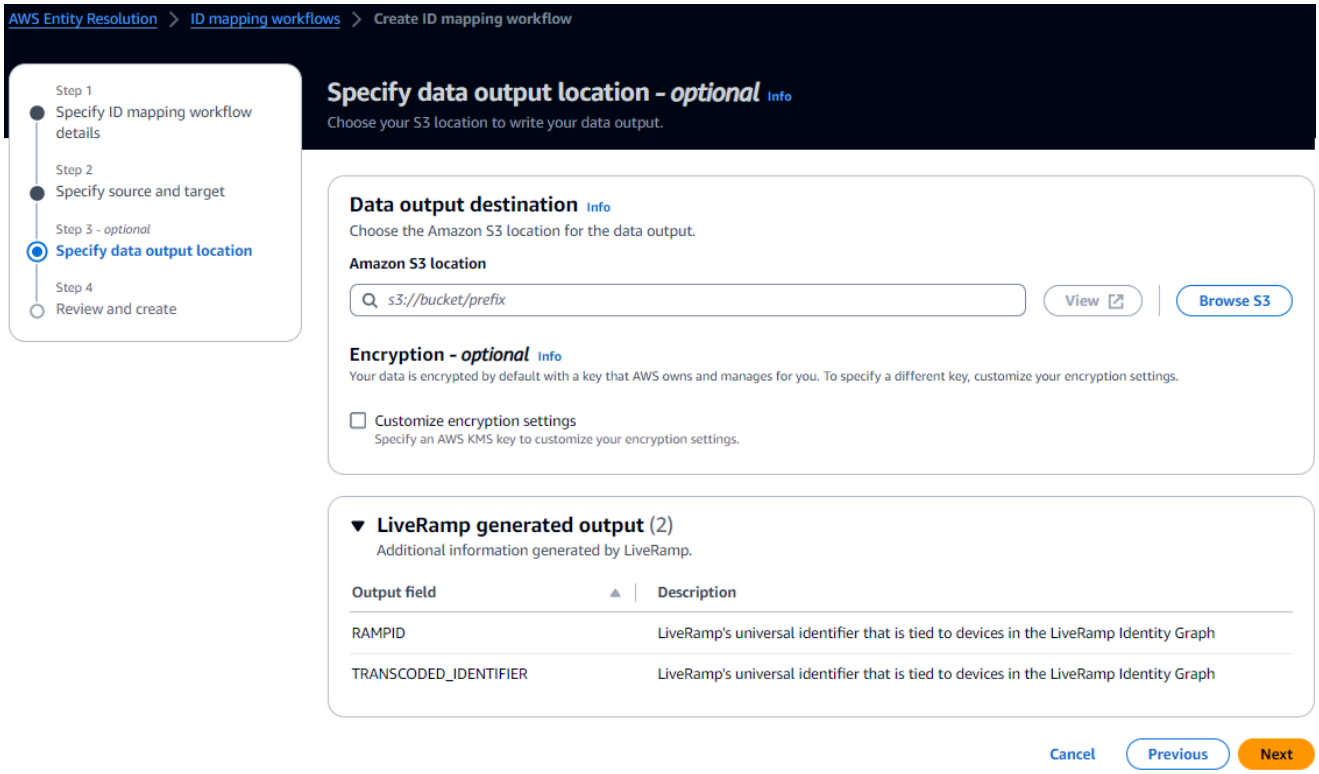
d. Para especificar as permissões de acesso ao serviço, escolha uma opção e execute a ação recomendada.



Opção	Ação recomendada
Criar e usar um novo perfil de serviço	• AWS Entity Resolution cria uma função de serviço com a política necessária para essa tabela. • O nome do perfil de serviço padrão é `entityresolution-id-mapping-workflow-<timestamp>` . • Você deve ter permissões para criar perfis e anexar políticas. • Se seus dados de entrada estiverem criptografados, escolha a opção Esses dados são criptografados por uma chave KMS. Em seguida, insira uma AWS KMS chave usada para descriptografar sua entrada de dados.

Opção	Ação recomendada
Use um perfil de serviço existente	1. Escolha um nome do perfil de serviço existente na lista suspensa. A lista de perfis é exibida se você tiver permissões para listar funções. Se você não tiver permissões para listar perfis, insira o nome do recurso da Amazon (ARN) do perfil que você deseja usar. Se não houver perfis de serviço existentes, a opção de Usar um perfil de serviço existente não estará disponível. 2. Para visualizar o perfil de serviço, selecione o link externo Visualizar no IAM. Por padrão, AWS Entity Resolution não tenta atualizar a política de função existente para adicionar as permissões necessárias.

6. Escolha Próximo.
7. Para a Etapa 3: Especifique o local de saída de dados — opcional, faça o seguinte.
   a. Para Destino de saída de dados, faça o seguinte:
      i. Escolha a localização do Amazon S3 para a saída de dados.
      ii. Em Criptografia, se você optar por Personalizar as configurações de criptografia, insira o ARN da AWS KMS chave ou escolha Criar uma AWS KMS chave.
   b. Visualize a saída LiveRamp gerada.
   c. Escolha Próximo.



8. Para a Etapa 4: revisar e criar, faça o seguinte.

   a. Revise as seleções feitas nas etapas anteriores e edite-as, se necessário.

   b. Escolha Criar.

      Uma mensagem aparece indicando que o fluxo de trabalho de mapeamento de ID foi criado.

9. Depois de criar o fluxo de trabalho de mapeamento de ID, você estará pronto para executar um fluxo de trabalho de mapeamento de ID.

Fluxo de trabalho de mapeamento de ID em dois Contas da AWS

Um fluxo de trabalho de mapeamento de ID em dois Contas da AWS permite que você execute o mapeamento de ID entre dois conjuntos de dados em dois Contas da AWS. Isso normalmente é feito entre você Conta da AWS e outro Conta da AWS.

Por exemplo, um editor pode criar um fluxo de trabalho de mapeamento de ID usando seu próprio namespace de ID de destino (no seu próprio Conta da AWS) e o namespace de ID de origem do anunciante (em outro). Conta da AWS

Antes de criar um fluxo de trabalho de mapeamento de ID entre dois Contas da AWS, você deve primeiro preencher os pré-requisitos.

Depois de criar um fluxo de trabalho de mapeamento de ID, você pode visualizar a saída (a tabela de mapeamento de ID) e usá-la para análise.

Os tópicos a seguir orientam você por um conjunto de etapas para criar um fluxo de trabalho de mapeamento de ID em duas Contas da AWS:

Tópicos

- Pré-requisitos
- Criação de um fluxo de trabalho de mapeamento de ID (baseado em regras)
- Criação de um fluxo de trabalho de mapeamento de ID (serviços do provedor)

Pré-requisitos

Antes de criar um fluxo de trabalho de mapeamento de ID entre dois Contas da AWS, você deve primeiro fazer o seguinte:

- Conclua as tarefas em Configurar AWS Entity Resolution.
- Crie uma fonte de namespace de ID.
- Crie um destino de namespace de ID.
- Adquira o ARN do namespace de ID se você estiver usando uma fonte de namespace de ID de outra. Conta da AWS
- (Somente serviços do provedor) A criação de um fluxo de trabalho de mapeamento de ID entre dois Contas da AWS exige permissão LiveRamp para acessar o bucket do S3 e a chave gerenciada pelo cliente AWS Key Management Service (AWS KMS).

  Antes de criar um fluxo de trabalho de mapeamento de ID entre dois Contas da AWS LiveRamp, adicione a seguinte política de permissão, que permite LiveRamp acessar o bucket do S3 e a chave gerenciada pelo cliente.

  JSON

```
{
    "Version": "2012-10-17",
    "Statement": [{
        "Effect": "Allow",
```

```
        "Principal": {
            "AWS": "arn:aws:iam::715724997226:root"
        },
        "Action": [
            "kms:Decrypt"
        ],
        "Resource": "<KMSKeyARN>",
        "Condition": {
            "StringEquals": {
                "kms:ViaService": "s3.amazonaws.com"
            }
        }
    }]
}
```

Na política de permissões anterior, substitua cada uma *<user input placeholder>* por suas próprias informações.

<KMSKeyARN>	O ARN de uma chave gerenciada pelo AWS KMS cliente.

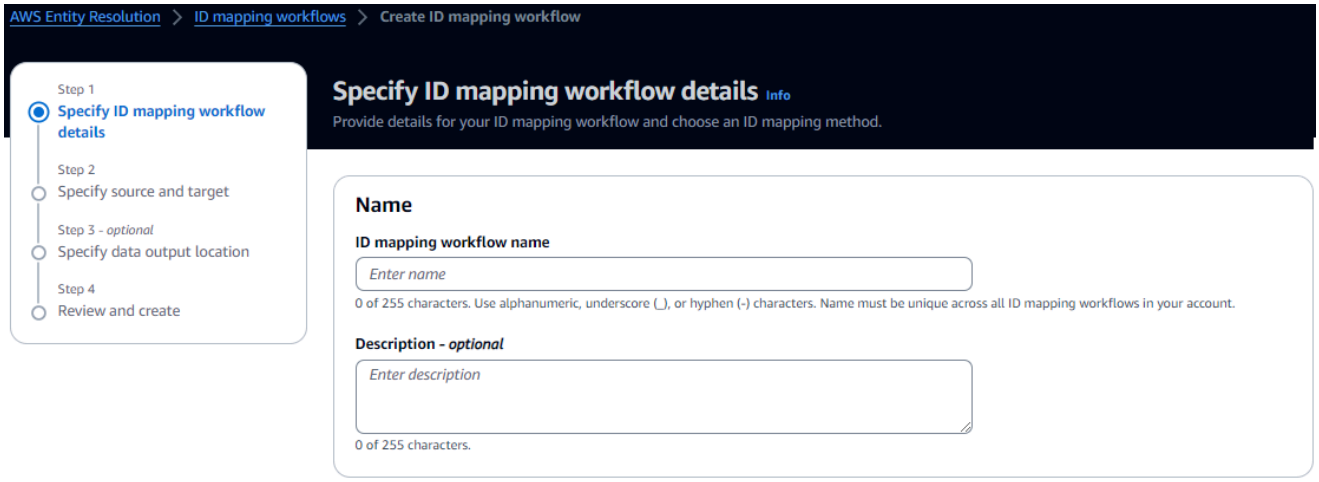
Criação de um fluxo de trabalho de mapeamento de ID (baseado em regras)

Depois de concluir os pré-requisitos, você pode criar um ou mais fluxos de trabalho de mapeamento de ID para usar regras de correspondência para traduzir dados primários de uma fonte para um destino.

Para criar um fluxo de trabalho de mapeamento de ID baseado em regras em dois Contas da AWS

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Mapeamento de ID.
3. Na página Fluxos de trabalho de mapeamento de ID, no canto superior direito, escolha Criar fluxo de trabalho de mapeamento de ID.
4. Para a Etapa 1: Especificar detalhes do fluxo de trabalho de mapeamento de ID, faça o seguinte.

    a. Insira um nome de fluxo de trabalho de mapeamento de ID e uma Descrição opcional.



b. Para o método de mapeamento de ID, escolha Baseado em regras.

c. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e, em seguida, insira o par de chave e valor.

d. Escolha Próximo.

5. Para a Etapa 2: Especificar a origem e o destino, faça o seguinte.

a. Ative as opções avançadas.

b. Em Origem, escolha Fluxo de trabalho correspondente e, em seguida, selecione o fluxo de trabalho correspondente existente na lista suspensa.

c. Para Target, escolha Fluxo de trabalho correspondente e, em seguida, selecione o fluxo de trabalho correspondente existente na lista suspensa.

d. Para parâmetros de regra, especifique os controles de regra escolhendo se uma origem ou um destino podem fornecer regras em um fluxo de trabalho de mapeamento de ID.

   Os controles de regras devem ser compatíveis entre a origem e o destino para serem usados em um fluxo de trabalho de mapeamento de ID. Por exemplo, se um namespace de ID de origem limitar as regras ao destino, mas o namespace de ID de destino limitar as regras à origem, isso vai gerar um erro.

e. Para parâmetros de comparação e correspondência, faça o seguinte.

   i. Especifique o tipo de comparação escolhendo uma opção com base em sua meta.

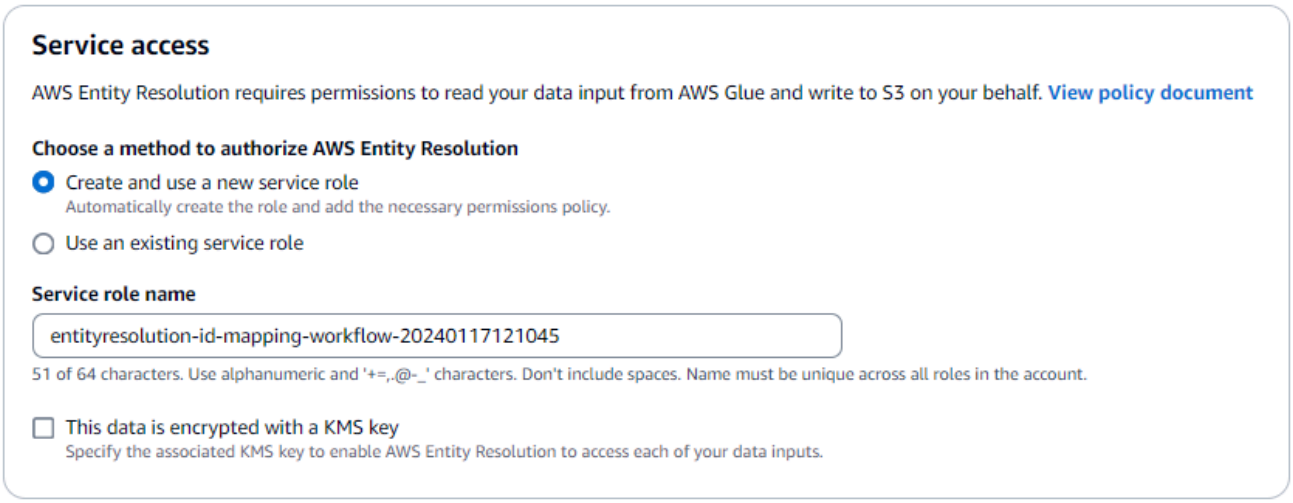
Seu objetivo	Opção recomendada
Encontre qualquer combinação de correspondências nos dados armazenados em vários campos de entrada, independentemente de os dados estarem no mesmo campo de entrada ou em um campo de entrada diferente.	Vários campos de entrada
Limite a comparação em um único campo de entrada, quando dados semelhantes armazenados em vários campos de entrada não devem ser correspondidos.	Campo de entrada único

ii. Especifique o tipo de correspondência de registros escolhendo uma opção com base em sua meta.

Seu objetivo	Opção recomendada
Limite o tipo de correspondência de registro para armazenar somente um registro correspondente na origem para cada registro correspondente no destino ao criar o fluxo de trabalho de mapeamento de ID.	Uma fonte para um alvo
Limite o tipo de correspondência de registro para armazenar todos os registros correspondentes na origem para cada registro correspondente no destino ao criar o fluxo de trabalho de mapeamento de ID.	Muitas fontes para um alvo

> **Note**
>
> Você deve especificar limitações compatíveis para os namespaces de ID de origem e destino.

f. Para especificar as permissões de acesso ao serviço, escolha uma opção e execute a ação recomendada.



Opção	Ação recomendada
Criar e usar um novo perfil de serviço	• AWS Entity Resolution cria uma função de serviço com a política necessária para essa tabela. • O nome do perfil de serviço padrão é `entityresolution-id-mapping-workflow-<timestamp>` . • Você deve ter permissões para criar perfis e anexar políticas. • Se seus dados de entrada estiverem criptografados, escolha a opção Esses dados são criptografados por uma chave KMS. Em seguida, insira uma AWS KMS chave usada para descriptografar sua entrada de dados.

Opção	Ação recomendada
Use um perfil de serviço existente	1. Escolha um nome do perfil de serviço existente na lista suspensa. A lista de perfis é exibida se você tiver permissões para listar funções. Se você não tiver permissões para listar perfis, insira o nome do recurso da Amazon (ARN) do perfil que você deseja usar. Se não houver perfis de serviço existentes, a opção de Usar um perfil de serviço existente não estará disponível. 2. Para visualizar o perfil de serviço, selecione o link externo Visualizar no IAM. Por padrão, AWS Entity Resolution não tenta atualizar a política de função existente para adicionar as permissões necessárias.

6. Escolha Próximo.
7. Para a Etapa 3: Especifique o local de saída de dados — opcional, faça o seguinte.
   a. Para Destino de saída de dados, faça o seguinte.
      i. Escolha a localização do Amazon S3 para a saída de dados.
      ii. Em Criptografia, se você optar por Personalizar as configurações de criptografia, insira o ARN da AWS KMS chave ou escolha Criar uma AWS KMS chave.
   b. Visualize a saída LiveRamp gerada.
   c. Escolha Próximo.
8. Para a Etapa 4: revisar e criar, faça o seguinte.
   a. Revise as seleções feitas nas etapas anteriores e edite-as, se necessário.

b. Escolha Criar.

Uma mensagem aparece indicando que o fluxo de trabalho de mapeamento de ID foi criado.

Depois de criar o fluxo de trabalho de mapeamento de ID, você estará pronto para executar um fluxo de trabalho de mapeamento de ID.

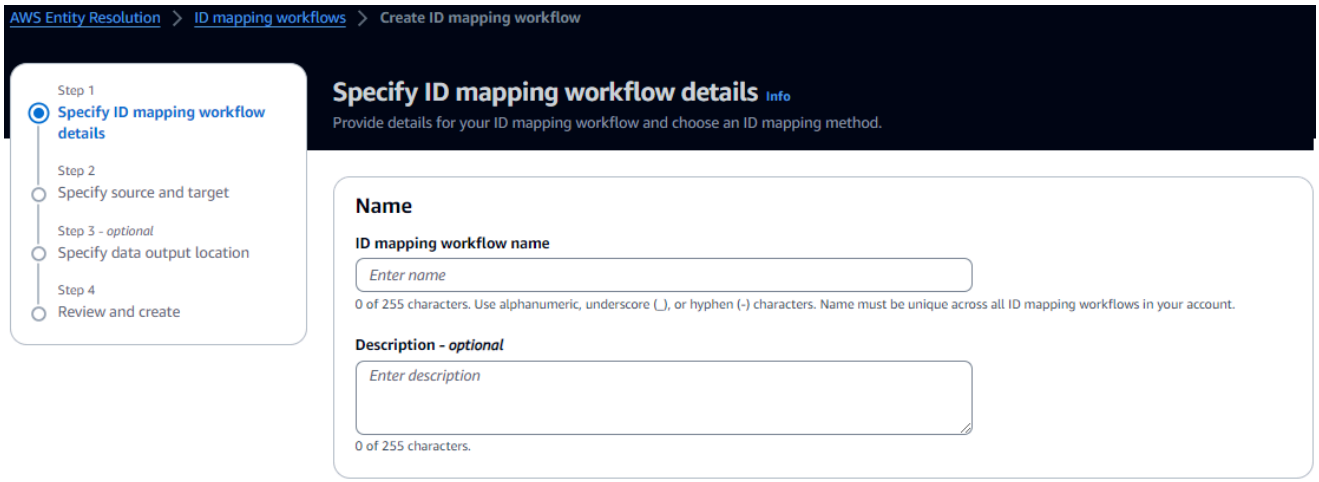
Criação de um fluxo de trabalho de mapeamento de ID (serviços do provedor)

Depois de concluir os pré-requisitos, você pode criar um ou mais fluxos de trabalho de mapeamento de ID usando o serviço do provedor. LiveRamp LiveRamp traduz um conjunto de rampa de origem IDs para outro conjunto usando rampa mantida ou derivada. IDs

Para criar um fluxo de trabalho de mapeamento de ID usando o serviço do provedor

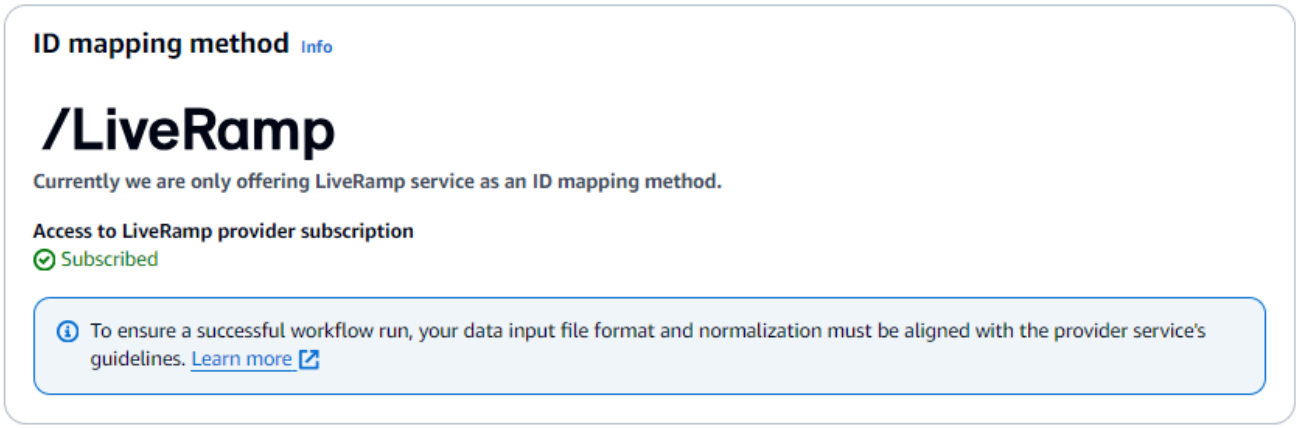
1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Mapeamento de ID.
3. Na página Fluxos de trabalho de mapeamento de ID, no canto superior direito, escolha Criar fluxo de trabalho de mapeamento de ID.
4. Para a Etapa 1: Especificar detalhes do fluxo de trabalho de mapeamento de ID, faça o seguinte.

   a. Insira um nome de fluxo de trabalho de mapeamento de ID e uma Descrição opcional.



   b. Para o método de mapeamento de ID, escolha Provider services.

AWS Entity Resolution atualmente oferece o serviço do LiveRamp provedor como um método de mapeamento de ID. Se você tiver uma assinatura LiveRamp, o status aparecerá como Assinado. Para obter mais informações sobre como assinar LiveRamp, consulteEtapa 1: Assine um serviço de provedor em AWS Data Exchange.



> **Note**
>
> Certifique-se de que o formato do arquivo de entrada de dados esteja alinhado com as diretrizes do serviço do provedor. Para obter mais informações sobre as diretrizes de formatação LiveRamp do arquivo de entrada, consulte Executar tradução por meio do ADX no site da LiveRamp documentação.

c. Para LiveRamp configuração, insira os seguintes valores que LiveRamp fornecem:

- Gerenciador de ID de cliente ARN
- Gerenciador secreto do cliente ARN

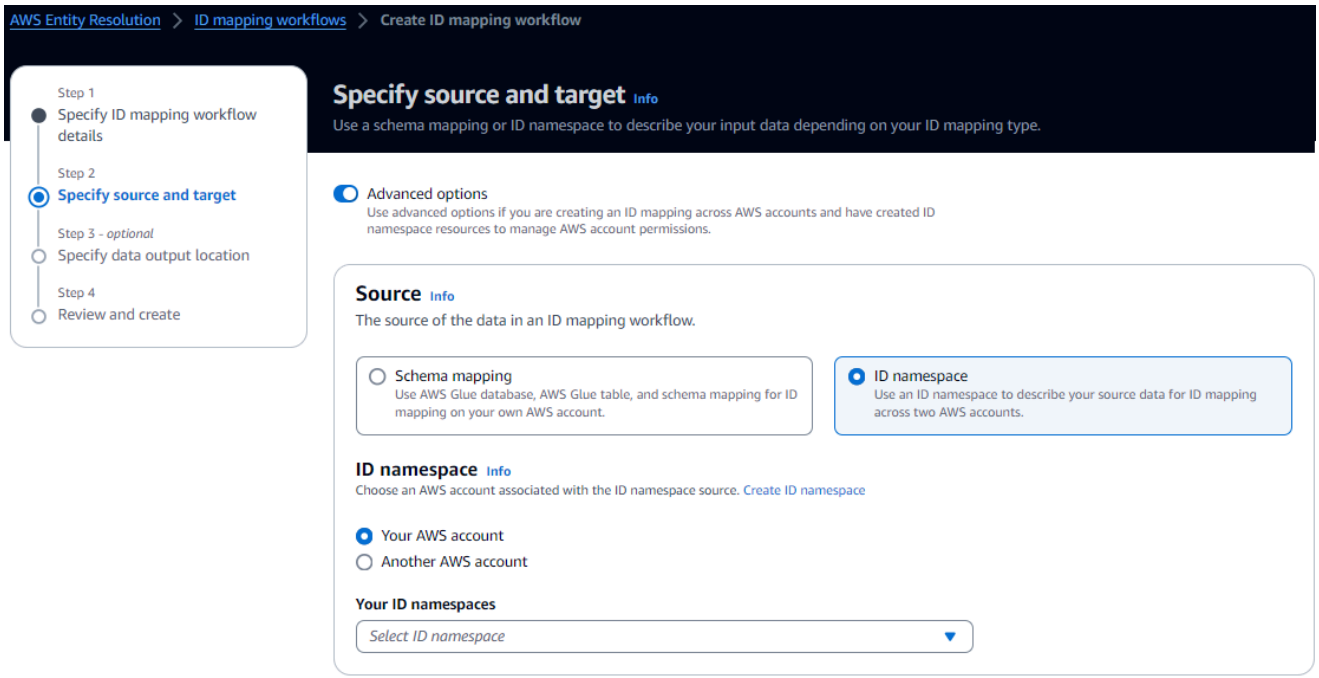
d. (Opcional) Para ativar tags para o recurso, escolha Adicionar nova tag e, em seguida, insira o par de chave e valor.

e. Escolha Próximo.

5. Para a Etapa 2: Especificar a origem e o destino, faça o seguinte.

a. Ative as opções avançadas.

