



用户指南

AWS Entity Resolution 数据匹配服务



AWS Entity Resolution 数据匹配服务: 用户指南

Copyright © 2024 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon 的商标和商业外观不得用于任何非 Amazon 的商品或服务，也不得以任何可能引起客户混淆、贬低或诋毁 Amazon 的方式使用。所有非 Amazon 拥有的其他商标均为各自所有者的财产，这些所有者可能附属于 Amazon、与 Amazon 有关联或由 Amazon 赞助，也可能不是如此。

Table of Contents

什么是 AWS Entity Resolution 数据匹配服务？	1
你是首次 AWS Entity Resolution 数据匹配服务 使用吗？	1
的特点 AWS Entity Resolution 数据匹配服务	1
相关服务	4
正在访问 AWS Entity Resolution 数据匹配服务	4
的定价 AWS Entity Resolution 数据匹配服务	5
设置	6
报名参加 AWS	6
创建管理员用户	6
为控制台用户创建IAM角色	7
创建工作流程工作角色	8
准备输入数据表	15
准备第一方输入数据	15
步骤 1：以支持的数据格式保存输入数据表	15
第 2 步：将您的输入数据表上传到 Amazon S3	15
步骤 3：创建一个 AWS Glue 表	16
准备第三方输入数据	17
步骤 1：在上订阅提供商服务 AWS Data Exchange	18
步骤 2：准备第三方数据表	19
步骤 3：以支持的数据格式保存输入数据表	22
步骤 4：将您的输入数据表上传到 Amazon S3	22
步骤 5：创建一个 AWS Glue 表	23
架构映射	25
创建架构映射	25
克隆架构映射	32
编辑架构映射	33
删除架构映射	33
ID 命名空间	35
ID 命名空间来源	35
创建 ID 命名空间源（基于规则）	36
创建 ID 命名空间源（提供者服务）	39
ID 命名空间目标	41
创建 ID 命名空间目标（基于规则的方法）	41
创建 ID 命名空间目标（提供者服务方法）	44

编辑 ID 命名空间	45
删除 ID 命名空间	45
添加或更新 ID 命名空间的资源策略	45
匹配工作流程	47
创建基于规则的匹配工作流程	48
创建基于机器学习的匹配工作流程	53
创建基于提供商服务的匹配工作流程	57
使用创建匹配的工作流程 LiveRamp	57
使用创建匹配的工作流程 TransUnion	64
使用 UID 2.0 创建匹配的工作流程	70
编辑匹配的工作流程	74
删除匹配的工作流程	74
为基于规则的匹配工作流程查找匹配 ID	74
从基于规则或基于 ML 的匹配工作流程中删除记录	75
故障排除	76
我在运行匹配的工作流程后收到了错误文件	76
身份映射工作流程	78
一个人的身份映射工作流程 AWS 账户	79
先决条件	79
创建 ID 映射工作流程 (基于规则)	81
创建 ID 映射工作流程 (提供者服务)	85
跨两个 ID 映射工作流程 AWS 账户	90
先决条件	91
创建 ID 映射工作流程 (基于规则)	92
创建 ID 映射工作流程 (提供者服务)	96
运行 ID 映射工作流程	101
使用新的输出目标运行 ID 映射工作流程	102
编辑 ID 映射工作流程	104
删除 ID 映射工作流程	104
为 ID 映射工作流程添加或更新资源策略	105
提供商集成	106
要求	106
在上列出提供商服务 AWS Data Exchange	106
确定你的属性	107
申请 AWS Entity Resolution 数据匹配服务 开放API规范	108
使用开放API规范	108

批处理集成	109
同步处理集成	111
测试提供商集成	113
安全性	120
数据保护	120
静态数据加密 AWS Entity Resolution 数据匹配服务	121
密钥管理	122
AWS PrivateLink	131
Identity and Access Management	133
受众	134
使用身份进行身份验证	134
使用策略管理访问	137
操作方法 AWS Entity Resolution 数据匹配服务 配合使用 IAM	139
基于身份的策略示例	144
AWS 托管策略	147
故障排除	152
合规性验证	153
AWS Entity Resolution 数据匹配服务 合规性最佳实践	154
弹性	155
监控	156
CloudTrail 日志	156
AWS Entity Resolution 数据匹配服务 信息在 CloudTrail	156
了解 AWS Entity Resolution 数据匹配服务 日志文件条目	157
AWS CloudFormation 资源	158
AWS实体分辨率和 AWS CloudFormation 模板	158
了解更多关于 AWS CloudFormation	159
配额	161
文档历史记录	164
术语表	167
亚马逊资源名称 (ARN)	167
自动处理	167
AWS KMS key ARN	167
明文	167
置信度 (ConfidenceLevel)	167
解密	167
加密	168

组名	168
哈希	168
哈希协议 (HashingProtocol)	168
身份映射方法	168
身份映射工作流程	169
ID 命名空间	169
输入字段	169
输入源 ARN (InputSourceARN)	169
输入类型	169
基于机器学习的匹配	170
手动处理	170
多对多匹配	170
比赛 ID (matchID)	170
匹配密钥 (MatchKey)	170
匹配密钥名称	171
匹配规则 (MatchRule)	171
匹配	171
匹配工作流程	171
匹配的工作流程描述	171
匹配工作流程名称	171
匹配工作流程元数据	172
标准化 (ApplyNormalization)	172
名称	172
Email	173
Phone	173
地址	173
经过哈希处理	175
来源_ID	175
一对一匹配	176
输出	176
outputs3Path	176
OutputSourceConfig	176
基于提供商服务的匹配	176
基于规则的匹配	177
架构	177
架构描述	177

架构名称	178
架构映射	178
架构映射 ARN	178
唯一标识	178
.....	clxxx

什么是 AWS Entity Resolution 数据匹配服务？

AWS Entity Resolution 数据匹配服务 是一项服务，可帮助您匹配、链接和增强存储在多个应用程序、渠道和数据存储中的相关记录。您可以开始使用灵活且可扩展的实体解析工作流程，并且可以连接到您现有的应用程序和数据服务提供商。

AWS Entity Resolution 数据匹配服务 提供高级匹配技术，例如基于规则的匹配、基于机器学习的匹配（机器学习匹配）和数据服务提供商主导的匹配。这些技术可以帮助您更准确地关联和增强客户信息、产品代码或业务数据代码的相关记录。

通过将最近发生的事件（例如广告点击、购物车放弃和购买）与来自数据服务提供商的匿名化信号关联到唯一的实体 ID，您可以使用 AWS Entity Resolution 数据匹配服务 创建统一的客户互动视图。您还可以更好地跟踪商店中使用不同代码（例如SKU，UPC）的产品。您可以使用 AWS Entity Resolution 数据匹配服务 来控制匹配精度，更好地保护数据安全，同时最大限度地减少数据移动。

主题

- [你是首次 AWS Entity Resolution 数据匹配服务 使用吗？](#)
- [的特点 AWS Entity Resolution 数据匹配服务](#)
- [相关服务](#)
- [正在访问 AWS Entity Resolution 数据匹配服务](#)
- [的定价 AWS Entity Resolution 数据匹配服务](#)

你是首次 AWS Entity Resolution 数据匹配服务 使用吗？

如果您是首次使用 AWS Entity Resolution 数据匹配服务，我们建议您先阅读以下章节：

- [的特点 AWS Entity Resolution 数据匹配服务](#)
- [正在访问 AWS Entity Resolution 数据匹配服务](#)
- [设置 AWS Entity Resolution 数据匹配服务](#)

的特点 AWS Entity Resolution 数据匹配服务

AWS Entity Resolution 数据匹配服务 包括以下功能：

- 灵活且可定制的数据准备

AWS Entity Resolution 数据匹配服务 从中读取您的数据 AWS Glue ，用作匹配处理的输入。您最多可以指定 20 个数据输入。AWS Entity Resolution 数据匹配服务 将数据输入表的每一行作为记录处理，并使用唯一的实体作为主键。AWS Entity Resolution 数据匹配服务 可以对加密的数据集进行操作。首先定义[架构映射](#) AWS Entity Resolution 数据匹配服务 ，以了解要在[匹配工作流程](#)中使用哪些输入字段。您可以从现有的数据输入中引入自己的 AWS Glue 数据架构或蓝图。或者，您可以使用交互式用户界面或JSON编辑器来构建自定义架构。默认情况下，AWS Entity Resolution 数据匹配服务 还会在匹配之前[对数据输入进行标准化](#)以改进匹配处理，例如删除特殊字符和多余空格，以及将文本格式化为小写。如果您的数据输入已经过标准化，则可以关闭标准化。我们还提供一个[GitHub 库](#)，您可以使用该库进一步自定义数据标准化过程以满足您的需求。

- 可配置的实体匹配工作流程

实体[匹配工作流程](#)是您设置的一系列步骤，用于说明 AWS Entity Resolution 数据匹配服务 如何匹配数据输入以及将合并数据输出写入何处。您可以设置一个或多个匹配工作流程来比较不同的数据输入，并使用不同的匹配技术，例如[基于规则的匹配](#)、[机器学习匹配](#)或[数据服务提供商主导的匹配](#)，无需实体解析或机器学习经验。您还可以查看现有匹配工作流程和指标的任务状态，例如资源编号、已处理的记录数和找到的匹配项数。

- 基于 Ready-to-use 规则的匹配

这种匹配方法在 o AWS Management Console r AWS Command Line Interface (AWS CLI) 中包含一组 ready-to-use 规则。您可以使用这些规则根据您的输入字段查找相关记录。您还可以通过添加或删除每条规则的输入字段、删除规则、重新排列规则优先级以及创建新规则来自定义规则。您也可以重置规则，使其恢复到其原始配置。您的亚马逊简单存储服务 (Amazon S3) Simple Service 存储桶中输出的数据包含使用基于[规则](#)的匹配技术生成的匹配组 AWS Entity Resolution 数据匹配服务 。每个匹配组都有用于生成与之关联的匹配项的规则编号，以帮助了解匹配项。例如，规则编号可以证明每个匹配组的精度，从而使规则一比规则二更精确。

- 预先配置的基于机器学习的匹配 (机器学习匹配)

这种匹配技术包括预先配置的机器学习模型，用于在所有数据输入中查找匹配项，尤其是基于消费者的记录。该模型使用与姓名、电子邮件地址、电话号码、地址和出生日期数据类型关联的所有输入字段。该模型生成由相关记录组成的匹配组，每个组中都有一个[置信度分数](#)，解释了与其他比赛组相比的比赛质量。该模型会考虑缺失的输入字段，并一起分析整个记录以表示实体。您的 Amazon S3 存储桶中的数据输出包含使用机器学习匹配 AWS Entity Resolution 数据匹配服务 生成的匹配组。在这里，每个匹配组的相关置信度分数为 0.0—1.0，这表示匹配的精度。

- 将记录与数据服务提供商进行匹配

借助，AWS Entity Resolution 数据匹配服务 您可以与领先的数据服务供应商和许可数据集进行匹配、关联和增强记录，从而扩大您了解、接触和服务客户的能力。例如，您可以为数据附加属性以增强记录，也可以提高所使用的系统和平台的互操作性以实现业务目标。您只需点击几下即可使用此匹配的工作流程，无需构建和维护复杂的专有集成。您必须与这些数据服务提供商签订许可协议才能利用这种匹配技术。

- 手动批量处理和自动增量处理

您可以使用数据处理来帮助将您的数据输入或输入转换为合并的数据输出表，该表包含使用实体匹配工作流程配置生成的公共匹配 ID 的类似记录。使用 [AWS Management Console](#) 或 [AWS CLI](#)，您可以根据现有提取、转换API和 load (ETL) 数据管道按需运行[手动批量处理](#)，该管道会重新处理所有新匹配项的数据，并更新现有匹配项。此外，对于基于规则的匹配方案，您可以启动[自动增量处理](#)，这样，只要您的 Amazon S3 存储桶中有新数据，该服务就会读取这些新记录并将其与现有记录进行比较。这样可以使您的匹配项与 Amazon S3 数据的任何变化保持同步。

- 近乎实时的查询

通过该[AWS Entity Resolution 数据匹配服务 GetMatchId API操作](#)查找任何实体字段可帮助您同步检索现有的匹配 ID。您可以使用通过不同来源和渠道获得的个人身份信息 (PII) 属性致电 AWS Entity Resolution 数据匹配服务。AWS Entity Resolution 数据匹配服务 对这些属性进行哈希处理以保护数据，并检索相应的匹配 ID 以关联和匹配客户。例如，您可以使用关联的姓名、电子邮件和邮寄地址进行网络注册。使用该 AWS Entity Resolution 数据匹配服务 [GetMatchId API操作](#)来确定存储在 S3 存储桶中的匹配结果中是否已存在该客户或实体，以及与之关联的相应实体匹配 ID。获得实体匹配 ID 后，您可以在源应用程序（例如客户关系管理 (CRM) 或客户数据平台 () 系统 () 中找到与之相关的交易信息。CDP

- 通过设计实现数据保护和区域化

AWS Entity Resolution 数据匹配服务 提供默认加密功能，可帮助您保护数据，并为输入到服务的每个数据提供加密密钥。例如，AWS Entity Resolution 数据匹配服务 允许您灵活地使用服务器端加密和哈希处理的数据来运行基于规则的匹配工作流程。AWS Entity Resolution 数据匹配服务 支持区域化，这意味着您的匹配工作流程以与您使用服务相同的位置 AWS 区域 来处理数据。在其他应用程序中使用已解析的数据之前，您还可以对 Amazon S3 中的数据输出进行加密和哈希处理。

- 多方转码

AWS Entity Resolution 数据匹配服务 帮助您定义数据源，并在想要使用数据协作的多方之间进行匹配配置，例如在 AWS Clean Rooms。

相关服务

以下内容与 AWS 服务 以下内容有关 AWS Entity Resolution 数据匹配服务：

- Amazon S3

将您导入的数据存储 AWS Entity Resolution 数据匹配服务 在 Amazon S3 中。

有关更多信息，请参阅[什么是 Amazon S3？](#) 在 Amazon 简单存储服务用户指南中。

- AWS Glue

根据您在 Amazon S3 中的数据创建 AWS Glue 表以供在中使用 AWS Entity Resolution 数据匹配服务。

有关更多信息，请参阅[什么是 AWS Glue？](#) 在《AWS Glue 开发人员指南》中。

- AWS CloudTrail

AWS Entity Resolution 数据匹配服务 与 CloudTrail 日志配合使用可增强对 AWS 服务 活动的分析。

有关更多信息，请参阅[使用记录 AWS Entity Resolution 数据匹配服务 API 调用 AWS CloudTrail。](#)

- AWS CloudFormation

在中创建以下资源 AWS CloudFormation：AWS::EntityResolution::MatchingWorkflow, AWS::EntityResolution::SchemaMapping, AWS::EntityResolution::IdMappingWorkflow, AWS::EntityResolution::IdNamespace 和 AWS::EntityResolution::PolicyStatement

有关更多信息，请参阅[使用创建AWS实体解析资源 AWS CloudFormation。](#)

正在访问 AWS Entity Resolution 数据匹配服务

您可以 AWS Entity Resolution 数据匹配服务 通过以下选项进行访问：

- 直接通过 AWS Entity Resolution 数据匹配服务 控制台，网址为<https://console.aws.amazon.com/entityresolution/>。
- 以编程方式通过 . AWS Entity Resolution 数据匹配服务 API 有关更多信息，请参阅《[AWS Entity Resolution 数据匹配服务 API参考](#)》。

- 如果您计划 AWS Entity Resolution 数据匹配服务 API 在 AWS Lambda Runtime 中调用，请创建自己的部署包并包含所需的 AWS SDK 库版本。有关更多信息，请参阅《AWS Lambda 开发人员指南》中的以下示例：
 - [使用 .zip 或文件存档部署 Java Lambda 函数 JAR](#)
 - [使用 Python Lambda 函数的 .zip 文件存档](#)

的定价 AWS Entity Resolution 数据匹配服务

有关定价信息，请参阅 [AWS Entity Resolution 数据匹配服务 定价](#)。

设置 AWS Entity Resolution 数据匹配服务

使用之前 AWS Entity Resolution 数据匹配服务 第一次，注册 AWS 并创建一个管理员用户来创建角色。

报名参加 AWS

如果你已经有一个 AWS 账户，跳过此步骤。

如果你没有 AWS 账户，请完成以下步骤来创建一个。

要注册 AWS 账户

1. 打开<https://portal.aws.amazon.com/billing/注册>。
2. 按照屏幕上的说明进行操作。

在注册时，将接到一通电话，要求使用电话键盘输入一个验证码。

当你注册时 AWS 账户，一个 AWS 账户根用户已创建。root 用户可以访问所有内容 AWS 服务 以及账户中的资源。作为安全最佳实践，请为用户分配管理访问权限，并且只使用根用户来执行[需要根用户访问权限的任务](#)。