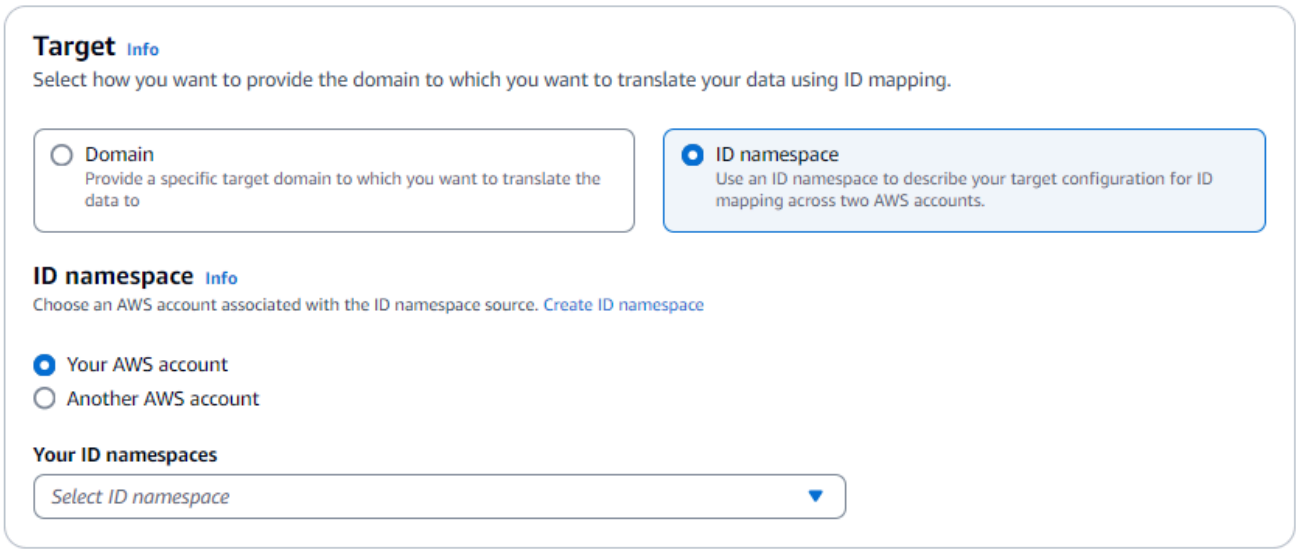
b. Em Source, escolha ID namespace.



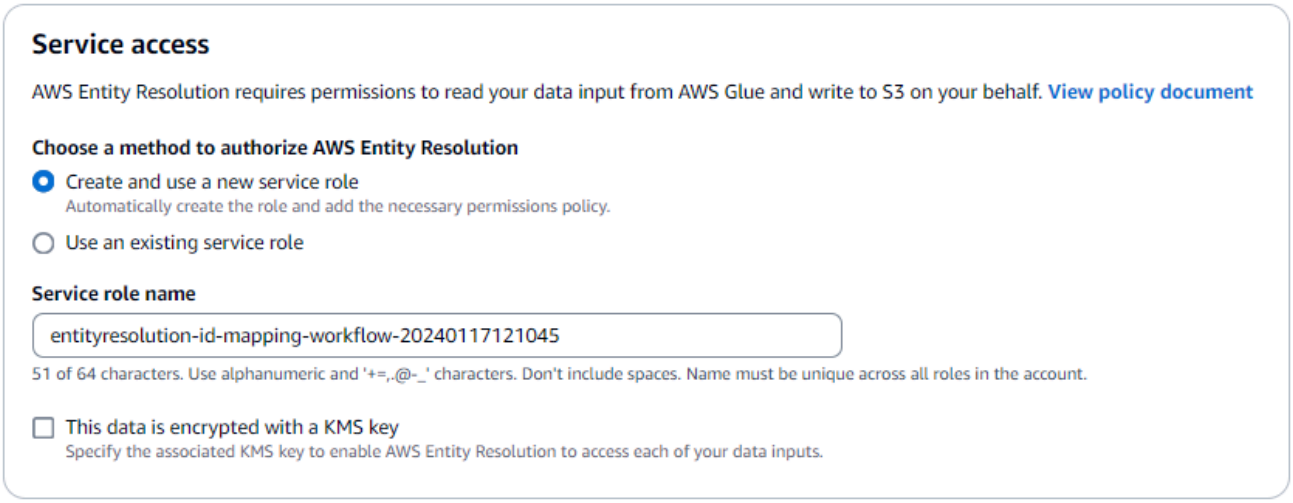
c. Para o namespace ID, identifique onde o namespace ID está localizado e, em seguida, execute a ação recomendada.

Localização do namespace de ID	Ação recomendada
Seu próprio Conta da AWS	1. Escolha o seu Conta da AWS. 2. Selecione o namespace ID na lista suspensa Seus namespaces de ID.
De outra pessoa Conta da AWS	1. Escolha outro Conta da AWS. 2. Insira o ARN do namespace ID.

d. Em Target, escolha o namespace ID.



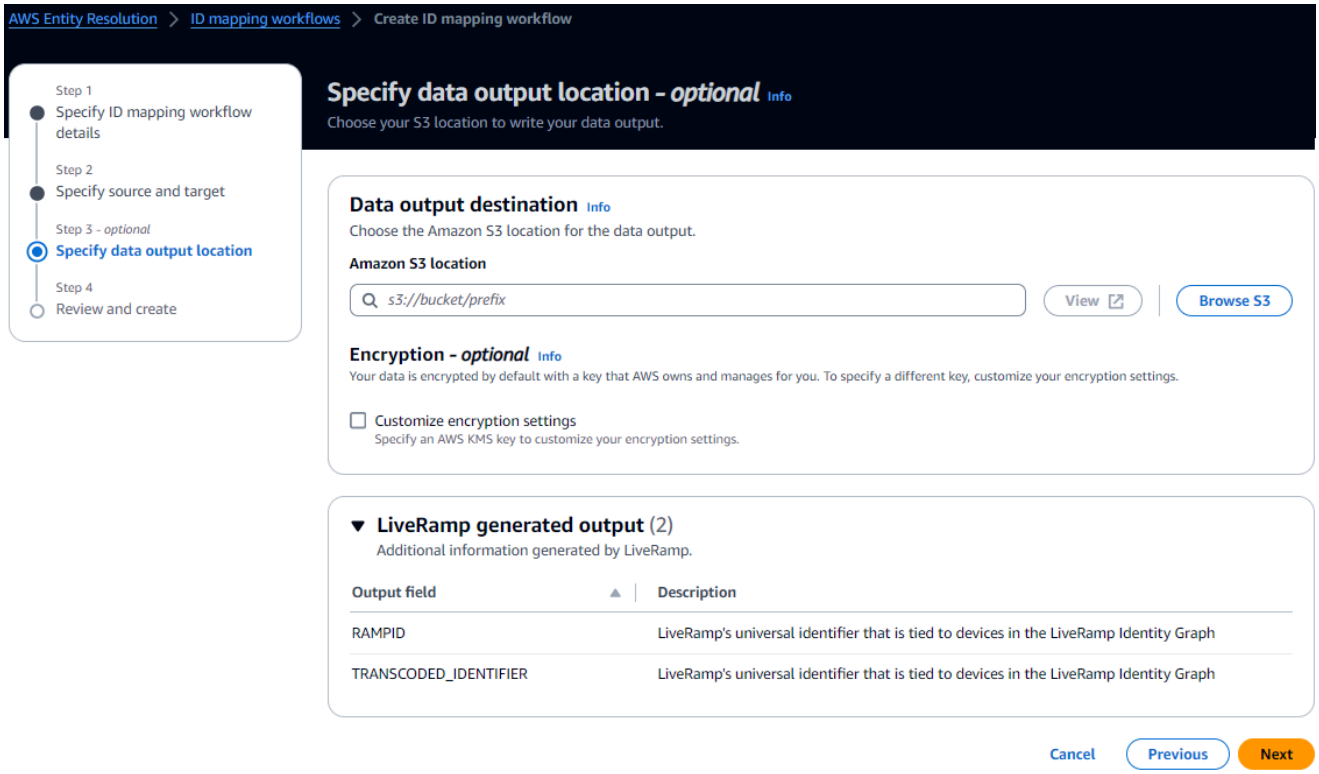
e. Para especificar as permissões de acesso ao serviço, escolha uma opção e execute a ação recomendada.



Opção	Ação recomendada
Criar e usar um novo perfil de serviço	• AWS Entity Resolution cria uma função de serviço com a política necessária para essa tabela. • O nome do perfil de serviço padrão é `entityresolution-id-mapping-workflow-<timestamp>` . • Você deve ter permissões para criar perfis e anexar políticas. • Se seus dados de entrada estiverem criptografados, escolha a opção Esses dados são criptografados por uma chave KMS. Em seguida, insira uma AWS KMS chave usada para descriptografar sua entrada de dados.

Opção	Ação recomendada
Use um perfil de serviço existente	1. Escolha um nome do perfil de serviço existente na lista suspensa. A lista de perfis é exibida se você tiver permissões para listar funções. Se você não tiver permissões para listar perfis, insira o nome do recurso da Amazon (ARN) do perfil que você deseja usar. Se não houver perfis de serviço existentes, a opção de Usar um perfil de serviço existente não estará disponível. 2. Para visualizar o perfil de serviço, selecione o link externo Visualizar no IAM. Por padrão, AWS Entity Resolution não tenta atualizar a política de função existente para adicionar as permissões necessárias.

6. Escolha Próximo.
7. Para a Etapa 3: Especifique o local de saída de dados — opcional, faça o seguinte.
   a. Para Destino de saída de dados, faça o seguinte.
      i. Escolha a localização do Amazon S3 para a saída de dados.
      ii. Em Criptografia, se você optar por Personalizar as configurações de criptografia, insira o ARN da AWS KMS chave ou escolha Criar uma AWS KMS chave.
   b. Visualize a saída LiveRamp gerada.
   c. Escolha Próximo.



8. Para a Etapa 4: revisar e criar, faça o seguinte.

   a. Revise as seleções feitas nas etapas anteriores e edite-as, se necessário.

   b. Escolha Criar.

      Uma mensagem aparece indicando que o fluxo de trabalho de mapeamento de ID foi criado.

Depois de criar o fluxo de trabalho de mapeamento de ID, você estará pronto para executar um fluxo de trabalho de mapeamento de ID.

Executar um fluxo de trabalho de mapeamento de ID

Depois de criar um fluxo de trabalho de mapeamento de ID para um Conta da AWS ou criar um fluxo de trabalho de mapeamento de ID em dois Contas da AWS, você pode executar o fluxo de trabalho de mapeamento de ID. O fluxo de trabalho de mapeamento de ID gera um arquivo CSV.

Para executar um fluxo de trabalho de mapeamento de ID

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.

2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Mapeamento de ID.
3. Escolha o fluxo de trabalho de mapeamento de ID.
4. Na página de detalhes do fluxo de trabalho de mapeamento de ID, no canto superior direito, escolha Executar.
5. Na página de detalhes do fluxo de trabalho correspondente, na guia Métricas, veja o seguinte em Métricas do último trabalho:

   - O Job ID
   - O tempo concluído para o trabalho de fluxo de trabalho
   - O status da tarefa de fluxo de trabalho correspondente: Em fila, em andamento, concluída, com falha
   - O número de registros processados
   - O número de registros não processados
   - O número de registros de entrada

   Em Histórico de tarefas, você também pode visualizar as métricas de tarefas de fluxo de trabalho de mapeamento de ID executadas anteriormente.
6. Após a conclusão do trabalho do fluxo de trabalho de mapeamento de ID (o status é Concluído), escolha Saída de dados e, em seguida, escolha sua localização no Amazon S3 para visualizar os resultados.

   Depois de obter seu arquivo CSV, você pode unir o. `RAMPID` com o. `TRANSCODED_ID`

Executando um fluxo de trabalho de mapeamento de ID com um novo destino de saída

Depois de criar um fluxo de trabalho de mapeamento de ID para um Conta da AWS ou criar um fluxo de trabalho de mapeamento de ID em dois Contas da AWS, você pode escolher um local diferente do S3 para gravar sua saída de dados.

Para executar um fluxo de trabalho de mapeamento de ID com um novo destino de saída

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Mapeamento de ID.

3. Escolha o fluxo de trabalho de mapeamento de ID.
4. Na página de detalhes do fluxo de trabalho de mapeamento de ID, no canto superior direito, escolha Executar com novo destino de saída na lista suspensa Executar fluxo de trabalho.
5. Para Destino de saída de dados, faça o seguinte.
   a. Escolha a localização do Amazon S3 para a saída de dados.
   b. Em Criptografia, se você optar por Personalizar as configurações de criptografia, insira o ARN da AWS KMS chave ou escolha Criar uma AWS KMS chave.
6. Para especificar as permissões de acesso ao serviço, escolha uma opção e execute a ação recomendada.

Opção	Ação recomendada
Criar e usar um novo perfil de serviço	• AWS Entity Resolution cria uma função de serviço com a política necessária para essa tabela. • O nome do perfil de serviço padrão é `entityresolution-id-mapping-workflow-<timestamp>` . • Você deve ter permissões para criar perfis e anexar políticas. • Se seus dados de entrada estiverem criptografados, escolha a opção Esses dados são criptografados por uma chave KMS. Em seguida, insira uma AWS KMS chave usada para descriptografar sua entrada de dados.
Use um perfil de serviço existente	1. Escolha um nome do perfil de serviço existente na lista suspensa. A lista de perfis é exibida se você tiver permissões para listar funções. Se você não tiver permissões para listar perfis, insira o nome do recurso da

Opção	Ação recomendada
	Amazon (ARN) do perfil que você deseja usar. Se não houver perfis de serviço existente s, a opção de Usar um perfil de serviço existente não estará disponível. 2. Para visualizar o perfil de serviço, selecione o link externo Visualizar no IAM. Por padrão, AWS Entity Resolution não tenta atualizar a política de função existente para adicionar as permissões necessárias.

7. Escolha Executar.
8. Na página de detalhes do fluxo de trabalho correspondente, na guia Métricas, veja o seguinte em Métricas do último trabalho:

   - O Job ID
   - O tempo concluído para o trabalho de fluxo de trabalho
   - O status da tarefa de fluxo de trabalho correspondente: Em fila, em andamento, concluída, com falha
   - O número de registros processados
   - O número de registros não processados
   - O número de registros de entrada

   Em Histórico de tarefas, você também pode visualizar as métricas de tarefas de fluxo de trabalho de mapeamento de ID executadas anteriormente.
9. Após a conclusão do trabalho do fluxo de trabalho de mapeamento de ID (o status é Concluído), escolha Saída de dados e, em seguida, escolha sua localização no Amazon S3 para visualizar os resultados.

   Depois de obter seu arquivo CSV, você pode unir o. `RAMPID` com o. `TRANSCODED_ID`

Editando um fluxo de trabalho de mapeamento de ID

A edição do fluxo de trabalho de mapeamento de ID permite que você mantenha seus recursos up-to-date de resolução de entidades alinhados às suas necessidades comerciais em evolução ao longo do tempo. Talvez você queira ajustar as regras, técnicas e parâmetros de mapeamento. Você pode otimizar o fluxo de trabalho para fornecer resultados de correspondência de ID mais precisos e confiáveis. Talvez você também queira adicionar novas fontes de dados, expandir os tipos de IDs mapeamento ou incorporar outros critérios de correspondência ao fluxo de trabalho. Se você identificar problemas ou erros nos resultados do mapeamento de ID, a edição com o fluxo de trabalho pode ajudá-lo a diagnosticar e resolver esses problemas.

Para editar um fluxo de trabalho de mapeamento de ID:

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Mapeamento de ID.
3. Escolha o fluxo de trabalho de mapeamento de ID.
4. Na página de detalhes do fluxo de trabalho de mapeamento de ID, no canto superior direito, escolha Editar.
5. Na página Especificar detalhes do fluxo de trabalho de mapeamento de ID, faça as alterações necessárias e escolha Avançar.
6. Na página Especificar saída de dados, faça as alterações necessárias e escolha Avançar.
7. Na página Revisar e salvar, faça as alterações necessárias e escolha Salvar.

Excluindo um fluxo de trabalho de mapeamento de ID

Se você não usa mais um fluxo de trabalho de mapeamento de ID, excluí-lo pode ajudar a simplificar o gerenciamento do fluxo de trabalho. Além disso, excluir fluxos de trabalho de mapeamento de ID redundantes ou menos eficientes que atendem a propósitos semelhantes pode ajudá-lo a consolidar seus processos.

Para excluir um fluxo de trabalho de mapeamento de ID:

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Mapeamento de ID.

3. Escolha o fluxo de trabalho de mapeamento de ID.
4. Na página de detalhes do fluxo de trabalho de mapeamento de ID, no canto superior direito, escolha Excluir.
5. Confirme a exclusão e escolha Excluir.

Adicionar ou atualizar uma política de recursos para um fluxo de trabalho de mapeamento de ID

Uma política de recursos permite que o criador do recurso de mapeamento de ID acesse seu recurso de fluxo de trabalho de mapeamento de ID.

Para adicionar ou atualizar uma política de recursos

1. Faça login no AWS Management Console e abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/.
2. No painel de navegação esquerdo, em Fluxos de trabalho, escolha Mapeamento de ID.
3. Escolha o fluxo de trabalho de mapeamento de ID.
4. Na página de detalhes do fluxo de trabalho de mapeamento de ID, escolha a guia Permissões.
5. Na seção Política de recursos, escolha Editar.
6. Adicione ou atualize a política no editor JSON.
7. Escolha Salvar alterações.

Integre-se AWS Entity Resolution como provedor

AWS Entity Resolution as integrações de fornecedores terceirizados ajudam os clientes a proteger a privacidade do consumidor e manter a conformidade com as leis de soberania de dados. Fornecedores terceirizados, como LiveRamp e TransUnion, traduzem identificadores de consumidores em publicidade IDs, como Ramp IDs e Fabrick. IDs Esses identificadores de publicidade são comumente usados em ferramentas de publicidade e marketing para impedir que os dados do consumidor sejam exportados para sistemas não AWS gerenciados. Esta seção fornece orientação para os provedores se integrarem para codificar ou transcodificar identificadores de consumidores em publicidade IDs para uso em um fluxo de trabalho de correspondência baseado em serviços de provedores. AWS Entity Resolution

Para obter mais informações sobre os serviços do provedor que estão atualmente integrados AWS Entity Resolution, consulteCriação de um fluxo de trabalho de correspondência baseado em serviços do provedor.

Tópicos

- Requisitos
- Usando a AWS Entity Resolution especificação OpenAPI
- Testando a integração de um provedor

Requisitos

Antes de se integrar como provedor de serviços com AWS Entity Resolution, preencha os seguintes requisitos.

Tópicos

- Listar um serviço de provedor em AWS Data Exchange
- Identifique seus atributos
- Solicite a AWS Entity Resolution especificação OpenAPI

Listar um serviço de provedor em AWS Data Exchange

Como fornecedor terceirizado, você deve publicar seu produto no catálogo de produtos do AWS Data Exchange (ADX). Depois que seu produto for listado no Catálogo de AWS Data Exchange Produtos, os assinantes poderão assinar seu produto por meio de uma oferta pública ou privada.

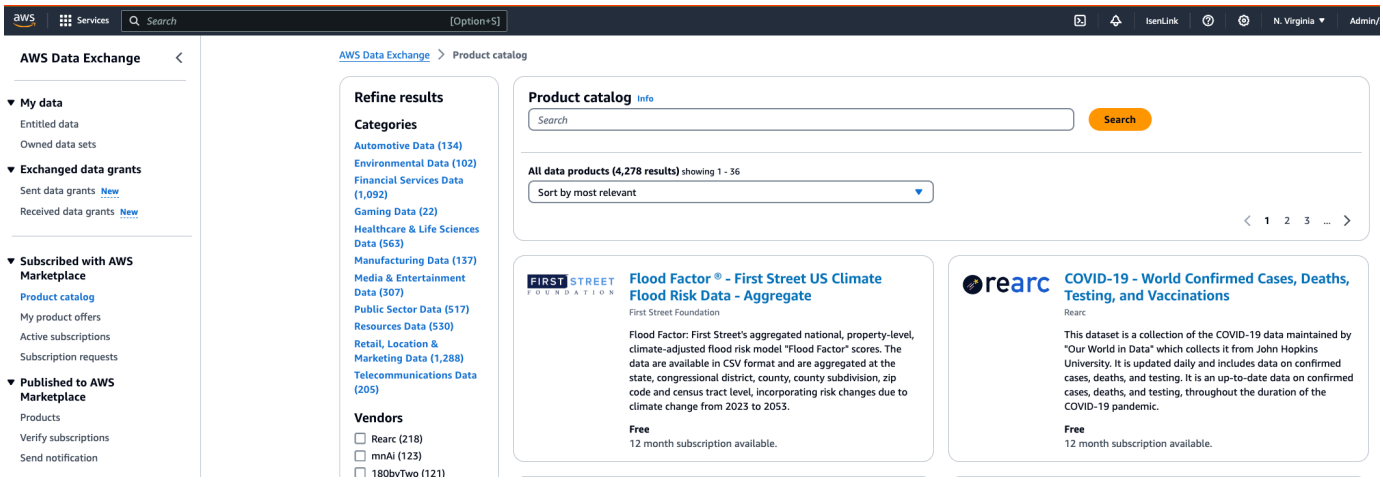
Para listar um serviço de provedor em AWS Data Exchange

1. Se você for um novo fornecedor de produtos de dados em AWS Data Exchange, conclua as etapas na seção intitulada Começando como provedor no Guia do AWS Data Exchange usuário.
2. Crie um conjunto de dados da API REST e publique um novo produto que contenha APIs on AWS Data Exchange seguindo as etapas na seção intitulada Como publicar um produto APIs contido no Guia do AWS Data Exchange usuário. Você pode concluir o processo usando o AWS Data Exchange console ou AWS Command Line Interface o.

   Se você definiu a visibilidade do produto como pública, a oferta pública estará disponível para todos os assinantes.

   Se você definiu a visibilidade do produto como Privada, conclua as etapas na seção intitulada Criar ofertas personalizadas no Guia AWS Data Exchange do usuário, dependendo do seu caso de uso.

   A imagem a seguir mostra um exemplo de um produto disponível no Catálogo de AWS Data Exchange produtos.



3. Depois que o produto estiver disponível no Catálogo de AWS Data Exchange Produtos, o assinante poderá assinar o produto das seguintes formas.

- Assine o produto público.
- Use uma oferta privada (oferta personalizada) emitida pelo serviço do provedor.
- Use uma oferta de assinatura Bring Your Own (BYOS).

Para obter mais informações, consulte Inscrever-se e acessar um produto APIs contido no Guia AWS Data Exchange do usuário.

Identifique seus atributos

Os atributos dos dados de entrada são as definições de tipo das entidades a serem resolvidas em um fluxo de trabalho. Alguns exemplos de atributos são`FirstName`,`LastName`,`Email`, ou`Custom String`.

Ao identificar seus atributos, você deve observar quaisquer requisitos ou diretrizes.

Example Exemplo

Veja a seguir um exemplo de validações para identificar os atributos do provedor.

- O `LastName` atributo `FirstName` ou é obrigatório.
- Se o `Email` atributo estiver presente, ele deverá ser codificado.

Como provedor, você deve identificar os atributos em seu produto de serviço de provedor e, em seguida, comunicá-los à equipe de desenvolvimento de AWS Entity Resolution negócios em `<aws-entity-resolution-bd@amazon .com>` para validação adicional antes de continuar.

Solicite a AWS Entity Resolution especificação OpenAPI

AWS Entity Resolution tem uma especificação OpenAPI que você, como provedor, pode usar como um handshake que contém os APIs envolvidos na integração. Para obter mais informações, consulte Usando a AWS Entity Resolution especificação OpenAPI.

Para solicitar a definição da OpenAPI, entre em contato com a equipe de desenvolvimento AWS Entity Resolution de negócios em `<aws-entity-resolution-bd@amazon>` .com.

Usando a AWS Entity Resolution especificação OpenAPI

A especificação OpenAPI define todos os protocolos associados a. AWS Entity Resolution Essa especificação é necessária para implementar a integração.

A definição de OpenAPI contém as seguintes operações de API:

- `POST AssignIdentities`
- `POST CreateJob`
- `GET GetJob`
- `POST StartJob`
- `POST MapIdentities`
- `GET Schema`

Para solicitar a especificação da OpenAPI, entre em contato com a equipe AWS Entity Resolution de desenvolvimento de negócios em `<aws-entity-resolution-bd@amazon>` .com.

A especificação OpenAPI suporta dois tipos de integrações para codificação e transcodificação de identificadores de consumidores, processamento em lote e processamento síncrono. Depois de obter a especificação OpenAPI, implemente o tipo de integração de processamento para seu caso de uso.

Tópicos

- Integração de processamento em lote
- Integração de processamento síncrono

Integração de processamento em lote

A integração do processamento em lote segue um padrão de design assíncrono. Depois que um fluxo de trabalho é iniciado AWS Data Exchange, ele envia um trabalho por meio de um endpoint de integração do provedor e, em seguida, o fluxo de trabalho aguarda a conclusão do trabalho consultando periodicamente o status do trabalho. Essa solução é mais desejável para execuções de trabalhos que podem levar mais tempo e ter uma taxa de transferência menor do provedor. O provedor receberá a localização do conjunto de dados como um link do Amazon S3, que pode ser processado por conta própria e gravar os resultados em um local predeterminado do S3 de saída.

A integração do processamento em lote é ativada usando três definições de API. AWS Entity Resolution chamará o endpoint do provedor que está disponível AWS Data Exchange na seguinte ordem:

1. `POST CreateJob`: essa operação de API envia as informações do trabalho ao provedor para processamento. Essas informações são sobre o tipo de trabalho; codificação ou transcodificação, localizações do S3, esquema fornecido pelo cliente e quaisquer propriedades adicionais de trabalho necessárias.

   Essa API retorna um`JobId`, e o Status do Job será um dos seguintes:`PENDING`,`READY`,`IN_PROGRESS`,`COMPLETE`, ou`FAILED`.

   Solicitação de amostra para codificação

```
POST /jobs
{
  "actionType": "ID_ASSIGNMENT",
  "s3SourceLocation": "string",
  "s3TargetLocation": "string",
  "jobProperties": {
    "assignmentJobProperties": {
      "fieldMappings": [
        {
          "name": "string",
          "type": "NAME"
        }
      ]
    }
  },
  "customerSpecifiedJobProperties": {
    "property1": "string",
    "property2": "string"
  },
  "outputSourceConfiguration": {
    "KMSArn": "string"
  }
}
```

   Exemplo de resposta

```
{
```

```
  "jobId": "string",
  "status": "PENDING"
}
```

2. `POST StartJob`: essa API permite que o provedor saiba como iniciar o trabalho com base no `JobId` fornecido. Isso permite que o provedor realize todas as validações necessárias de `CreateJob` até`StartJob`.

   Essa API retorna a`JobId`, the `Status` for the Jobstatus`Message`, the `statusCode` e.

   Solicitação de amostra para codificação

```
POST/jobs/{jobId}
{
  "customerSpecifiedJobProperties": {
    "property1": "string",
    "property2": "string"
  }
}
```

   Exemplo de resposta

```
{
  "jobId": "string",
  "status": "PENDING",
  "statusMessage": "string",
  "statusCode": 200
}
```

3. `GET GetJob`: essa API informa AWS Entity Resolution se o trabalho foi concluído ou se há algum outro status.

   Essa API retorna a`JobId`, the `Status` for the Jobstatus`Message`, the `statusCode` e.

   Solicitação de amostra para codificação

```
GET /jobs/{jobId}
```

   Exemplo de resposta

```
{
```

```
  "jobId": "string",
  "status": "PENDING",
  "statusMessage": "string",
  "statusCode": 200
}
```

A definição completa deles APIs é fornecida na especificação AWS Entity Resolution OpenAPI.

Integração de processamento síncrono

A solução de processamento síncrono é mais desejável para os provedores que têm um tempo de resposta quase em tempo real com tempo de resposta em tempo real com maior taxa de transferência e maior TPS. Esse AWS Entity Resolution fluxo de trabalho particiona o conjunto de dados e faz várias solicitações de API em paralelo. O AWS Entity Resolution fluxo de trabalho então gerencia a gravação dos resultados no local de saída desejado.

Esse processo é ativado usando uma das definições da API. AWS Entity Resolution chama o endpoint do provedor, que está disponível por meio de AWS Data Exchange:

`POST AssignIdentities`: essa API envia dados ao provedor usando um `source_id` identificador `recordFields` associado a esse registro.

Essa API retorna `assignedRecords` o.

Solicitação de amostra para codificação

```
POST /assignment
{
  "sourceRecords": [
    {
      "sourceId": "string",
      "recordFields": [
        {
          "name": "string",
          "type": "NAME",
          "value": "string"
        }
      ]
    }
  ]
}
```

Exemplo de resposta

```
{
  "assignedRecords": [
    {
      "sourceRecord": {
        "sourceId": "string",
        "recordFields": [
          {
            "name": "string",
            "type": "NAME",
            "value": "string"
          }
        ]
      },
      "identity": any
    }
  ]
}
```

A definição completa deles APIs é fornecida na especificação AWS Entity Resolution OpenAPI.

Dependendo da abordagem escolhida pelo provedor, AWS Entity Resolution criará uma configuração para esse provedor que será usada para iniciar a codificação ou transcodificação. Além disso, essas configurações estão disponíveis para os clientes que usam as APIs fornecidas pelo AWS Entity Resolution.

Essa configuração pode ser acessada usando um Amazon Resource Name (ARN), que é derivado de onde a oferta de serviços do provedor AWS Data Exchange está hospedada e do tipo de serviço do provedor. AWS Entity Resolution refere-se a esse ARN como o. `providerServiceARN`

Testando a integração de um provedor

Embora AWS Entity Resolution hospede serviços de correspondência de dados, a integração de um provedor é um componente terceirizado crucial para o fluxo de trabalho de end-to-end correspondência. Existem vários testes definidos para os provedores que adicionam uma proteção quando essa integração falha. AWS Entity Resolution Essa abordagem oferece uma oportunidade para os provedores monitorarem a integridade de seus serviços de acordo com esses casos end-to-end de teste.

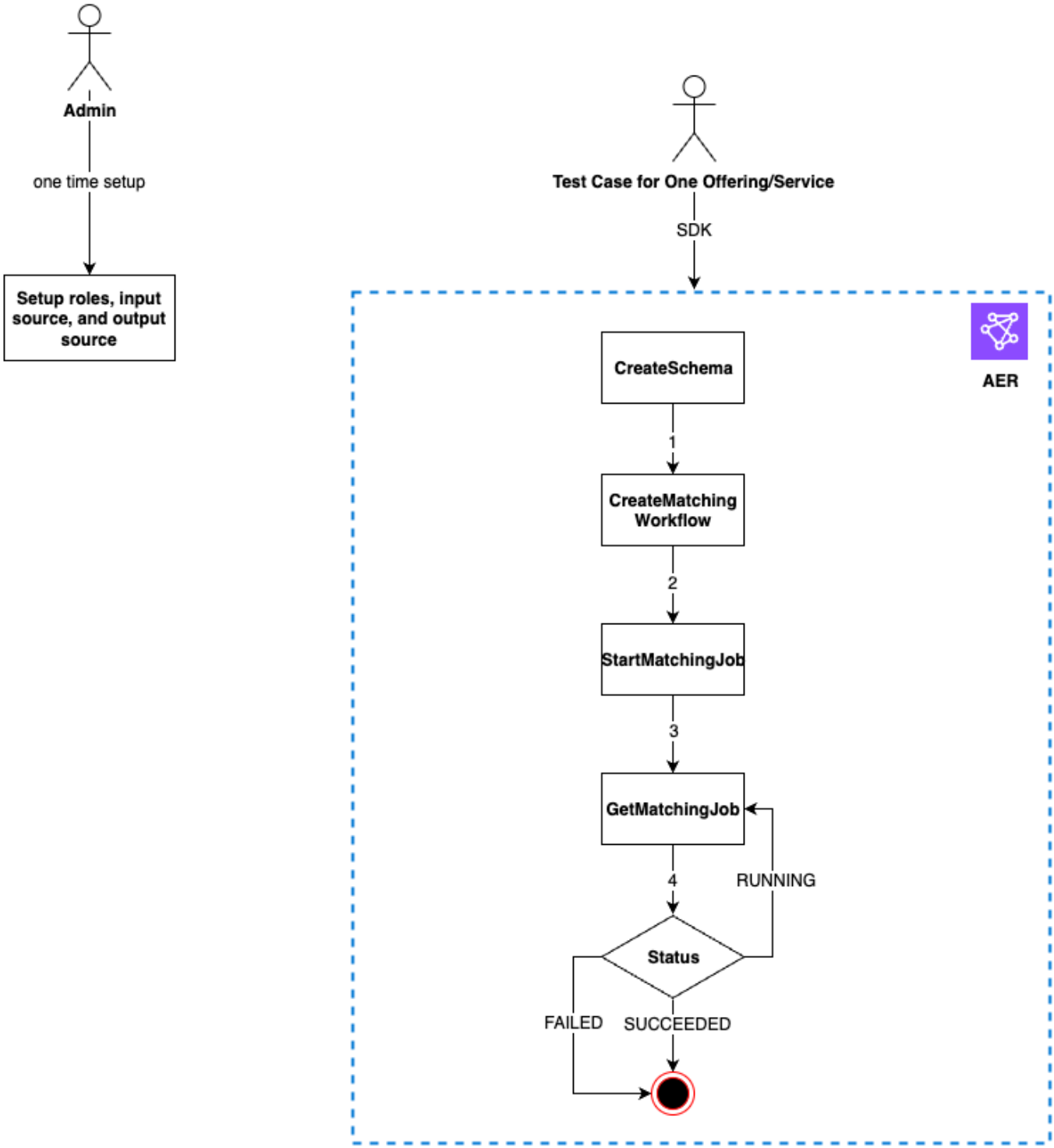
Os provedores podem usar suas contas de teste e seus próprios dados para executar esses casos de end-to-end teste usando o AWS Entity Resolution Software Development Kit (SDK). Se houver algum problema dos fornecedores, AWS Entity Resolution use o caminho de escalonamento preferido para escalar o problema. Além disso, os provedores precisam implementar seu próprio monitoramento dos resultados dos testes. Os provedores precisam compartilhar com Conta da AWS IDs quem estão acostumados a executar esses testes AWS Entity Resolution.

Uma execução bem-sucedida significa que um provedor pode configurar seus dados, usar seu próprio serviço e retornar o status do trabalho como Concluído sem erros. AWS Entity Resolution Isso pode ser feito programaticamente usando o APIs fornecido por. AWS Entity Resolution

Por exemplo, os provedores podem configurar seu bucket do S3, fonte de entrada, funções, esquema e fluxos de trabalho de acordo com seus serviços. Depois que essas configurações forem concluídas, os provedores poderão executar esses fluxos de trabalho uma vez por dia com 200 registros para testar seus serviços. Nessa abordagem, os provedores usam o SDK de sua escolha e realizam um end-to-end teste para os serviços oferecidos por meio do AWS Data Exchange uso de suas contas de teste. Espera-se que os provedores executem esses testes para cada uma de suas ofertas ou serviços.

> **Note**
>
> Os provedores precisam fornecer AWS Entity Resolution o Conta da AWS ID () `accountId`) que eles usam para executar esses fluxos de trabalho para testes. Além disso, os provedores precisam monitorar esses testes e garantir que eles sejam aprovados, o que significa que os provedores precisam habilitar a notificação em caso de falhas e resolver o problema adequadamente.

O diagrama a seguir mostra um caso típico end-to-end de teste de fluxo de trabalho.



Para testar a integração de um provedor

1. (Configuração única) Configure recursos para AWS Entity Resolution seguindo os procedimentos emConfigurar AWS Entity Resolution.

Depois de concluir os procedimentos de configuração únicos, você deverá ter suas funções, dados e fonte de dados prontos. Agora você está pronto para testar a integração do provedor usando o AWS Entity Resolution console ou APIs.

2. Teste a integração do provedor usando o console AWS Entity Resolution APIs ou.

API

Para testar a integração de um provedor usando o AWS Entity Resolution APIs

1. Crie um mapeamento de esquema usando a CreateSchemaMapping API. Para obter uma lista completa das linguagens de programação compatíveis, consulte a seção Consulte também da CreateSchemaMapping API.

   O mapeamento do esquema é o processo pelo qual você explica AWS Entity Resolution como interpretar seus dados para fins de correspondência. Você define o esquema da tabela de dados de entrada que deseja que o AWS Entity Resolution leia em um fluxo de trabalho correspondente.

   Ao criar um mapeamento de esquema, um identificador exclusivo deve ser designado e atribuído a cada linha de dados de entrada que o AWS Entity Resolution lê. Por exemplo: `Primary_key`, `Row_ID`, `Record_ID`.

   Example Criação de um mapeamento de esquema para fonte de dados contendo **`id`** e **`email`**

   Veja a seguir um exemplo de mapeamento de esquema para uma fonte de dados que contém `id` e`email`:

```
[
  {
    "fieldName": "id",
    "type": "UNIQUE_ID"
  },
  {
    "fieldName": "email",
    "type": "EMAIL_ADDRESS"
  }
]
```

Example Criação de um mapeamento de esquema para fonte de dados contendo **`id`** e **`email`** usando o Java SDK

Veja a seguir um exemplo de mapeamento de esquema para uma fonte de dados que contém `id` e `email` usa o Java SDK:

```
EntityResolutionClient.createSchemaMapping(
            CreateSchemaMappingRequest.builder()
                                        .schemaName(<schema-name>)
                                        .mappedInputFields([

 SchemaInputAttribute.builder().fieldName("id").type("UNIQUE_ID").build(),

 SchemaInputAttribute.builder().fieldName("email").type("EMAIL_ADDRESS").build()
                                         ])
                                        .build()

)
```

2. Crie um fluxo de trabalho correspondente usando a CreateMatchingWorkflow API. Para obter uma lista completa das linguagens de programação compatíveis, consulte a seção Consulte também da CreateMatchingWorkflow API.

Example Criação de um fluxo de trabalho correspondente usando o Java SDK

Veja a seguir um exemplo de um fluxo de trabalho correspondente usando o Java SDK:

```
EntityResolutionClient.createMatchingWorkflow(
            CreateMatchingWorkflowRequest.builder()
                                         .workflowName(<workflow-name>)
                                         .inputSourceConfig(

 InputSource.builder().inputSourceARN(<glue-inputsource-from-
step1>).schemaName(<schema-name-from-step2>).build()
                                         )

 .outputSourceConfig(OutputSource.builder().outputS3Path(<output-s3-
path>).output(<output-1>, <output-2>, <output-3>).build())

 .resolutionTechniques(ResolutionTechniques.builder()

   .resolutionType(PROVIDER)
```

```
   .providerProperties(ProviderProperties.builder()
                                          .providerServiceArn(<provider-arn>)
                                           .providerConfiguration(<configuration-
depending-on-service>)

 .intermediateSourceConfiguration(<intermedaite-s3-path>)
                                          .build())
   .build()
                                          .roleArn(<role-from-step1>)
                                          .build()

)
```

Depois que o fluxo de trabalho correspondente for configurado, você poderá executar um fluxo de trabalho.

3. Execute um fluxo de trabalho correspondente usando a StartMatchingJob API. Para executar um fluxo de trabalho correspondente, você deve ter criado um fluxo de trabalho correspondente usando o `CreateMatchingWorkflow` endpoint.

   Para obter uma lista completa das linguagens de programação compatíveis, consulte a seção Consulte também da StartMatchingJob API.

   Example Executando um fluxo de trabalho correspondente usando o Java SDK

   Veja a seguir um exemplo de um fluxo de trabalho correspondente em execução usando o Java SDK:

   ```
   EntityResolutionClient.startMatchingJob(StartMatchingJobRequest.builder()
                   .workflowName(<name-of-workflow-from-step3)
                   .build()
   )
   ```

4. Monitore o status de um fluxo de trabalho usando a GetMatchingJob API.

   Essa API retorna o status, as métricas e os erros (se houver) associados a um trabalho.

Example Monitorando um fluxo de trabalho correspondente usando o Java SDK

Veja a seguir um exemplo de monitoramento de um trabalho de fluxo de trabalho correspondente usando o Java SDK:

```
EntityResolutionClient.getMatchingJob(GetMatchingJobRequest.builder()
                .workflowName(<name-of-workflow-from-step3)
                .jobId(jobId-from-startMatchingJob)
                .build()
)
```

O end-to-end teste será concluído se o fluxo de trabalho for concluído com êxito.

Console

Para testar a integração de um provedor usando o AWS Entity Resolution console

1. Crie um mapeamento de esquema seguindo as etapas emCriação de um mapeamento de esquema.

   O mapeamento do esquema é o processo pelo qual você explica AWS Entity Resolution como interpretar seus dados para fins de correspondência. Você define o esquema da tabela de dados de entrada que AWS Entity Resolution deseja ler em um fluxo de trabalho correspondente.

   Ao criar um mapeamento de esquema, um identificador exclusivo deve ser designado e atribuído a cada linha de dados de entrada AWS Entity Resolution lida. Por exemplo: `Primary_key`, `Row_ID`, `Record_ID`.

   Example Mapeamento de esquema para fonte de dados contendo **`id`** e **`email`**

   Veja a seguir um exemplo de mapeamento de esquema para uma fonte de dados que contém `id` e`email`:

```
[
  {
    "fieldName": "id",
    "type": "UNIQUE_ID"
  },
```

```
  {
    "fieldName": "email",
    "type": "EMAIL_ADDRESS"
  }
]
```

2. Crie e execute o fluxo de trabalho correspondente seguindo as etapas emCriação de um fluxo de trabalho de correspondência baseado em serviços do provedor.

   Criar um fluxo de trabalho correspondente é o processo que você configura para especificar os dados de entrada a serem combinados e como a correspondência deve ser realizada. No fluxo de trabalho baseado em provedor, se uma conta tiver uma assinatura com um serviço de provedor AWS Data Exchange, você poderá combinar seus identificadores conhecidos com seu provedor preferido. Dependendo do provedor e do serviço que você está usando para realizar um teste de ponta a ponta, você pode configurar seu fluxo de trabalho correspondente adequadamente.

   O AWS Entity Resolution console combina as ações de criar e executar em um único botão. Depois de selecionar Criar e executar, aparece uma mensagem indicando que o fluxo de trabalho correspondente foi criado e que o trabalho foi iniciado.

3. Monitore o status do fluxo de trabalho na página Fluxos de trabalho correspondentes.

   O end-to-end teste será concluído se o fluxo de trabalho for concluído com êxito (o status do trabalho é Concluído).

   Na guia Métricas da página de detalhes do fluxo de trabalho correspondente, você pode ver o seguinte em Métricas do último trabalho:

   - O Job ID.
   - O status da tarefa de fluxo de trabalho correspondente: Em fila, em andamento, concluída, com falha
   - O tempo concluído para o trabalho do fluxo de trabalho.
   - O número de registros processados.
   - O número de registros não processados.
   - A partida única IDs gerada.
   - O número de registros de entrada.

Você também pode visualizar as métricas de trabalho para trabalhos de fluxo de trabalho correspondentes que foram executados anteriormente no Histórico de trabalhos.

Segurança em AWS Entity Resolution

A segurança na nuvem AWS é a maior prioridade. Como AWS cliente, você se beneficia de data centers e arquiteturas de rede criados para atender aos requisitos das organizações mais sensíveis à segurança.

A segurança é uma responsabilidade compartilhada entre você AWS e você. O modelo de responsabilidade compartilhada descreve isso como segurança da nuvem e segurança na nuvem:

- Segurança da nuvem — AWS é responsável por proteger a infraestrutura que funciona Serviços da AWS no Nuvem AWS. AWS também fornece serviços que você pode usar com segurança. Auditores terceirizados testam e verificam regularmente a eficácia de nossa segurança como parte dos Programas de Conformidade Programas de AWS de . Para saber mais sobre os programas de conformidade aplicáveis AWS Entity Resolution, consulte AWS Serviços no escopo do programa de conformidade AWS .
- Segurança na nuvem — Sua responsabilidade é determinada pelo AWS service (Serviço da AWS) que você usa. Você também é responsável por outros fatores, incluindo a confidencialidade de seus dados, os requisitos da empresa e as leis e regulamentos aplicáveis.

Esta documentação ajuda você a entender como aplicar o modelo de responsabilidade compartilhada ao usar AWS Entity Resolution. Os tópicos a seguir mostram como configurar para atender AWS Entity Resolution aos seus objetivos de segurança e conformidade. Você também aprenderá a usar outros Serviços da AWS que o ajudem a monitorar e proteger seus AWS Entity Resolution recursos.

Tópicos

- Proteção de dados em AWS Entity Resolution
- Gerenciamento de identidade e acesso para AWS Entity Resolution
- Validação de conformidade para AWS Entity Resolution
- Resiliência em AWS Entity Resolution

Proteção de dados em AWS Entity Resolution

O modelo de responsabilidade AWS compartilhada modelo se aplica à proteção de dados em AWS Entity Resolution. Conforme descrito neste modelo, AWS é responsável por proteger a

infraestrutura global que executa todos os Nuvem AWS. Você é responsável por manter o controle sobre o conteúdo hospedado nessa infraestrutura. Você também é responsável pelas tarefas de configuração e gerenciamento de segurança dos Serviços da AWS que usa. Para obter mais informações sobre a privacidade de dados, consulte as Data Privacy FAQ. Para obter mais informações sobre a proteção de dados na Europa, consulte a postagem do blog AWS Shared Responsibility Model and RGPD no Blog de segurança da AWS .

Para fins de proteção de dados, recomendamos que você proteja Conta da AWS as credenciais e configure usuários individuais com AWS IAM Identity Center ou AWS Identity and Access Management (IAM). Dessa maneira, cada usuário receberá apenas as permissões necessárias para cumprir suas obrigações de trabalho. Recomendamos também que você proteja seus dados das seguintes formas:

- Use uma autenticação multifator (MFA) com cada conta.
- Use SSL/TLS para se comunicar com AWS os recursos. Exigimos TLS 1.2 e recomendamos TLS 1.3.
- Configure a API e o registro de atividades do usuário com AWS CloudTrail. Para obter informações sobre o uso de CloudTrail trilhas para capturar AWS atividades, consulte Como trabalhar com CloudTrail trilhas no Guia AWS CloudTrail do usuário.
- Use soluções de AWS criptografia, juntamente com todos os controles de segurança padrão Serviços da AWS.
- Use serviços gerenciados de segurança avançada, como o Amazon Macie, que ajuda a localizar e proteger dados sigilosos armazenados no Amazon S3.
- Se você precisar de módulos criptográficos validados pelo FIPS 140-3 ao acessar AWS por meio de uma interface de linha de comando ou de uma API, use um endpoint FIPS. Para obter mais informações sobre os endpoints FIPS disponíveis, consulte Federal Information Processing Standard (FIPS) 140-3.

É altamente recomendável que nunca sejam colocadas informações confidenciais ou sigilosas, como endereços de e-mail de clientes, em tags ou campos de formato livre, como um campo Nome. Isso inclui quando você trabalha com AWS Entity Resolution ou Serviços da AWS usa o console, a API ou AWS SDKs. AWS CLI Quaisquer dados inseridos em tags ou em campos de texto de formato livre usados para nomes podem ser usados para logs de faturamento ou de diagnóstico. Se você fornecer um URL para um servidor externo, é fortemente recomendável que não sejam incluídas informações de credenciais no URL para validar a solicitação nesse servidor.

Criptografia de dados em repouso para AWS Entity Resolution

AWS Entity Resolution fornece criptografia por padrão para proteger dados confidenciais do cliente em repouso usando chaves AWS de criptografia próprias.

Chaves de propriedade da AWS — AWS Entity Resolution usa essas chaves por padrão para criptografar automaticamente dados de identificação pessoal. Você não pode visualizar, gerenciar ou usar chaves de propriedade da AWS nem auditar seu uso. No entanto, você não precisa realizar nenhuma ação para proteger as chaves que criptografam seus dados. Para obter mais informações, consulte AWS owned keys no Guia do desenvolvedor do AWS Key Management Service .

Por padrão, a criptografia de dados em repouso ajuda a reduzir a sobrecarga operacional e a complexidade envolvidas na proteção de dados confidenciais. Ao mesmo tempo, você pode usá-lo para criar aplicativos seguros que atendam aos rigorosos requisitos regulamentares e de conformidade de criptografia.

Como alternativa, você também pode fornecer uma chave KMS gerenciada pelo cliente para criptografia ao criar seu recurso de fluxo de trabalho correspondente.

Chaves gerenciadas pelo cliente — AWS Entity Resolution suporta o uso de uma chave KMS simétrica gerenciada pelo cliente que você cria, possui e gerencia para permitir a criptografia de seus dados confidenciais. Como você tem controle total dessa camada de criptografia, você pode realizar tarefas como:

- Estabelecer e manter as políticas de chave
- Estabelecer e manter subsídios e IAM policies
- Habilitar e desabilitar políticas de chaves
- Alternar os materiais de criptografia de chave
- Adicionar etiquetas
- Criar réplicas de chaves
- Programar chaves para exclusão

Para obter mais informações, consulte a chave gerenciada pelo cliente no Guia do AWS Key Management Service desenvolvedor.

Para obter mais informações sobre AWS KMS, consulte O que é o AWS Key Management Service?

Gerenciamento de chaves

Como AWS Entity Resolution usa subsídios em AWS KMS

AWS Entity Resolution exige uma concessão para usar sua chave gerenciada pelo cliente. Quando você cria um fluxo de trabalho correspondente criptografado com uma chave gerenciada pelo cliente, AWS Entity Resolution cria uma concessão em seu nome enviando uma CreateGrantsolicitação para AWS KMS. As concessões AWS KMS são usadas para dar AWS Entity Resolution acesso a uma chave KMS em uma conta de cliente. AWS Entity Resolution exige que a concessão use sua chave gerenciada pelo cliente para as seguintes operações internas:

- Envie GenerateDataKeysolicitações AWS KMS para gerar chaves de dados criptografadas pela chave gerenciada pelo cliente.
- Envie solicitações de descriptografia para AWS KMS descriptografar as chaves de dados criptografadas para que elas possam ser usadas para criptografar seus dados.

É possível revogar o acesso à concessão, ou remover o acesso do serviço à chave gerenciada pelo cliente a qualquer momento. Se você fizer isso, AWS Entity Resolution não conseguirá acessar nenhum dos dados criptografados pela chave gerenciada pelo cliente, o que afeta as operações que dependem desses dados. Por exemplo, se você remover o acesso ao serviço à sua chave por meio da concessão e tentar iniciar um trabalho para um fluxo de trabalho correspondente criptografado com uma chave de cliente, a operação retornará um `AccessDeniedException` erro.

Criar uma chave gerenciada pelo cliente

Você pode criar uma chave simétrica gerenciada pelo cliente usando o AWS Management Console, ou o. AWS KMS APIs

Para criar uma chave simétrica gerenciada pelo cliente

AWS Entity Resolution suporta criptografia usando chaves KMS de criptografia simétrica. Siga as etapas de Criar uma chave simétrica gerenciada pelo cliente no Guia do desenvolvedor do AWS Key Management Service .

Declaração de política chave

As políticas de chaves controlam o acesso à chave gerenciada pelo cliente. Cada chave gerenciada pelo cliente deve ter exatamente uma política de chaves, que contém declarações que determinam quem pode usar a chave e como pode usá-la. Ao criar a chave gerenciada pelo cliente, você pode

especificar uma política de chaves. Para obter mais informações, consulte Gerenciamento do acesso às chaves gerenciadas pelo cliente no Guia do AWS Key Management Service desenvolvedor.

Para usar sua chave gerenciada pelo cliente com seus AWS Entity Resolution recursos, as seguintes operações de API devem ser permitidas na política de chaves:

- `kms:DescribeKey`— fornece informações como o ARN da chave, a data de criação (e a data de exclusão, se aplicável), o estado da chave e a data de origem e expiração (se houver) do material da chave. Ele inclui campos, como`KeySpec`, que ajudam você a distinguir diferentes tipos de chaves KMS. Ele também exibe o uso da chave (criptografia, assinatura ou geração e verificação MACs) e os algoritmos que a chave KMS suporta. AWS Entity Resolution valida que `KeySpec` é `SYMMETRIC_DEFAULT` e `KeyUsage` é`ENCRYPT_DECRYPT`.
- `kms:CreateGrant`: adiciona uma concessão a uma chave gerenciada pelo cliente. Concede acesso de controle a uma chave KMS especificada, o que permite o acesso AWS Entity Resolution necessário às operações de concessão. Para obter mais informações sobre Utilizar concessões, consulte o Guia do desenvolvedor do AWS Key Management Service .

Isso permite AWS Entity Resolution fazer o seguinte:

- Ligar para `GenerateDataKey` para gerar uma chave de dados criptografada e armazená-la, porque a chave de dados não é usada imediatamente para criptografar.
- Ligar para `Decrypt` para usar a chave de dados criptografada armazenada para acessar os dados criptografados.
- Configure uma entidade principal aposentada para permitir que o serviço para `RetireGrant`.

Veja a seguir exemplos de declarações de política que você pode adicionar para AWS Entity Resolution:

```
{
      "Sid" : "Allow access to principals authorized to use AWS Entity Resolution",
      "Effect" : "Allow",
      "Principal" : {
        "AWS" : "*"
      },
      "Action" : ["kms:DescribeKey","kms:CreateGrant"],
      "Resource" : "*",
      "Condition" : {
        "StringEquals" : {
```

```
          "kms:ViaService" : "entityresolution.region.amazonaws.com",
          "kms:CallerAccount" : "111122223333"
        }
      }
}
```

Permissões para usuários

Quando você configura uma chave KMS como a chave padrão para criptografia, a política de chave KMS padrão permite que qualquer usuário com acesso às ações necessárias do KMS use essa chave KMS para criptografar ou descriptografar recursos. Você deve conceder permissão aos usuários para executar as seguintes ações para usar a criptografia de chave KMS gerenciada pelo cliente:

- `kms:CreateGrant`
- `kms:Decrypt`
- `kms:DescribeKey`
- `kms:GenerateDataKey`

Durante uma `CreateMatchingWorkflow`solicitação, AWS Entity Resolution enviará uma DescribeKeye uma CreateGrantsolicitação para AWS KMS em seu nome. Isso exigirá que a entidade do IAM que faz a `CreateMatchingWorkflow` solicitação com uma chave KMS gerenciada pelo cliente tenha as `kms:DescribeKey` permissões na política de chaves do KMS.