创建管理员用户

要创建管理员用户，请选择以下选项之一。

选择一种方法来管理您的管理员	目的	方式	您也可以
在 IAM 身份中心中	使用短期凭证进行访问 AWS。 这符合安全最佳实践。有关最佳实践的信	按照《 入门 》中的说明进行操作 AWS IAM Identity Center 用户指南。	通过配置来 配置编程访问权限 AWS CLI 要使用 AWS IAM Identity Center 中的AWS

选择一种方法来管理您的管理员	目的	方式	您也可以
(建议)	息，请参阅《IAM用户指南》IAM中的 安全最佳实践 。		Command Line Interface 用户指南。
在 IAM (不推荐使用)	使用长期凭证进行访问 AWS.	按照《用户指南》中 创建您的第一个IAM管理员用户和用户组 中的IAM说明进行操作。	通过在《用户指南》中 管理IAM用户的访问密钥 来配置编程访问权限。IAM

为控制台用户创建IAM角色

如果您使用的是，请完成以下步骤 AWS Entity Resolution 数据匹配服务 console。

创建 IAM 角色

1. 使用您的管理员帐户登录IAM控制台 (<https://console.aws.amazon.com/iam/>)。
2. 在 Access management (访问管理) 下，请选择 Roles (角色)。

您可以使用角色创建短期证书，建议使用此方法来提高安全性。您也可以选择用户来创建长期凭证。

3. 选择 Create role (创建角色)。
4. 在创建角色向导中，对于可信实体类型，选择 AWS 帐户。
5. 保持“此帐户”选项处于选中状态，然后选择“下一步”。
6. 对于添加权限，请选择创建策略。

将打开一个新选项卡。

- a. 选择该JSON选项卡，然后根据授予控制台用户的权限添加策略。AWS Entity Resolution 数据匹配服务 根据常见用例提供以下托管策略：

- [AWS 托管策略：AWSEntityResolutionConsoleFullAccess](#)

- [AWS 托管策略：AWSEntityResolutionConsoleReadOnlyAccess](#)

- 选择下一步: 标签，添加标签（可选），然后选择下一步: 审核。
- 对于查看策略，输入名称和描述，然后查看摘要。
- 选择创建策略。

您已经为协作成员创建了策略。

- 返回原始选项卡，在“添加权限”下，输入您刚刚创建的策略的名称。（您可能需要重新加载页面。）
 - 选中您创建的策略名称旁边的复选框，然后选择下一步。
- 对于命名、查看和创建，输入角色名称和描述。
 - 查看选择可信实体，输入 AWS 账户 适用于将担任该角色的一名或多名人员（如有必要）。
 - 在添加权限中查看权限，并在必要时进行编辑。
 - 查看标签，并在必要时添加标签。
 - 选择 Create role（创建角色）。

为创建工作流工作角色 AWS Entity Resolution 数据匹配服务

AWS Entity Resolution 数据匹配服务 使用工作流作业角色来运行工作流。如果您拥有必要的IAM 权限，则可以使用控制台创建此角色。如果您没有CreateRole权限，请让您的管理员创建该角色。

为创建工作流工作角色 AWS Entity Resolution 数据匹配服务

- 使用您的管理员帐户登录IAM控制台。<https://console.aws.amazon.com/iam/>
- 在 Access management（访问管理）下，请选择 Roles（角色）。

您可以使用角色创建短期证书，建议使用此方法来提高安全性。您也可以选择用户来创建长期凭证。

- 选择 Create role（创建角色）。
- 在创建角色向导中，对于可信实体类型，选择自定义信任策略。
- 将以下自定义信任策略复制并粘贴到JSON编辑器中。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
```

```

        "Effect": "Allow",
        "Principal": {
            "Service": [
                "entityresolution.amazonaws.com"
            ]
        },
        "Action": "sts:AssumeRole"
    }
]
}

```

6. 选择下一步。
7. 对于添加权限，请选择创建策略。

此时会出现一个新选项卡。

- a. 将以下策略复制并粘贴到JSON编辑器中。

Note

以下示例策略支持读取相应数据资源所需的权限，例如 Amazon S3 和 AWS Glue。但是，您可能需要修改此策略，具体取决于您设置数据源的方式。您的 AWS Glue 资源和底层 Amazon S3 资源必须相同 AWS 区域 如同 AWS Entity Resolution 数据匹配服务。你不需要授予 AWS KMS 权限（如果您的数据源未加密或解密）。

```

{
  "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:{{aws-region}}:{{accountId}}:database/{{input-databases}}",
            "arn:aws:glue:{{aws-region}}:{{accountId}}:table/{{input-database}}/{{input-tables}}",
            "arn:aws:glue:{{aws-region}}:{{accountId}}:catalog"
        ]
    }
}

```

```

    ]
  }
]
}

```

替换每个 `{{user input placeholder}}` 用你自己的信息。

aws-region

AWS 区域 你的资源。您的 AWS Glue 资源、底层 Amazon S3 资源和 AWS KMS 资源必须相同 AWS 区域 如AWS Entity Resolution ##### 。

accountId

您的 AWS 账户 同上。

input-buckets

Amazon S3 存储桶，其中包含的底层数据对象 AWS Glue AWS Entity Resolution ##### 将在哪里读取。

output-buckets

Amazon S3 存储桶AWS Entity Resolution ##### 将在其中生成输出数据。

input-databases

AWS Glue AWS Entity Resolution #####将从中读取的数据库。

- b. (可选) 如果使用客户的KMS密钥对输入的 Amazon S3 存储桶进行加密，请添加以下内容：