Durante uma `StartIdMappingJob`solicitação `CreateIdMappingWorkflow`e, AWS Entity Resolution enviará uma CreateGrantsolicitação DescribeKeye uma para AWS KMS em seu nome. Isso exigirá que a entidade do IAM que faz a `StartIdMappingJob` solicitação `CreateIdMappingWorkflow` e com uma chave KMS gerenciada pelo cliente tenha as `kms:DescribeKey` permissões na política de chaves do KMS. Os provedores poderão acessar a chave gerenciada pelo cliente para descriptografar os dados no bucket do Amazon S3 AWS Entity Resolution .

A seguir estão exemplos de declarações de política que você pode adicionar para que os provedores descriptografem os dados no bucket do Amazon S3: AWS Entity Resolution

JSON

```
{
```

```
    "Version": "2012-10-17",
    "Statement": [{
        "Effect": "Allow",
        "Principal": {
            "AWS": "arn:aws:iam::715724997226:root"
        },
        "Action": [
            "kms:Decrypt"
        ],
        "Resource": "<KMSKeyARN>",
        "Condition": {
            "StringEquals": {
                "kms:ViaService": "s3.amazonaws.com"
            }
        }
    }]
}
```

Substitua cada *<user input placeholder>* por suas próprias informações.

<KMSKeyARN>	AWS KMS Nome do recurso da Amazon.

Da mesma forma, a entidade do IAM que invoca a `StartMatchingJobAPI` não deve ter `kms:Decrypt` nenhuma `kms:GenerateDataKey` permissão na chave KMS gerenciada pelo cliente fornecida no fluxo de trabalho correspondente.

Para obter mais informações sobre a especificação de permissões em uma política, consulte o Guia do AWS Key Management Service desenvolvedor.

Para obter mais informações sobre como solucionar problemas de acesso por chave, consulte o Guia do AWS Key Management Service desenvolvedor.

Especificando uma chave gerenciada pelo cliente para AWS Entity Resolution

Você pode especificar uma chave gerenciada pelo cliente para fornecer uma segunda camada de criptografia para os seguintes recursos:

Fluxo de trabalho correspondente — Ao criar um recurso de fluxo de trabalho correspondente, você pode especificar a chave de dados inserindo uma KMSArn, que é AWS Entity Resolution usada para criptografar os dados pessoais identificáveis armazenados pelo recurso.

KMSArn— Insira um ARN de chave, que é um identificador de chave para uma chave gerenciada pelo AWS KMS cliente.

Você pode especificar uma chave gerenciada pelo cliente como uma criptografia de segunda camada para os seguintes recursos se estiver criando ou executando um fluxo de trabalho de mapeamento de ID em dois Contas da AWS:

Fluxo de trabalho de mapeamento de ID ou Iniciar fluxo de trabalho de mapeamento de ID — Ao criar um recurso de fluxo de trabalho de mapeamento de ID ou iniciar um trabalho de fluxo de trabalho de mapeamento de ID KMSArn, você pode especificar a chave de dados inserindo um, que AWS Entity Resolution usa para criptografar os dados pessoais identificáveis armazenados pelo recurso.

KMSArn— Insira um ARN de chave, que é um identificador de chave para uma chave gerenciada pelo AWS KMS cliente.

Monitorando suas chaves de criptografia para o AWS Entity Resolution serviço

Ao usar uma chave gerenciada pelo AWS KMS cliente com seus recursos AWS Entity Resolution de serviço, você pode usar a AWS CloudTrail ou a Amazon CloudWatch Logs para rastrear solicitações AWS Entity Resolution enviadas para AWS KMS.

Os exemplos a seguir são AWS CloudTrail eventos para`CreateGrant`, `GenerateDataKeyDecrypt`, e `DescribeKey` para monitorar AWS KMS operações chamadas por AWS Entity Resolution para acessar dados criptografados pela chave gerenciada pelo cliente:

Tópicos

- CreateGrant
- DescribeKey
- GenerateDataKey
- Decrypt

CreateGrant

Quando você usa uma chave gerenciada pelo AWS KMS cliente para criptografar seu recurso de fluxo de trabalho correspondente, AWS Entity Resolution envia uma `CreateGrant` solicitação em seu nome para acessar a chave KMS no seu. Conta da AWS A concessão AWS Entity Resolution criada é específica para o recurso associado à chave gerenciada pelo AWS KMS cliente. Além disso,

AWS Entity Resolution usa a `RetireGrant` operação para remover uma concessão quando você exclui um recurso.

O evento de exemplo a seguir registra a operação `CreateGrant`:


```
{
    "eventVersion": "1.08",
    "userIdentity": {
        "type": "AssumedRole",
        "principalId": "AROAIGDTESTANDEXAMPLE:Sampleuser01",
        "arn": "arn:aws:sts::111122223333:assumed-role/Admin/Sampleuser01",
        "accountId": "111122223333",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE3",
        "sessionContext": {
            "sessionIssuer": {
                "type": "Role",
                "principalId": "AROAIGDTESTANDEXAMPLE:Sampleuser01",
                "arn": "arn:aws:sts::111122223333:assumed-role/Admin/Sampleuser01",
                "accountId": "111122223333",
                "userName": "Admin"
            },
            "webIdFederationData": {},
            "attributes": {
                "mfaAuthenticated": "false",
                "creationDate": "2021-04-22T17:02:00Z"
            }
        },
        "invokedBy": "entityresolution.amazonaws.com"
    },
    "eventTime": "2021-04-22T17:07:02Z",
    "eventSource": "kms.amazonaws.com",
    "eventName": "CreateGrant",
    "awsRegion": "us-west-2",
    "sourceIPAddress": "172.12.34.56",
    "userAgent": "ExampleDesktop/1.0 (V1; OS)",
    "requestParameters": {
        "retiringPrincipal": "entityresolution.region.amazonaws.com",
        "operations": [
            "GenerateDataKey",
            "Decrypt",
        ],
        "keyId": "arn:aws:kms:us-
west-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE",
```

```
        "granteePrincipal": "entityresolution.region.amazonaws.com"
    },
    "responseElements": {
        "grantId":
 "0ab0ac0d0b000f00ea00cc0a0e00fc00bce000c000f0000000c0bc0a0000aaafSAMPLE",
        "keyId": "arn:aws:kms:us-
west-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE",
    },
    "requestID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "readOnly": false,
    "resources": [
        {
            "accountId": "111122223333",
            "type": "AWS::KMS::Key",
            "ARN": "arn:aws:kms:us-
west-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
        }
    ],
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "eventCategory": "Management",
    "recipientAccountId": "111122223333"
}
```

DescribeKey

AWS Entity Resolution usa a `DescribeKey` operação para verificar se a chave gerenciada pelo AWS KMS cliente associada ao seu recurso correspondente existe na conta e na região.

O evento de exemplo a seguir registra a operação `DescribeKey`:

```
{
    "eventVersion": "1.08",
    "userIdentity": {
        "type": "AssumedRole",
        "principalId": "AROAIGDTESTANDEXAMPLE:Sampleuser01",
        "arn": "arn:aws:sts::111122223333:assumed-role/Admin/Sampleuser01",
        "accountId": "111122223333",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE3",
        "sessionContext": {
            "sessionIssuer": {
                "type": "Role",
```

```
                "principalId": "AROAIGDTESTANDEXAMPLE:Sampleuser01",
                "arn": "arn:aws:sts::111122223333:assumed-role/Admin/Sampleuser01",
                "accountId": "111122223333",
                "userName": "Admin"
            },
            "webIdFederationData": {},
            "attributes": {
                "mfaAuthenticated": "false",
                "creationDate": "2021-04-22T17:02:00Z"
            }
        },
        "invokedBy": "entityresolution.amazonaws.com"
    },
    "eventTime": "2021-04-22T17:07:02Z",
    "eventSource": "kms.amazonaws.com",
    "eventName": "DescribeKey",
    "awsRegion": "us-west-2",
    "sourceIPAddress": "172.12.34.56",
    "userAgent": "ExampleDesktop/1.0 (V1; OS)",
    "requestParameters": {
        "keyId": "00dd0db0-0000-0000-ac00-b0c000SAMPLE"
    },
    "responseElements": null,
    "requestID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "readOnly": true,
    "resources": [
        {
            "accountId": "111122223333",
            "type": "AWS::KMS::Key",
            "ARN": "arn:aws:kms:us-
west-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
        }
    ],
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "eventCategory": "Management",
    "recipientAccountId": "111122223333"
}
```

GenerateDataKey

Quando você habilita uma chave gerenciada pelo AWS KMS cliente para seu recurso de fluxo de trabalho correspondente, AWS Entity Resolution envia uma `GenerateDataKey` solicitação por meio do Amazon Simple Storage Service (Amazon S3) AWS KMS para especificar a chave gerenciada AWS KMS pelo cliente para o recurso.

O evento de exemplo a seguir registra a operação `GenerateDataKey`:


```
{
    "eventVersion": "1.08",
    "userIdentity": {
        "type": "AWSService",
        "invokedBy": "s3.amazonaws.com"
    },
    "eventTime": "2021-04-22T17:07:02Z",
    "eventSource": "kms.amazonaws.com",
    "eventName": "GenerateDataKey",
    "awsRegion": "us-west-2",
    "sourceIPAddress": "172.12.34.56",
    "userAgent": "ExampleDesktop/1.0 (V1; OS)",
    "requestParameters": {
        "keySpec": "AES_256",
        "keyId": "arn:aws:kms:us-
west-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
    },
    "responseElements": null,
    "requestID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "readOnly": true,
    "resources": [
        {
            "accountId": "111122223333",
            "type": "AWS::KMS::Key",
            "ARN": "arn:aws:kms:us-
west-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
        }
    ],
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "eventCategory": "Management",
    "recipientAccountId": "111122223333",
    "sharedEventID": "57f5dbee-16da-413e-979f-2c4c6663475e"
```

```
}
```

Decrypt

Quando você habilita uma chave gerenciada pelo AWS KMS cliente para seu recurso de fluxo de trabalho correspondente, AWS Entity Resolution envia uma `Decrypt` solicitação por meio do Amazon Simple Storage Service (Amazon S3) AWS KMS para especificar a chave gerenciada AWS KMS pelo cliente para o recurso.

O evento de exemplo a seguir registra a operação `Decrypt`:

```
{
    "eventVersion": "1.08",
    "userIdentity": {
        "type": "AWSService",
        "invokedBy": "s3.amazonaws.com"
    },
    "eventTime": "2021-04-22T17:10:51Z",
    "eventSource": "kms.amazonaws.com",
    "eventName": "Decrypt",
    "awsRegion": "us-west-2",
    "sourceIPAddress": "172.12.34.56",
    "userAgent": "ExampleDesktop/1.0 (V1; OS)",
    "requestParameters": {
        "keyId": "arn:aws:kms:us-
west-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE",
        "encryptionAlgorithm": "SYMMETRIC_DEFAULT"
    },
    "responseElements": null,
    "requestID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "readOnly": true,
    "resources": [
        {
            "accountId": "111122223333",
            "type": "AWS::KMS::Key",
            "ARN": "arn:aws:kms:us-
west-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
        }
    ],
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "eventCategory": "Management",
```

```
    "recipientAccountId": "111122223333",
    "sharedEventID": "dc129381-1d94-49bd-b522-f56a3482d088"
}
```

Considerações

AWS Entity Resolution não oferece suporte à atualização de um fluxo de trabalho correspondente com uma nova chave KMS gerenciada pelo cliente. Nesses casos, você pode criar um novo fluxo de trabalho com a chave KMS gerenciada pelo cliente.

Saiba mais

Os recursos a seguir fornecem mais informações sobre a criptografia de dados em repouso.

Para obter mais informações sobre os conceitos básicos do AWS Key Management Service, consulte o Guia do AWS Key Management Service desenvolvedor.

Para obter mais informações sobre as melhores práticas de segurança do AWS Key Management Service, consulte o Guia do AWS Key Management Service desenvolvedor.

Acesso AWS Entity Resolution usando um endpoint de interface ()AWS PrivateLink

Você pode usar AWS PrivateLink para criar uma conexão privada entre sua VPC e. AWS Entity Resolution Você pode acessar AWS Entity Resolution como se estivesse em sua VPC, sem o uso de um gateway de internet, dispositivo NAT, conexão VPN ou conexão. AWS Direct Connect As instâncias na sua VPC não precisam de endereços IP públicos para acessar o AWS Entity Resolution.

Estabeleça essa conectividade privada criando um endpoint de interface, habilitado pelo AWS PrivateLink. Criaremos um endpoint de interface de rede em cada sub-rede que você habilitar para o endpoint de interface. Estas são interfaces de rede gerenciadas pelo solicitante que servem como ponto de entrada para o tráfego destinado ao AWS Entity Resolution.

Para obter mais informações, consulte Acesso Serviços da AWS por meio AWS PrivateLink do AWS PrivateLink Guia.

Considerações para AWS Entity Resolution

Antes de configurar um endpoint de interface para AWS Entity Resolution, consulte Considerações no AWS PrivateLink Guia.

AWS Entity Resolution suporta fazer chamadas para todas as suas ações de API por meio do endpoint da interface.

Há suporte para políticas de endpoint de VPC. AWS Entity Resolution Por padrão, o acesso total a AWS Entity Resolution é permitido por meio do endpoint da interface. Como alternativa, você pode associar um grupo de segurança às interfaces de rede do endpoint para controlar o tráfego a AWS Entity Resolution por meio do endpoint da interface.

Crie um endpoint de interface para AWS Entity Resolution

Você pode criar um endpoint de interface para AWS Entity Resolution usar o console Amazon VPC ou AWS Command Line Interface o AWS CLI(). Para obter mais informações, consulte Criar um endpoint de interface no Guia do usuário do AWS PrivateLink .

Crie um endpoint de interface para AWS Entity Resolution usar o seguinte nome de serviço:

```
com.amazonaws.region.entityresolution
```

Se você habilitar o DNS privado para o endpoint da interface, poderá fazer solicitações de API a AWS Entity Resolution usando seu nome DNS regional padrão. Por exemplo, `.entityresolution.us-east-1.amazonaws.com`

Crie uma política de endpoint para seu endpoint de interface.

Uma política de endpoint é um recurso do IAM que você pode anexar ao endpoint de interface. A política de endpoint padrão permite acesso total AWS Entity Resolution por meio do endpoint da interface. Para controlar o acesso AWS Entity Resolution permitido pela sua VPC, anexe uma política de endpoint personalizada ao endpoint da interface.

Uma política de endpoint especifica as seguintes informações:

- As entidades principais que podem realizar ações (Contas da AWS, usuários do IAM e perfis do IAM).
- As ações que podem ser realizadas.
- Os recursos nos quais as ações podem ser executadas.

Para obter mais informações, consulte Controlar o acesso aos serviços usando políticas de endpoint no Guia do AWS PrivateLink .

Exemplo: política de VPC endpoint para ações AWS Entity Resolution

Veja a seguir um exemplo de uma política de endpoint personalizado. Quando você anexa essa política ao seu endpoint de interface, ela concede acesso às AWS Entity Resolution ações listadas para todos os diretores em todos os recursos.

```
{
   "Statement": [
      {
         "Principal": "*",
         "Effect": "Allow",
         "Action": [
            "entityresolution:CreateMatchingWorkflow",
            "entityresolution:StartMatchingJob",
            "entityresolution:GetMatchingJob"
         ],
         "Resource":"*"
      }
   ]
}
```

Gerenciamento de identidade e acesso para AWS Entity Resolution

AWS Identity and Access Management (IAM) é uma ferramenta AWS service (Serviço da AWS) que ajuda o administrador a controlar com segurança o acesso aos AWS recursos. Os administradores do IAM controlam quem pode ser autenticado (conectado) e autorizado (tem permissões) a usar AWS Entity Resolution os recursos. O IAM é um AWS service (Serviço da AWS) que você pode usar sem custo adicional.

> **Note**
>
> AWS Entity Resolution suporta políticas de contas cruzadas. Consulte mais informações em Acesso a recursos entre contas no IAM no Guia do usuário do IAM.

Tópicos

- Público
- Autenticar com identidades

- Gerenciar o acesso usando políticas
- Como AWS Entity Resolution funciona com o IAM
- Exemplos de políticas baseadas em identidade para o AWS Entity Resolution
- AWS políticas gerenciadas para AWS Entity Resolution
- Solução de problemas AWS Entity Resolution de identidade e acesso

Público

A forma como você usa AWS Identity and Access Management (IAM) difere, dependendo do trabalho que você faz AWS Entity Resolution.

Usuário do serviço — Se você usar o AWS Entity Resolution serviço para realizar seu trabalho, seu administrador fornecerá as credenciais e as permissões de que você precisa. À medida que você usa mais AWS Entity Resolution recursos para fazer seu trabalho, talvez precise de permissões adicionais. Entender como o acesso é gerenciado pode ajudá-lo a solicitar as permissões corretas ao seu administrador. Se não for possível acessar um atributo no AWS Entity Resolution, consulte Solução de problemas AWS Entity Resolution de identidade e acesso.

Administrador de serviços — Se você é responsável pelos AWS Entity Resolution recursos da sua empresa, provavelmente tem acesso total AWS Entity Resolution a. É seu trabalho determinar quais AWS Entity Resolution recursos e recursos seus usuários do serviço devem acessar. Envie as solicitações ao administrador do IAM para alterar as permissões dos usuários de serviço. Revise as informações nesta página para compreender os conceitos básicos do IAM. Para saber mais sobre como sua empresa pode usar o IAM com AWS Entity Resolution, consulteComo AWS Entity Resolution funciona com o IAM.

Administrador do IAM: se você for um administrador do IAM, talvez queira saber detalhes sobre como pode gravar políticas para gerenciar acesso ao AWS Entity Resolution. Para ver exemplos de políticas AWS Entity Resolution baseadas em identidade que você pode usar no IAM, consulte. Exemplos de políticas baseadas em identidade para o AWS Entity Resolution

Autenticar com identidades

A autenticação é a forma como você faz login AWS usando suas credenciais de identidade. Você deve estar autenticado (conectado AWS) como o Usuário raiz da conta da AWS, como usuário do IAM ou assumindo uma função do IAM.

Você pode entrar AWS como uma identidade federada usando credenciais fornecidas por meio de uma fonte de identidade. AWS IAM Identity Center Usuários (IAM Identity Center), a autenticação de login único da sua empresa e suas credenciais do Google ou do Facebook são exemplos de identidades federadas. Quando você faz login como identidade federada, o administrador já configurou anteriormente a federação de identidades usando perfis do IAM. Ao acessar AWS usando a federação, você está assumindo indiretamente uma função.

Dependendo do tipo de usuário que você é, você pode entrar no AWS Management Console ou no portal de AWS acesso. Para obter mais informações sobre como fazer login AWS, consulte Como fazer login Conta da AWS no Guia do Início de Sessão da AWS usuário.

Se você acessar AWS programaticamente, AWS fornece um kit de desenvolvimento de software (SDK) e uma interface de linha de comando (CLI) para assinar criptograficamente suas solicitações usando suas credenciais. Se você não usa AWS ferramentas, você mesmo deve assinar as solicitações. Para obter mais informações sobre como usar o método recomendado para designar solicitações por conta própria, consulte Versão 4 do AWS Signature para solicitações de API no Guia do usuário do IAM.

Independente do método de autenticação usado, também pode ser necessário fornecer informações adicionais de segurança. Por exemplo, AWS recomenda que você use a autenticação multifator (MFA) para aumentar a segurança da sua conta. Para saber mais, consulte Autenticação multifator no Guia do usuário do AWS IAM Identity Center e Usar a autenticação multifator da AWS no IAM no Guia do usuário do IAM.

Conta da AWS usuário root

Ao criar uma Conta da AWS, você começa com uma identidade de login que tem acesso completo a todos Serviços da AWS os recursos da conta. Essa identidade é chamada de usuário Conta da AWS raiz e é acessada fazendo login com o endereço de e-mail e a senha que você usou para criar a conta. É altamente recomendável não usar o usuário-raiz para tarefas diárias. Proteja as credenciais do usuário-raiz e use-as para executar as tarefas que somente ele puder executar. Para obter a lista completa das tarefas que exigem login como usuário-raiz, consulte Tarefas que exigem credenciais de usuário-raiz no Guia do Usuário do IAM.

Identidade federada

Como prática recomendada, exija que usuários humanos, incluindo usuários que precisam de acesso de administrador, usem a federação com um provedor de identidade para acessar Serviços da AWS usando credenciais temporárias.

Uma identidade federada é um usuário do seu diretório de usuários corporativo, de um provedor de identidade da web AWS Directory Service, do diretório do Identity Center ou de qualquer usuário que acesse usando credenciais fornecidas Serviços da AWS por meio de uma fonte de identidade. Quando as identidades federadas acessam Contas da AWS, elas assumem funções, e as funções fornecem credenciais temporárias.

Para o gerenciamento de acesso centralizado, é recomendável usar o AWS IAM Identity Center. Você pode criar usuários e grupos no IAM Identity Center ou pode se conectar e sincronizar com um conjunto de usuários e grupos em sua própria fonte de identidade para uso em todos os seus Contas da AWS aplicativos. Para obter mais informações sobre o Centro de Identidade do IAM, consulte O que é o Centro de Identidade do IAM? no Guia do Usuário do AWS IAM Identity Center .

Usuários e grupos do IAM

Um usuário do IAM é uma identidade dentro da sua Conta da AWS que tem permissões específicas para uma única pessoa ou aplicativo. Sempre que possível, é recomendável contar com credenciais temporárias em vez de criar usuários do IAM com credenciais de longo prazo, como senhas e chaves de acesso. No entanto, se você tiver casos de uso específicos que exijam credenciais de longo prazo com usuários do IAM, é recomendável alternar as chaves de acesso. Para obter mais informações, consulte Alternar as chaves de acesso regularmente para casos de uso que exijam credenciais de longo prazo no Guia do Usuário do IAM.

Um grupo do IAM é uma identidade que especifica uma coleção de usuários do IAM. Não é possível fazer login como um grupo. É possível usar grupos para especificar permissões para vários usuários de uma vez. Os grupos facilitam o gerenciamento de permissões para grandes conjuntos de usuários. Por exemplo, você pode ter um grupo chamado IAMAdminse conceder a esse grupo permissões para administrar recursos do IAM.

Usuários são diferentes de perfis. Um usuário é exclusivamente associado a uma pessoa ou a uma aplicação, mas um perfil pode ser assumido por qualquer pessoa que precisar dele. Os usuários têm credenciais permanentes de longo prazo, mas os perfis fornecem credenciais temporárias. Para saber mais, consulte Casos de uso para usuários do IAM no Guia do usuário do IAM.

Perfis do IAM

Uma função do IAM é uma identidade dentro da sua Conta da AWS que tem permissões específicas. Ele é semelhante a um usuário do IAM, mas não está associado a uma pessoa específica. Para assumir temporariamente uma função do IAM no AWS Management Console, você pode alternar de um usuário para uma função do IAM (console). Você pode assumir uma função chamando uma

operação de AWS API AWS CLI ou usando uma URL personalizada. Para obter mais informações sobre métodos para usar perfis, consulte Métodos para assumir um perfil no Guia do usuário do IAM.

Perfis do IAM com credenciais temporárias são úteis nas seguintes situações:

- Acesso de usuário federado: para atribuir permissões a identidades federadas, é possível criar um perfil e definir permissões para ele. Quando uma identidade federada é autenticada, essa identidade é associada ao perfil e recebe as permissões definidas por ele. Para ter mais informações sobre perfis para federação, consulte Criar um perfil para um provedor de identidade de terceiros (federação) no Guia do usuário do IAM. Se usar o Centro de Identidade do IAM, configure um conjunto de permissões. Para controlar o que suas identidades podem acessar após a autenticação, o Centro de Identidade do IAM correlaciona o conjunto de permissões a um perfil no IAM. Para obter informações sobre conjuntos de permissões, consulte Conjuntos de Permissões no Guia do Usuário do AWS IAM Identity Center .
- Permissões temporárias para usuários do IAM: um usuário ou um perfil do IAM pode presumir um perfil do IAM para obter temporariamente permissões diferentes para uma tarefa específica.
- Acesso entre contas: é possível usar um perfil do IAM para permitir que alguém (uma entidade principal confiável) em outra conta acesse recursos em sua conta. Os perfis são a principal forma de conceder acesso entre contas. No entanto, com alguns Serviços da AWS, você pode anexar uma política diretamente a um recurso (em vez de usar uma função como proxy). Para conhecer a diferença entre perfis e políticas baseadas em recurso para acesso entre contas, consulte Acesso a recursos entre contas no IAM no Guia do usuário do IAM.
- Acesso entre serviços — Alguns Serviços da AWS usam recursos em outros Serviços da AWS. Por exemplo, quando você faz uma chamada em um serviço, é comum que esse serviço execute aplicativos na Amazon EC2 ou armazene objetos no Amazon S3. Um serviço pode fazer isso usando as permissões da entidade principal da chamada, usando um perfil de serviço ou um perfil vinculado ao serviço.
  - Sessões de acesso direto (FAS) — Quando você usa um usuário ou uma função do IAM para realizar ações AWS, você é considerado principal. Ao usar alguns serviços, você pode executar uma ação que inicia outra ação em um serviço diferente. O FAS usa as permissões do diretor chamando um AWS service (Serviço da AWS), combinadas com a solicitação AWS service (Serviço da AWS) para fazer solicitações aos serviços posteriores. As solicitações do FAS são feitas somente quando um serviço recebe uma solicitação que requer interações com outros Serviços da AWS ou com recursos para ser concluída. Nesse caso, você precisa ter permissões para executar ambas as ações. Para obter detalhes da política ao fazer solicitações de FAS, consulte Sessões de acesso direto.

  - Perfil de serviço: um perfil de serviço é um perfil do IAM que um serviço assume para executar ações em seu nome. Um administrador do IAM pode criar, modificar e excluir um perfil de serviço do IAM. Para obter mais informações, consulte Criar um perfil para delegar permissões a um AWS service (Serviço da AWS) no Guia do Usuário do IAM.
  - Função vinculada ao serviço — Uma função vinculada ao serviço é um tipo de função de serviço vinculada a um. AWS service (Serviço da AWS) O serviço pode presumir o perfil de executar uma ação em seu nome. As funções vinculadas ao serviço aparecem em você Conta da AWS e são de propriedade do serviço. Um administrador do IAM pode visualizar, mas não editar as permissões para perfis vinculados ao serviço.
- Aplicativos em execução na Amazon EC2 — Você pode usar uma função do IAM para gerenciar credenciais temporárias para aplicativos que estão sendo executados em uma EC2 instância e fazendo solicitações AWS CLI de AWS API. Isso é preferível a armazenar chaves de acesso na EC2 instância. Para atribuir uma AWS função a uma EC2 instância e disponibilizá-la para todos os aplicativos, você cria um perfil de instância anexado à instância. Um perfil de instância contém a função e permite que programas em execução na EC2 instância recebam credenciais temporárias. Para obter mais informações, consulte Usar uma função do IAM para conceder permissões a aplicativos executados em EC2 instâncias da Amazon no Guia do usuário do IAM.

Gerenciar o acesso usando políticas

Você controla o acesso AWS criando políticas e anexando-as a AWS identidades ou recursos. Uma política é um objeto AWS que, quando associada a uma identidade ou recurso, define suas permissões. AWS avalia essas políticas quando um principal (usuário, usuário raiz ou sessão de função) faz uma solicitação. As permissões nas políticas determinam se a solicitação será permitida ou negada. A maioria das políticas é armazenada AWS como documentos JSON. Para obter mais informações sobre a estrutura e o conteúdo de documentos de políticas JSON, consulte Visão geral das políticas JSON no Guia do usuário do IAM.

Os administradores podem usar políticas AWS JSON para especificar quem tem acesso ao quê. Ou seja, qual entidade principal pode executar ações em quais recursos e em que condições.

Por padrão, usuários e perfis não têm permissões. Para conceder permissão aos usuários para executar ações nos recursos que eles precisam, um administrador do IAM pode criar políticas do IAM. O administrador pode então adicionar as políticas do IAM aos perfis e os usuários podem assumir os perfis.

As políticas do IAM definem permissões para uma ação independentemente do método usado para executar a operação. Por exemplo, suponha que você tenha uma política que permite a ação `iam:GetRole`. Um usuário com essa política pode obter informações de função da AWS Management Console AWS CLI, da ou da AWS API.

Políticas baseadas em identidade

As políticas baseadas em identidade são documentos de políticas de permissões JSON que você pode anexar a uma identidade, como usuário, grupo de usuários ou perfil do IAM. Essas políticas controlam quais ações os usuários e perfis podem realizar, em quais recursos e em que condições. Para saber como criar uma política baseada em identidade, consulte Definir permissões personalizadas do IAM com as políticas gerenciadas pelo cliente no Guia do Usuário do IAM.

As políticas baseadas em identidade podem ser categorizadas como políticas em linha ou políticas gerenciadas. As políticas em linha são anexadas diretamente a um único usuário, grupo ou perfil. As políticas gerenciadas são políticas autônomas que você pode associar a vários usuários, grupos e funções em seu Conta da AWS. As políticas AWS gerenciadas incluem políticas gerenciadas e políticas gerenciadas pelo cliente. Para saber como escolher entre uma política gerenciada ou uma política em linha, consulte Escolher entre políticas gerenciadas e políticas em linha no Guia do usuário do IAM.

Políticas baseadas em recursos

Políticas baseadas em recursos são documentos de políticas JSON que você anexa a um recurso. São exemplos de políticas baseadas em recursos as políticas de confiança de perfil do IAM e as políticas de bucket do Amazon S3. Em serviços compatíveis com políticas baseadas em recursos, os administradores de serviço podem usá-las para controlar o acesso a um recurso específico. Para o atributo ao qual a política está anexada, a política define quais ações uma entidade principal especificado pode executar nesse atributo e em que condições. Você deve especificar uma entidade principal em uma política baseada em recursos. Os diretores podem incluir contas, usuários, funções, usuários federados ou. Serviços da AWS

Políticas baseadas em recursos são políticas em linha localizadas nesse serviço. Você não pode usar políticas AWS gerenciadas do IAM em uma política baseada em recursos.

Listas de controle de acesso (ACLs)

As listas de controle de acesso (ACLs) controlam quais diretores (membros da conta, usuários ou funções) têm permissões para acessar um recurso. ACLs são semelhantes às políticas baseadas em recursos, embora não usem o formato de documento de política JSON.

O Amazon S3 e o AWS WAF Amazon VPC são exemplos de serviços que oferecem suporte. ACLs Para saber mais ACLs, consulte a visão geral da lista de controle de acesso (ACL) no Guia do desenvolvedor do Amazon Simple Storage Service.

Outros tipos de política

AWS oferece suporte a tipos de políticas adicionais menos comuns. Esses tipos de política podem definir o máximo de permissões concedidas a você pelos tipos de política mais comuns.

- Limites de permissões: um limite de permissões é um recurso avançado no qual você define o máximo de permissões que uma política baseada em identidade pode conceder a uma entidade do IAM (usuário ou perfil do IAM). É possível definir um limite de permissões para uma entidade. As permissões resultantes são a interseção das políticas baseadas em identidade de uma entidade com seus limites de permissões. As políticas baseadas em recurso que especificam o usuário ou o perfil no campo `Principal` não são limitadas pelo limite de permissões. Uma negação explícita em qualquer uma dessas políticas substitui a permissão. Para obter mais informações sobre limites de permissões, consulte Limites de permissões para identidades do IAM no Guia do usuário do IAM.
- Políticas de controle de serviço (SCPs) — SCPs são políticas JSON que especificam as permissões máximas para uma organização ou unidade organizacional (OU) em AWS Organizations. AWS Organizations é um serviço para agrupar e gerenciar centralmente vários Contas da AWS que sua empresa possui. Se você habilitar todos os recursos em uma organização, poderá aplicar políticas de controle de serviço (SCPs) a qualquer uma ou a todas as suas contas. O SCP limita as permissões para entidades nas contas dos membros, incluindo cada uma Usuário raiz da conta da AWS. Para obter mais informações sobre Organizations e SCPs, consulte Políticas de controle de serviços no Guia AWS Organizations do Usuário.
- Políticas de controle de recursos (RCPs) — RCPs são políticas JSON que você pode usar para definir o máximo de permissões disponíveis para recursos em suas contas sem atualizar as políticas do IAM anexadas a cada recurso que você possui. O RCP limita as permissões para recursos nas contas dos membros e pode afetar as permissões efetivas para identidades, incluindo a Usuário raiz da conta da AWS, independentemente de pertencerem à sua organização. Para obter mais informações sobre Organizations e RCPs, incluindo uma lista Serviços da AWS desse suporte RCPs, consulte Políticas de controle de recursos (RCPs) no Guia AWS Organizations do usuário.
- Políticas de sessão: são políticas avançadas que você transmite como um parâmetro quando cria de forma programática uma sessão temporária para um perfil ou um usuário federado. As permissões da sessão resultante são a interseção das políticas baseadas em identidade do

usuário ou do perfil e das políticas de sessão. As permissões também podem ser provenientes de uma política baseada em recursos. Uma negação explícita em qualquer uma dessas políticas substitui a permissão. Para obter mais informações, consulte Políticas de sessão no Guia do usuário do IAM.

Vários tipos de política

Quando vários tipos de política são aplicáveis a uma solicitação, é mais complicado compreender as permissões resultantes. Para saber como AWS determinar se uma solicitação deve ser permitida quando vários tipos de políticas estão envolvidos, consulte Lógica de avaliação de políticas no Guia do usuário do IAM.

Como AWS Entity Resolution funciona com o IAM

Antes de usar o IAM para gerenciar o acesso AWS Entity Resolution, saiba com quais recursos do IAM estão disponíveis para uso AWS Entity Resolution.

Recursos do IAM que você pode usar com AWS Entity Resolution

Atributo do IAM	AWS Entity Resolution apoio
Políticas baseadas em identidade	Sim
Políticas baseadas em atributos	Sim
Ações de políticas	Sim
Recursos de políticas	Sim
Chaves de condição de políticas	Sim
ACLs	Não
ABAC (tags em políticas)	Parcial
Credenciais temporárias	Sim
Sessões de acesso direto (FAS)	Sim
Perfis de serviço	Sim

Atributo do IAM	AWS Entity Resolution apoio
Perfis vinculados a serviço	Não

Para ter uma visão de alto nível de como AWS Entity Resolution e outros AWS serviços funcionam com a maioria dos recursos do IAM, consulte AWS os serviços que funcionam com o IAM no Guia do usuário do IAM.

Políticas baseadas em identidade para AWS Entity Resolution

Compatível com políticas baseadas em identidade: sim

As políticas baseadas em identidade são documentos de políticas de permissões JSON que você pode anexar a uma identidade, como usuário do IAM, grupo de usuários ou perfil. Essas políticas controlam quais ações os usuários e perfis podem realizar, em quais recursos e em que condições. Para saber como criar uma política baseada em identidade, consulte Definir permissões personalizadas do IAM com as políticas gerenciadas pelo cliente no Guia do Usuário do IAM.

Com as políticas baseadas em identidade do IAM, é possível especificar ações e recursos permitidos ou negados, assim como as condições sob as quais as ações são permitidas ou negadas. Você não pode especificar a entidade principal em uma política baseada em identidade porque ela se aplica ao usuário ou perfil ao qual ela está anexada. Para saber mais sobre todos os elementos que podem ser usados em uma política JSON, consulte Referência de elemento de política JSON do IAM no Guia do usuário do IAM.

Exemplos de políticas baseadas em identidade para AWS Entity Resolution

Para ver exemplos de políticas AWS Entity Resolution baseadas em identidade, consulte. Exemplos de políticas baseadas em identidade para o AWS Entity Resolution

Políticas baseadas em recursos dentro AWS Entity Resolution

Compatível com políticas baseadas em recursos: sim

Políticas baseadas em recursos são documentos de políticas JSON que você anexa a um recurso. São exemplos de políticas baseadas em recursos as políticas de confiança de perfil do IAM e as políticas de bucket do Amazon S3. Em serviços compatíveis com políticas baseadas em recursos, os administradores de serviço podem usá-las para controlar o acesso a um recurso específico. Para o atributo ao qual a política está anexada, a política define quais ações uma entidade principal

especificado pode executar nesse atributo e em que condições. Você deve especificar uma entidade principal em uma política baseada em recursos. Os diretores podem incluir contas, usuários, funções, usuários federados ou. Serviços da AWS

Para permitir o acesso entre contas, você pode especificar uma conta inteira ou as entidades do IAM em outra conta como a entidade principal em uma política baseada em recursos. Adicionar uma entidade principal entre contas à política baseada em recurso é apenas metade da tarefa de estabelecimento da relação de confiança. Quando o principal e o recurso são diferentes Contas da AWS, um administrador do IAM na conta confiável também deve conceder permissão à entidade principal (usuário ou função) para acessar o recurso. Eles concedem permissão ao anexar uma política baseada em identidade para a entidade. No entanto, se uma política baseada em recurso conceder acesso a uma entidade principal na mesma conta, nenhuma política baseada em identidade adicional será necessária. Consulte mais informações em Acesso a recursos entre contas no IAM no Guia do usuário do IAM.

Ações políticas para AWS Entity Resolution

Compatível com ações de políticas: sim

Os administradores podem usar políticas AWS JSON para especificar quem tem acesso ao quê. Ou seja, qual entidade principal pode executar ações em quais recursos e em que condições.

O elemento `Action` de uma política JSON descreve as ações que podem ser usadas para permitir ou negar acesso em uma política. As ações de política geralmente têm o mesmo nome da operação de AWS API associada. Existem algumas exceções, como ações somente de permissão, que não têm uma operação de API correspondente. Algumas operações também exigem várias ações em uma política. Essas ações adicionais são chamadas de ações dependentes.

Incluem ações em uma política para conceder permissões para executar a operação associada.

Para ver uma lista de AWS Entity Resolution ações, consulte Ações definidas por AWS Entity Resolution na Referência de autorização de serviço.

As ações de política AWS Entity Resolution usam o seguinte prefixo antes da ação:

```
entityresolution
```

Para especificar várias ações em uma única declaração, separe-as com vírgulas.

```
"Action": [
```

```
      "entityresolution:action1",
      "entityresolution:action2"
         ]
```

Para ver exemplos de políticas AWS Entity Resolution baseadas em identidade, consulte. Exemplos de políticas baseadas em identidade para o AWS Entity Resolution

Recursos políticos para AWS Entity Resolution

Compatível com recursos de políticas: sim

Os administradores podem usar políticas AWS JSON para especificar quem tem acesso ao quê. Ou seja, qual entidade principal pode executar ações em quais recursos e em que condições.

O elemento de política JSON `Resource` especifica o objeto ou os objetos aos quais a ação se aplica. As instruções devem incluir um elemento `Resource` ou `NotResource`. Como prática recomendada, especifique um recurso usando seu nome do recurso da Amazon (ARN). Isso pode ser feito para ações que oferecem compatibilidade com um tipo de recurso específico, conhecido como permissões em nível de recurso.

Para ações que não oferecem compatibilidade com permissões em nível de recurso, como operações de listagem, use um curinga (*) para indicar que a instrução se aplica a todos os recursos.

```
"Resource": "*"
```

Para ver uma lista dos tipos de AWS Entity Resolution recursos e seus ARNs, consulte Recursos definidos por AWS Entity Resolution na Referência de autorização de serviço. Para saber com quais ações você pode especificar o ARN de cada recurso, consulte Ações definidas pelo AWS Entity Resolution.

Para ver exemplos de políticas AWS Entity Resolution baseadas em identidade, consulte. Exemplos de políticas baseadas em identidade para o AWS Entity Resolution

Chaves de condição de política para AWS Entity Resolution

Compatível com chaves de condição de política específicas de serviço: sim

Os administradores podem usar políticas AWS JSON para especificar quem tem acesso ao quê. Ou seja, qual entidade principal pode executar ações em quais recursos e em que condições.

O elemento `Condition` (ou bloco `Condition`) permite que você especifique condições nas quais uma instrução estiver em vigor. O elemento `Condition` é opcional. É possível criar expressões condicionais que usem agentes de condição, como “igual a” ou “menor que”, para fazer a condição da política corresponder aos valores na solicitação.

Se você especificar vários elementos de `Condition` em uma declaração ou várias chaves em um único elemento de `Condition`, a AWS os avaliará usando uma operação lógica `AND`. Se você especificar vários valores para uma única chave de condição, AWS avalia a condição usando uma `OR` operação lógica. Todas as condições devem ser atendidas antes que as permissões da instrução sejam concedidas.

Você também pode usar variáveis de espaço reservado ao especificar condições. Por exemplo, é possível conceder a um usuário do IAM permissão para acessar um recurso somente se ele estiver marcado com seu nome de usuário do IAM. Para obter mais informações, consulte Elementos da política do IAM: variáveis e tags no Guia do usuário do IAM.

AWS suporta chaves de condição globais e chaves de condição específicas do serviço. Para ver todas as chaves de condição AWS globais, consulte as chaves de contexto de condição AWS global no Guia do usuário do IAM.

Para ver uma lista de chaves de AWS Entity Resolution condição, consulte Chaves de condição AWS Entity Resolution na Referência de autorização de serviço. Para saber com quais ações e recursos você pode usar uma chave de condição, consulte Ações definidas por AWS Entity Resolution.

Para ver exemplos de políticas AWS Entity Resolution baseadas em identidade, consulte. Exemplos de políticas baseadas em identidade para o AWS Entity Resolution

ACLs in AWS Entity Resolution

Suportes ACLs: Não

As listas de controle de acesso (ACLs) controlam quais diretores (membros da conta, usuários ou funções) têm permissões para acessar um recurso. ACLs são semelhantes às políticas baseadas em recursos, embora não usem o formato de documento de política JSON.

ABAC com AWS Entity Resolution

Compatível com ABAC (tags em políticas): parcial

O controle de acesso por atributo (ABAC) é uma estratégia de autorização que define as permissões com base em atributos. Em AWS, esses atributos são chamados de tags. Você pode anexar tags a entidades do IAM (usuários ou funções) e a vários AWS recursos. Marcar de entidades e atributos é a primeira etapa do ABAC. Em seguida, você cria políticas de ABAC para permitir operações quando a tag da entidade principal corresponder à tag do recurso que ela estiver tentando acessar.

O ABAC é útil em ambientes que estão crescendo rapidamente e ajuda em situações em que o gerenciamento de políticas se torna um problema.

Para controlar o acesso baseado em tags, forneça informações sobre as tags no elemento de condição de uma política usando as `aws:ResourceTag/key-name`, `aws:RequestTag/key-name` ou chaves de condição `aws:TagKeys`.

Se um serviço for compatível com as três chaves de condição para cada tipo de recurso, o valor será Sim para o serviço. Se um serviço for compatível com as três chaves de condição somente para alguns tipos de recursos, o valor será Parcial

Para obter mais informações sobre o ABAC, consulte Definir permissões com autorização do ABAC no Guia do usuário do IAM. Para visualizar um tutorial com etapas para configurar o ABAC, consulte Usar controle de acesso baseado em atributos (ABAC) no Guia do usuário do IAM.

Usando credenciais temporárias com AWS Entity Resolution

Compatível com credenciais temporárias: sim

Alguns Serviços da AWS não funcionam quando você faz login usando credenciais temporárias. Para obter informações adicionais, incluindo quais Serviços da AWS funcionam com credenciais temporárias, consulte Serviços da AWS trabalhar com o IAM no Guia do usuário do IAM.

Você está usando credenciais temporárias se fizer login AWS Management Console usando qualquer método, exceto um nome de usuário e senha. Por exemplo, quando você acessa AWS usando o link de login único (SSO) da sua empresa, esse processo cria automaticamente credenciais temporárias. Você também cria automaticamente credenciais temporárias quando faz login no console como usuário e, em seguida, alterna perfis. Para obter mais informações sobre como alternar funções, consulte Alternar para um perfil do IAM (console) no Guia do usuário do IAM.

Você pode criar manualmente credenciais temporárias usando a AWS API AWS CLI ou. Em seguida, você pode usar essas credenciais temporárias para acessar AWS. AWS recomenda que você gere credenciais temporárias dinamicamente em vez de usar chaves de acesso de longo prazo. Para obter mais informações, consulte Credenciais de segurança temporárias no IAM.

Sessões de acesso direto para AWS Entity Resolution

Compatibilidade com o recurso de encaminhamento de sessões de acesso (FAS): sim

Quando você usa um usuário ou uma função do IAM para realizar ações AWS, você é considerado principal. Ao usar alguns serviços, você pode executar uma ação que inicia outra ação em um serviço diferente. O FAS usa as permissões do diretor chamando um AWS service (Serviço da AWS), combinadas com a solicitação AWS service (Serviço da AWS) para fazer solicitações aos serviços posteriores. As solicitações do FAS são feitas somente quando um serviço recebe uma solicitação que requer interações com outros Serviços da AWS ou com recursos para ser concluída. Nesse caso, você precisa ter permissões para executar ambas as ações. Para obter detalhes da política ao fazer solicitações de FAS, consulte Sessões de acesso direto.

Funções de serviço para AWS Entity Resolution

Compatível com perfis de serviço: sim

O perfil de serviço é um perfil do IAM que um serviço assume para executar ações em seu nome. Um administrador do IAM pode criar, modificar e excluir um perfil de serviço do IAM. Para obter mais informações, consulte Criar um perfil para delegar permissões a um AWS service (Serviço da AWS) no Guia do Usuário do IAM.

> **Warning**
>
> Alterar as permissões de uma função de serviço pode interromper AWS Entity Resolution a funcionalidade. Edite as funções de serviço somente quando AWS Entity Resolution fornecer orientação para fazer isso.

Funções vinculadas a serviços para AWS Entity Resolution

Compatível com perfis vinculados ao serviço: Não

Uma função vinculada ao serviço é um tipo de função de serviço vinculada a um. AWS service (Serviço da AWS) O serviço pode presumir o perfil de executar uma ação em seu nome. As funções vinculadas ao serviço aparecem em você Conta da AWS e são de propriedade do serviço. Um administrador do IAM pode visualizar, mas não editar as permissões para funções vinculadas ao serviço.

Para obter detalhes sobre como criar ou gerenciar perfis vinculados a serviços, consulte Serviços da AWS que funcionam com o IAM. Encontre um serviço na tabela que inclua um `Yes` na coluna Perfil vinculado ao serviço. Escolha o link Sim para visualizar a documentação do perfil vinculado a esse serviço .

Exemplos de políticas baseadas em identidade para o AWS Entity Resolution

Por padrão, usuários e perfis não têm permissão para criar ou modificar recursos do AWS Entity Resolution . Eles também não podem realizar tarefas usando a AWS API AWS Management Console, AWS Command Line Interface (AWS CLI) ou. Para conceder permissão aos usuários para executar ações nos recursos que eles precisam, um administrador do IAM pode criar políticas do IAM. O administrador pode então adicionar as políticas do IAM aos perfis e os usuários podem assumir os perfis.

Para aprender a criar uma política baseada em identidade do IAM ao usar esses documentos de política em JSON de exemplo, consulte Criar políticas do IAM (console) no Guia do usuário do IAM.

Para obter detalhes sobre ações e tipos de recursos definidos por AWS Entity Resolution, incluindo o formato do ARNs para cada um dos tipos de recursos, consulte Ações, recursos e chaves de condição AWS Entity Resolution na Referência de autorização de serviço.

Tópicos

- Práticas recomendadas de política
- Usar o console do AWS Entity Resolution
- Permitir que os usuários visualizem suas próprias permissões

Práticas recomendadas de política

As políticas baseadas em identidade determinam se alguém pode criar, acessar ou excluir AWS Entity Resolution recursos em sua conta. Essas ações podem incorrer em custos para sua Conta da AWS. Ao criar ou editar políticas baseadas em identidade, siga estas diretrizes e recomendações:

- Comece com as políticas AWS gerenciadas e passe para as permissões de privilégios mínimos — Para começar a conceder permissões aos seus usuários e cargas de trabalho, use as políticas AWS gerenciadas que concedem permissões para muitos casos de uso comuns. Eles estão disponíveis no seu Conta da AWS. Recomendamos que você reduza ainda mais as permissões definindo políticas gerenciadas pelo AWS cliente que sejam específicas para seus casos de uso.

Para obter mais informações, consulte Políticas gerenciadas pela AWS ou Políticas gerenciadas pela AWS para funções de trabalho no Guia do usuário do IAM.

- Aplique permissões de privilégio mínimo: ao definir permissões com as políticas do IAM, conceda apenas as permissões necessárias para executar uma tarefa. Você faz isso definindo as ações que podem ser executadas em recursos específicos sob condições específicas, também conhecidas como permissões de privilégio mínimo. Para obter mais informações sobre como usar o IAM para aplicar permissões, consulte Políticas e permissões no IAM no Guia do usuário do IAM.
- Use condições nas políticas do IAM para restringir ainda mais o acesso: você pode adicionar uma condição às políticas para limitar o acesso a ações e recursos. Por exemplo, você pode escrever uma condição de política para especificar que todas as solicitações devem ser enviadas usando SSL. Você também pode usar condições para conceder acesso às ações de serviço se elas forem usadas por meio de uma ação específica AWS service (Serviço da AWS), como AWS CloudFormation. Para obter mais informações, consulte Elementos da política JSON do IAM: condição no Guia do usuário do IAM.
- Use o IAM Access Analyzer para validar suas políticas do IAM a fim de garantir permissões seguras e funcionais: o IAM Access Analyzer valida as políticas novas e existentes para que elas sigam a linguagem de política do IAM (JSON) e as práticas recomendadas do IAM. O IAM Access Analyzer oferece mais de cem verificações de política e recomendações práticas para ajudar a criar políticas seguras e funcionais. Para obter mais informações, consulte Validação de políticas do IAM Access Analyzer no Guia do Usuário do IAM.
- Exigir autenticação multifator (MFA) — Se você tiver um cenário que exija usuários do IAM ou um usuário root, ative Conta da AWS a MFA para obter segurança adicional. Para exigir MFA quando as operações de API forem chamadas, adicione condições de MFA às suas políticas. Para obter mais informações, consulte Configuração de acesso à API protegido por MFA no Guia do Usuário do IAM.

Para obter mais informações sobre as práticas recomendadas do IAM, consulte Práticas recomendadas de segurança no IAM no Guia do usuário do IAM.

Usar o console do AWS Entity Resolution

Para acessar o AWS Entity Resolution console, você deve ter um conjunto mínimo de permissões. Essas permissões devem permitir que você liste e visualize detalhes sobre os AWS Entity Resolution recursos em seu Conta da AWS. Caso crie uma política baseada em identidade mais restritiva que as permissões mínimas necessárias, o console não funcionará como pretendido para entidades (usuários ou perfis) com essa política.

Você não precisa permitir permissões mínimas do console para usuários que estão fazendo chamadas somente para a API AWS CLI ou para a AWS API. Em vez disso, permita o acesso somente a ações que correspondam à operação de API que estiverem tentando executar.

Para garantir que usuários e funções ainda possam usar o AWS Entity Resolution console, anexe também a política AWS Entity Resolution *ConsoleAccess* ou a política *ReadOnly* AWS gerenciada às entidades. Para obter informações, consulte Adicionar permissões a um usuário no Guia do usuário do IAM.

Permitir que os usuários visualizem suas próprias permissões

Este exemplo mostra como criar uma política que permita que os usuários do IAM visualizem as políticas gerenciadas e em linha anexadas a sua identidade de usuário. Essa política inclui permissões para concluir essa ação no console ou programaticamente usando a API AWS CLI ou AWS .