```

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

```

替换每个 `{{user input placeholder}}` 用你自己的信息。

aws-region

AWS 区域 你的资源。您的 AWS Glue 资源、底层 Amazon S3 资源和 AWS KMS 资源必须相同 AWS 区域 如AWS Entity Resolution ##### 。

accountId

您的 AWS 账户 同上。

inputKeys

中的托管密钥 AWS Key Management Service。如果您的输入源已加密，则AWS Entity Resolution ##### 必须使用您的密钥解密数据。

- c. (可选) 如果写入输出 Amazon S3 存储桶的数据需要加密，请添加以下内容：

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

替换每个 *{{user input placeholder}}* 用你自己的信息。

aws-region

AWS 区域 你的资源。您的 AWS Glue 资源、底层 Amazon S3 资源和 AWS KMS 资源必须相同 AWS 区域 如AWS Entity Resolution ##### 。

accountId

您的 AWS 账户 同上。

outputKeys

中的托管密钥 AWS Key Management Service。如果您需要对输出源进行加密，则AWS Entity Resolution ##### 必须使用您的密钥对输出数据进行加密。

- d. (可选) 如果您通过以下方式订阅了提供商服务 AWS Data Exchange，并希望将现有角色用于基于提供者服务的工作流程，请添加以下内容：

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

替换每个 *{{user input placeholder}}* 用你自己的信息。

aws-region

这些区域有：AWS 区域 在那里授予提供者资源。你可以在 ARN 上面的资产中找到这个价值 AWS Data Exchange console。例如：arn:aws:dataexchange:us-east-2::data-sets/111122223333/revisions/339ffc64444examplef3bc15cf0b2346b/assets/546468b8dexamplea37bfc73b8f79fefa

datasetId

数据集的 ID，可在 AWS Data Exchange console。

revisionId

数据集的修订版，可在 AWS Data Exchange console。

assetId

资产的 ID，可在 AWS Data Exchange console。

- 返回原始选项卡，在“添加权限”下，输入您刚刚创建的策略的名称。（您可能需要重新加载页面。）
- 选中您创建的策略名称旁边的复选框，然后选择下一步。
- 对于命名、查看和创建，输入角色名称和描述。

Note

角色名称必须与授予成员的passRole权限模式相匹配，该成员可以传递权限workflow job role以创建匹配的工作流程。

例如，如果您使用的是AWSEntityResolutionConsoleFullAccess托管策略，请记得在角色名称中加entityresolution入。

- a. 查看选择受信任的实体，并在必要时进行编辑。
- b. 在添加权限中查看权限，并在必要时进行编辑。
- c. 查看标签，并在必要时添加标签。
- d. 选择 Create role (创建角色)。

的工作流工作角色 AWS Entity Resolution 数据匹配服务 已创建。

准备输入数据表

In AWS Entity Resolution 数据匹配服务，您的每个输入数据表都包含源记录。这些记录包含消费者标识符，例如名字、姓氏、电子邮件地址或电话号码。这些源记录可以与您在相同或其他输入数据表中提供的其他源记录相匹配。每条记录都必须有一个唯一的记录 ID ([唯一标识](#))，并且在创建架构映射时必须将其定义为主键 AWS Entity Resolution 数据匹配服务。

每个输入数据表都可用作 AWS Glue 由亚马逊 S3 支持的表。您可以使用已存在于 Amazon S3 中的第一方数据，也可以将来自其他第三方 SaaS 提供商的数据表导入 Amazon S3。将数据上传到 Amazon S3 后，您可以使用 AWS Glue 用于在中创建数据表的 Crawler AWS Glue Data Catalog。然后，您可以使用数据表作为输入 AWS Entity Resolution 数据匹配服务。

以下各节介绍如何准备第一方数据和第三方数据。

主题

- [准备第一方输入数据](#)
- [准备第三方输入数据](#)

准备第一方输入数据

[以下步骤描述了如何准备第一方数据，以用于基于规则的匹配工作流程、基于机器学习的匹配工作流程或身份映射工作流程。](#)

步骤 1：以支持的数据格式保存输入数据表

如果您已经以支持的数据格式保存了第一方输入数据，则可以跳过此步骤。

要将 AWS Entity Resolution 数据匹配服务，则输入数据的格式必须为 AWS Entity Resolution 数据匹配服务支持。AWS Entity Resolution 数据匹配服务支持以下数据格式：

- 逗号分隔值 (,) CSV
- Parquet

第 2 步：将您的输入数据表上传到 Amazon S3

如果您在 Amazon S3 中已经有了第一方数据表，则可以跳过此步骤。

Note

输入数据必须存储在亚马逊简单存储服务 (Amazon S3) 中 AWS 账户 以及 AWS 区域 你想在其中运行匹配的工作流程。

将您的输入数据表上传到 Amazon S3

1. 登录 AWS Management Console 然后打开 Amazon S3 控制台，网址为<https://console.aws.amazon.com/s3/>。
2. 选择 Buckets，然后选择一个存储桶来存储您的数据表。
3. 选择上传，然后按照提示进行操作。
4. 选择对象选项卡，查看存储数据的前缀。记下文件夹的名称。

您可以选择要查看数据表的文件夹。

步骤 3：创建一个 AWS Glue 表

Amazon S3 中的输入数据必须编入目录 AWS Glue 并表示为 AWS Glue 桌子。有关如何创建的更多信息 AWS Glue 以 Amazon S3 为输入的表，请参阅上的“[使用爬虫](#)” [AWS Glue 控制台](#) 在 AWS Glue 开发者指南。

Note

AWS Entity Resolution 数据匹配服务 不支持分区表。

在此步骤中，您将在中设置爬虫 AWS Glue 它会抓取 S3 存储桶中的所有文件并创建一个 AWS Glue 桌子。

Note

AWS Entity Resolution 数据匹配服务 目前不支持注册到的 Amazon S3 地点 AWS Lake Formation.

要创建 AWS Glue 表

1. 登录 AWS Management Console 然后打开 AWS Glue 控制台位于<https://console.aws.amazon.com/glue/>。
2. 从导航栏中，选择爬网程序。
3. 从列表中选择您的 S3 存储桶，然后选择添加爬网程序。
4. 在添加爬网程序页面上，输入爬网程序名称，然后选择下一步。
5. 继续浏览添加爬网程序页面，指定详细信息。
6. 在“选择IAM角色”页面上，选择“选择现有IAM角色”，然后选择“下一步”。

如果需要，您也可以选择创建IAM角色或让管理员创建IAM角色。

7. 对于为此爬网程序创建计划，请保留默认频率（按需运行），然后选择下一步。
8. 在“配置爬虫的输出”中，输入 AWS Glue 数据库，然后选择“下一步”。
9. 查看所有详细信息，然后选择“完成”。
10. 在爬网程序页面上，选中 S3 存储桶旁边的复选框，然后选择运行爬网程序。
11. 爬网器运行完毕后，在 AWS Glue 导航栏，选择数据库，然后选择您的数据库名称。
12. 在数据库页面上，选择 {your database name} 中的表。
 - a. 查看中的表格 AWS Glue 数据库。
 - b. 要查看表的架构，请选择一个特定的表。
 - c. 记下来 AWS Glue 数据库名称和 AWS Glue 表名。

现在，您可以创建架构映射了。有关更多信息，请参阅 [创建架构映射](#)。

准备第三方输入数据

第三方数据服务提供的标识符可以与您的已知标识符相匹配。

AWS Entity Resolution 数据匹配服务 目前支持以下第三方数据提供商服务：

数据提供商服务

公司名	可用 AWS 区域	标识符
LiveRamp	美国东部（弗吉尼亚北部） （us-east-1）、美国东部（俄	坡道 ID

公司名	可用 AWS 区域	标识符
	亥俄州) (us-east-2) 和美国西部 (俄勒冈) (us-west-2)	
TransUnion	美国东部 (弗吉尼亚北部) (us-east-1)、美国东部 (俄亥俄州) (us-east-2) 和美国西部 (俄勒冈) (us-west-2)	TransUnion 个人和家庭 IDs
统一身份证 2.0	美国东部 (弗吉尼亚北部) (us-east-1)、美国东部 (俄亥俄州) (us-east-2) 和美国西部 (俄勒冈) (us-west-2)	raw UID 2

以下步骤介绍如何准备第三方数据，以使用基于[提供商服务的匹配工作流程](#)或基于[提供商服务的身份映射](#)工作流程。

主题

- [步骤 1：在上订阅提供商服务 AWS Data Exchange](#)
- [步骤 2：准备第三方数据表](#)
- [步骤 3：以支持的数据格式保存输入数据表](#)
- [步骤 4：将您的输入数据表上传到 Amazon S3](#)
- [步骤 5：创建一个 AWS Glue 表](#)

步骤 1：在上订阅提供商服务 AWS Data Exchange

如果您通过以下方式订阅了提供商服务 AWS Data Exchange，您可以使用以下提供商服务之一运行匹配的工作流程，将您的已知标识符与您的首选提供商进行匹配。您的数据将与您的首选提供商定义的一组输入相匹配。

要在上订阅提供商服务 AWS Data Exchange

1. 在上查看提供商列表 AWS Data Exchange。以下提供商列表可用：
 - LiveRamp
 - [LiveRamp 身份解析](#)

- [LiveRamp转码](#)
 - TransUnion
 - TransUnion TruAudience 无需转移的身份解析和充实
 - TransUnion TruAudience 无需转移的身份解析
 - 统一身份证 2.0
 - [统一 ID 2.0 身份解析](#)
2. 根据您的报价类型，完成以下步骤之一。
- 私人报价 — 如果您与提供商存在关系，请按照中的[私人产品和报价](#)程序进行操作 AWS Data Exchange 接受私人报价的用户指南 AWS Data Exchange.
 - 自带订阅 — 如果您已经向提供商订阅了现有的数据，请按照中的[自带订阅 \(BYOS\) 优惠](#)程序进行操作 AWS Data Exchange 接受报BYOS价的用户指南 AWS Data Exchange.
3. 在您订阅了提供商服务之后 AWS Data Exchange，然后您可以使用该提供商服务创建匹配的工作流程或 ID 映射工作流程。

有关如何访问包含以下内容的提供商产品的更多信息APIs，请参阅中的[访问API产品](#) AWS Data Exchange 用户指南。

步骤 2：准备第三方数据表

每种第三方服务都有一套不同的建议和指南，以帮助确保成功的匹配工作流程。

要准备第三方数据表，请参阅下表：

数据提供商服务指南

提供者服务	需要唯一的身份证吗？	操作
LiveRamp	是	请确保以下几点： <ul style="list-style-type: none"> • 唯一 ID 可以是您自己的匿名标识符，也可以是行 ID。 • 您的数据输入文件格式和标准化符合 LiveRamp 指导方针。

提供者服务	需要唯一的身份证吗？	操作
		<p>有关匹配工作流程的输入文件格式指南的更多信息，请参阅 LiveRamp 文档ADX中的通过执行身份解析。</p> <p>有关 ID 映射工作流程的输入文件格式指南的更多信息，请参阅 LiveRamp 文档ADX中的通过执行转码。</p>
TransUnion	是	<p>请确保以下几点：</p> <ul style="list-style-type: none"> 存在用于 TransUnion 数据扩充的唯一 ID。 <div data-bbox="883 747 1507 1016" style="border: 1px solid #add8e6; border-radius: 10px; padding: 10px; margin: 10px 0;"> <p> Note</p> <p>允许传递属性在输入和输出中保持不变 TransUnion。家用 E 密钥和特定 HHID 于客户端命名空间。</p> </div> <ul style="list-style-type: none"> Phone number 应为 10 位数字，不含任何特殊字符，例如空格或连字符。 Addresses 应该分成 <ul style="list-style-type: none"> 单个地址行（如果有，则合并地址行 1 和 2） city zip（或 zip plus4），不含任何特殊字符，例如空格或连字符 州，指定为 2 个字母的代码 3 Email addresses 应为纯文本。 First Name 可以是小写或大写，支持昵称，但应排除标题和后缀。 Last Name 可以是小写或大写，中间的首字母可以排除在外。

提供者服务	需要唯一的身份证吗？	操作
统一身份证 2.0	是	<p>请确保以下几点：</p> <ul style="list-style-type: none"> • 唯一 ID 不能是哈希。 • UID2支持UID2生成电子邮件和电话号码。但是，如果两个值都存在于架构映射中，则工作流会复制输出中的每条记录。一条记录使用电子邮件生UID2成，第二条记录使用电话号码。如果您的数据混合包含电子邮件和电话号码，并且您不希望在输出中出现这种重复的记录，那么最好的方法是为每个数据创建一个单独的工作流程，并使用不同的架构映射。在这种情况下，请执行两次步骤——为电子邮件创建一个工作流程，为电话号码创建一个单独的工作流程。 <div data-bbox="852 947 1508 1692" style="border: 1px solid #add8e6; border-radius: 10px; padding: 10px; margin-top: 10px;"> <p> Note</p> <p>无论是谁提出请求，特定的电子邮件或电话号码在任何特定时间都会产生相同的原始UID2价值。</p> <p>生盐UID2s是通过添加盐桶中的盐来制成的，这些盐桶大约每年轮换一次，这样生的盐也会UID2随之旋转。不同的盐桶在一年中的不同时间轮换。AWS Entity Resolution 数据匹配服务 目前无法跟踪旋转盐桶和未加工盐桶的情况UID2s，因此建议您每天重新生成未加工UID2s的盐桶。有关更多信息，请参阅增量更新UID2s应多久刷新一次？在 UID 2.0 文档中。</p> </div>

步骤 3：以支持的数据格式保存输入数据表

如果您已经以支持的数据格式保存了第三方输入数据，则可以跳过此步骤。

要将 AWS Entity Resolution 数据匹配服务，则输入数据的格式必须为 AWS Entity Resolution 数据匹配服务 支持。AWS Entity Resolution 数据匹配服务 支持以下数据格式：

- 逗号分隔值 () CSV

Note

LiveRamp 仅支持CSV文件。

- Parquet

步骤 4：将您的输入数据表上传到 Amazon S3

如果您在 Amazon S3 中已有第三方数据表，则可以跳过此步骤。

Note

输入数据必须存储在亚马逊简单存储服务 (Amazon S3) 中 AWS 账户 以及 AWS 区域 你想在其中运行匹配的工作流程。

将您的输入数据表上传到 Amazon S3

1. 登录 AWS Management Console 然后打开 Amazon S3 控制台，网址为<https://console.aws.amazon.com/s3/>。
2. 选择 Buckets，然后选择一个存储桶来存储您的数据表。
3. 选择上传，然后按照提示进行操作。
4. 选择对象选项卡，查看存储数据的前缀。记下文件夹的名称。

您可以选择要查看数据表的文件夹。

步骤 5：创建一个 AWS Glue 表

Amazon S3 中的输入数据必须编入目录 AWS Glue 并表示为 AWS Glue 桌子。有关如何创建的更多信息 AWS Glue 以 Amazon S3 为输入的表，请参阅上的“[使用爬虫](#)” [AWS Glue 控制台](#)在 AWS Glue 开发者指南。

Note

AWS Entity Resolution 数据匹配服务 不支持分区表。

在此步骤中，您将在中设置爬虫 AWS Glue 它会抓取 S3 存储桶中的所有文件并创建一个 AWS Glue 桌子。

Note

AWS Entity Resolution 数据匹配服务 目前不支持注册到的 Amazon S3 地点 AWS Lake Formation.

要创建 AWS Glue 表

1. 登录 AWS Management Console 然后打开 AWS Glue 控制台位于<https://console.aws.amazon.com/glue/>。
2. 从导航栏中，选择爬网程序。
3. 从列表中选择您的 S3 存储桶，然后选择添加爬网程序。
4. 在添加爬网程序页面上，输入爬网程序名称，然后选择下一步。
5. 继续浏览添加爬网程序页面，指定详细信息。
6. 在“选择IAM角色”页面上，选择“选择现有IAM角色”，然后选择“下一步”。

如果需要，您也可以选择创建IAM角色或让管理员创建IAM角色。

7. 对于为此爬网程序创建计划，请保留默认频率（按需运行），然后选择下一步。
8. 在“配置爬虫的输出”中，输入 AWS Glue 数据库，然后选择“下一步”。
9. 检查所有详细信息，然后选择完成。
10. 在爬网程序页面上，选中 S3 存储桶旁边的复选框，然后选择运行爬网程序。
11. 爬网程序运行完毕后，在 AWS Glue 导航栏，选择数据库，然后选择您的数据库名称。

12. 在数据库页面上，选择 {your database name} 中的表。
 - a. 查看中的表格 AWS Glue 数据库。
 - b. 要查看表的架构，请选择一个特定的表。
 - c. 记下来 AWS Glue 数据库名称和 AWS Glue 表名。

使用架构映射定义输入数据

架构映射定义了您要解析的输入数据。它还提供有关输入数据的元数据，例如列的属性类型（输入类型）以及要匹配的列。

创建架构映射时，首先要定义输入字段和输入类型，然后定义匹配键和分组相关数据。下图总结了如何创建架构映射。



Define your data

Import columns from an AWS Glue table, build a custom schema, or use a JSON editor.



Select input types

Assign a pre-defined input type for each input field to classify your data.



Assign match keys

Define a match key for each input field to enable comparison for your matching workflow.



Create data groups

Group related data that is separated into two or more input fields.

在创建架构映射之前，必须先进行设置 [AWS Entity Resolution 数据匹配服务](#) 并准备好您的数据表。有关更多信息，请参阅 [设置 AWS Entity Resolution 数据匹配服务](#) 和 [准备输入数据表](#)。

创建架构映射后，您可以执行以下操作之一：

- [创建匹配的工作流程](#) 以查找不同数据输入之间的匹配项。
- [创建 ID 命名空间源](#)，您可以在 ID 映射工作流程中使用该源将数据转换为目标。
- [在同一个工作流程中创建 ID 映射工作流程 AWS 账户](#) 使用您的架构映射作为源。

主题

- [创建架构映射](#)
- [克隆架构映射](#)
- [编辑架构映射](#)
- [删除架构映射](#)

创建架构映射

此过程描述了使用创建架构映射的过程 [AWS Entity Resolution 数据匹配服务 控制台](#)。

有三种方法可以创建架构映射：

- 使用“导入自”导入现有输入数据 AWS Glueoption — 使用此创建方法定义输入字段，从中预填充的列开始 AWS Glue 使用引导流程的表格。
- 使用“构建自定义架构”选项手动定义输入数据-使用此创建方法通过引导流程手动定义输入字段。
- 使用“使用JSON编辑器”选项手动创建-使用JSON编辑器手动创建、使用样本或导入现有输入数据。

Note

此选项不可用“唯一 ID”和“输入”字段。

Import from AWS Glue

通过从中导入现有输入数据来创建架构映射 AWS Glue

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“数据准备”下，选择“架构映射”。
3. 在架构映射页面的右上角，选择创建架构映射。
4. 对于步骤 1：指定架构详细信息，请执行以下操作：
 - a. 在名称和创建方法中，输入架构映射名称和可选的描述。
 - b. 在“创建方法”中，选择“导入来源”AWS Glue.
 - c. 选择 AWS Glue 从下拉列表中选择 database，然后选择 AWS Glue 下拉列表中的表格。

要创建新表，请转至 AWS Glue 控制台 <https://console.aws.amazon.com/glue/>。有关更多信息，请参阅 [AWS Glue](#) 中的桌子 AWS Glue 用户指南。

- d. 对于“唯一 ID”，请指定明确引用每行数据的列。

Example

例如，**Primary_key**、**Row_ID** 或 **Record_ID**。

Note

“唯一 ID”列为必填字段。唯一 ID 必须是单个表中的唯一标识符。但是，在不同的表中，唯一 ID 可能有重复的值。如果未指定唯一 ID、在同一来源中不是唯一的，或者不同源的属性名称存在重叠，则 AWS Entity Resolution 数据匹配服务 在运行

匹配的工作流程时拒绝该记录。如果您在基于规则的匹配工作流程中使用此架构映射，则唯一 ID 不得超过 38 个字符。

- e. 对于输入字段，选择要用于匹配和可选传递的列。

您最多可以选择 34 列进行匹配和直通。

- i. 在“匹配”下，选择要用作匹配输入字段的列。

您最多可以选择 24 列进行匹配。

- ii. 如果要指定不用于匹配的列，请选择“添加直通列”。

- iii. （可选）在“直通”下，选择要包含为直通列的列。

- f. （可选）如果要为资源启用标签，请选择添加新标签，然后输入密钥和值对。

- g. 选择下一步。

5. 对于“步骤 2：映射输入字段”，定义要用于匹配和可选传递的输入字段。

- a. 对于要匹配的输入字段，为每个输入字段指定输入类型、匹配键和哈希状态。

输入类型可帮助您对数据进行分类。Match 键允许将输入字段与您的匹配工作流程进行比较。哈希状态表示该输入字段的列值是经过哈希处理还是明文。

Note

如果您要创建架构映射以与基于 LiveRamp 提供者服务的匹配技术一起使用，则可以：

- 将输入类型指定为 LiveRampID。
- 将名称字段指定为多个字段（例如 **first_name**、**last_name**）或在一个字段中。
- 将街道地址字段指定为多个字段（例如 **address1**、**address2**）或一个字段。

如果与地址匹配，则需要邮政编码。

- 包括带有姓名的电子邮件或电话，这些字段可以与街道地址匹配。

Note

如果您要创建架构映射以用于基于机器学习的匹配工作流程，则您的数据集必须至少包含以下属性之

一：**phonenumber**、**emailaddress**、**fullname**、**addresses**、**birthdate**

或。**birthdate**
请勿将其中任何属性的输入类型指定为自定义字符串。

- b. (可选) 对于直通输入字段，添加不匹配的输入字段及其相应的哈希状态。

哈希状态表示该输入字段的列值是经过哈希处理还是明文。

- c. 选择下一步。

6. 对于“步骤 3：对数据进行分组”，请执行以下操作：

- a. 选择相关的“名称”字段，然后输入组名和匹配键。

Example

例如，选择输入字段**First name**、**Middle name**、和**Last name**。然后输入名为“**Full name**”的组名和名为“**Full name**”的匹配键以启用比较。

- b. 选择相关的地址字段，然后输入群组名称和匹配键。

Example

例如，选择输入字段**Home street address 1**、**Home street address 2**、和**Home city**。然后输入名为“**Shipping address**”的组名和名为“**Shipping address**”的匹配键以启用比较。

- c. 选择相关的电话号码字段，然后输入群组名称和匹配键。


Example

例如，选择输入字段**Home phone 1**、**Home phone 2**、和**Cell phone**。然后输入名为“**Shipping phone number**”的组名和名为“**Shipping phone number**”的匹配键以启用比较。

如果您有多种类型的数据，则可以添加更多组。

- d. 选择下一步。

7. 对于“步骤 4：查看并创建”，请执行以下操作：
 - a. 查看您在之前的步骤中所做的选择，并在必要时进行编辑。
 - b. 选择创建架构映射。

 Note

将架构映射与工作流程关联后，您无法对其进行修改。如果要使用现有配置来创建新的架构映射，则可以克隆架构映射。

创建架构映射后，就可以[创建匹配的工作流程](#)或[创建 ID 命名空间](#)了。


Build custom schema

使用“构建自定义架构”选项创建架构映射

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“数据准备”下，选择“架构映射”。
3. 在架构映射页面的右上角，选择创建架构映射。
4. 对于步骤 1：指定架构详细信息，请执行以下操作：
 - a. 在名称和创建方法中，输入架构映射名称和可选的描述。
 - b. 在“创建方法”中，选择“生成自定义架构”。
 - c. 在“唯一 ID”中，输入唯一的 ID 以识别您的每一行数据。

Example

例如，**Primary_key**、**Row_ID** 或 **Record_ID**。

 Note

“唯一 ID”列为必填字段。唯一 ID 必须是单个表中的唯一标识符。但是，在不同的表中，唯一 ID 可能有重复的值。如果未指定唯一 ID、在同一来源中不是唯一的，或者不同源的属性名称存在重叠，则 AWS Entity Resolution 数据匹配服务 在运行匹配的工作流程时拒绝该记录。如果您在基于规则的匹配工作流程中使用此架构映射，则唯一 ID 不得超过 38 个字符。


- d. (可选) 如果要为资源启用标签, 请选择添加新标签, 然后输入密钥和值对。
 - e. 选择下一步。
5. 对于“步骤 2: 映射输入字段”, 定义要用于匹配和可选传递的输入字段。

您最多可以为匹配和传递定义总共 34 列。


- a. 对于要匹配的输入字段, 添加一个输入字段及其对应的输入类型、匹配键和哈希状态。

您最多可以添加 24 个输入字段进行匹配。

输入类型可帮助您对数据进行分类。Match 键允许将输入字段与您的匹配工作流程进行比较。哈希状态表示该输入字段的列值是经过哈希处理还是明文。

 Note

如果您要创建架构映射以与基于 LiveRamp 提供者服务的匹配技术一起使用, 则可以将输入类型指定为 LiveRamp ID。如果要在输出中包含 PII 数据, 则必须将输入类型指定为自定义字符串。

 Note

如果您要创建架构映射以用于基于机器学习的匹配工作流程, 则您的数据集必须至少包含以下属性之一: **phonenumber**、**emailaddressfullnameaddresses**、或 **birthdate**
请勿将其中任何属性的输入类型指定为自定义字符串。

- b. (可选) 对于直通输入字段, 添加不匹配的输入字段及其相应的哈希状态。
 - c. 选择下一步。
6. 对于步骤 3: 对数据进行分组:
- a. 选择相关的“名称”字段, 然后输入组名和匹配键。

Example

例如, 选择输入字段 **First name**、**Middle name**、和 **Last name**。然后输入名为“**Full name**”的组名和名为“**Full name**”的匹配键以启用比较。

- b. 选择相关的地址字段，然后输入群组名称和匹配键。

Example

例如，选择输入字段**Home street address 1**、**Home street address 2**、和**Home city**。然后输入名为“**Shipping address**”的组名和名为“**Shipping address**”的匹配键以启用比较。

- c. 选择相关的电话号码字段，然后输入群组名称和匹配键。

Example

例如，选择输入字段**Home phone 1**、**Home phone 2**、和**Cell phone**。然后输入名为“**Shipping phone number**”的组名和名为“**Shipping phone number**”的匹配键以启用比较。

如果您有多种类型的数据，则可以添加更多组。

- d. 选择下一步。
7. 对于“步骤 4：查看并创建”，请执行以下操作：
 - a. 查看您在之前的步骤中所做的选择，并在必要时进行编辑。
 - b. 选择创建架构映射。

Note

将架构映射与工作流程关联后，您无法对其进行修改。如果要使用现有配置来创建新的架构映射，则可以克隆架构映射。

创建架构映射后，就可以[创建匹配的工作流程](#)或[创建 ID 命名空间](#)了。

Use JSON editor

使用JSON编辑器创建架构映射

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“数据准备”下，选择“架构映射”。
3. 在架构映射页面的右上角，选择创建架构映射。
4. 对于步骤 1：指定架构详细信息，请执行以下操作：

- a. 在名称和创建方法中，输入架构映射名称和可选的描述。
 - b. 在“创建方法”中，选择“使用JSON编辑器”。
 - c. （可选）如果要为资源启用标签，请选择添加新标签，然后输入密钥和值对。
 - d. 选择下一步。
5. 对于步骤 2：指定映射：
- a. 开始在JSON编辑器中构建架构，或者根据您的目标选择以下选项之一：

你的目标	推荐选项
开始构建架构映射	插入示例，JSON然后根据需要进行编辑。
使用现有JSON文件	从文件导入

- b. 选择下一步。
6. 对于步骤 3：查看并创建：
- a. 查看您在之前的步骤中所做的选择，并在必要时进行编辑。
 - b. 选择创建架构映射。

Note

将架构映射与工作流程关联后，您无法对其进行修改。如果要使用现有配置来创建新的架构映射，则可以克隆架构映射。

创建架构映射后，就可以[创建匹配的工作流程](#)或[创建 ID 命名空间](#)了。

克隆架构映射

如果要使用现有配置来创建新的架构映射，则可以克隆架构映射。

要克隆架构映射，请执行以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。

2. 在左侧导航窗格的“数据准备”下，选择“架构映射”。
3. 选择架构映射。
4. 选择克隆。
5. 在“指定架构详细信息”页面上，进行任何必要的更改，然后选择“下一步”。
6. 在“选择匹配技术”页面上，进行必要的更改，然后选择“下一步”。
7. 在地图输入字段页面上，进行任何必要的更改，然后选择下一步。
8. 在“分组数据”页面上，进行必要的更改，然后选择“下一步”。
9. 在“查看并保存”页面上，进行任何必要的更改，然后选择“克隆架构映射”。

编辑架构映射

在将架构映射关联到工作流程之前，您只能对其进行编辑。将架构映射与工作流程关联后，便无法对其进行编辑。如果要使用现有配置来创建新的架构映射，则可以克隆架构映射。

要编辑架构映射，请执行以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“数据准备”下，选择“架构映射”。
3. 选择架构映射。
4. 选择编辑。
5. 在“指定架构详细信息”页面上，进行任何必要的更改，然后选择“下一步”。
6. 在“选择匹配技术”页面上，进行必要的更改，然后选择“下一步”。
7. 在地图输入字段页面上，进行任何必要的更改，然后选择下一步。
8. 在“分组数据”页面上，进行必要的更改，然后选择“下一步”。
9. 在“查看并保存”页面上，进行任何必要的更改，然后选择“编辑架构映射”。

删除架构映射

当架构映射与匹配的工作流程关联时，您无法将其删除。必须先从所有关联的匹配工作流中移除架构映射，然后才能将其删除。

要删除架构映射，请执行以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“数据准备”下，选择“架构映射”。
3. 选择架构映射。
4. 选择删除。
5. 确认删除，然后选择删除。

使用 ID 命名空间定义输入数据

ID 命名空间是围绕输入数据表的包装。您可以使用 ID 命名空间来提供元数据，解释您的输入数据和匹配技术，以及如何在 [ID 映射工作流程](#) 中使用它们。

ID 命名空间有两种类型：源命名空间和目标空间。

- 源包含源数据的配置 AWS Entity Resolution 数据匹配服务 ID 映射工作流程中的进程。
- Target 包含所有源解析的目标数据配置。

您可以定义要在两个之间解析的输入数据 AWS 账户 在 ID 映射工作流程中。一个参与者创建 ID 命名空间源，另一个参与者创建 ID 命名空间目标。参与者创建源和目标后，您可以运行 ID 映射工作流程将数据从源转换为目标。

下图总结了如何创建用于 ID 映射工作流程的 ID 命名空间。



Prerequisite

An ID namespace that is a source requires a data input: [schema mapping](#) and an associated AWS Glue database. An ID namespace that is the target requires a target domain.



Create ID namespace

Provide the name and description, and then choose the type: source or target.



Configure your data

Select the configuration method and enter your source or target information.



Use in ID mapping workflows

Use your ID namespace as either a source or a target in an ID mapping workflow across two AWS accounts.

以下各节介绍如何创建 ID 命名空间源和 ID 命名空间目标。

主题

- [ID 命名空间来源](#)
- [ID 命名空间目标](#)
- [编辑 ID 命名空间](#)
- [删除 ID 命名空间](#)
- [添加或更新 ID 命名空间的资源策略](#)

ID 命名空间来源

ID 命名空间源是 [ID 映射工作流程](#) 中的数据源。

在创建 ID 命名空间源之前，必须先创建架构映射或匹配的工作流程，具体取决于您的用例。有关更多信息，请参阅[创建架构映射](#)和[使用匹配的工作流程匹配输入数据](#)。