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "ViewOwnUserInfo",
            "Effect": "Allow",
            "Action": [
                "iam:GetUserPolicy",
                "iam:ListGroupsForUser",
                "iam:ListAttachedUserPolicies",
                "iam:ListUserPolicies",
                "iam:GetUser"
            ],
            "Resource": ["arn:aws:iam::*:user/${aws:username}"]
        },
        {
            "Sid": "NavigateInConsole",
            "Effect": "Allow",
            "Action": [
                "iam:GetGroupPolicy",
                "iam:GetPolicyVersion",
                "iam:GetPolicy",
                "iam:ListAttachedGroupPolicies",
                "iam:ListGroupPolicies",
                "iam:ListPolicyVersions",
                "iam:ListPolicies",
```

```
                "iam:ListUsers"
            ],
            "Resource": "*"
        }
    ]
}
```

AWS políticas gerenciadas para AWS Entity Resolution

Uma política AWS gerenciada é uma política autônoma criada e administrada por AWS. AWS as políticas gerenciadas são projetadas para fornecer permissões para muitos casos de uso comuns, para que você possa começar a atribuir permissões a usuários, grupos e funções.

Lembre-se de que as políticas AWS gerenciadas podem não conceder permissões de privilégio mínimo para seus casos de uso específicos porque estão disponíveis para uso de todos os AWS clientes. Recomendamos que você reduza ainda mais as permissões definindo as políticas gerenciadas pelo cliente que são específicas para seus casos de uso.

Você não pode alterar as permissões definidas nas políticas AWS gerenciadas. Se AWS atualizar as permissões definidas em uma política AWS gerenciada, a atualização afetará todas as identidades principais (usuários, grupos e funções) às quais a política está anexada. AWS é mais provável que atualize uma política AWS gerenciada quando uma nova AWS service (Serviço da AWS) é lançada ou novas operações de API são disponibilizadas para serviços existentes.

Para obter mais informações, consulte Políticas gerenciadas pela AWS no Guia do usuário do IAM.

AWS política gerenciada: AWSEntity ResolutionConsoleFullAccess

É possível anexar a política `AWSEntityResolutionConsoleFullAccess` às identidades do IAM.

Essa política concede acesso total aos AWS Entity Resolution endpoints e recursos.

Essa política também permite determinado acesso de leitura a informações relacionadas, Serviços da AWS como S3 AWS Glue, Marcação, AWS KMS para que o console possa exibir opções e usar as selecionadas para realizar ações de resolução de entidades. Alguns recursos são reduzidos para conter o nome `entityresolution` do serviço.

Como AWS Entity Resolution depende de uma função passada para realizar ações em AWS recursos relacionados, essa política também concede as permissões para selecionar e transmitir a função desejada.

Detalhes das permissões

Esta política inclui as seguintes permissões.

- `EntityResolutionAccess`— Permite que os diretores tenham acesso total aos AWS Entity Resolution endpoints e recursos.
- `GlueSourcesConsoleDisplay`— Concede acesso às AWS Glue tabelas de listagem como opções de fonte de dados e ao esquema de importação da tabela de uma fonte de dados para a experiência do usuário.
- `S3BucketsConsoleDisplay`— Concede acesso para listar todos os buckets do S3 como opções de fonte de dados.
- `S3SourcesConsoleDisplay`— Concede acesso para exibir buckets do S3 como opções de fonte de dados.
- `TaggingConsoleDisplay`— Concede acesso à leitura de chaves e valores de marcação.
- `KMSConsoleDisplay`— Concede acesso para descrever chaves e listar aliases para AWS Key Management Service descriptografar e criptografar fontes de dados.
- `ListRolesToPickForPassing`— Concede acesso para listar todas as funções para que o usuário possa escolher a função a ser passada.
- `PassRoleToEntityResolutionService`— Concede acesso para passar uma função restrita ao AWS Entity Resolution serviço.
- `ManageEventBridgeRules`— Concede acesso para criar, atualizar e excluir a EventBridge regra da Amazon para receber notificações do S3.
- `ADXReadAccess`— Concede acesso AWS Data Exchange para verificar se o cliente tem um direito ou uma assinatura.

Para visualizar as permissões para esta política, consulte AWSEntityResolutionConsoleFullAccess na Referência de políticas gerenciadas pela AWS .

AWS política gerenciada: AWSEntity ResolutionConsoleReadOnlyAccess

Você pode anexar `AWSEntityResolutionConsoleReadOnlyAccess` às entidades do IAM.

Essa política concede acesso somente para leitura a AWS Entity Resolution endpoints e recursos.

Detalhes das permissões

Esta política inclui as seguintes permissões.

- `EntityResolutionRead`— Permite que os diretores tenham acesso somente de leitura aos AWS Entity Resolution endpoints e recursos.

Para visualizar as permissões para esta política, consulte AWSEntityResolutionConsoleReadOnlyAccess na Referência de políticas gerenciadas pela AWS .

AWS Entity Resolution atualizações nas políticas AWS gerenciadas

Veja detalhes sobre as atualizações das políticas AWS gerenciadas AWS Entity Resolution desde que esse serviço começou a rastrear essas alterações. Para receber alertas automáticos sobre alterações nessa página, assine o feed RSS na página Histórico do AWS Entity Resolution documento.

Alteração	Descrição	Data
`AWSEntityResolutionConsoleFullAccess` : atualizar para uma política existente.	Adicionado `ADXReadAccess` e `ManageEventBridgeRules` para ativar a opção de serviços do provedor no fluxo de trabalho correspondente.	16 de outubro de 2023
AWS Entity Resolution começou a rastrear alterações	AWS Entity Resolution começou a rastrear as mudanças em suas políticas AWS gerenciadas.	18 de agosto de 2023

Solução de problemas AWS Entity Resolution de identidade e acesso

Use as informações a seguir para ajudá-lo a diagnosticar e corrigir problemas comuns que você pode encontrar ao trabalhar com AWS Entity Resolution um IAM.

Tópicos

- Não estou autorizado a realizar uma ação em AWS Entity Resolution
- Não estou autorizado a realizar iam: PassRole
- Quero permitir que pessoas fora da minha Conta da AWS acessem meus AWS Entity Resolution recursos

Não estou autorizado a realizar uma ação em AWS Entity Resolution

Se isso AWS Management Console indicar que você não está autorizado a realizar uma ação, entre em contato com o administrador para obter ajuda. O administrador é a pessoa que forneceu o seu nome de usuário e senha.

O erro do exemplo a seguir ocorre quando o usuário do IAM `mateojackson` tenta usar o console para visualizar detalhes sobre um recurso do *`my-example-widget`* fictício, mas não tem as permissões fictícias do `entityresolution:`*`GetWidget`*.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
 entityresolution:GetWidget on resource: my-example-widget
```

Neste caso, Mateo pede ao administrador para atualizar suas políticas para permitir a ele o acesso ao recurso *`my-example-widget`* usando a ação `entityresolution:`*`GetWidget`*.

Não estou autorizado a realizar iam: PassRole

Se você receber uma mensagem de erro informando que não está autorizado a executar a ação `iam:PassRole`, as suas políticas devem ser atualizadas para permitir que você passe uma função para o AWS Entity Resolution.

Alguns Serviços da AWS permitem que você passe uma função existente para esse serviço em vez de criar uma nova função de serviço ou uma função vinculada ao serviço. Para fazê-lo, você deve ter permissões para passar o perfil para o serviço.

O exemplo de erro a seguir ocorre quando uma usuária do IAM chamada `marymajor` tenta utilizar o console para executar uma ação no AWS Entity Resolution. No entanto, a ação exige que o serviço tenha permissões concedidas por um perfil de serviço. Mary não tem permissões para passar o perfil para o serviço.

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
 iam:PassRole
```

Nesse caso, as políticas de Mary devem ser atualizadas para permitir que ela realize a ação `iam:PassRole`.

Se precisar de ajuda, entre em contato com seu AWS administrador. Seu administrador é a pessoa que forneceu suas credenciais de login.

Quero permitir que pessoas fora da minha Conta da AWS acessem meus AWS Entity Resolution recursos

É possível criar um perfil que os usuários de outras contas ou pessoas fora da sua organização podem usar para acessar seus recursos. É possível especificar quem é confiável para assumir o perfil. Para serviços que oferecem suporte a políticas baseadas em recursos ou listas de controle de acesso (ACLs), você pode usar essas políticas para conceder às pessoas acesso aos seus recursos.

Para saber mais, consulte:

- Para saber se é AWS Entity Resolution compatível com esses recursos, consulteComo AWS Entity Resolution funciona com o IAM.
- Para saber como fornecer acesso aos seus recursos em todos os Contas da AWS que você possui, consulte Como fornecer acesso a um usuário do IAM em outro Conta da AWS que você possui no Guia do usuário do IAM.
- Para saber como fornecer acesso aos seus recursos a terceiros Contas da AWS, consulte Como fornecer acesso Contas da AWS a terceiros no Guia do usuário do IAM.
- Para saber como conceder acesso por meio da federação de identidades, consulte Conceder acesso a usuários autenticados externamente (federação de identidades) no Guia do usuário do IAM.
- Para conhecer a diferença entre perfis e políticas baseadas em recurso para acesso entre contas, consulte Acesso a recursos entre contas no IAM no Guia do usuário do IAM.

Validação de conformidade para AWS Entity Resolution

Para saber se um AWS service (Serviço da AWS) está dentro do escopo de programas de conformidade específicos, consulte Serviços da AWS Escopo por Programa de Conformidade Serviços da AWS e escolha o programa de conformidade em que você está interessado. Para obter informações gerais, consulte Programas de AWS conformidade Programas AWS de .

Você pode baixar relatórios de auditoria de terceiros usando AWS Artifact. Para obter mais informações, consulte Baixar relatórios em AWS Artifact .

Sua responsabilidade de conformidade ao usar Serviços da AWS é determinada pela confidencialidade de seus dados, pelos objetivos de conformidade de sua empresa e pelas leis e regulamentos aplicáveis. AWS fornece os seguintes recursos para ajudar na conformidade:

- Governança e conformidade de segurança: esses guias de implementação de solução abordam considerações sobre a arquitetura e fornecem etapas para implantar recursos de segurança e conformidade.
- Referência de serviços qualificados para HIPAA: lista os serviços qualificados para HIPAA. Nem todos Serviços da AWS são elegíveis para a HIPAA.
- AWS Recursos de https://aws.amazon.com/compliance/resources/ de conformidade — Essa coleção de pastas de trabalho e guias pode ser aplicada ao seu setor e local.
- AWS Guias de conformidade do cliente — Entenda o modelo de responsabilidade compartilhada sob a ótica da conformidade. Os guias resumem as melhores práticas de proteção Serviços da AWS e mapeiam as diretrizes para controles de segurança em várias estruturas (incluindo o Instituto Nacional de Padrões e Tecnologia (NIST), o Conselho de Padrões de Segurança do Setor de Cartões de Pagamento (PCI) e a Organização Internacional de Padronização (ISO)).
- Avaliação de recursos com regras no Guia do AWS Config desenvolvedor — O AWS Config serviço avalia o quão bem suas configurações de recursos estão em conformidade com as práticas internas, as diretrizes e os regulamentos do setor.
- AWS Security Hub— Isso AWS service (Serviço da AWS) fornece uma visão abrangente do seu estado de segurança interno AWS. O Security Hub usa controles de segurança para avaliar os recursos da AWS e verificar a conformidade com os padrões e as práticas recomendadas do setor de segurança. Para obter uma lista dos serviços e controles aceitos, consulte a Referência de controles do Security Hub.
- Amazon GuardDuty — Isso AWS service (Serviço da AWS) detecta possíveis ameaças às suas cargas de trabalho Contas da AWS, contêineres e dados monitorando seu ambiente em busca de atividades suspeitas e maliciosas. GuardDuty pode ajudá-lo a atender a vários requisitos de conformidade, como o PCI DSS, atendendo aos requisitos de detecção de intrusões exigidos por determinadas estruturas de conformidade.
- AWS Audit Manager— Isso AWS service (Serviço da AWS) ajuda você a auditar continuamente seu AWS uso para simplificar a forma como você gerencia o risco e a conformidade com as regulamentações e os padrões do setor.

AWS Entity Resolution melhores práticas de conformidade

Esta seção fornece as melhores práticas e recomendações para conformidade quando você usa AWS Entity Resolution.

Padrões de segurança de dados do setor de cartão de pagamento (PCI DSS)

AWS Entity Resolution suporta o processamento, armazenamento e transmissão de dados de cartão de crédito por um comerciante ou provedor de serviços e foi validado como compatível com o Padrão de Segurança de Dados (DSS) do Setor de Cartões de Pagamento (PCI). Para obter mais informações sobre o PCI DSS, incluindo como solicitar uma cópia do PCI AWS Compliance Package, consulte PCI DSS Nível 1.

Controles do Sistema e da Organização (CSO)

AWS Entity Resolution está em conformidade com as medidas de Controles do Sistema e da Organização (SOC), incluindo SOC 1, SOC 2 e SOC 3. Os relatórios do SOC são relatórios de exame independentes e terceirizados que demonstram como AWS alcança os principais controles e objetivos de conformidade. Essas auditorias garantem que os procedimentos e as proteções devidos sejam estabelecidos para minimizar os riscos que podem afetar a segurança, a confidencialidade, e a disponibilidade dos dados dos clientes e da empresa. Os resultados dessas auditorias terceirizadas estão disponíveis no site da AWS SOC Compliance, onde você pode ver os relatórios publicados para obter mais informações sobre os controles que dão suporte às AWS operações e à conformidade.

Resiliência em AWS Entity Resolution

A infraestrutura AWS global é construída em torno Regiões da AWS de zonas de disponibilidade. Regiões da AWS fornecem várias zonas de disponibilidade fisicamente separadas e isoladas, conectadas a redes de baixa latência, alta taxa de transferência e alta redundância. Com as zonas de disponibilidade, é possível projetar e operar aplicações e bancos de dados que automaticamente executam o failover entre as zonas sem interrupção. As zonas de disponibilidade são altamente disponíveis, tolerantes a falhas e escaláveis que uma ou várias infraestruturas de data center tradicionais.

Para obter mais informações sobre zonas de disponibilidade Regiões da AWS e zonas de disponibilidade, consulte Infraestrutura AWS global.

Além da infraestrutura AWS global, AWS Entity Resolution oferece vários recursos para ajudar a suportar suas necessidades de resiliência e backup de dados.

Monitoramento AWS Entity Resolution

O monitoramento é uma parte importante da manutenção da confiabilidade, disponibilidade e desempenho de AWS Entity Resolution suas outras AWS soluções. AWS fornece as seguintes ferramentas de monitoramento para observar AWS Entity Resolution, relatar quando algo está errado e realizar ações automáticas quando apropriado:

- AWS CloudTrailcaptura chamadas de API e eventos relacionados feitos por você ou em seu nome Conta da AWS e entrega os arquivos de log em um bucket do Amazon S3 que você especificar. Você pode identificar quais usuários e contas ligaram AWS, o IP de origem discute de onde as chamadas foram feitas e quando elas ocorreram. Para obter mais informações, consulte o Guia do usuário do AWS CloudTrail.
- O Amazon CloudWatch Logs permite que você verifique, armazene e acesse seus registros de EC2 instâncias da Amazon e de outras fontes. CloudTrail CloudWatch Os registros podem verificar as informações nos arquivos de log e informar quando determinados limites foram atingidos. É possível também arquivar seus dados de log em armazenamento resiliente. Para obter mais informações, consulte o Guia do usuário do Amazon CloudWatch Logs.

Tópicos

- Registrando chamadas de AWS Entity Resolution API usando AWS CloudTrail
- Monitoramento e registro de fluxos de trabalho usando Amazon CloudWatch Logs

Registrando chamadas de AWS Entity Resolution API usando AWS CloudTrail

AWS Entity Resolution é integrado com AWS CloudTrail, um serviço que fornece um registro das ações realizadas por um usuário, função ou AWS serviço em AWS Entity Resolution. CloudTrail captura todas as chamadas de API AWS Entity Resolution como eventos. As chamadas capturadas incluem chamadas do AWS Entity Resolution console e chamadas de código para as operações AWS Entity Resolution da API. Se você criar uma trilha, poderá habilitar a entrega contínua de CloudTrail eventos para um bucket do Amazon S3, incluindo eventos para. AWS Entity Resolution Se você não configurar uma trilha, ainda poderá ver os eventos mais recentes no CloudTrail console no Histórico de eventos. Usando as informações coletadas por CloudTrail, você pode determinar a

solicitação que foi feita AWS Entity Resolution, o endereço IP do qual a solicitação foi feita, quem fez a solicitação, quando ela foi feita e detalhes adicionais.

Para saber mais sobre isso CloudTrail, consulte o Guia AWS CloudTrail do usuário.

AWS Entity Resolution informações em CloudTrail

CloudTrail é ativado no seu Conta da AWS quando você cria a conta. Quando a atividade ocorre em AWS Entity Resolution, essa atividade é registrada em um CloudTrail evento junto com outros eventos AWS de serviço no histórico de eventos. Você pode exibir, pesquisar e baixar eventos recentes em sua Conta da AWS. Para obter mais informações, consulte Visualização de eventos com histórico de CloudTrail eventos.

Para um registro contínuo dos eventos em sua Conta da AWS, incluindo eventos para AWS Entity Resolution, crie uma trilha. Uma trilha permite CloudTrail entregar arquivos de log para um bucket do Amazon S3. Por padrão, quando você cria uma trilha no console, ela é aplicada a todas as Regiões da AWS. A trilha registra eventos de todas as regiões na AWS partição e entrega os arquivos de log ao bucket do Amazon S3 que você especificar. Além disso, você pode configurar outros AWS serviços para analisar e agir com base nos dados de eventos coletados nos CloudTrail registros. Para obter mais informações, consulte:

- Visão geral da criação de uma trilha
- CloudTrail serviços e integrações suportados
- Configurando notificações do Amazon SNS para CloudTrail
- Recebendo arquivos de CloudTrail log de várias regiões e Recebendo arquivos de CloudTrail log de várias contas

Todas AWS Entity Resolution as ações são registradas CloudTrail e documentadas na Referência da AWS Entity Resolution API.

Cada entrada de log ou evento contém informações sobre quem gerou a solicitação. As informações de identidade ajudam a determinar o seguinte:

- Se a solicitação foi feita com credenciais de usuário root ou AWS Identity and Access Management (IAM).
- Se a solicitação foi feita com credenciais de segurança temporárias de uma função ou de um usuário federado.

- Se a solicitação foi feita por outro AWS serviço.

Para obter mais informações, consulte Elemento userIdentity do CloudTrail .

Entendendo as entradas do arquivo de AWS Entity Resolution log

Uma trilha é uma configuração que permite a entrega de eventos como arquivos de log para um bucket do Amazon S3 que você especificar. CloudTrail os arquivos de log contêm uma ou mais entradas de log. Um evento representa uma única solicitação de qualquer fonte e inclui informações sobre a ação solicitada, a data e a hora da ação, os parâmetros da solicitação e assim por diante. CloudTrail os arquivos de log não são um rastreamento de pilha ordenado das chamadas públicas de API, portanto, eles não aparecem em nenhuma ordem específica.

Monitoramento e registro de fluxos de trabalho usando Amazon CloudWatch Logs

AWS Entity Resolution fornece recursos abrangentes de registro que ajudam você a verificar e analisar seus fluxos de trabalho de correspondência e mapeamento de ID. Por meio da integração com o Amazon CloudWatch Logs, você pode capturar informações detalhadas sobre a execução do fluxo de trabalho, incluindo tipos de eventos, registros de data e hora, estatísticas de processamento e contagens de erros. Você pode optar por entregar esses registros aos destinos CloudWatch Logs, Amazon S3 ou Amazon Data Firehose. Ao analisar esses registros, você pode avaliar o desempenho do serviço, solucionar problemas, obter informações sobre sua base de clientes e entender melhor seu AWS Entity Resolution uso e cobrança. Embora o registro esteja desativado por padrão, você pode ativá-lo para fluxos de trabalho novos e existentes por meio do console ou da API.

As taxas de CloudWatch venda padrão da Amazon se aplicam quando você ativa o registro em registros para AWS Entity Resolution fluxos de trabalho, incluindo custos associados à ingestão, armazenamento e análise de registros. Para obter informações detalhadas sobre preços, acesse a CloudWatch página de preços. .

Tópicos

- Configurando a entrega do log
- Desativando o registro (console)
- Lendo os registros

Configurando a entrega do log

Esta seção explicará as permissões necessárias para usar o AWS Entity Resolution registro e como habilitar a entrega de registros usando o console APIs e.

Tópicos

- Permissões
- Habilitando o registro em um novo fluxo de trabalho (console)
- Habilitando o registro para um novo fluxo de trabalho (API)
- Habilitando o registro em um fluxo de trabalho existente (console)

Permissões

AWS Entity Resolution usa registros CloudWatch vendidos para fornecer registros de fluxo de trabalho. Para entregar registros de fluxo de trabalho, você precisa de permissões para o destino de registro que você especificar.

Para ver as permissões necessárias para cada destino de registro, escolha um dos seguintes AWS serviços no Guia do usuário do Amazon CloudWatch Logs.

- CloudWatch Registros da Amazon
- Amazon Simple Storage Service (Amazon S3)
- Amazon Data Firehose

Para criar, visualizar ou alterar a configuração de login AWS Entity Resolution, você deve ter as permissões necessárias. Sua função do IAM deve incluir as seguintes permissões mínimas para gerenciar o registro do fluxo de trabalho no AWS Entity Resolution console.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowLogDeliveryActionsConsoleCWL",
            "Effect": "Allow",
            "Action": [
```

```
                "logs:DescribeLogGroups"
            ],
            "Resource": [
                "arn:aws:logs:us-east-1:111122223333:log-group:*"
            ]
        },
        {
            "Sid": "AllowLogDeliveryActionsConsoleS3",
            "Effect": "Allow",
            "Action": [
                "s3:ListAllMyBuckets",
                "s3:ListBucket",
                "s3:GetBucketLocation"
            ],
            "Resource": [
                "arn:aws:s3:::*"
            ]
        },
        {
            "Sid": "AllowLogDeliveryActionsConsoleFH",
            "Effect": "Allow",
            "Action": [
                "firehose:ListDeliveryStreams",
                "firehose:DescribeDeliveryStream"
            ],
            "Resource": [
                "*"
            ]
        }
    ]
}
```

Para obter mais informações sobre permissões para gerenciar o registro do fluxo de trabalho, consulte Habilitar o registro de AWS serviços no Guia do usuário do Amazon CloudWatch Logs.

Habilitando o registro em um novo fluxo de trabalho (console)

Depois de configurar as permissões para o destino do registro, você pode habilitar o registro para um novo fluxo de trabalho AWS Entity Resolution usando o console.

Para habilitar o registro em um novo fluxo de trabalho (console)

1. Abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/casa.
2. Em Fluxos de trabalho, selecione Fluxos de trabalho correspondentes ou fluxos de trabalho de mapeamento de ID.
3. Siga as etapas para criar um dos seguintes fluxos de trabalho:
   - Fluxo de trabalho de correspondência baseado em regras
   - Fluxo de trabalho de correspondência baseado em aprendizado de máquina
   - Fluxo de trabalho de correspondência baseado em serviços do provedor
   - Fluxo de trabalho de mapeamento de ID para uma conta
   - Fluxo de trabalho de mapeamento de ID em duas contas
4. Na Etapa 1 Especificar detalhes do fluxo de trabalho correspondente, em Entregas de registros — Registros EntityResolution de fluxo de trabalho, escolha Adicionar.
   - Escolha um dos seguintes destinos de registro.
     - Para Amazon CloudWatch Logs
     - Para o Amazon S3
     - Para o Amazon Data Firehose

     > **Tip**
     >
     > Se você escolher o Amazon S3 ou o Firehose, poderá enviar seus registros para uma conta Cross ou uma conta corrente In.
     > Para habilitar a entrega entre contas, ambas Contas da AWS devem ter as permissões necessárias. Para obter mais informações, consulte o exemplo de entrega entre contas no Guia do usuário do Amazon CloudWatch Logs.
5. Para o grupo de registros de destino, os grupos de registros prefixados com '/aws/vendedlogs/' são criados automaticamente. Se você estiver usando outros grupos de registros, use-os antes de configurar uma entrega de registros. Para obter mais informações, consulte Como trabalhar com grupos e fluxos de registros no Guia do usuário do Amazon CloudWatch Logs.
6. Para Mais configurações - opcional, escolha o seguinte:

a. Em Seleção de campo, selecione os campos de registro a serem incluídos em cada registro de registro.

b. (CloudWatch Registros) Em Formato de saída, escolha o formato de saída para o registro.

c. Em Delimitador de campo, escolha como separar cada campo de registro.

d. (Amazon S3) Em Suffix, especifique o caminho do sufixo para particionar seus dados.

e. (Amazon S3) Para ser compatível com o Hive, escolha Enable se quiser usar caminhos do S3 compatíveis com o Hive.

7. Para criar outro destino de registro, escolha Adicionar e repita as etapas 4 a 6.
8. Conclua as etapas restantes para configurar e executar o fluxo de trabalho.
9. Depois que os trabalhos do fluxo de trabalho forem concluídos, verifique os registros do fluxo de trabalho no destino de entrega do registro que você especificou.

Habilitando o registro para um novo fluxo de trabalho (API)

Depois de configurar as permissões para o destino de registro, você pode habilitar o registro para um novo fluxo de trabalho AWS Entity Resolution usando o Amazon CloudWatch Logs APIs.

Para habilitar o registro em um novo fluxo de trabalho (API)

1. Depois de criar um fluxo de trabalho no AWS Entity Resolution console, obtenha o Amazon Resource Name (ARN) do fluxo de trabalho.

   Você pode encontrar o ARN na página do fluxo de trabalho no AWS Entity Resolution console ou chamar a operação `GetMatchingWorkflow` ou `GetIdMappingWorkflow` API.

   O ARN do fluxo de trabalho segue esse formato:

   ```
   arn:(aws|aws-us-gov|aws-cn):entityresolution:[a-z]{2}-[a-z]{1,10}-
   [0-9]:[0-9]{12}:(matchingworkflow/[a-zA-Z_0-9-]{1,255})
   ```

   Um ARN de mapeamento de ID segue esse formato:

   ```
   arn:(aws|aws-us-gov|aws-cn):entityresolution:[a-z]{2}-[a-z]{1,10}-
   [0-9]:[0-9]{12}:(idmappingworkflow/[a-zA-Z_0-9-]{1,255})
   ```

   Para obter mais informações, consulte GetMatchingWorkflow ou GetIdMappingWorkflow na Referência de APIs do AWS Entity Resolution .

2. Use a operação da `PutDeliverySource` API CloudWatch Logs para criar uma fonte de entrega para os registros do fluxo de trabalho.

   Para obter mais informações, consulte PutDeliverySourcea Referência da API Amazon CloudWatch Logs.

   a. Passe `resourceArn` o.

   b. Para`logType`, os tipos de registros que são coletados são`WORKFLOW_LOGS`:

   Example

   Exemplo de operação de `PutDeliverySource` API

```
{
    "logType": "WORKFLOW_LOGS",
    "name": "my-delivery-source",
    "resourceArn": "arn:aws:entityresolution:region:accoungId:matchingworkflow/
XXXWorkflow"
}
```

3. Use a operação `PutDeliveryDestination` da API para configurar onde armazenar seus registros.

   Você pode escolher CloudWatch Logs, Amazon S3 ou Firehose como destino. Você deve especificar o ARN de uma das opções de destino para onde seus registros serão armazenados.

   Para obter mais informações, consulte PutDeliveryDestinationa Referência da API Amazon CloudWatch Logs.

   Example

   Exemplo de operação de `PutDeliveryDestination` API

```
{
   "delivery-destination-configuration": {
      "destinationResourceArn": "arn:aws:logs:region:accountId:log-group:my-log-
group"
   },
   "name": "my-delivery-destination",
   "outputFormat": "json",
   }
```

```
}
```

> **Note**
>
> Se você estiver entregando registros entre contas, deverá usar a PutDeliveryDestinationPolicyAPI para atribuir uma política AWS Identity and Access Management (IAM) à conta de destino. A política do IAM permite a entrega de uma conta para outra.

4. Use a operação da `CreateDelivery` API para vincular a fonte de entrega ao destino que você criou nas etapas anteriores. Essa operação de API associa a fonte de entrega ao destino final.

   Para obter mais informações, consulte PutDeliveryDestinationa Referência da API Amazon CloudWatch Logs.

   Example

   Exemplo de operação de `CreateDelivery` API

   ```
   {
      "delivery-destination-arn": "arn:aws:logs:region:accountId:log-group:my-log-group",
      "delivery-source-name": "my-delivery-source",
      "tags": {
         "string" : "string"
      }
   }
   ```

5. Executar o fluxo de trabalho.
6. Depois que os trabalhos do fluxo de trabalho forem concluídos, verifique os registros do fluxo de trabalho no destino de entrega do registro que você especificou.

Habilitando o registro em um fluxo de trabalho existente (console)

Depois de configurar as permissões para o destino do registro, você pode habilitar o registro para um fluxo de trabalho existente AWS Entity Resolution usando a guia Entregas de registros no console.

Para habilitar o registro em um fluxo de trabalho existente usando a guia Entregas de registros (console)

1. Abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/casa.
2. Em Fluxos de trabalho, selecione Fluxos de trabalho correspondentes ou fluxos de trabalho de mapeamento de ID e, em seguida, selecione seu fluxo de trabalho existente.
3. Na guia Entregas de registros, em Entrega de registros, selecione Adicionar e escolha um dos seguintes destinos de registro.

   - Para Amazon CloudWatch Logs
   - Para o Amazon S3
     - Entre contas
     - Na conta atual
   - Para o Amazon Data Firehose
     - Entre contas
     - Na conta atual

   > **Tip**
   >
   > Se você escolher o Amazon S3 ou o Firehose, poderá enviar seus registros para uma conta Cross ou uma conta corrente In.
   > Para habilitar a entrega entre contas, ambas Contas da AWS devem ter as permissões necessárias. Para obter mais informações, consulte o exemplo de entrega entre contas no Guia do usuário do Amazon CloudWatch Logs.

4. No modal, faça o seguinte, dependendo do tipo de entrega de log que você escolheu.

   a. Veja o tipo de registro: WORKFLOW_LOGS.

      O tipo de registro não pode ser alterado.

   b. (CloudWatch Registros) Para o grupo de registros de destino, os grupos de registros prefixados com '/aws/vendedlogs/' são criados automaticamente. Se você estiver usando outros grupos de registros, use-os antes de configurar uma entrega de registros. Para obter mais informações, consulte Como trabalhar com grupos e fluxos de registros no Guia do usuário do Amazon CloudWatch Logs.

(Amazon S3 na conta corrente) Para o bucket S3 de destino, selecione um bucket ou insira um ARN.

(Conta cruzada do Amazon S3) Para ARN de destino de entrega, insira um ARN de destino de entrega.

(Firehose na conta atual) Em Fluxo de entrega de destino, insira o ARN do recurso de destino de entrega que foi criado em outra conta.

(Conta cruzada Firehose) Em ARN de destino de entrega, insira um ARN de destino de entrega.

5. Para Mais configurações - opcional, escolha o seguinte:
   a. Em Seleção de campo, selecione os campos de registro a serem incluídos em cada registro de registro.
   b. (CloudWatch Registros) Em Formato de saída, escolha o formato de saída para o registro.
   c. Em Delimitador de campo, escolha como separar cada campo de registro.
   d. (Amazon S3) Em Suffix, especifique o caminho do sufixo para particionar seus dados.
   e. (Amazon S3) Para ser compatível com o Hive, escolha Enable se quiser usar caminhos do S3 compatíveis com o Hive.
6. Escolha Adicionar.
7. Na página do fluxo de trabalho, escolha Executar.
8. Depois que os trabalhos do fluxo de trabalho forem concluídos, verifique os registros do fluxo de trabalho no destino de entrega do registro que você especificou.

Desativando o registro (console)

Você pode desativar o registro em log para seu AWS Entity Resolution fluxo de trabalho a qualquer momento no console.

Para desativar o registro do fluxo de trabalho (console)

1. Abra o AWS Entity Resolution console em https://console.aws.amazon.com/entityresolution/casa.
2. Em Fluxos de trabalho, selecione Fluxos de trabalho correspondentes ou fluxos de trabalho de mapeamento de ID e, em seguida, selecione seu fluxo de trabalho.

3. Na guia Registrar entregas, em Registrar entrega, selecione o destino e escolha Excluir.
4. Revise suas alterações e, em seguida, navegue até a próxima etapa para salvá-las.

Lendo os registros

A leitura do Amazon CloudWatch Logs ajuda você a manter AWS Entity Resolution fluxos de trabalho eficientes. Os registros fornecem visibilidade detalhada da execução do fluxo de trabalho, incluindo métricas importantes, como o número de registros processados e quaisquer erros encontrados, ajudando você a garantir que o processamento de dados ocorra sem problemas. Além disso, os registros oferecem rastreamento em tempo real da progressão do fluxo de trabalho por meio de registros de data e hora e tipos de eventos, permitindo que você identifique rapidamente gargalos ou problemas em seu pipeline de processamento de dados. As informações abrangentes de rastreamento de erros e contagem de registros ajudam a manter a qualidade e a integridade dos dados, mostrando exatamente quantos registros foram processados com sucesso e se algum permaneceu sem processamento.

Se você estiver usando o CloudWatch Logs como destino, poderá usar o CloudWatch Logs Insights para ler os registros do fluxo de trabalho. Aplicam-se taxas típicas de CloudWatch registros. Para obter mais informações, consulte Análise de dados de log com o CloudWatch Logs Insights no Guia do usuário do Amazon CloudWatch Logs.

> **Note**
>
> Os registros do fluxo de trabalho podem levar alguns minutos para aparecer no seu destino. Se você não vê os registros, aguarde alguns minutos e atualize a página.

Os registros do fluxo de trabalho consistem em uma sequência de registros de log formatados, em que cada registro representa um fluxo de trabalho. A ordem dos campos dentro do log pode variar.

```
{
  "resource_arn": "arn:aws:ses:us-east-1:1234567890:mailmanager-ingress-point/inp-
xxxxx",
  "event_type": "JOB_START",
  "event_timestamp": 1728562395042,
  "job_id": "b01eea4678d4423a4b43eeada003f6",
  "workflow_name": "TestWorkflow",
  "workflow_start_time": "2025-03-11 10:19:56",
```

```
  "data_procesing_progression": "Matching Job Starts ...",
  "total_records_processed": 1500,
  "total_records_unprocessed": 0,
  "incremental_records_processed": 0,
  "error_message": "sample error that caused workflow failure"
}
```

A lista a seguir descreve os campos de registro de log, em ordem:

`resource_arn`

O Amazon Resource Name (ARN) que identifica de forma exclusiva o AWS recurso que está sendo usado no fluxo de trabalho.

`event_type`

O tipo de evento que ocorreu durante a execução do fluxo de trabalho. AWS Entity Resolution atualmente suporta:

`JOB_START`

`DATA_PROCESSING_STEP_START`

`DATA_PROCESSING_STEP_END`

`JOB_SUCCESS`

`JOB_FAILURE`

`event_timestamp`

O timestamp Unix indicando quando o evento ocorreu durante o fluxo de trabalho.

`job_id`

Um identificador exclusivo atribuído à execução específica da tarefa do fluxo de trabalho.

`workflow_name`

O nome dado ao fluxo de trabalho que está sendo executado.

`workflow_start_time`

A data e a hora em que a execução do fluxo de trabalho começou.

`data_procesing_progression`

Uma descrição do estágio atual no fluxo de trabalho de processamento de dados. Exemplos: `"Matching Job Starts"`, `"Loading Step Starts"`, `"ID_Mapping Job Ends Successfully"`.

`total_records_processed`

O número total de registros que foram processados com sucesso durante o fluxo de trabalho.

`total_records_unprocessed`

O número de registros que não foram processados durante a execução do fluxo de trabalho.

`incremental_records_processed`

O número de novos registros processados em uma atualização incremental do fluxo de trabalho.

`error_message`

A causa raiz da falha no fluxo de trabalho.

Crie recursos de resolução de entidades da AWS com AWS CloudFormation

O AWS Entity Resolution é integrado com AWS CloudFormation um serviço que ajuda você a modelar e configurar seus AWS recursos para que você possa gastar menos tempo criando e gerenciando seus recursos e infraestrutura. Você cria um modelo que descreve todos os AWS recursos que você deseja (como AWS::EntityResolution::MatchingWorkflow, AWS::EntityResolution::SchemaMapping, AWS::EntityResolution:IdMappingWorkflow, AWS::EntityResolution::IdNamespace e AWS::EntityResolution::PolicyStatement) e AWS CloudFormation provisiona e configura esses recursos para você.

Ao usar AWS CloudFormation, você pode reutilizar seu modelo para configurar seus recursos de resolução de entidades da AWS de forma consistente e repetida. Descreva seus recursos uma vez e, em seguida, provisione os mesmos recursos repetidamente em várias Contas da AWS regiões.

Resolução e AWS CloudFormation modelos de entidades da AWS

Para provisionar e configurar recursos para a resolução de entidades da AWS e serviços relacionados, você deve entender AWS CloudFormation os modelos. Os modelos são arquivos de texto formatados em JSON ou YAML. Esses modelos descrevem os recursos que você deseja provisionar em suas AWS CloudFormation pilhas. Se você não estiver familiarizado com JSON ou YAML, você pode usar o AWS CloudFormation Designer para ajudá-lo a começar a usar modelos. AWS CloudFormation Para obter mais informações, consulte O que é o AWS CloudFormation Designer? no Guia do usuário do AWS CloudFormation .

A resolução de entidades da AWS oferece suporte à criação AWS::EntityResolution::MatchingWorkflow, AWS::EntityResolution::SchemaMapping, AWS::EntityResolution:IdMappingWorkflow, AWS::EntityResolution::IdNamespace e AWS::EntityResolution::PolicyStatement à entrada AWS CloudFormation. Para obter mais informações, incluindo exemplos de modelos JSON e YAML para AWS::EntityResolution::MatchingWorkflow, AWS::EntityResolution::SchemaMapping, AWS::EntityResolution:IdMappingWorkflow, AWS::EntityResolution::IdNamespace e AWS::EntityResolution::PolicyStatement, consulte a referência do tipo de recurso de resolução de entidades da AWS no Guia do AWS CloudFormation usuário.

Os seguintes modelos estão disponíveis:

- Fluxo de trabalho correspondente

  Crie um `MatchingWorkflow` objeto que armazene a configuração da tarefa de processamento de dados a ser executada.

  Para obter mais informações, consulte os tópicos a seguir.

  AWS::EntityResolution::MatchingWorkflow no AWS CloudFormation Guia do usuário

  CreateMatchingWorkflow na Referência de API do AWS Entity Resolution
- Mapeamento de esquemas

  Crie um mapeamento de esquema, que define o esquema da tabela de registros do cliente de entrada.

  Para obter mais informações, consulte os tópicos a seguir.

  AWS::EntityResolution::SchemaMapping no AWS CloudFormation Guia do usuário

  CreateSchemaMapping na Referência de API do AWS Entity Resolution
- Workflow de mapeamento de ID

  Crie um `IdMappingWorkflow` objeto que armazene a configuração da tarefa de processamento de dados a ser executada.

  Para obter mais informações, consulte os tópicos a seguir.

  AWS::EntityResolution::IdMappingWorkflow no AWS CloudFormation Guia do usuário

  CreateIdMappingWorkflow na Referência de API do AWS Entity Resolution
- Namespace de ID

  Crie um `IdNamespace` objeto, que armazene os metadados explicando o conjunto de dados e como usá-lo.

  Para obter mais informações, consulte os tópicos a seguir.

  AWS::EntityResolution::IdNamespace no AWS CloudFormation Guia do usuário

  CreateIdNamespace na Referência de API do AWS Entity Resolution
- PolicyStatement

Crie um objeto `PolicyStatement`.

Para obter mais informações, consulte os tópicos a seguir.

AWS::EntityResolution::PolicyStatement no AWS CloudFormation Guia do usuário

AddPolicyStatement na Referência de API do AWS Entity Resolution

Saiba mais sobre AWS CloudFormation

Para saber mais sobre isso AWS CloudFormation, consulte os seguintes recursos:

- AWS CloudFormation
- AWS CloudFormation Guia do usuário
- Referência de API do AWS CloudFormation
- AWS CloudFormation Guia do usuário da interface de linha de comando

Cotas para AWS Entity Resolution

Você Conta da AWS tem cotas padrão, anteriormente chamadas de limites, para cada um. AWS service (Serviço da AWS) A menos que especificado de outra forma, cada cota é específica da região. Você pode solicitar aumentos para algumas cotas, mas outras não podem ser aumentadas.

Para ver as cotas de AWS Entity Resolution, abra o console Service Quotas. No painel de navegação, escolha Serviços AWS e selecione AWS Entity Resolution.

Para solicitar o aumento da cota, consulte Solicitar um aumento de cota no Guia do usuário do Service Quotas. Se a cota ainda não estiver disponível no serviço de cotas, use o formulário de aumento de limite.

Você Conta da AWS tem as seguintes cotas relacionadas a. AWS Entity Resolution

Name	Padrão	Ajustável	Descrição
Trabalhos simultâneos de mapeamento de ID	1	Não	O número máximo de trabalhos de mapeamento de ID que podem ser processados simultaneamente no atual Região da AWS.
Trabalhos correspondentes simultâneos	1	Não	O número máximo de trabalhos correspondentes que podem ser processados simultaneamente no atual Região da AWS.
Trabalhos simultâneos de correspondência de serviços do provedor	1	Não	O número máximo de trabalhos correspondentes ao serviço do provedor que podem ser processados simultaneamente no atual Região da AWS.
Entrada de dados	20	Não	Esta é a lista de tabelas de entrada que você deseja usar em um fluxo de trabalho de correspondência. Cada entrada corresponde a uma coluna em sua tabela AWS Glue de dados de

Name	Padrão	Ajustável	Descrição
			entrada, que contém o nome da coluna e informações adicionais que são AWS Entity Resolution usadas para fins de correspondência. As entradas devem conter uma ID exclusiva mais pelo menos um campo de entrada adicional.
Dados de saída	750	Não	Essa é uma lista de `OutputAttribute` objetos, cada um com os campos Nome e Hashed. Cada um desses objetos representa uma coluna a ser incluída na tabela AWS Glue de saída e se você deseja que os valores na coluna sejam criptografados.
Esquema de dados	25	Não	O número máximo de campos de entrada do esquema de dados.
fluxos de trabalho de mapeamento de ID	10	Sim	O número máximo de fluxos de trabalho de mapeamento de ID que você pode criar Conta da AWS neste momento Região da AWS.
namespaces de ID	10	Sim	O número máximo de namespaces de ID que você pode criar Conta da AWS neste momento. Região da AWS
Partida IDs	500	Não	O número máximo de registros que podem ser consolidados em um MatchID por carga de trabalho.

Name	Padrão	Ajustável	Descrição
Regra de correspondência	15	Não	Para correspondência baseada em regras, esse é o número da regra aplicada que gerou um conjunto de registros correspondente. Isso faz parte da correspondência dos metadados do fluxo de trabalho que serão incluídos na saída.
Fluxos de trabalho correspondentes	10	Sim	O número máximo de fluxos de trabalho de correspondência.
Taxa de solicitações de API `GetMatchId`	50	Sim	O número máximo de solicitações de `GetCustomerID` API por segundo.
Registros por fluxo de trabalho baseado em aprendizado de máquina	250 M	Sim	O número máximo de registros que podem ser processados por um fluxo de trabalho de correspondência baseado em aprendizado de máquina.
Registros por fluxo de trabalho de correspondência baseado em regras	100M	Sim	O número máximo de registros que podem ser processados por um fluxo de trabalho de correspondência baseado em regras.
Regras por fluxo de trabalho	15	Não	O número máximo de regras por fluxo de trabalho de correspondência.
Mapeamentos de esquema	50	Sim	O número máximo de mapeamentos de esquema que você pode criar nessa conta na região atual. AWS

Name	Padrão	Ajustável	Descrição
Chaves de correspondência exclusivas por conjunto de regras	15	Não	O número máximo de chaves de correspondência exclusivas por conjunto de regras. Uma chave de correspondência instrui AWS Entity Resolution quais campos de entrada devem ser considerados como dados semelhantes e quais devem ser considerados como dados diferente s. Isso ajuda a configurar AWS Entity Resolution automaticamente as regras de correspondência baseadas em regras e a comparar dados semelhant es armazenados em diferentes campos de entrada.

Cotas de controle de utilização da API

Recurso	Limite de taxa	Descrição
Taxa de solicitações `CreateMatchingWork flow`	5 TPS	Número máximo de chamadas de `CreateMatchingWork flow` API por segundo.
Taxa de solicitações `DeleteMatchingWork flow`	5 TPS	Número máximo de chamadas de `DeleteMatchingWork flow` API por segundo.
Taxa de solicitações `GetMatchingWorkflow`	5 TPS	Número máximo de chamadas de `GetMatchingWorkflo w` API por segundo.
Taxa de solicitações `ListMatchingWorkfl ows`	5 TPS	Número máximo de chamadas de `ListMatchingWorkfl ows` API por segundo.

Recurso	Limite de taxa	Descrição
Taxa de solicitações `UpdateMatchingWorkflow`	5 TPS	Número máximo de chamadas de `UpdateMatchingWorkflow` API por segundo.
Taxa de solicitações `CreateSchemaMapping`	5 TPS	Número máximo de chamadas de `CreateSchemaMapping` API por segundo.
Taxa de solicitações `DeleteSchemaMapping`	5 TPS	Número máximo de chamadas de `DeleteSchemaMapping` API por segundo.
Taxa de solicitações `GetSchemaMapping`	5 TPS	Número máximo de chamadas de `GetSchemaMapping` API por segundo.
Taxa de solicitações `ListSchemaMappings`	5 TPS	Número máximo de chamadas de `ListSchemaMappings` API por segundo.
Taxa de solicitações `UpdateSchemaMapping`	5 TPS	Número máximo de chamadas de `UpdateSchemaMapping` API por segundo.
Taxa de solicitações `GetPartnerComponent`	5 TPS	Número máximo de chamadas de `GetPartnerComponent` API por segundo.
Taxa de solicitações `ListPartnerComponents`	5 TPS	Número máximo de chamadas de `ListPartnerComponents` API por segundo.
Taxa de solicitações `TagResource`	5 TPS	Número máximo de chamadas de `TagResource` API por segundo.

Recurso	Limite de taxa	Descrição
Taxa de solicitações `UntagResource`	5 TPS	Número máximo de chamadas de `UntagResource` API por segundo.
Taxa de solicitações `ListTagsForResource`	5 TPS	Número máximo de chamadas de `ListTagsForResource` API por segundo.
Taxa de solicitações `CreateIdMappingWorkflow`	5 TPS	Número máximo de chamadas de `CreateIdMappingWorkflow` API por segundo.
Taxa de solicitações `DeleteIdMappingWorkflow`	5 TPS	Número máximo de chamadas de `DeleteIdMappingWorkflow` API por segundo.
Taxa de solicitações `GetIdMappingWorkflow`	5 TPS	Número máximo de chamadas de `GetIdMappingWorkflow` API por segundo.
Taxa de solicitações `ListIdMappingWorkflow`	5 TPS	Número máximo de chamadas de `ListIdMappingWorkflow` API por segundo.
Taxa de solicitações `UpdateIdMappingWorkflow`	5 TPS	Número máximo de chamadas de `UpdateIdMappingWorkflow` API por segundo.
Taxa de solicitações `ListProviderServices`	5 TPS	Número máximo de chamadas de `ListProviderServices` API por segundo.
Taxa de solicitações `GetProviderService`	5 TPS	Número máximo de chamadas de `GetProviderService` API por segundo.

Recurso	Limite de taxa	Descrição
Taxa de solicitações `CreateIdNamespace`	5 TPS	Número máximo de chamadas de `CreateIdNamespace` API por segundo.
Taxa de solicitações `DeleteIdNamespace`	5 TPS	Número máximo de chamadas de `DeleteIdNamespace` API por segundo.
Taxa de solicitações `GetIdNamespace`	5 TPS	Número máximo de chamadas de `GetIdNamespace` API por segundo.
Taxa de solicitações `ListIdNamespaces`	5 TPS	Número máximo de chamadas de `ListIdNamespaces` API por segundo.
Taxa de solicitações `UpdateIdNamespace`	5 TPS	Número máximo de chamadas de `UpdateIdNamespace` API por segundo.
Taxa de solicitações `AddPolicyStatement`	5 TPS	Número máximo de chamadas de `AddPolicyStatement` API por segundo.
Taxa de solicitações `DeletePolicyStatement`	5 TPS	Número máximo de chamadas de `DeletePolicyStatement` API por segundo.
Taxa de solicitações `GetPolicy`	5 TPS	Número máximo de chamadas de `GetPolicy` API por segundo.
Taxa de solicitações `PutPolicy`	5 TPS	Número máximo de chamadas de `PutPolicy` API por segundo.

Recurso	Limite de taxa	Descrição
Taxa de solicitações `GetMatchingJob`	10 TPS	Número máximo de chamadas de `GetMatchingJob` API por segundo.
Taxa de solicitações `ListMatchingJobs`	5 TPS	Número máximo de chamadas de `ListMatchingJobs` API por segundo.
Taxa de solicitações `StartMatchingJob`	5 TPS	Número máximo de chamadas de `StartMatchingJob` API por segundo.
Taxa de solicitações `GetMatchId`	50 TPS	Número máximo de chamadas de `GetMatchId` API por segundo.
Taxa de solicitações `GetIdMappingJob`	10 TPS	Número máximo de chamadas de `GetIdMappingJob` API por segundo.
Taxa de solicitações `ListIdMappingJobs`	5 TPS	Número máximo de chamadas de `ListIdMappingJobs` API por segundo.
Taxa de solicitações `StartIdMappingJob`	5 TPS	Número máximo de chamadas de `StartIdMappingJob` API por segundo.
Taxa de solicitações `BatchDeleteUniqueId`	5 TPS	Número máximo de chamadas de `BatchDeleteUniqueId` API por segundo.

Histórico de documentos para o Guia AWS Entity Resolution do usuário

A tabela a seguir descreve as versões de documentação do AWS Entity Resolution.

Para receber notificações sobre atualizações dessa documentação, você pode se inscrever em o feed RSS. Para assinar as atualizações de RSS, você deve ter um plug-in de RSS habilitado para o navegador que está usando.

Alteração	Descrição	Data
Support para condições de regras aprimoradas e exclusões incrementais	Agora, os clientes podem usar condições de regra com operadores booleanos e novas funções de correspondência ExactManyToMany, como, por exemplo, permitir critérios de correspondência mais precisos com combinações de correspondência exata e difusa. Além disso, os clientes podem excluir registros de forma incremental em fluxos de trabalho de correspon dência avançada usando um arquivo Amazon S3.	30 de julho de 2025
Esclarecimento sobre o processamento do Match ID	Foi adicionado o esclareci mento de que as opções Modificar ou gerar ID de correspondência e Pesquisar ID de correspondência exigem cadência de processamento automática em fluxos de trabalho correspondentes.	17 de julho de 2025

Gere um novo ID de partida	Agora, os clientes podem pesquisar e modificar uma ID de correspondência existente ou gerar uma nova ID de correspondência ao usar um fluxo de trabalho de correspondência baseado em regras.	2 de junho de 2025
Fluxo de trabalho de correspondência baseado em serviços do provedor — atualização	Agora, os clientes podem usar identificadores digitais como IPV4 IPV6, e MAID ao usar o fluxo de trabalho de correspondência baseado em serviços do TransUnion provedor.	21 de abril de 2025
CloudWatch Registros da Amazon	AWS Entity Resolution agora oferece suporte à integração de CloudWatch registros, permitindo que você habilite o registro detalhado do fluxo de trabalho que captura métricas de execução de trabalhos, tempo e estatísticas de processamento que podem ser entregues aos destinos do CloudWatch Logs, Amazon S3 ou Amazon Data Firehose.	14 de abril de 2025
Fluxo de trabalho de mapeamento de ID — atualização	Agora, os clientes podem configurar o AWS Glue particionamento ao usar um fluxo de trabalho de mapeamento de ID.	25 de março de 2025

Cotas — atualização	Atualização somente da documentação. Os fluxos de trabalho de correspon dência baseados em regras podem processar até 100 milhões de registros, enquanto os fluxos de trabalho de correspondência baseados em aprendizado de máquina podem processar até 250 milhões de registros. Os clientes que precisam de limites mais altos devem entrar em contato com a equipe de atendimento.	7 de fevereiro de 2025
Mapeamento do esquema — atualização	Atualização somente com documentação para esclarece r que a normalização é compatível com os tipos de atributos Nome completo, Endereço completo e Telefone completo.	17 de janeiro de 2025
Integração com provedores	Atualização somente da documentação. Os clientes podem aprender como se integrar como provedor de serviços com AWS Entity Resolution.	8 de agosto de 2024
Fluxo de trabalho de mapeamento de ID — atualização	Agora, os clientes podem usar regras de correspondência para traduzir dados primários em um fluxo de trabalho de mapeamento de ID.	23 de julho de 2024

Fluxo de trabalho correspondente — atualização	Agora, os clientes podem excluir os registros de um fluxo de trabalho de correspondência baseado em regras ou em ML para ajudar a cumprir os regulamentos de gerenciamento de dados.	8 de abril de 2024
Fluxo de trabalho de mapeamento de ID — atualização	Agora, os clientes podem usar um fluxo de trabalho de mapeamento de ID em vários Contas da AWS.	2 de abril de 2024
AWS CloudFormation Recursos - Recursos novos e atualizados	AWS Entity Resolution adicionou os seguintes recursos: `AWS::EntityResolution::IdNamespace` `AWS::EntityResolution::PolicyStatement` e atualizou o seguinte recurso:`AWS::EntityResolution::IdMappingWorkflow` .	2 de abril de 2024
Encontre o ID da partida	Agora, os clientes podem encontrar o Match ID correspondente e a regra associada para um fluxo de trabalho processado baseado em regras.	25 de março de 2024

Fluxo de trabalho correspondente — atualização	AWS Entity Resolution agora oferece suporte à atribuição de RAMPID baseada em PII no fluxo de trabalho de correspondência baseado em LiveRamp serviços do provedor.	12 de fevereiro de 2024
AWS PrivateLink	AWS Entity Resolution agora oferece suporte adicional à segurança de dados, o AWS PrivateLink que ajuda os clientes a acessar de forma privada os serviços hospedados em AWS.	20 de outubro de 2023
AWS CloudFormation Recursos — Recursos novos e atualizados	AWS Entity Resolution adicionou o seguinte recurso: `AWS::EntityResolution:IdMappingWorkflow` e atualizou os seguintes recursos: `AWS::EntityResolution::MatchingWorkflow` `AWS::EntityResolution::Schemamapping` e.	19 de outubro de 2023
Atualizar a política existente	As seguintes novas permissões foram adicionadas à política gerenciada: `AWSEntityResolutionConsoleFullAccess` `ADXReadAccess` `ManageEventBridgeRules` e.	16 de outubro de 2023

Mapeamento do esquema — atualização	Agora, os clientes podem editar e atualizar um esquema de dados existente.	16 de outubro de 2023
Fluxo de trabalho correspondente — atualização	Agora, os clientes podem selecionar um serviço de provedor de dados preferencial para ajudar a combinar e vincular seus dados.	16 de outubro de 2023
Workflow de mapeamento de ID	Os clientes podem usar esse novo fluxo de trabalho para especificar detalhes do mapeamento de ID, escolher o método de mapeamento de ID desejado e especificar campos de entrada e saída de dados.	16 de outubro de 2023
AWS CloudFormation integração	AWS Entity Resolution agora se integra com AWS CloudFormation.	24 de agosto de 2023
AWS atualização de política gerenciada - Novas políticas	AWS Entity Resolution adicionou duas novas políticas gerenciadas.	18 de agosto de 2023
Lançamento inicial	Versão inicial do Guia AWS Entity Resolution do usuário	26 de julho de 2023

AWS Entity Resolution Glossário

Nome do recurso da Amazon (ARN)

Um identificador exclusivo para AWS recursos. ARNs são necessários quando você precisa especificar um recurso de forma inequívoca em todos eles AWS Entity Resolution, como em AWS Entity Resolution políticas, tags do Amazon Relational Database Service (Amazon RDS) e chamadas de API.