创建 ID 命名空间源后，您可以在 ID 映射工作流程中将其与 ID 命名空间目标一起使用。有关更多信息，请参阅[使用 ID 映射工作流程映射输入数据](#)。

有两种方法可以在中创建 ID 命名空间源 AWS Entity Resolution 数据匹配服务 控制台：[基于规则的方法](#)或[提供者服务方法](#)。

主题

- [创建 ID 命名空间源 \(基于规则 \)](#)
- [创建 ID 命名空间源 \(提供者服务 \)](#)

创建 ID 命名空间源 (基于规则)

本主题介绍使用基于规则的方法创建 ID 命名空间源的过程。此方法使用匹配规则在 ID 映射工作流程中将第一方数据从源转换为目标。

Note

如果输入数据是源，则它必须具有架构映射和关联的 AWS Glue 数据库。

创建 ID 命名空间源 (基于规则)

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“数据准备”下，选择 ID 命名空间。
3. 在 ID 命名空间页面的右上角，选择创建 ID 命名空间。
4. 有关详细信息，请执行以下操作：
 - a. 在 ID 命名空间名称中，输入一个唯一的名称。
 - b. (可选) 在描述中，输入可选描述。
 - c. 对于 ID 命名空间类型，请选择来源。
5. 对于 ID 命名空间方法，请选择基于规则。
6. 对于数据输入，请选择要使用的输入类型，然后采取建议的操作。

提供者服务	建议的操作
现有的架构映射	<ol style="list-style-type: none"> 1. 选择架构映射。 2. 选择 AWS Glue 数据库，AWS Glue 表，以及下拉列表中的架构映射。 <p>您最多可以添加 20 个数据输入。</p>
现有的匹配工作流程	<ol style="list-style-type: none"> 1. 选择匹配工作流程。 2. 选择与 ID 命名空间关联的账户：要么是 AWS 账户或者另一个 AWS 账户。 3. 根据账户类型，选择匹配工作流程名称或输入匹配工作流程 ARN。

7. 对于规则参数，请执行以下操作。

a. 根据您的目标选择以下选项之一，指定规则控件。

你的目标	推荐选项
允许来自源和目标的规则	没有偏好
选择来源、目标或两者是否可以在 ID 映射工作流程中提供规则	有限的规则

规则控件必须在源和目标之间兼容，才能在 ID 映射工作流程中使用。例如，如果源 ID 命名空间将规则限制在目标上，但目标 ID 命名空间将规则限制在源上，则会导致错误。

b. 根据您的数据输入类型选择以下选项之一，指定匹配规则。

数据输入类型	推荐操作
架构映射	<p>选择“添加其他规则”以添加匹配的规则。</p> <p>您最多可以应用 25 个匹配规则来定义匹配标准。</p>

数据输入类型	推荐操作
匹配工作流程	选择“使用匹配工作流程中的规则”或“提供新规则”来定义您的匹配规则。

8. 对于比较和匹配参数，请执行以下操作。

a. 根据您的目标选择以下选项之一，指定比较类型。

你的目标	推荐选项
在创建 ID 映射工作流程时，允许使用任何比较类型。	没有偏好
查找存储在多个输入字段中的数据的任意匹配组合，无论数据位于相同还是不同的输入字段中。	多个输入字段
当存储在多个输入字段中的相似数据不应匹配时，请限制在单个输入字段内进行比较。	单个输入字段

b. 根据您的目标选择以下选项之一，指定“记录”匹配类型。

你的目标	推荐选项
在创建 ID 映射工作流程时，允许使用任何比较类型。	没有偏好
创建 ID 映射工作流程时，将记录匹配类型限制为仅在源中为目标中的每条匹配记录存储一条匹配记录。	有限的记录匹配 以及 一个源对一个目标
创建 ID 映射工作流程时，将记录匹配类型限制为在源中存储目标中每条匹配记录的所有匹配记录。	有限的记录匹配 以及 一个目标有多个来源

Note

您必须为源 ID 和目标 ID 命名空间指定兼容限制。例如，如果源 ID 命名空间将规则限制在目标上，但目标 ID 命名空间将规则限制在源上，则会导致错误。

9. 通过从下拉列表中选择现有服务角色名称来指定服务访问权限。
10. (可选) 要为资源启用标签，请选择添加新标签，然后输入密钥和值对。
11. 选择创建 ID 命名空间。

ID 命名空间源已创建。现在，您可以[创建 ID 命名空间目标](#)了。

创建 ID 命名空间源 (提供者服务)

本主题介绍使用提供者服务方法创建 ID 命名空间源的过程。此方法使用名为的提供者服务 LiveRamp。LiveRamp 在 ID 映射工作流程中，将第三方编码的数据从源转换为目标。

Note

如果输入数据是源，则它必须具有架构映射和关联的 AWS Glue 数据库。

创建 ID 命名空间源 (提供者服务)

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“数据准备”下，选择 ID 命名空间。
3. 在 ID 命名空间页面的右上角，选择创建 ID 命名空间。
4. 有关详细信息，请执行以下操作：
 - a. 在 ID 命名空间名称中，输入一个唯一的名称。
 - b. (可选) 在描述中，输入可选描述。
 - c. 对于 ID 命名空间类型，请选择来源。
5. 对于 ID 命名空间方法，请选择提供者服务。

Note

AWS Entity Resolution 数据匹配服务 目前 LiveRamp 提供提供者服务作为 ID 命名空间方法。如果您订阅了 LiveRamp，则状态将显示为“已订阅”。有关如何订阅的更多信息 LiveRamp，请参阅[步骤 1：在上订阅提供商服务 AWS Data Exchange](#)。

- 对于数据输入，选择 AWS Glue 数据库，AWS Glue 表，以及下拉列表中的架构映射。

您最多可以添加 20 个数据输入。

- 要指定服务访问权限，请选择一个选项并采取建议的操作。

选项	推荐操作
创建并使用新的服务角色	<ul style="list-style-type: none"> • AWS Entity Resolution 数据匹配服务 使用此表所需的策略创建服务角色。 • 默认的服务角色名称为 <code>entityresolution-id-mapping-workflow- <timestamp></code>。 • 您必须拥有创建角色并附加策略的权限。 • 如果您的输入数据已加密，请选择“此数据由 KMS 密钥加密”选项。然后，输入 AWS KMS 用于解密您的数据输入的密钥。
使用现有服务角色	<ol style="list-style-type: none"> 1. 从下拉列表中选择一个现有服务角色名称。 如果您有列出角色的权限，则会显示角色列表。 如果您无权列出角色，则可以输入要使用的角色的 Amazon 资源名称 (ARN)。 如果没有现有的服务角色，则使用现有服务角色选项不可用。 2. 通过选择“在 IAM 外部查看”链接来查看服务角色。

选项	推荐操作
	默认情况下，AWS Entity Resolution 数据匹配服务不会尝试更新现有角色策略以添加必要的权限。

- （可选）要为资源启用标签，请选择添加新标签，然后输入密钥和值对。
- 选择创建 ID 命名空间。

ID 命名空间源已创建。现在，您可以[创建 ID 命名空间目标](#)了。

ID 命名空间目标

ID 命名空间目标是 [ID 映射工作流程](#)中数据的目标。所有来源都解析到目标。

在创建 ID 命名空间目标之前，必须先创建匹配的工作流程或订阅提供者服务 (LiveRamp)，具体取决于您的用例。有关更多信息，请参阅[使用匹配的工作流程匹配输入数据](#)和[步骤 1：在上订阅提供商服务 AWS Data Exchange](#)。

创建 ID 命名空间目标后，您可以在 ID 映射工作流程中将其与 ID 命名空间源一起使用。有关更多信息，请参阅[使用 ID 映射工作流程映射输入数据](#)。

有两种方法可以在中创建 ID 命名空间目标 AWS Entity Resolution 数据匹配服务 控制台：[基于规则的方法](#)或[提供者服务方法](#)。

主题

- [创建 ID 命名空间目标（基于规则的方法）](#)
- [创建 ID 命名空间目标（提供者服务方法）](#)

创建 ID 命名空间目标（基于规则的方法）

本主题介绍使用基于规则的方法创建 ID 命名空间目标的过程。在 ID 映射工作流程中，此方法使用匹配规则将第一方数据从源转换为目标。

创建 ID 命名空间目标（基于规则）

- 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。

2. 在左侧导航窗格的“数据准备”下，选择 ID 命名空间。
3. 在 ID 命名空间页面的右上角，选择创建 ID 命名空间。
4. 有关详细信息，请执行以下操作：
 - a. 在 ID 命名空间名称中，输入一个唯一的名称。
 - b. （可选）在描述中，输入可选描述。
 - c. 对于 ID 命名空间类型，请选择目标。
5. 对于 ID 命名空间方法，请选择基于规则。
6. 对于数据输入，在“匹配工作流程”下，执行以下操作。
 - a. 选择与 ID 命名空间关联的账户：要么是 AWS 账户或者另一个 AWS 账户。
 - b. 根据账户类型，选择匹配工作流程名称或输入匹配工作流程 ARN。
7. 对于规则参数，请执行以下操作。
 - a. 根据您的目标选择以下选项之一，指定规则控件。

你的目标	推荐选项
允许来自源和目标的规则	没有偏好
选择来源、目标或两者是否可以在 ID 映射工作流程中提供规则	有限的规则

规则控件必须在源和目标之间兼容，才能在 ID 映射工作流程中使用。例如，如果源 ID 命名空间将规则限制在目标上，但目标 ID 命名空间将规则限制在源上，则会导致错误。


- b. 对于匹配规则，AWS Entity Resolution 数据匹配服务 自动添加匹配工作流程中的规则。
8. 对于比较和匹配参数，请执行以下操作。
 - a. 根据您的目标选择以下选项之一，指定比较类型。

你的目标	推荐选项
在创建 ID 映射工作流程时，允许使用任何比较类型。	没有偏好

你的目标	推荐选项
查找存储在多个输入字段中的数据的任意匹配组合，无论数据位于相同还是不同的输入字段中。	多个输入字段
当存储在多个输入字段中的相似数据不应匹配时，请限制在单个输入字段内进行比较。	单个输入字段

- b. 根据您的目标选择以下选项之一，指定“记录”匹配类型。

你的目标	推荐选项
在创建 ID 映射工作流程时，允许使用任何比较类型。	没有偏好
创建 ID 映射工作流程时，将记录匹配类型限制为仅在源中为目标中的每条匹配记录存储一条匹配记录。	有限的记录匹配 以及 一个源对一个目标
创建 ID 映射工作流程时，将记录匹配类型限制为在源中存储目标中每条匹配记录的所有匹配记录。	有限的记录匹配 以及 一个目标有多个来源

 Note

您必须为源 ID 和目标 ID 命名空间指定兼容限制。例如，如果源 ID 命名空间将规则限制在目标上，但目标 ID 命名空间将规则限制在源上，则会导致错误。

9. 通过从下拉列表中选择现有服务角色名称来指定服务访问权限。
10. (可选) 要为资源启用标签，请选择添加新标签，然后输入密钥和值对。
11. 选择创建 ID 命名空间。

ID 命名空间目标已创建。创建 ID 映射工作流程所需的 ID 命名空间（源和目标）后，就可以[创建 ID 映射](#)工作流程了。

创建 ID 命名空间目标（提供者服务方法）

本主题介绍使用提供者服务方法创建 ID 命名空间目标的过程。此方法使用名为的提供者服务 LiveRamp。LiveRamp 在 ID 映射工作流程中，将第三方编码的数据从源转换为目标。

创建 ID 命名空间目标（提供者服务）

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“数据准备”下，选择 ID 命名空间。
3. 在 ID 命名空间页面的右上角，选择创建 ID 命名空间。
4. 有关详细信息，请执行以下操作：
 - a. 在 ID 命名空间名称中，输入一个唯一的名称。
 - b. （可选）在描述中，输入可选描述。
 - c. 对于 ID 命名空间类型，请选择目标。
5. 对于 ID 命名空间方法，请选择提供者服务。

Note

AWS Entity Resolution 数据匹配服务 目前 LiveRamp 提供提供者服务作为 ID 命名空间方法。

如果您订阅了 LiveRamp，则状态将显示为“已订阅”。

有关如何订阅的更多信息 LiveRamp，请参阅[步骤 1：在上订阅提供商服务 AWS Data Exchange](#)。

6. 在目标域中，输入 LiveRamp 提供的转码目标 LiveRamp 客户机域标识符。
7. （可选）要为资源启用标签，请选择添加新标签，然后输入密钥和值对。
8. 选择创建 ID 命名空间。

ID 命名空间目标已创建。创建 ID 映射工作流程所需的 ID 命名空间（源和目标）后，即可开始[创建 ID 映射](#)工作流程。

编辑 ID 命名空间

在将 ID 命名空间关联到 ID 映射工作流程之前，您只能对其进行编辑。将 ID 命名空间与 ID 映射工作流程关联后，便无法对其进行编辑。

要编辑 ID 命名空间，请执行以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户（如果你还没有这样做的话）。
2. 在左侧导航窗格的“数据准备”下，选择 ID 命名空间。
3. 选择 ID 命名空间。
4. 选择编辑。
5. 在“编辑 ID 命名空间”页面上，进行必要的更改，然后选择“保存”。

删除 ID 命名空间

当 ID 命名空间与 ID 映射工作流程关联时，您无法将其删除。必须先从所有关联的 ID 映射工作流中移除架构映射，然后才能将其删除。

要删除 ID 命名空间，请执行以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户（如果你还没有这样做的话）。
2. 在左侧导航窗格的“数据准备”下，选择 ID 命名空间。
3. 选择 ID 命名空间。
4. 选择删除。
5. 确认删除，然后选择删除。

添加或更新 ID 命名空间的资源策略

资源策略允许 ID 映射资源的创建者访问您的 ID 命名空间资源。

添加或更新资源策略

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。

2. 在左侧导航窗格的“工作流程”下，选择 ID 命名空间。
3. 选择 ID 命名空间。
4. 在 ID 命名空间详细信息页面上，选择权限选项卡。
5. 在资源策略部分，选择编辑。
6. 在JSON编辑器中添加或更新策略。
7. 选择 Save changes (保存更改)。

使用匹配的工作流程匹配输入数据

匹配 workflow 是一种数据处理作业，它合并和比较来自不同输入源的数据，并根据不同的匹配技术确定哪些数据匹配。它生成一个数据输出表。

创建匹配 workflow 时，首先要指定数据输入、标准化步骤，然后选择所需的匹配技术和数据输出。AWS Entity Resolution 数据匹配服务 从您指定的一个或多个位置读取您的数据，并在您的数据中找到两条或多条记录之间的匹配项。然后，它为 [匹配的数据集中的记录分配一个匹配 ID](#)。AWS Entity Resolution 数据匹配服务 然后将数据输出文件写入您选择的位置。您可以使用 ... AWS Entity Resolution 数据匹配服务 如果需要，可以对输出数据进行哈希处理，从而帮助您保持对数据的控制。

匹配的工作流程可以有 multiple 运行，结果（成功或错误）将写入名称 jobId 为的文件夹。

数据输出包含成功匹配的文件和错误的文件。数据输出可以包含多个字段。成功结果将写入包含多个文件的文件 success 夹，每个文件都包含成功记录的子集。同样，错误会写入包含多个字段 error 的文件夹，每个字段都包含错误记录的子集。有关故障排除的更多信息，请参阅 [匹配 workflow 疑难解答](#)。

下图总结了如何创建匹配的工作流程。



Complete prerequisite

Create a schema mapping to define your data.



Choose your data input

Select the AWS Glue database and table that contains your data and the associated schema mapping.



Set up matching techniques

Configure rule-based matching, use machine learning matching, or choose a provider service.



Specify data output

Choose your data output fields and format to write to your S3 location.

在创建匹配的工作流程之前，必须先创建架构映射。有关更多信息，请参阅 [创建架构映射](#)。

[基于匹配技术创建匹配 workflow 的方法有三种：基于规则、基于机器学习或基于提供商服务。](#)

创建并运行匹配的工作流程后，您可以执行以下操作：

- 在您指定的 S3 位置查看结果。对数据进行索引 IDs 后会生成匹配的工作流程。
- 使用 [基于规则的匹配](#) 或 [机器学习 \(ML\) 匹配](#) 的输出作为 [基于提供商服务的匹配](#) 的输入，或者反过来满足您的业务需求。

例如，为了节省提供商订阅成本，您可以先运行 [基于规则的匹配](#) 来查找数据上的匹配项。然后，您可以将不匹配记录的子集发送给 [基于提供商服务的匹配](#)。

主题

- [创建基于规则的匹配工作流程](#)
- [创建基于机器学习的匹配工作流程](#)
- [创建基于提供商服务的匹配工作流程](#)
- [编辑匹配的工作流程](#)
- [删除匹配的工作流程](#)
- [为基于规则的匹配工作流程查找匹配 ID](#)
- [从基于规则或基于 ML 的匹配工作流程中删除记录](#)
- [匹配工作流程疑难解答](#)

创建基于规则的匹配工作流程

[基于规则的匹配](#)是一组分层的瀑布匹配规则，建议由 AWS Entity Resolution 数据匹配服务，基于您输入的数据，并且完全可以由您进行配置。基于规则的匹配工作流程使您可以比较明文数据或哈希数据，以根据您自定义的条件找到精确的匹配项。

时间 AWS Entity Resolution 数据匹配服务 在数据中的两条或多条记录之间找到匹配项，它会分配：

- 与[匹配数据集中的记录的匹配 ID](#)
- 生成[匹配项的匹配规则](#)。

创建基于规则的匹配工作流程

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户（如果你还没有这样做的话）。
2. 在左侧导航窗格的“工作流程”下，选择“匹配”。
3. 在匹配工作流程页面的右上角，选择创建匹配工作流程。
4. 对于“步骤 1：指定匹配的工作流程详细信息”，请执行以下操作：
 - a. 输入匹配的工作流程名称和可选的描述。
 - b. 对于数据输入，选择一个 AWS Glue 从下拉列表中选择 AWS Glue 表，然后是相应的架构映射。

您最多可以添加 19 个数据输入。
 - c. 默认情况下，“标准化数据”选项处于选中状态，以便在匹配之前对数据输入进行标准化。如果您不想对数据进行标准化处理，请取消选择“标准化数据”选项。

d. 要指定服务访问权限，请选择一个选项并采取建议的操作。

选项	推荐操作
创建并使用新的服务角色	<ul style="list-style-type: none"> • AWS Entity Resolution 数据匹配服务 使用此表所需的策略创建服务角色。 • 默认服务角色名称为 <code>entityresolution-matching-workflow-<timestamp></code>。 • 您必须拥有创建角色并附加策略的权限。 • 如果您的输入数据已加密，请选择“此数据由KMS密钥加密”选项。然后，输入 AWS KMS 用于解密您的数据输入的密钥。
使用现有服务角色	<ol style="list-style-type: none"> 1. 从下拉列表中选择一个现有服务角色名称。 <ul style="list-style-type: none"> 如果您有列出角色的权限，则会显示角色列表。 如果您无权列出角色，则可以输入要使用的角色的 Amazon 资源名称 (ARN)。 如果没有现有的服务角色，则使用现有服务角色选项不可用。 2. 通过选择“在IAM外部查看”链接来查看服务角色。 <ul style="list-style-type: none"> 默认情况下，AWS Entity Resolution 数据匹配服务 不会尝试更新现有角色策略以添加必要的权限。

e. (可选) 要为资源启用标签，请选择添加新标签，然后输入密钥和值对。

f. 选择下一步。

5. 对于步骤 2：选择匹配技术：

- a. 在“匹配方法”中，选择“基于规则的匹配”。

[AWS Entity Resolution](#) > [Matching workflows](#) > Create matching workflow

Choose matching technique info
Specify how you want your data to be matched or choose a provider service.

Matching method

Rule-based matching
Use customized rules to find exact matches.

Machine learning-based matching
Use our machine learning model to help find a broader range of matches.

Provider services
Use this option if you have a subscription to a preferred provider through AWS Data Exchange.

Rule-based matching info
Your data will be evaluated against a set of rules to find exact matches.

- Match keys are used as a basis for comparison and rules are automatically created based on your match keys.
- You can customize the rules for matching by editing the **Matching rules** section.

Processing cadence info
Determine how often to run your matching workflow job. The first job runs after you create the matching workflow. [See pricing](#)

Manual
Your matching workflow job is run on demand. Useful for bulk processing.

Automatic
Your matching workflow job is run automatically when you add or update your data inputs. Useful for incremental updates. This option is available only for rule-based matching.

Index only for ID mapping - new

Turn on
By default, matching workflows generate IDs after the data is indexed. If you want to use the matching workflow as a source or a target in an ID mapping workflow, choose to only index the data and not generate IDs.

- b. 对于处理节奏，请根据您的目标选择以下选项之一。

你的目标	推荐选项
按需运行工作流程以进行批量更新	手册
S3 存储桶中有新数据后立即运行工作流程	自动

Note

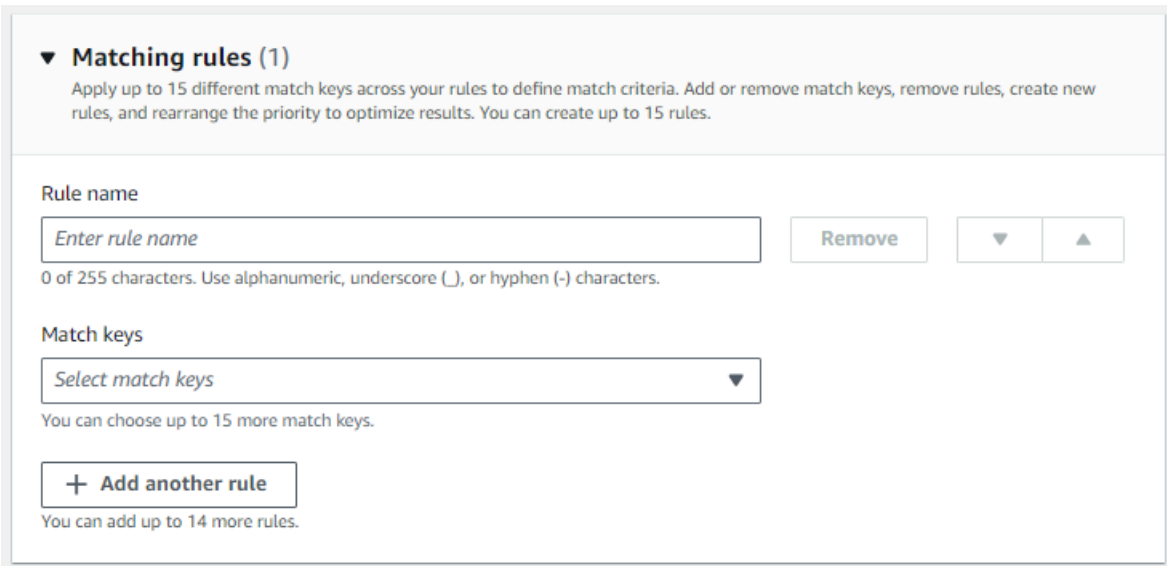
如果您选择“自动”，请确保您的 S3 存储桶已启用 Amazon EventBridge 通知。有关 EventBridge 使用 S3 控制台启用亚马逊的说明，请参阅 [Amazon S3 用户指南 EventBridge 中的启用亚马逊](#)。

- c. (可选) 对于仅适用于 ID 映射的索引，您可以选择启用仅索引数据而不生成数据的功能 IDs。

默认情况下，匹配的工作流程会在数据编制索引 IDs 后生成。

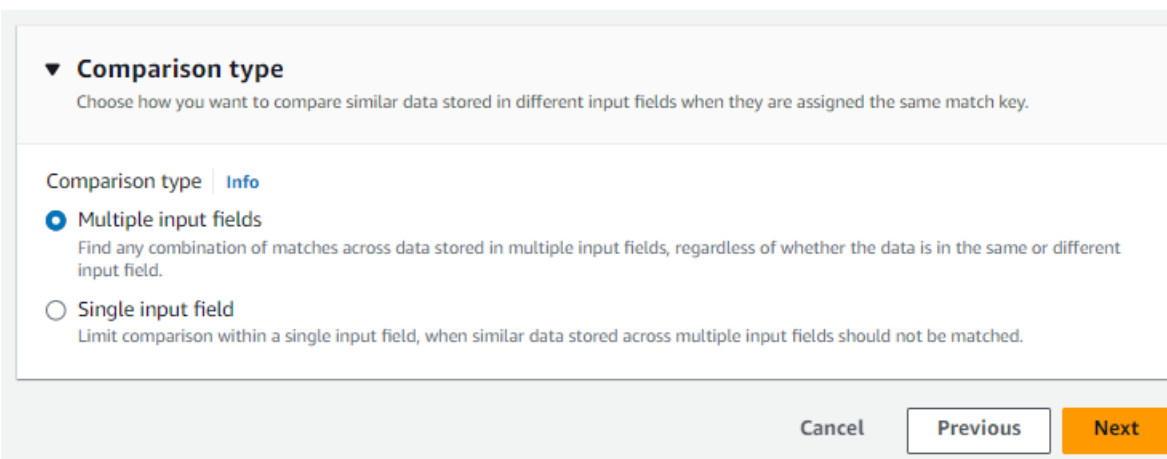
- d. 在匹配规则中，输入规则名称，然后为该规则选择匹配密钥。

您最多可以创建 15 个规则，并且可以在规则中应用最多 15 个不同的匹配密钥来定义匹配条件。



e. 对于“比较类型”，请根据您的目标选择以下选项之一。

你的目标	推荐选项
在存储在多个输入字段中的数据中查找任意匹配项组合	多个输入字段
将比较限制为单个输入字段	单一输入字段



f. 选择下一步。

6. 对于步骤 3：指定数据输出和格式：

- a. 对于数据输出目标和格式，选择数据输出的 Amazon S3 位置，以及数据格式是标准化数据还是原始数据。
- b. 对于加密，如果您选择自定义加密设置，请输入 AWS KMS 密钥ARN。
- c. 查看系统生成的输出。
- d. 对于数据输出，请决定要包含、隐藏或掩盖哪些字段，然后根据目标采取建议的操作。

你的目标	推荐选项
包括字段	将输出状态保持为“已包含”。
隐藏字段（从输出中排除）	选择“输出”字段，然后选择“隐藏”。
掩码字段	选择“输出”字段，然后选择“哈希输出”。
重置之前的设置	选择 重置。

- e. 选择下一步。

7. 对于步骤 4：查看并创建：

- a. 查看您在之前的步骤中所做的选择，并在必要时进行编辑。
- b. 选择创建并运行。

将出现一条消息，表示匹配的工作流程已创建且作业已启动。

8. 在匹配的工作流程详细信息页面的指标选项卡上，在“上次作业指标”下查看以下内容：

- Job ID。
- 匹配工作流作业的状态：已排队、进行中、已完成、失败
- 工作流作业的完成时间。
- 已处理的记录数。
- 未处理的记录数。
- IDs生成的唯一匹配项。
- 输入记录的数量。

您还可以查看任务历史记录下先前运行过的匹配工作流作业的作业指标。

9. 匹配的工作流程任务完成（状态为已完成）后，您可以转到数据输出选项卡，然后选择您的 Amazon S3 位置以查看结果。
10. （仅限手动处理类型）如果您创建了手动处理类型的基于规则的匹配 workflow，则可以在匹配的工作流详细信息页面上选择“运行 workflow”，随时运行匹配 workflow。

创建基于机器学习的匹配 workflow

[基于机器学习的匹配](#)是一个预设过程，它尝试匹配您输入的所有数据的记录。基于机器学习的匹配 workflow 使您能够使用机器学习模型比较明文数据以找到广泛的匹配项。

Note

机器学习模型不支持哈希数据的比较。

时间 AWS Entity Resolution 数据匹配服务 在数据中的两条或多条记录之间找到匹配项，它会分配：

- 与[匹配数据集中的记录的匹配 ID](#)
- 匹配[置信度](#)百分比。

您可以使用基于 ML 的匹配 workflow 的输出作为数据服务提供商匹配的输入，反之亦然，以实现您的特定目标。例如，您可以运行基于 ML 的匹配，先在自己的记录中查找数据源的匹配项。如果子集未匹配，则可以运行[基于提供商服务的匹配](#)来查找其他匹配项。

要创建基于 ML 的匹配 workflow，请执行以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户（如果你还没有这样做的话）。
2. 在左侧导航窗格的“workflow”下，选择“匹配”。
3. 在匹配 workflow 页面的右上角，选择创建匹配 workflow。
4. 对于“步骤 1：指定匹配的 workflow 详细信息”，请执行以下操作：
 - a. 输入匹配的 workflow 名称和可选的描述。
 - b. 对于数据输入，选择一个 AWS Glue 从下拉列表中选择 AWS Glue 表，然后是相应的架构映射。

您最多可以添加 20 个数据输入。

- c. 默认情况下，“标准化数据”选项处于选中状态，以便在匹配之前对数据输入进行标准化。如果您不想对数据进行标准化处理，请取消选择“标准化数据”选项。
- d. 要指定服务访问权限，请选择一个选项并采取建议的操作。

选项	推荐操作
创建并使用新的服务角色	<ul style="list-style-type: none"> • AWS Entity Resolution 数据匹配服务 使用此表所需的策略创建服务角色。 • 默认服务角色名称为 <code>entityresolution-matching-workflow-<timestamp></code> 。 • 您必须拥有创建角色并附加策略的权限。 • 如果您的输入数据已加密，请选择“此数据由KMS密钥加密”选项。然后，输入 AWS KMS 用于解密您的数据输入的密钥。
使用现有服务角色	<ol style="list-style-type: none"> 1. 从下拉列表中选择现有服务角色名称。 如果您有列出角色的权限，则会显示角色列表。 如果您无权列出角色，则可以输入要使用的角色的 Amazon 资源名称 (ARN)。 如果没有现有的服务角色，则使用现有服务角色选项不可用。 2. 通过选择“在IAM外部查看”链接来查看服务角色。 默认情况下，AWS Entity Resolution 数据匹配服务 不会尝试更新现有角色策略以添加必要的权限。

- e. (可选) 要为资源启用标签，请选择添加新标签，然后输入密钥和值对。
- f. 选择下一步。

5. 对于步骤 2：选择匹配技术：

a. 对于匹配方法，选择基于机器学习的匹配。

The screenshot shows the AWS Entity Resolution console interface for creating a matching workflow. The breadcrumb trail is 'AWS Entity Resolution > Matching workflows > Create matching workflow'. The left sidebar shows the progress: Step 1 (Specify matching workflow details), Step 2 (Choose matching technique), Step 3 (Specify data output), and Step 4 (Review and create). The main content area is titled 'Choose matching technique' and includes an 'Info' icon. Below the title, it says 'Specify how you want your data to be matched or choose a provider service.' There are three radio button options for the 'Matching method': 'Rule-based matching' (unselected), 'Machine learning-based matching' (selected), and 'Provider services' (unselected). Below these options, there is a section for 'Machine learning-based matching' with an 'Info' icon, explaining that data will be evaluated against a set of rules. Underneath, the 'Processing cadence' is set to 'Manual' (selected), with 'Automatic' as an alternative. A warning box at the bottom states: 'Using hashed data may limit matching functionality. Rule-based matching is recommended when comparing hashed data. The machine learning model is unable to compare hashed data. Learn more.' At the bottom right, there are 'Cancel', 'Previous', and 'Next' buttons.

b. 对于处理节奏，选择了“手动”选项。

此选项允许您按需运行工作流程以进行批量更新。

c. 选择下一步。

6. 对于步骤 3：指定数据输出和格式：

- 对于数据输出目标和格式，选择数据输出的 Amazon S3 位置，以及数据格式是标准化数据还是原始数据。
- 对于加密，如果您选择自定义加密设置，请输入 AWS KMS 密钥ARN。
- 查看系统生成的输出。
- 对于数据输出，请决定要包含、隐藏或掩盖哪些字段，然后根据目标采取建议的操作。

你的目标	推荐选项
包括字段	将输出状态保持为“已包含”。
隐藏字段 (从输出中排除)	选择“输出”字段，然后选择“隐藏”。
掩码字段	选择“输出”字段，然后选择“哈希输出”。
重置之前的设置	选择 重置。

e. 选择下一步。

7. 对于步骤 4：查看并创建：

- a. 查看您在之前的步骤中所做的选择，并在必要时进行编辑。
- b. 选择创建并运行。

将出现一条消息，表示匹配的工作流程已创建且作业已启动。

8. 在匹配的工作流程详细信息页面的指标选项卡上，在“上次作业指标”下查看以下内容：

- Job ID。
- 匹配 workflow 作业的状态：已排队、进行中、已完成、失败
- 工作流作业的完成时间。
- 已处理的记录数。
- 未处理的记录数。
- IDs 生成的唯一匹配项。
- 输入记录的数量。

您还可以查看任务历史记录下先前运行过的匹配 workflow 作业的作业指标。

9. 匹配的工作流程任务完成 (状态为已完成) 后，您可以转到数据输出选项卡，然后选择您的 Amazon S3 位置以查看结果。

10. (仅限手动处理类型) 如果您创建了手动处理类型的基于机器学习的匹配 workflow，则可以在匹配 workflow 详细信息页面上选择“运行 workflow”，随时运行匹配 workflow。

创建基于提供商服务的匹配工作流程

[基于提供商服务的匹配](#)使您能够将已知标识符与首选数据服务提供商进行匹配。

AWS Entity Resolution 数据匹配服务 目前支持以下数据提供商服务：

- LiveRamp
- TransUnion
- 统一身份证 2.0

有关支持的提供商服务的更多信息，请参阅[准备第三方输入数据](#)。

您可以通过以下方式使用这些提供商的公开订阅 AWS Data Exchange 或者直接与数据提供商协商私人报价。有关创建新订阅或重复使用提供商服务的现有订阅的更多信息，请参阅[步骤 1：在上订阅提供商服务 AWS Data Exchange](#)。

以下各节介绍如何创建基于提供者的匹配工作流程。

主题

- [使用创建匹配的工作流程 LiveRamp](#)
- [使用创建匹配的工作流程 TransUnion](#)
- [使用 UID 2.0 创建匹配的工作流程](#)

使用创建匹配的工作流程 LiveRamp

如果您订阅了该 LiveRamp 服务，则可以创建与该 LiveRamp 服务匹配的工作流程来执行身份解析。

该 LiveRamp 服务提供了一个名为 rampID 的标识符。RampID 是需求方平台 IDs 中最常用于为广告活动吸引受众的平台之一。使用与匹配的工作流程 LiveRamp，您可以将经过哈希处理的电子邮件地址解析为 RAMPIDs。

Note

AWS Entity Resolution 数据匹配服务 支持 PII 基于 rampID 的分配。

此工作流程需要一个 Amazon S3 数据暂存存储桶，您希望在其中临时写入匹配的工作流程输出。在使用创建 ID 映射工作流程之前 LiveRamp，请向数据暂存存储桶添加以下权限。


```
{
  "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>/*"
      ]
    }
  ]
}
```

替换每个 *<user input placeholder>* 用你自己的信息。

staging-bucket

Amazon S3 存储桶，用于在运行基于提供商服务的工作流程时临时存储您的数据。

要创建匹配的工作流程，请执行 LiveRamp 以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户（如果你还没有这样做的话）。
2. 在左侧导航窗格的“工作流程”下，选择“匹配”。
3. 在匹配工作流程页面的右上角，选择创建匹配工作流程。
4. 对于“步骤 1：指定匹配的工作流程详细信息”，请执行以下操作：
 - a. 输入匹配的工作流程名称和可选的描述。
 - b. 对于数据输入，选择一个 AWS Glue 从下拉列表中选择 AWS Glue 表，然后选择相应的架构映射。

您最多可以添加 20 个数据输入。

- c. 默认情况下，“标准化数据”选项处于选中状态，以便在匹配之前对数据输入进行标准化。


如果您使用的是仅限电子邮件的解析流程，请取消选择“标准化数据”选项，因为只有经过哈希处理的电子邮件才用于输入数据。

- d. 要指定服务访问权限，请选择一个选项并采取建议的操作。

选项	推荐操作
创建并使用新的服务角色	<ul style="list-style-type: none"> • AWS Entity Resolution 数据匹配服务 使用此表所需的策略创建服务角色。 • 默认服务角色名称为 <code>entityresolution-matching-workflow-<timestamp></code>。 • 您必须拥有创建角色并附加策略的权限。 • 如果您的输入数据已加密，请选择“此数据由KMS密钥加密”选项。然后，输入 AWS KMS 用于解密您的数据输入的密钥。

选项	推荐操作
使用现有服务角色	<p>1. 从下拉列表中选择一个现有服务角色名称。</p> <p>如果您有列出角色的权限，则会显示角色列表。</p> <p>如果您无权列出角色，则可以输入要使用的角色的 Amazon 资源名称 (ARN)。</p> <p>如果没有现有的服务角色，则使用现有服务角色选项不可用。</p> <p>2. 通过选择“在IAM外部查看”链接来查看服务角色。</p> <p>默认情况下，AWS Entity Resolution 数据匹配服务不会尝试更新现有角色策略以添加必要的权限。</p>

- e. (可选) 要为资源启用标签，请选择添加新标签，然后输入密钥和值对。
 - f. 选择下一步。
5. 对于步骤 2：选择匹配技术：
- a. 在“匹配方法”中，选择“提供者服务”。
 - b. 对于提供商服务，请选择LiveRamp。

 Note

确保您的数据输入文件格式和标准化符合提供商服务的指南。
有关匹配工作流程的输入文件格式指南的更多信息，请参阅 LiveRamp 文档ADX中的[通过执行身份解析](#)。

- c. 对于LiveRamp 产品，请从下拉列表中选择产品。

Matching method

Rule-based matching
Use customized rules to find exact matches.


Machine learning-based matching
Use our machine learning model to help find a broader range of matches.

Provider services
Use this option if you have a subscription to a preferred provider through AWS Data Exchange.

Provider services [Info](#)

You must have a provider agreement to use a provider service. Your data will be matched with a set of inputs defined by your preferred provider. Some information may be required and shared between you and your provider service.

LiveRamp
/LiveRamp

TransUnion
TransUnion 

Unified ID 2.0
Unified iD _{2.0}

LiveRamp products

Choose from available products from LiveRamp.

[Cancel](#) [Previous](#) [Next](#)

Note

如果选择“分配”PII，则在执行实体解析时必须至少提供一个非标识符列。例如，GENDER。

- d. 要进行LiveRamp 配置，请输入客户端 ID 管理ARN器和客户机密管理器ARN。

LiveRamp configuration
These are the required fields to use the LiveRamp service.

Client ID manager ARN
Enter the Client ID manager ARN provided by LiveRamp.
arn:aws:secretsmanager:us-east-1: [redacted] :secret:[redacted]
83 of 2,048 characters.

Client secret manager ARN
Enter the Client secret manager ARN provided by LiveRamp.
arn:aws:secretsmanager:us-east-1: [redacted] :secret:[redacted]
87 of 2,048 characters.

Data staging [Info](#)
Choose the Amazon S3 location for temporarily storing your data while it processes. Your information will not be saved permanently.

Amazon S3 location
s3:// [redacted] X View [View](#) [Browse S3](#)

Cancel Previous **Next**

- e. 对于数据暂存，请选择 Amazon S3 位置，以便在处理数据时临时存储数据。

您必须拥有访问 Amazon S3 数据暂存位置的权限。有关更多信息，请参阅 [为创建工作流程工作角色 AWS Entity Resolution 数据匹配服务](#)。

- f. 选择下一步。

6. 对于步骤 3：指定数据输出：

- 对于数据输出目标和格式，选择数据输出的 Amazon S3 位置，以及数据格式是标准化数据还是原始数据。
- 对于加密，如果您选择自定义加密设置，请输入 AWS KMS 密钥ARN。
- 查看LiveRamp 生成的输出。

这是由生成的其他信息 LiveRamp。

- 对于数据输出，请决定要包含、隐藏或掩盖哪些字段，然后根据目标采取建议的操作。

Note

如果您已选择 LiveRamp，则由于 LiveRamp 隐私过滤器会删除个人身份信息 (PII)，某些字段将显示“不可用”的输出状态。

你的目标	推荐选项
包括字段	将输出状态保持为“已包含”。
隐藏字段（从输出中排除）	选择“输出”字段，然后选择“隐藏”。
掩码字段	选择“输出”字段，然后选择“哈希输出”。
重置之前的设置	选择 重置。

AWS Entity Resolution > ID mapping workflows > Create ID mapping workflow

Step 1 Specify ID mapping workflow details

Step 2 Specify source and target

Step 3 - optional

Step 4 Specify data output location

Step 4 Review and create

Specify data output location - optional Info

Choose your S3 location to write your data output.

Data output destination Info

Choose the Amazon S3 location for the data output.

Amazon S3 location

View 🔗 Browse S3

Encryption - optional Info

Your data is encrypted by default with a key that AWS owns and manages for you. To specify a different key, customize your encryption settings.

Customize encryption settings
Specify an AWS KMS key to customize your encryption settings.

▼ LiveRamp generated output (2)

Additional information generated by LiveRamp.

Output field	Description
RAMPID	LiveRamp's universal identifier that is tied to devices in the LiveRamp Identity Graph
TRANSCODED_IDENTIFIER	LiveRamp's universal identifier that is tied to devices in the LiveRamp Identity Graph

Cancel Previous Next

e. 选择下一步。

7. 对于步骤 4：查看并创建：

- a. 查看您在之前的步骤中所做的选择，并在必要时进行编辑。
- b. 选择创建并运行。

将出现一条消息，表示匹配的工作流程已创建且作业已启动。

8. 在匹配的工作流程详细信息页面的指标选项卡上，在“上次作业指标”下查看以下内容：

- Job ID。
- 匹配工作流作业的状态：已排队、进行中、已完成、失败
- 工作流作业的完成时间。
- 已处理的记录数。
- 未处理的记录数。
- IDs生成的唯一匹配项。
- 输入记录的数量。

您还可以查看任务历史记录下先前运行过的匹配工作流作业的作业指标。

9. 匹配的工作流程任务完成（状态为已完成）后，您可以转到数据输出选项卡，然后选择您的 Amazon S3 位置以查看结果。

使用创建匹配的工作流程 TransUnion

如果您订阅了该 TransUnion 服务，则可以通过使用 TransUnion 个人和家庭电子密钥以及200多个数据属性链接、匹配和增强存储在不同渠道上的客户相关记录来提高对客户理解。

该 TransUnion 服务提供名为“TransUnion 个人和家庭”的标识符IDs。TransUnion 提供已知标识符（例如姓名、地址、电话号码和电子邮件地址）的 ID 分配（也称为编码）。

此工作流程需要一个 Amazon S3 数据暂存存储桶，您希望在其中临时写入匹配的工作流程输出。在使用创建匹配的工作流程之前 TransUnion，请向数据暂存存储桶添加以下权限。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::103054336026: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::103054336026:root"
    },
    "Action": [
      "s3:ListBucket",
      "s3:GetBucketLocation",
      "s3:GetBucketPolicy",
      "s3:ListBucketVersions",
      "s3:GetBucketAcl"
    ],
    "Resource": [
      "arn:aws:s3:::<staging-bucket>",
      "arn:aws:s3:::<staging-bucket>/*"
    ]
  }
]
```