Tipo de atributo

O tipo do atributo para o campo de entrada. Ao criar um mapeamento de esquema, você seleciona o tipo de atributo em uma lista pré-configurada de valores, como nome, endereço, número de telefone ou endereço de e-mail. AWS Entity Resolution O tipo de atributo informa que tipo de dados você está apresentando, permitindo que sejam classificados e normalizados adequadamente.

Processamento automático

Uma opção de cadência de processamento para uma tarefa de fluxo de trabalho correspondente que permite que ela seja executada automaticamente quando a entrada de dados é alterada.

Essa opção está disponível somente para correspondência baseada em regras.

Por padrão, a cadência de processamento de uma tarefa de fluxo de trabalho correspondente é definida como Manual, o que permite que ela seja executada sob demanda. Você pode configurar o processamento automático para executar automaticamente sua tarefa de fluxo de trabalho correspondente quando a entrada de dados for alterada. Isso mantém a saída correspondente do fluxo de trabalho up-to-date.

AWS KMS key ARN

Este é o seu nome de recurso AWS KMS da Amazon (ARN) para criptografia em repouso. Se não for fornecido, o sistema usará uma chave KMS AWS Entity Resolution gerenciada.

Texto não criptografado

Dados que não estão protegidos criptograficamente.

Nível de confiança (ConfidenceLevel)

Para correspondência de ML, esse é o nível de confiança aplicado AWS Entity Resolution quando o ML identifica um conjunto de registros correspondente. Isso faz parte dos metadados correspondentes do fluxo de trabalho que serão incluídos na saída.

Descriptografia

O processo de transformar dados criptografados de volta à sua forma original. Só será possível realizar se você tiver acesso à chave secreta.

Criptografia

O processo de codificação de dados em um formato que parece aleatório usando um valor secreto chamado chave. É impossível determinar o texto sem formatação original sem acesso à chave.

Group name

O nome do grupo faz referência a todo o grupo de campos de entrada e pode ajudá-lo a agrupar dados analisados para fins de correspondência.

Por exemplo, se houver três campos de entrada: **`first_namemiddle_name`**,, e**`last_name`**, você pode agrupá-los inserindo o nome do grupo **`full_name`** para correspondência e saída.

Hash

O hashing significa aplicar um algoritmo criptográfico que produz uma sequência irreversível e exclusiva de caracteres de tamanho fixo, chamada de hash. AWS Entity Resolution usa o protocolo de hash Secure Hash Algorithm de 256 bits (SHA256) e produzirá uma cadeia de caracteres de 32 bytes. Em AWS Entity Resolution, você pode escolher se deseja fazer o hash dos valores de dados em sua saída.

Protocolo de hash () HashingProtocol

AWS Entity Resolution usa o protocolo de hash Secure Hash Algorithm de 256 bits (SHA256) e produzirá uma cadeia de caracteres de 32 bytes. Isso faz parte dos metadados correspondentes do fluxo de trabalho que serão incluídos na saída.

Método de mapeamento de ID

Como você deseja que o mapeamento de ID seja executado.

Há dois métodos de mapeamento de ID:

- Baseado em regras — O método pelo qual você usa regras de correspondência para traduzir dados primários de uma fonte para um destino em um fluxo de trabalho de mapeamento de ID.
- Serviços do provedor — O método pelo qual você usa um serviço do provedor para traduzir dados codificados por terceiros de uma fonte para um destino em um fluxo de trabalho de mapeamento de ID.

  AWS Entity Resolution atualmente é compatível com LiveRamp o método de mapeamento de ID baseado em serviços do provedor. Você deve ter uma assinatura LiveRamp até o AWS Data Exchange fim para usar esse método. Para obter mais informações, consulte Etapa 1: Assine um serviço de provedor em AWS Data Exchange.

Fluxo de trabalho de mapeamento de ID

Um trabalho de processamento de dados que mapeia dados de uma fonte de dados de entrada para um destino de dados de entrada com base no método de mapeamento de ID especificado. Ele produz uma tabela de mapeamento de ID. Esse fluxo de trabalho exige que você especifique o método de mapeamento de ID e os dados de entrada que você deseja traduzir de uma origem para um destino.

Você pode configurar um fluxo de trabalho de mapeamento de ID para ser executado sozinho Conta da AWS ou em dois Contas da AWS.

namespace de ID

Um recurso AWS Entity Resolution que contém metadados que explicam conjuntos de dados em vários Contas da AWS e como usar esses conjuntos de dados em um fluxo de trabalho de mapeamento de ID.

Há dois tipos de namespaces de ID: e. `SOURCE` `TARGET` O `SOURCE` contém configurações para os dados de origem que serão processados em um fluxo de trabalho de mapeamento de ID. O `TARGET` contém uma configuração dos dados de destino para os quais todas as fontes resolverão. Para definir os dados de entrada que você deseja resolver em dois Contas da AWS, crie uma fonte de

namespace de ID e um destino de namespace de ID para traduzir seus dados de um set () para outro ()SOURCE. TARGET

Depois que você e outro membro criarem namespaces de ID e executarem um fluxo de trabalho de mapeamento de ID, você poderá participar de uma colaboração AWS Clean Rooms para executar uma união de várias tabelas na tabela de mapeamento de ID e analisar os dados.

Para obter mais informações, consulte o Guia do usuário do AWS Clean Rooms.

Campo de entrada

Um campo de entrada corresponde ao nome de uma coluna da sua tabela AWS Glue de dados de entrada.

ARN da fonte de entrada (ARN) InputSource

O Amazon Resource Name (ARN) que foi gerado para uma entrada de AWS Glue tabela. Isso faz parte da correspondência dos metadados do fluxo de trabalho que serão incluídos na saída.

Correspondência baseada em aprendizado de máquina

A correspondência baseada em aprendizado de máquina (correspondência de ML) encontra correspondências em seus dados que podem estar incompletas ou podem não ter a mesma aparência. A correspondência de ML é um processo predefinido que tentará combinar registros em todos os dados inseridos. A correspondência de ML retorna uma ID de correspondência e um nível de confiança para cada conjunto de dados correspondente.

Processamento manual

Uma opção de cadência de processamento para uma tarefa de fluxo de trabalho correspondente que permite que ela seja executada sob demanda.

Essa opção é definida por padrão e está disponível tanto para correspondência baseada em regras quanto para correspondência baseada em aprendizado de máquina.

Many-to-Many combinando

Many-to-many a correspondência compara várias instâncias de dados semelhantes. Os valores nos campos de entrada aos quais foi atribuída a mesma chave de correspondência serão comparados

entre si, independentemente de estarem no mesmo campo de entrada ou em campos de entrada diferentes.

Por exemplo, você pode ter vários campos de entrada de número de telefone, como `mobile_phone` e `home_phone` que tenham a mesma tecla de correspondência “Telefone”. Use a many-to-many correspondência para comparar dados no campo `mobile_phone` de entrada com dados no campo `mobile_phone` de entrada e dados no campo `home_phone` de entrada.

As regras de correspondência avaliam dados em vários campos de entrada com a mesma chave de correspondência com uma operação (ou), e a one-to-many correspondência compara valores em vários campos de entrada. Isso significa que, se alguma combinação de `mobile_phone` ou `home_phone` corresponder entre dois registros, a tecla de correspondência “Telefone” retornará uma correspondência. Para combinar, tecle “Telefone” para encontrar uma correspondência, `Record One mobile_phone = Record Two mobile_phone` OR `Record One mobile_phone = Record Two home_phone` OR `Record One home_phone = Record Two home_phone` OR`Record One home_phone = Record Two mobile_phone`.

ID da partida (MatchID)

Para correspondência baseada em regras e correspondência de ML, essa é a ID gerada AWS Entity Resolution e aplicada a cada conjunto de registros correspondente. Isso faz parte dos metadados correspondentes do fluxo de trabalho que serão incluídos na saída.

Tecla de correspondência (MatchKey)

A chave de correspondência instrui AWS Entity Resolution quais campos de entrada devem ser considerados como dados semelhantes e quais devem ser considerados como dados diferentes. Isso ajuda a configurar AWS Entity Resolution automaticamente as regras de correspondência baseadas em regras e a comparar dados semelhantes armazenados em diferentes campos de entrada.

Se houver vários tipos de informações de número de telefone, como um `mobile_phone` campo de `home_phone` entrada e um campo de entrada em seus dados, que você gostaria de comparar, forneça a ambos a tecla de correspondência “Telefone”. Em seguida, a correspondência baseada em regras pode ser configurada para comparar dados usando instruções “ou” em todos os campos de entrada com a tecla de correspondência “Telefone” (consulte Definições de One-to-One correspondência e Many-to-Many correspondência na seção Fluxo de trabalho correspondente).

Se você quiser que a correspondência baseada em regras considere diferentes tipos de informações de números de telefone de forma completamente separada, você pode criar chaves de correspondência mais específicas, como “Celular_Telefone” e “Home_Phone”. Em seguida, ao configurar um fluxo de trabalho de correspondência, você pode especificar como cada chave de correspondência telefônica será usada na correspondência baseada em regras.

Se não MatchKey for especificado para um campo de entrada específico, ele não poderá ser usado na correspondência, mas poderá ser realizado pelo processo de fluxo de trabalho correspondente e poderá ser gerado, se desejado.

Nome da chave de correspondência

O nome atribuído a uma chave Match.

Regra de partida (MatchRule)

Para correspondência baseada em regras, esse é o número da regra aplicada que gerou um conjunto de registros correspondente. Isso faz parte dos metadados correspondentes do fluxo de trabalho que serão incluídos na saída.

Correspondência

O processo de combinar e comparar dados de diferentes campos de entrada, tabelas ou bancos de dados e determinar quais deles são semelhantes — ou “coincidem” — com base na satisfação de determinados critérios de correspondência (por exemplo, por meio de regras ou modelos de correspondência).

Fluxo de trabalho correspondente

O processo que você configurou para especificar os dados de entrada a serem combinados e como a correspondência deve ser realizada.

Descrição do fluxo de trabalho correspondente

Uma descrição opcional do fluxo de trabalho correspondente que você pode optar por inserir. As descrições ajudam a diferenciar os fluxos de trabalho correspondentes se você criar mais de um.

Nome do fluxo de trabalho correspondente

O nome do fluxo de trabalho correspondente que você especifica.

> **Note**
>
> Os nomes de fluxo de trabalho correspondentes devem ser exclusivos. Eles não podem ter o mesmo nome ou um erro será retornado.

Metadados de fluxo de trabalho correspondentes

Informações geradas e enviadas AWS Entity Resolution durante um trabalho de fluxo de trabalho correspondente. Essas informações são necessárias na saída.

Normalização () ApplyNormalization

Escolha se deseja normalizar os dados de entrada conforme definido no esquema. A normalização padroniza os dados removendo espaços extras e caracteres especiais e padronizando para o formato minúsculo.

Por exemplo, se um campo de entrada tiver um tipo de atributo de Telefone completo e os valores na tabela de entrada estiverem formatados como`(123) 456-7890`, os valores AWS Entity Resolution serão normalizados para. `1234567890`

> **Note**
>
> A normalização só é suportada no tipo de grupo para nome, endereço, telefone e e-mail.

As seções a seguir descrevem nossas regras de normalização padrão.

Para correspondência baseada em ML especificamente, consulte. Normalização (ApplyNormalization) — somente com base em ML

Tópicos

- Name
- E-mail

- Telefone
- Endereço
- Hashado
- ID de origem

Name

> **Note**
>
> A normalização só é suportada para o tipo de grupo Nome.
> O tipo de grupo Nome aparece como Nome completo no console e `NAME` na API.
> Se você quiser normalizar os subtipos do tipo de grupo Nome:
>
> - No console, atribua os seguintes subtipos ao grupo Nome completo: Nome, segundo nome e sobrenome.
> - Na CreateSchemaMappingAPI, atribua os seguintes tipos ao `NAME` groupName`NAME_FIRST`:`NAME_MIDDLE`, e. `NAME_LAST`

- TRIM = Remove os espaços em branco à esquerda e à direita
- MINÚSCULAS = Coloca em minúsculas todos os caracteres alfa
- CONVERT_ACCENT = Letra acentuada oculta em letra normal
- REMOVE_ALL_NON_ALPHA = Remove todos os caracteres não alfa [A-zA-z]

E-mail

> **Note**
>
> A normalização é suportada para o tipo de grupo de e-mail.
> O tipo de grupo de e-mail aparece como endereço de e-mail no console e como **`EMAIL_ADDRESS`** na API.

- TRIM = Remove os espaços em branco à esquerda e à direita
- MINÚSCULAS = Coloca em minúsculas todos os caracteres alfa

- CONVERT_ACCENT = Letra acentuada oculta em letra normal
- EMAIL_ADDRESS_UTIL_NORM = Remove todos os pontos (.) do nome de usuário, remove qualquer coisa após um sinal de adição (+) no nome de usuário e padroniza variações comuns de domínio
- REMOVE_ALL_NON_EMAIL_CHARS = Remove todos os caracteres [a-zA-z0-9] e [.@-] non-alpha-numeric

Telefone

> **Note**
>
> A normalização é suportada somente para o tipo de grupo de telefone.
> O tipo de grupo Telefone aparece como Telefone completo no console e como `PHONE` na API.
> Se você quiser normalizar os subtipos do tipo de grupo de telefone:
>
> - No console, atribua os seguintes subtipos ao grupo de telefone completo: Número de telefone e Código do país do telefone.
> - Na CreateSchemaMappingAPI, atribua os seguintes tipos ao `PHONE` **`PHONE_NUMBER`**GroupName: e. `PHONE_COUNTRYCODE`

- TRIM = Remove os espaços em branco à esquerda e à direita
- REMOVE_ALL_NON_NUMERIC = Remove todos os caracteres não numéricos [0-9]
- REMOVE_ALL_LEADING_ZEROES=Remove todos os zeros iniciais
- ENSURE_PREFIX_WITH_MAP, "phonePrefixMap" = Examina cada número de telefone e tenta compará-lo com os padrões do. phonePrefixMap Se uma correspondência for encontrada, a regra adicionará ou modificará o prefixo do número de telefone para garantir que ele esteja em conformidade com o formato padronizado especificado no mapa.

Endereço

> **Note**
>
> A normalização é suportada somente para o tipo de grupo de endereços.

O tipo de grupo de endereços aparece como Endereço completo no console e `ADDRESS` na API.

Se você quiser normalizar os subtipos do tipo de grupo de endereços:

- No console, atribua os seguintes subtipos ao grupo Endereço completo: Endereço 1, Endereço 2: nome do endereço 3, nome da cidade, estado, país e código postal t
- Na CreateSchemaMappingAPI, atribua os seguintes tipos ao `ADDRESS` GroupName`ADDRESS_STREET1`:`ADDRESS_STREET2`,,,`ADDRESS_STREET3`, `ADDRESS_CITY``ADDRESS_STATE`, e. `ADDRESS_COUNTRY` `ADDRESS_POSTALCODE`

- TRIM = Remove os espaços em branco à esquerda e à direita
- MINÚSCULAS = Coloca em minúsculas todos os caracteres alfa
- CONVERT_ACCENT = Letra acentuada oculta em letra normal
- REMOVE_ALL_NON_ALPHA = Remove todos os caracteres não alfa [A-zA-z]
- RENAME_WORDS usando ADDRESS_RENAME_WORD_MAP = substituir palavras na string de endereço por palavras de ADDRESS_RENAME_WORD_MAP
- RENAME_DELIMITERS usando ADDRESS_RENAME_DELIMITER_MAP = substituir delimitadores na string de endereço pela string de ADDRESS_RENAME_DELIMITER_MAP
- RENAME_DIRECTIONS usando ADDRESS_RENAME_DIRECTION_MAP = substituir delimitadores na string de endereço pela string de ADDRESS_RENAME_DIRECTION_MAP
- RENAME_NUMBERS usando ADDRESS_RENAME_NUMBER_MAP = substituir números na string de endereço pela string de ADDRESS_RENAME_NUMBER_MAP
- RENAME_SPECIAL_CHARS usando ADDRESS_RENAME_SPECIAL_CHAR_MAP = substituir caracteres especiais na string de endereço pela string de ADDRESS_RENAME_SPECIAL_CHAR_MAP

ENDEREÇO_RENOME_MAPA_PALAVRA_DE_ENDEREÇO

Essas são as palavras que serão renomeadas ao normalizar a string de endereço.

```
"avenue": "ave",
 "bouled": "blvd",
 "circle": "cir",
 "circles": "cirs",
 "court": "ct",
```

```
"centre": "ctr",
"center": "ctr",
"drive": "dr",
"freeway": "fwy",
"frwy": "fwy",
"highway": "hwy",
"lane": "ln",
"parks": "park",
"parkways": "pkwy",
"pky": "pkwy",
"pkway": "pkwy",
"pkwys": "pkwy",
"parkway": "pkwy",
"parkwy": "pkwy",
"place": "pl",
"plaza": "plz",
"plza": "plz",
"road": "rd",
"square": "sq",
"squ": "sq",
"sqr": "sq",
"street": "st",
"str": "st",
"str.": "strasse"
```

MAPA_DELIMITADOR_DELIMITADOR DE ENDEREÇOS

Esses são os delimitadores que serão renomeados ao normalizar a string de endereço.

```
",": " ",
".": " ",
"[": " ",
"]": " ",
"/": " ",
"-": " ",
"#": " number "
```

ENDEREÇO_RENOME_MAPA_DIREÇÃO_DE_ENDEREÇO

Esses são os identificadores de direção que serão renomeados ao normalizar a string de endereço.

```
"east": "e",
```

```
"north": "n",
"south": "s",
"west": "w",
"northeast": "ne",
"northwest": "nw",
"southeast": "se",
"southwest": "sw"
```

ENDEREÇO_RENOME_NÚMERO_MAPA_DO_ENDEREÇO

Essas são as sequências numéricas que serão renomeadas ao normalizar a sequência de endereço.

```
"número": "number",
 "numero": "number",
 "no": "number",
 "núm": "number",
 "num": "number"
```

ADDRESS_RENAME_SPECIAL_CHAR_MAP

Essas são as cadeias de caracteres especiais que serão renomeadas ao normalizar a cadeia de endereços.

```
"ß": "ss",
 "ä": "ae",
 "ö": "oe",
 "ü": "ue",
 "ø": "o",
 "æ": "ae"
```

Hashado

- TRIM = Remove os espaços em branco à esquerda e à direita

ID de origem

- TRIM = Remove os espaços em branco à esquerda e à direita

Normalização (ApplyNormalization) — somente com base em ML

Escolha se deseja normalizar os dados de entrada conforme definido no esquema. A normalização padroniza os dados removendo espaços extras e caracteres especiais e padronizando para o formato minúsculo.

Por exemplo, se um campo de entrada tiver um tipo de `NAME` atributo e os valores na tabela de entrada estiverem formatados como`Johns Smith`, AWS Entity Resolution normalizará os valores para. `john smith`

As seções a seguir descrevem as regras de normalização para fluxos de trabalho de correspondência baseados em aprendizado de máquina.

Tópicos

- Name
- E-mail
- Telefone

Name

- TRIM = Remove os espaços em branco à esquerda e à direita
- MINÚSCULAS = Coloca em minúsculas todos os caracteres alfa

E-mail

- MINÚSCULAS = Coloca em minúsculas todos os caracteres alfa
- Substitui somente (at) (com distinção entre maiúsculas e minúsculas) por um símbolo @
- Remove todos os espaços em branco, em qualquer lugar no valor
- Remove tudo o que está fora do primeiro, "< >" se existir

Telefone

- TRIM = Remove os espaços em branco à esquerda e à direita
- REMOVE_ALL_NON_NUMERIC = Remove todos os caracteres não numéricos [0-9]
- REMOVE_ALL_LEADING_ZEROES=Remove todos os zeros iniciais

- ENSURE_PREFIX_WITH_MAP, "phonePrefixMap" = Examina cada número de telefone e tenta compará-lo com os padrões do. phonePrefixMap Se uma correspondência for encontrada, a regra adicionará ou modificará o prefixo do número de telefone para garantir que ele esteja em conformidade com o formato padronizado especificado no mapa.

One-to-One combinando

One-to-one a correspondência compara instâncias únicas de dados semelhantes. Os campos de entrada com a mesma chave de correspondência e valores no mesmo campo de entrada serão comparados entre si.

Por exemplo, você pode ter vários campos de entrada de número de telefone, como `mobile_phone` e `home_phone` que tenham a mesma tecla de correspondência “Telefone”. Use a one-to-one correspondência para comparar dados no campo `mobile_phone` de entrada com dados no campo `mobile_phone` de entrada e para comparar dados no campo `home_phone` de entrada com dados no campo `home_phone` de entrada. Os dados no campo `mobile_phone` de entrada não serão comparados com os dados no campo `home_phone` de entrada.

As regras de correspondência avaliam dados em vários campos de entrada com a mesma chave de correspondência com uma operação (ou), e a one-to-many correspondência compara valores em um único campo de entrada. Isso significa que se `mobile_phone` ou `home_phone` corresponder entre dois registros, a tecla de correspondência “Telefone” retornará uma correspondência. Para combinar, tecle “Telefone” para encontrar uma correspondência, `Record One mobile_phone = Record Two mobile_phone` OU`Record One home_phone = Record Two home_phone`.

As regras de correspondência avaliam dados em campos de entrada com chaves de correspondência diferentes com uma operação (e). Se você quiser que a correspondência baseada em regras considere diferentes tipos de informações de números de telefone de forma completamente separada, você pode criar chaves de correspondência mais específicas, como “mobile_phone” e “home_phone”. Se você quiser usar as duas teclas de correspondência em uma regra para encontrar correspondências, `Record One mobile_phone = Record Two mobile_phone` AND`Record One home_phone = Record Two home_phone`.

Saída

Uma lista de OutputAttributeobjetos, cada um com os campos Nome e Hashed. Cada um desses objetos representa uma coluna a ser incluída na tabela AWS Glue de saída e se você deseja que os valores na coluna sejam criptografados.

Saídas 3Path

O destino do S3 no qual AWS Entity Resolution gravará a tabela de saída.

OutputSourceConfig

Uma lista de OutputSource objetos, cada um com os campos Outputs3Path e Output. ApplyNormalization

Correspondência baseada em serviços de provedores

A correspondência baseada em serviços de provedores é um processo projetado para combinar, vincular e aprimorar seus registros com provedores de serviços de dados preferenciais e conjuntos de dados licenciados. Você deve ter uma assinatura AWS Data Exchange com o serviço do provedor para usar essa técnica de correspondência.

AWS Entity Resolution atualmente se integra aos seguintes provedores de serviços de dados:

- LiveRamp
- TransUnion
- UID 2.0

Correspondência baseada em regras

A correspondência baseada em regras é um processo projetado para encontrar correspondências exatas. A correspondência baseada em regras é um conjunto hierárquico de regras de correspondência em cascata, sugerido por AWS Entity Resolution, com base nos dados que você insere e totalmente configurável por você. Todas as chaves de correspondência fornecidas nos critérios da regra devem corresponder exatamente para que os dados comparados sejam declarados como correspondências e para que os metadados associados sejam gerados. A correspondência baseada em regras retorna uma ID de correspondência e um número de regra para cada conjunto de dados correspondente.

Recomendamos definir regras que possam identificar uma entidade de forma exclusiva. Ordene suas regras para encontrar combinações mais precisas primeiro.

Por exemplo, digamos que você tenha duas regras, Regra 1 e Regra 2.

Essas regras têm as seguintes chaves de correspondência:

- A regra 1 inclui nome completo e endereço
- A regra 2 inclui nome completo, endereço e telefone

Como a Regra 1 é executada primeiro, nenhuma correspondência será encontrada pela Regra 2 porque todas teriam sido encontradas pela Regra 1.

Para encontrar correspondências diferenciadas por telefone, reordene as regras, assim:

- A regra 2 inclui nome completo, endereço e telefone
- A regra 1 inclui nome completo e endereço

Schema

O termo usado para uma estrutura ou layout que define como um conjunto de dados é organizado e conectado.

Descrição do esquema

Uma descrição opcional do esquema que você pode escolher inserir. As descrições ajudam a diferenciar os mapeamentos de esquema se você criar mais de um.

Nome do esquema

O nome do esquema.

> **Note**
>
> Os nomes dos esquemas devem ser exclusivos. Eles não podem ter o mesmo nome ou um erro será retornado.

Mapeamento de esquemas

O mapeamento de esquemas AWS Entity Resolution é o processo pelo qual você informa AWS Entity Resolution como interpretar seus dados para fins de correspondência. Você define o esquema

da tabela de dados de entrada que AWS Entity Resolution deseja ler em um fluxo de trabalho correspondente.

ARN de mapeamento de esquema

O Amazon Resource Name (ARN) gerado para o mapeamento do esquema.

ID exclusivo

Um identificador exclusivo que você designa e que deve ser atribuído a cada linha de dados de entrada AWS Entity Resolution lida.

Example

Por exemplo: **`Primary_key`**, **`Row_ID`** ou **`Record_ID`**.

A coluna ID exclusiva é obrigatória.

O ID exclusivo deve ser um identificador exclusivo em uma única tabela.

O ID exclusivo deve satisfazer esse padrão: `[a-zA-Z0-9_-]`

Em tabelas diferentes, o ID exclusivo pode ter valores duplicados.

O tamanho máximo do ID exclusivo é 38 para um fluxo de trabalho correspondente

O tamanho máximo do ID exclusivo é de 257 caracteres para um Fluxo de trabalho de mapeamento de ID

Quando o fluxo de trabalho correspondente for executado, o registro será rejeitado se a ID exclusiva:

- não está especificado
- não é exclusivo na mesma tabela
- sobreposições em termos de nome de atributo em todas as fontes
- excede 38 caracteres (somente fluxos de trabalho correspondentes baseados em regras)

As traduções são geradas por tradução automática. Em caso de conflito entre o conteúdo da tradução e da versão original em inglês, a versão em inglês prevalecerá.