替换每个 *<user input placeholder>* 用你自己的信息。

staging-bucket

Amazon S3 存储桶，用于在运行基于提供商服务的工作流程时临时存储您的数据。

要创建匹配的工作流程，请执行 TransUnion 以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户（如果你还没有这样做的话）。
2. 在左侧导航窗格的“工作流程”下，选择“匹配”。
3. 在匹配工作流程页面的右上角，选择创建匹配工作流程。
4. 对于“步骤 1：指定匹配的工作流程详细信息”，请执行以下操作：
 - a. 输入匹配的工作流程名称和可选的描述。
 - b. 对于数据输入，选择一个 AWS Glue 从下拉列表中选择 AWS Glue 表，然后选择相应的架构映射。


您最多可以添加 20 个数据输入。

- c. 默认情况下，“标准化数据”选项处于选中状态，以便在匹配之前对数据输入进行标准化。如果您不想对数据进行标准化处理，请取消选择“标准化数据”选项。
- d. 要指定服务访问权限，请选择一个选项并采取建议的操作。

选项	推荐操作
创建并使用新的服务角色	<ul style="list-style-type: none"> • AWS Entity Resolution 数据匹配服务 使用此表所需的策略创建服务角色。 • 默认服务角色名称为 <code>entityresolution-matching-workflow- <timestamp></code>。 • 您必须拥有创建角色并附加策略的权限。 • 如果您的输入数据已加密，请选择“此数据由KMS密钥加密”选项。然后，输入 AWS KMS 用于解密您的数据输入的密钥。
使用现有服务角色	<ol style="list-style-type: none"> 1. 从下拉列表选择一个现有服务角色名称。 <p>如果您有列出角色的权限，则会显示角色列表。</p>

选项	推荐操作
	<p>如果您无权列出角色，则可以输入要使用的角色的 Amazon 资源名称 (ARN)。</p> <p>如果没有现有的服务角色，则使用现有服务角色选项不可用。</p> <p>2. 通过选择“在IAM外部查看”链接来查看服务角色。</p> <p>默认情况下，AWS Entity Resolution 数据匹配服务 不会尝试更新现有角色策略以添加必要的权限。</p>

- e. (可选) 要为资源启用标签，请选择添加新标签，然后输入密钥和值对。
 - f. 选择下一步。
5. 对于步骤 2：选择匹配技术：
- a. 在“匹配方法”中，选择“提供者服务”。
 - b. 对于提供商服务，请选择TransUnion。

 Note

确保您的数据输入文件格式和标准化符合提供商服务的指南。

- c. 对于TransUnion 产品，请从下拉列表中选择产品。

[AWS Entity Resolution](#) > [Matching workflows](#) > Create matching workflow

Step 1
[Specify matching workflow details](#)

Step 2
Choose matching technique

Step 3
Specify data output

Step 4
Review and create

Choose matching technique [Info](#)

Specify how you want your data to be matched or choose a provider service.

Matching method

Rule-based matching
Use customized rules to find exact matches.

Machine learning-based matching
Use our machine learning model to help find a broader range of matches.

Provider services
Use this option if you have a subscription to a preferred provider through AWS Data Exchange.

Provider services [Info](#)

You must have a provider agreement in order to use a provider service. Your data will be matched with a set of inputs defined by your preferred provider. Some information may be required and shared between you and your provider service.

LiveRamp

TransUnion

Unified ID 2.0

TransUnion products
Choose from available products from TransUnion.

Choose product ▼

Cancel Previous **Next**

- d. 对于数据暂存，请选择 Amazon S3 位置，以便在处理数据时临时存储数据。

您必须拥有访问 Amazon S3 数据暂存位置的权限。有关更多信息，请参阅 [the section called “创建工作流程工作角色”](#)。

6. 选择下一步。
7. 对于步骤 3：指定数据输出：
- 对于数据输出目标和格式，选择数据输出的 Amazon S3 位置，以及数据格式是标准化数据还是原始数据。
 - 对于加密，如果您选择自定义加密设置，请输入 AWS KMS 密钥ARN。
 - 查看TransUnion 生成的输出。

这是由生成的其他信息 TransUnion。

- d. 对于数据输出，请决定要包含、隐藏或掩盖哪些字段，然后根据目标采取建议的操作。

你的目标	推荐选项
包括字段	将输出状态保持为“已包含”。
隐藏字段（从输出中排除）	选择“输出”字段，然后选择“隐藏”。
掩码字段	选择“输出”字段，然后选择“哈希输出”。
重置之前的设置	选择 重置。

- e. 对于系统生成的输出，请查看包含的所有字段。

- f. 选择下一步。

8. 对于步骤 4：查看并创建：

- 查看您在之前的步骤中所做的选择，并在必要时进行编辑。
- 选择创建并运行。

将出现一条消息，表示匹配的工作流程已创建且作业已启动。

9. 在匹配的工作流程详细信息页面的指标选项卡上，在“上次作业指标”下查看以下内容：

- Job ID。
- 匹配工作流作业的状态：已排队、进行中、已完成、失败
- 工作流作业的完成时间。
- 已处理的记录数。
- 未处理的记录数。
- IDs生成的唯一匹配项。
- 输入记录的数量。

您还可以查看任务历史记录下先前运行过的匹配工作流作业的作业指标。

10. 匹配的工作流程任务完成（状态为已完成）后，您可以转到数据输出选项卡，然后选择您的 Amazon S3 位置以查看结果。

使用 UID 2.0 创建匹配的工作流程

如果您订阅了 Unified ID 2.0 服务，则可以激活具有确定性身份的广告活动，并依靠与广告生态系统中许多 UID2 支持参与者的互操作性。有关更多信息，请参阅 [Unified ID 2.0 概述](#)。

Unified ID 2.0 服务提供 raw UID 2，用于在 The Trade Desk 平台中制作广告活动。UID2.0 是使用开源框架生成的。

在一个工作流程中，您可以将其中一个 **Email Address** 或 **Phone number** 用于原始生 UID2 成，但不能同时使用两者。如果两者都存在于架构映射中，则工作流将选择，**Email Address** 然后 **Phone number** 将是直通字段。要同时支持这两者，请创建一个新的架构映射 **Phone number**，其中已映射但 **Email Address** 未映射。然后，使用这个新的架构映射创建第二个工作流程。

Note

生盐 UID2s 是通过添加盐桶中的盐来制成的，这些盐桶大约每年轮换一次，这样生的盐也会 UID2 随之旋转。因此，建议您 UID2s 每天刷新 raw。有关更多信息，请参阅 <https://unifiedid.com/docs/getting-started/gs-faqs# 2-incremental-how-often-should-uid-updates. s-be-refreshed-for>

要使用 UID 2.0 创建匹配的工作流程，请执行以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户（如果你还没有这样做的话）。
2. 在左侧导航窗格的“工作流程”下，选择“匹配”。
3. 在匹配工作流程页面的右上角，选择创建匹配工作流程。
4. 对于“步骤 1：指定匹配的工作流程详细信息”，请执行以下操作：
 - a. 输入匹配的工作流程名称和可选的描述。
 - b. 对于数据输入，选择一个 AWS Glue 从下拉列表中选择 AWS Glue 表，然后选择相应的架构映射。

您最多可以添加 20 个数据输入。
 - c. 保持“标准化数据”选项处于选中状态，以便在匹配之前对数据输入（**Email Address** 或 **Phone number**）进行标准化。

有关 **Email Address** 标准化的更多信息，请参阅 UID 2.0 文档中的 [电子邮件地址标准化](#)。

有关 **Phone number** 标准化的更多信息，请参阅 UID 2.0 文档中的 [电话号码标准化](#)。

d. 要指定服务访问权限，请选择一个选项并采取建议的操作。

选项	推荐操作
创建并使用新的服务角色	<ul style="list-style-type: none"> • AWS Entity Resolution 数据匹配服务 使用此表所需的策略创建服务角色。 • 默认服务角色名称为 <code>entityresolution-matching-workflow-<timestamp></code>。 • 您必须拥有创建角色并附加策略的权限。 • 如果您的输入数据已加密，请选择“此数据由KMS密钥加密”选项。然后，输入 AWS KMS 用于解密您的数据输入的密钥。
使用现有服务角色	<ol style="list-style-type: none"> 1. 从下拉列表中选择一个现有服务角色名称。 <ul style="list-style-type: none"> 如果您有列出角色的权限，则会显示角色列表。 如果您无权列出角色，则可以输入要使用的角色的 Amazon 资源名称 (ARN)。 如果没有现有的服务角色，则使用现有服务角色选项不可用。 2. 通过选择“在IAM外部查看”链接来查看服务角色。 <ul style="list-style-type: none"> 默认情况下，AWS Entity Resolution 数据匹配服务 不会尝试更新现有角色策略以添加必要的权限。

e. (可选) 要为资源启用标签，请选择添加新标签，然后输入密钥和值对。

f. 选择下一步。

5. 对于步骤 2：选择匹配技术：
 - a. 在“匹配方法”中，选择“提供者服务”。
 - b. 对于提供商服务，请选择统一 ID 2.0。

The screenshot shows the AWS Entity Resolution console interface for creating a matching workflow. The breadcrumb trail is 'AWS Entity Resolution > Matching workflows > Create matching workflow'. The left sidebar shows the progress: Step 1 (Specify matching workflow details), Step 2 (Choose matching technique), Step 3 (Specify data output), and Step 4 (Review and create). The main content area is titled 'Choose matching technique' with an 'Info' link. Below the title is the instruction: 'Specify how you want your data to be matched or choose a provider service.' Under the 'Matching method' section, three options are available: 'Rule-based matching' (unselected), 'Machine learning-based matching' (unselected), and 'Provider services' (selected). Below this, the 'Provider services' section explains that a provider agreement is required. Three provider options are shown: 'LiveRamp' (unselected), 'TransUnion' (unselected), and 'Unified ID 2.0' (selected). Below the 'Unified ID 2.0' option, it indicates 'Access to Unified ID 2.0 provider subscription' and shows a green checkmark with the text 'Subscribed'. At the bottom right, there are 'Cancel', 'Previous', and 'Next' buttons.

- c. 选择下一步。
6. 对于步骤 3：指定数据输出：
 - a. 对于数据输出目标和格式，选择数据输出的 Amazon S3 位置，以及数据格式是标准化数据还是原始数据。
 - b. 对于加密，如果您选择自定义加密设置，请输入 AWS KMS 密钥ARN。
 - c. 查看 Unified ID 2.0 生成的输出。

这是 UID 2.0 生成的所有其他信息的列表

- d. 对于数据输出，请决定要包含、隐藏或掩盖哪些字段，然后根据目标采取建议的操作。

你的目标	推荐选项
包括字段	将输出状态保持为“已包含”。
隐藏字段 (从输出中排除)	选择“输出”字段，然后选择“隐藏”。
掩码字段	选择“输出”字段，然后选择“哈希输出”。
重置之前的设置	选择 重置。

- e. 对于系统生成的输出，请查看包含的所有字段。
 - f. 选择下一步。
7. 对于步骤 4：查看并创建：
- a. 查看您在之前的步骤中所做的选择，并在必要时进行编辑。
 - b. 选择创建并运行。

将出现一条消息，表示匹配的工作流程已创建且作业已启动。

8. 在匹配的工作流程详细信息页面的指标选项卡上，在“上次作业指标”下查看以下内容：
- Job ID。
 - 匹配 workflow 作业的状态：已排队、进行中、已完成、失败
 - 工作流作业的完成时间。
 - 已处理的记录数。
 - 未处理的记录数。
 - IDs 生成的唯一匹配项。
 - 输入记录的数量。

您还可以查看任务历史记录下先前运行过的匹配 workflow 作业的作业指标。

9. 匹配的工作流程任务完成 (状态为已完成) 后，您可以转到数据输出选项卡，然后选择您的 Amazon S3 位置以查看结果。

编辑匹配的工作流程

要编辑匹配的工作流程，请执行以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“工作流程”下，选择“匹配”。
3. 选择匹配的工作流程。
4. 在匹配的工作流程详细信息页面上，选择右上角的编辑。
5. 在“指定匹配的工作流程详细信息”页面上，进行任何必要的更改，然后选择“下一步”。
6. 在“选择匹配技术”页面上，进行必要的更改，然后选择“下一步”。
7. 在“指定数据输出”页面上，进行必要的更改，然后选择“下一步”。
8. 在“查看并保存”页面上，进行必要的更改，然后选择“保存”。

删除匹配的工作流程

要删除匹配的工作流程，请执行以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“工作流程”下，选择“匹配”。
3. 选择匹配的工作流程。
4. 在匹配的工作流程详细信息页面的右上角，选择删除。
5. 确认删除，然后选择删除。

为基于规则的匹配工作流程查找匹配 ID

运行基于规则的匹配工作流程后，您可以找到相应的 Match ID 和已处理记录的关联规则。

要查找基于规则的匹配工作流程的匹配 ID，请执行以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“工作流程”下，选择“匹配”。

3. 选择已处理的基于规则的匹配工作流程 (Job 状态为“已完成”)。
4. 在匹配工作流程详细信息页面上，选择查找匹配 ID 选项卡。
5. 请执行以下操作之一：

如果...	操作...
只有一个架构映射与此工作流程相关联。	查看默认情况下选择的架构映射。
有多个架构映射与此工作流程相关联。	从下拉列表中选择架构映射。

6. 展开“匹配规则”。
7. 为每个匹配键输入一个值。

默认情况下，“标准化数据”选项处于选中状态，以便在匹配之前对数据输入进行标准化。如果您不想对数据进行标准化处理，请取消选择“标准化数据”选项。

Tip

输入尽可能多的值以帮助找到匹配 ID。

8. 选择查找。
9. 查看相应的匹配 ID 和用于匹配的关联规则。

从基于规则或基于 ML 的匹配工作流程中删除记录

如果您需要遵守数据管理法规，则可以从基于规则或基于机器学习的匹配工作流程中删除记录。

从基于规则或基于 ML 的匹配工作流程中删除记录

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“工作流程”下，选择“匹配”。
3. 选择基于规则或基于 ML 的匹配工作流程。
4. 在匹配的工作流程详细信息页面上，IDs 从“操作”下拉列表中选择“删除唯一”。
5. 在“唯一”IDs 部分输入要删除的唯一 ID。

您最多可以输入 10 个独一无二的 IDs。

6. 指定要从中删除唯一值的输入源IDs。

如果工作流程只有一个输入源，则默认情况下会列出输入源。

如果您只指定一个输入源，则其他输入源IDs中的唯一输入源不会受到影响。

7. 选择“删除唯一”IDs。

匹配工作流程疑难解答

使用以下信息来帮助您诊断和修复运行匹配工作流程时可能遇到的常见问题。

我在运行匹配的工作流程后收到了错误文件

常见原因

匹配的工作流程可以有多个运行，结果（成功或错误）将写入名称jobId为的文件夹。

匹配工作流程的成功结果将写入包含多个文件的文件success夹，每个文件都包含成功记录的子集。

匹配工作流程的错误将写入包含多个字段error的文件夹，每个字段都包含错误记录的子集。

创建错误文件的原因如下：

- [唯一 ID](#) 是：
 - null
 - 一行数据中缺失
 - 数据表中的一条记录中缺失
 - 在数据表的另一行数据中重复
 - 未指定
 - 在同一个来源中不是唯一的
 - 在多个来源中不是唯一的
 - 跨源重叠
 - 超过 38 个字符（仅限基于规则的匹配工作流程）
- [架构映射](#) 中的一个字段包含一个保留名称：
 - EmailAddress
 - InputSourceARN

- MatchRule
- matchID
- HashingProtocol
- ConfidenceLevel
- 来源

Note

如果错误文件中的记录是由于前面列出的原因而创建的，则需要向您收费，因为这会产生服务的处理成本。如果错误文件中的记录是由于内部服务器错误造成的，则无需向您收费。

解决方案

要解决这个问题

1. 检查[唯一 ID](#) 是否有效。

如果[唯一 ID](#) 无效，请更新数据表中的唯一 ID，保存新的数据表，创建新的架构映射，然后再次运行匹配的工作流程。

2. 检查[架构映射](#)中的一个字段是否包含保留名称。

如果其中一个字段包含保留名称，请使用新名称创建新的架构映射，然后再次运行匹配的工作流程。

使用 ID 映射工作流程映射输入数据

ID 映射工作流程是一种数据处理作业，它根据指定的 ID 映射方法将输入数据源中的数据映射到输入数据目标。它生成一个 ID 映射表。

ID 映射工作流程需要输入数据源和输入数据目标。您的数据输入源和目标取决于您要执行的 ID 映射类型。执行 ID 映射的方法有两种：基于规则的服务或提供商的服务：

- 基于规则 ID 映射-您可以使用匹配规则将第一方数据从源转换为目标。
- 提供商服务 ID 映射-您可以使用 LiveRamp 提供者服务将第三方数据从源转换为目标。

Note

中的提供商服务 ID 映射工作流程 AWS Entity Resolution 数据匹配服务 目前已与 LiveRamp。如果您订阅了该 LiveRamp 服务，则可以使用创建 ID 映射工作流程 LiveRamp 来执行转码。通过 LiveRamp 转码，您可以将一组源 RampIDs 转换为任何目标目标 ramPid。通过使用 RamPid 作为代表客户的代币，您可以避免直接与广告平台共享客户数据。

有关更多信息，请参阅 LiveRamp 文档网站 ADX 上的 [执行直通翻译](#)。

在以下任一方案中，您可以在两个数据集之间执行 ID 映射：

- 在你自己的内心深处 AWS 账户
- 跨越两个不同的 AWS 账户

下图总结了如何设置 ID 映射工作流程。



Complete prerequisite

Create a [schema mapping](#) for ID mapping in your AWS account or an [ID namespace](#) for ID mapping across AWS accounts to define your data.



Specify ID mapping details

Provide details for your ID mapping workflow and choose an ID mapping method.



Specify source and target

Use a schema mapping or ID namespace to describe your input data depending on your ID mapping type.



Specify data output location - optional

Choose your S3 location to write your data output.

主题

- [一个人的身份映射工作流程 AWS 账户](#)
- [跨两个 ID 映射工作流程 AWS 账户](#)
- [运行 ID 映射工作流程](#)
- [使用新的输出目标运行 ID 映射工作流程](#)
- [编辑 ID 映射工作流程](#)
- [删除 ID 映射工作流程](#)
- [为 ID 映射工作流程添加或更新资源策略](#)

一个人的身份映射工作流程 AWS 账户

一个人的身份映射工作流程 AWS 账户使您能够自己在两个数据集之间执行 ID 映射 AWS 账户。

在您自己创建 ID 映射工作流程之前 AWS 账户，您必须先完成[先决条件](#)。

创建并运行 ID 映射工作流程后，您可以查看输出（ID 映射表）并将其用于分析。

以下主题将引导您完成一系列步骤，以便在同一工作流程中创建 ID 映射工作流程 AWS 账户。

主题

- [先决条件](#)
- [创建 ID 映射工作流程（基于规则）](#)
- [创建 ID 映射工作流程（提供者服务）](#)

先决条件

在为一个人创建 ID 映射工作流程之前 AWS 账户 使用基于规则或提供商服务 ID 映射方法，必须首先执行以下操作：

- 完成[设置AWS实体解析](#)中的任务。
- [创建架构映射](#)或[创建匹配的工作流程](#)。
- （仅限提供商服务 ID 映射）在使用创建 ID 映射工作流程之前 LiveRamp，必须选择要临时写入身份映射工作流程输出的亚马逊简单存储服务 (Amazon S3) Service 数据暂存存储桶。

如果您使用 LiveRamp 提供商服务来翻译第三方数据，请添加以下权限策略，该策略允许您访问数据暂存存储桶。

```
{
  "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>/*"
      ]
    }
  ]
}
```

在前面的权限策略中，替换每个 *<user input placeholder>* 用你自己的信息。

staging-bucket

Amazon S3 存储桶，用于在运行基于提供商服务的工作流程时临时存储您的数据。

创建 ID 映射工作流程（基于规则）

本主题介绍为一个人创建 ID 映射工作流的过程 AWS 账户 它使用匹配规则将第一方数据从来源转换为目标。

为一个人创建基于规则的 ID 映射工作流程 AWS 账户

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“工作流程”下，选择 ID 映射。
3. 在 ID 映射工作流程页面的右上角，选择创建 ID 映射工作流程。
4. 对于步骤 1：指定 ID 映射工作流程详细信息，请执行以下操作。
 - a. 输入 ID 映射工作流程名称和可选描述。

AWS Entity Resolution > ID mapping workflows > Create ID mapping workflow

Step 1
Specify ID mapping workflow details

Step 2
Specify source and target

Step 3 - optional
Specify data output location

Step 4
Review and create

Specify ID mapping workflow details Info

Provide details for your ID mapping workflow and choose an ID mapping method.

Name

ID mapping workflow name

Enter name

0 of 255 characters. Use alphanumeric, underscore (_), or hyphen (-) characters. Name must be unique across all ID mapping workflows in your account.

Description - optional

Enter description

0 of 255 characters.

- b. 对于 ID 映射方法，请选择基于规则。
 - c. （可选）要为资源启用标签，请选择添加新标签，然后输入密钥和值对。
 - d. 选择下一步。
5. 对于“步骤 2：指定源和目标”，请执行以下操作。
 - a. 对于 S ourc e，选择适用于您的场景，然后采取建议的操作。

场景	推荐操作
在 ID 映射工作流程中使用自己AWS的 Glue 数据库、AWS Glue 表和架构映射。	<ol style="list-style-type: none"> 1. 选择架构映射。 2. 选择一个 AWS Glue从下拉列表中选择 AWS Glue 表，然后选择相应的架构映射。 <p>您最多可以添加 19 个数据输入。</p>
使用现有的匹配工作流程，该工作流程指向要在 ID 映射工作流程中使用的记录数据。	<ol style="list-style-type: none"> 1. 选择“匹配工作流程”。 2. 从下拉列表中选择现有的匹配工作流程。

- b. 对于 Target，从下拉列表选择一个现有的匹配工作流程。
- c. 对于规则参数，请执行以下操作。
 - i. 根据您的源类型选择以下选项之一，指定规则控件。


源类型	推荐操作
匹配工作流程	<p>通过选择来源、目标或两者是否可以在 ID 映射工作流中提供规则来指定规则控件。</p> <p>规则控件必须在源和目标之间兼容，才能在 ID 映射工作流程中使用。</p> <p>例如，如果源 ID 命名空间将规则限制在目标上，但目标 ID 命名空间将规则限制在源上，则会导致错误。</p>
架构映射	跳过此步骤。

- ii. 对于比较和匹配参数，比较类型自动设置为多个输入字段。

这是因为两个参与者之前都选择了此选项。

- d. 根据您的目标选择以下选项之一，指定“记录”匹配类型。

你的目标	推荐选项
创建 ID 映射工作流程时，将记录匹配类型限制为仅在源中为目标中的每条匹配记录存储一条匹配记录。	一个源对一个目标
创建 ID 映射工作流程时，将记录匹配类型限制为在源中存储目标中每条匹配记录的所有匹配记录。	同一个目标有多个来源

 Note

您必须为源 ID 和目标 ID 命名空间指定兼容限制。

- e. 要指定服务访问权限，请选择一个选项并采取建议的操作。

Service access

AWS Entity Resolution requires permissions to read your data input from AWS Glue and write to S3 on your behalf. [View policy document](#)

Choose a method to authorize AWS Entity Resolution

- Create and use a new service role
Automatically create the role and add the necessary permissions policy.
- Use an existing service role

Service role name

entityresolution-id-mapping-workflow-20240117121045

51 of 64 characters. Use alphanumeric and '+=, @-_' characters. Don't include spaces. Name must be unique across all roles in the account.

- This data is encrypted with a KMS key
Specify the associated KMS key to enable AWS Entity Resolution to access each of your data inputs.

选项	推荐操作
创建并使用新的服务角色	<ul style="list-style-type: none"> • AWS Entity Resolution 数据匹配服务 使用此表所需的策略创建服务角色。 • 默认服务角色名称为 <code>entityresolution-id-mapping-workflow-<timestamp></code> 。 • 您必须拥有创建角色并附加策略的权限。 • 如果您的输入数据已加密，请选择“此数据由KMS密钥加密”选项。然后，输入 AWS KMS 用于解密您的数据输入的密钥。
使用现有服务角色	<ol style="list-style-type: none"> 1. 从下拉列表中选择一個现有服务角色名称。 <p>如果您有列出角色的权限，则会显示角色列表。</p> <p>如果您无权列出角色，则可以输入要使用的角色的 Amazon 资源名称 (ARN)。</p> <p>如果没有现有的服务角色，则使用现有服务角色选项不可用。</p> 2. 通过选择“在IAM外部查看”链接来查看服务角色。 <p>默认情况下，AWS Entity Resolution 数据匹配服务 不会尝试更新现有角色策略以添加必要的权限。</p>

6. 选择下一步。

7. 对于步骤 3：指定数据输出位置（可选），请执行以下操作。

a. 对于数据输出目标，请执行以下操作：

i. 选择数据输出的 Amazon S3 位置。

- ii. 对于加密，如果您选择自定义加密设置，则输入 AWS KMS 键ARN或选择“创建” AWS KMS 钥匙。
 - b. 选择下一步。
8. 对于“步骤 4：查看并创建”，请执行以下操作。
 - a. 查看您在之前的步骤中所做的选择，并在必要时对其进行编辑。
 - b. 选择创建。

将出现一条消息，表明身份映射工作流程已创建。

创建 ID 映射工作流程后，就可以[运行身份映射工作流程](#)了。

创建 ID 映射工作流程（提供者服务）

本主题介绍为一个人创建 ID 映射工作流的过程 AWS 账户 使用名为的提供者服务 LiveRamp。LiveRamp 使用维护的或派生的 RampIDs 将一组源 RampIDs 转换为另一个集合。

为一个人创建基于提供者服务的身份映射工作流程 AWS 账户

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“工作流程”下，选择 ID 映射。
3. 在 ID 映射工作流程页面的右上角，选择创建 ID 映射工作流程。
4. 对于步骤 1：指定 ID 映射工作流程详细信息，请执行以下操作。
 - a. 输入 ID 映射工作流程名称和可选描述。

The screenshot shows the AWS Entity Resolution console interface for creating an ID mapping workflow. The breadcrumb navigation at the top reads: AWS Entity Resolution > ID mapping workflows > Create ID mapping workflow. On the left, a vertical progress bar shows four steps: Step 1 (Specify ID mapping workflow details, currently active), Step 2 (Specify source and target), Step 3 - optional (Specify data output location), and Step 4 (Review and create). The main content area is titled 'Specify ID mapping workflow details' with an 'Info' icon. Below the title is the instruction: 'Provide details for your ID mapping workflow and choose an ID mapping method.' The form contains two fields: 'Name' with the label 'ID mapping workflow name' and a text input field containing 'Enter name'. Below the name field is the text: '0 of 255 characters. Use alphanumeric, underscore (_), or hyphen (-) characters. Name must be unique across all ID mapping workflows in your account.' The second field is 'Description - optional' with a text input field containing 'Enter description'. Below the description field is the text: '0 of 255 characters.'

- b. 对于 ID 映射方法，请选择提供商服务。

AWS Entity Resolution 数据匹配服务 目前提供 LiveRamp 提供者服务作为 ID 映射方法。如果您订阅了 LiveRamp，则状态将显示为“已订阅”。有关如何订阅的更多信息 LiveRamp，请参阅 [步骤 1：在上订阅提供商服务 AWS Data Exchange](#)。

ID mapping method Info

/LiveRamp

Currently we are only offering LiveRamp service as an ID mapping method.

Access to LiveRamp provider subscription

✔ Subscribed

i To ensure a successful workflow run, your data input file format and normalization must be aligned with the provider service's guidelines. [Learn more](#)

i Note

确保您的数据输入文件格式符合提供商服务的指南。有关输入文件格式指南 LiveRamp 的更多信息，请参阅 LiveRamp 文档网站 ADX 上的 [执行直通翻译](#)。

- c. 要进行 LiveRamp 配置，请输入以下 LiveRamp 提供的值：

- 客户端 ID 管理器 ARN
- 客户机密管理器 ARN

LiveRamp configuration Info

Client ID manager ARN

Enter the Client ID manager ARN provided by LiveRamp.

0 of 2,048 characters.

Client secret manager ARN

Enter the Client secret manager ARN provided by LiveRamp.

0 of 2,048 characters.

- d. (可选) 要为资源启用标签，请选择添加新标签，然后输入密钥和值对。

- e. 选择下一步。
5. 对于“步骤 2：指定源和目标”，请执行以下操作。
- a. 对于 Source，选择适用于您的场景，然后采取建议的操作。

场景	推荐操作
在 ID 映射工作流程中使用自己 AWS 的 Glue 数据库、AWS Glue 表和架构映射。	<ol style="list-style-type: none"> 1. 选择架构映射。 2. 选择一个 AWS Glue 从下拉列表中选择 AWS Glue 表，然后选择相应的架构映射。 <p>您最多可以添加 19 个数据输入。</p>
使用现有的匹配工作流程，该工作流程指向要在 ID 映射工作流程中使用的记录数据。	<ol style="list-style-type: none"> 1. 选择“匹配工作流程”。 2. 从下拉列表中选择现有的匹配工作流程。

- b. 对于 Target，请根据您选择的 ID 映射方法执行以下操作之一。

身份映射方法	推荐操作
基于规则	从下拉列表中选择现有的匹配工作流程。
提供商服务	<p>输入目标域中 LiveRamp 提供的转码目标 LiveRamp 客户端域标识符。</p> 

- c. 对于数据暂存，请选择您要临时写入 ID 映射工作流程输出的 Amazon S3 位置。

Data staging [Info](#)

Choose the Amazon S3 location for temporarily storing your data while it processes. Your information will not be saved permanently.

Amazon S3 location

[View](#) [Browse S3](#)

d. 要指定服务访问权限，请选择一个选项并采取建议的操作。

Service access

AWS Entity Resolution requires permissions to read your data input from AWS Glue and write to S3 on your behalf. [View policy document](#)

Choose a method to authorize AWS Entity Resolution

Create and use a new service role
Automatically create the role and add the necessary permissions policy.

Use an existing service role

Service role name

entityresolution-id-mapping-workflow-20240117121045

51 of 64 characters. Use alphanumeric and '+,=, @, -, _' characters. Don't include spaces. Name must be unique across all roles in the account.

This data is encrypted with a KMS key
Specify the associated KMS key to enable AWS Entity Resolution to access each of your data inputs.

选项	推荐操作
创建并使用新的服务角色	<ul style="list-style-type: none"> • AWS Entity Resolution 数据匹配服务 使用此表所需的策略创建服务角色。 • 默认服务角色名称为 entityresolution-id-mapping-workflow-<timestamp>。 • 您必须拥有创建角色并附加策略的权限。 • 如果您的输入数据已加密，请选择“此数据由KMS密钥加密”选项。然后，输入 AWS KMS 用于解密您的数据输入的密钥。

选项	推荐操作
使用现有服务角色	<ol style="list-style-type: none">1. 从下拉列表中选择一个现有服务角色名称。 如果您有列出角色的权限，则会显示角色列表。 如果您无权列出角色，则可以输入要使用的角色的 Amazon 资源名称 (ARN)。 如果没有现有的服务角色，则使用现有服务角色选项不可用。2. 通过选择“在IAM外部查看”链接来查看服务角色。 默认情况下，AWS Entity Resolution 数据匹配服务不会尝试更新现有角色策略以添加必要的权限。

6. 选择下一步。

7. 对于步骤 3：指定数据输出位置（可选），请执行以下操作。

a. 对于数据输出目标，请执行以下操作：

i. 选择数据输出的 Amazon S3 位置。

ii. 对于加密，如果您选择自定义加密设置，则输入 AWS KMS 键ARN或选择“创建”AWS KMS 钥匙。

b. 查看LiveRamp 生成的输出。

c. 选择下一步。

AWS Entity Resolution > ID mapping workflows > Create ID mapping workflow

Step 1
Specify ID mapping workflow details

Step 2
Specify source and target

Step 3 - optional
Specify data output location

Step 4
Review and create

Specify data output location - optional Info

Choose your S3 location to write your data output.

Data output destination Info
Choose the Amazon S3 location for the data output.

Amazon S3 location

Q s3://bucket/prefix View Browse S3

Encryption - optional Info
Your data is encrypted by default with a key that AWS owns and manages for you. To specify a different key, customize your encryption settings.

Customize encryption settings
Specify an AWS KMS key to customize your encryption settings.

▼ LiveRamp generated output (2)
Additional information generated by LiveRamp.

Output field	Description
RAMPID	LiveRamp's universal identifier that is tied to devices in the LiveRamp Identity Graph
TRANSCODED_IDENTIFIER	LiveRamp's universal identifier that is tied to devices in the LiveRamp Identity Graph

Cancel Previous Next

8. 对于“步骤 4：查看并创建”，请执行以下操作。
 - a. 查看您在之前的步骤中所做的选择，并在必要时对其进行编辑。
 - b. 选择创建。

将出现一条消息，表明身份映射工作流程已创建。

9. 创建 ID 映射工作流程后，就可以[运行身份映射工作流程](#)了。

跨两个 ID 映射工作流程 AWS 账户

跨两个 ID 映射工作流程 AWS 账户使您能够在两个数据集之间的两个数据集之间执行 ID 映射 AWS 账户。这通常是在你自己之间完成的 AWS 账户 还有另一个 AWS 账户。

例如，发布者可以使用自己的目标 ID 命名空间（在自己的命名空间中）创建 ID 映射工作流程 AWS 账户和广告商的来源 ID 命名空间（在另一个命名空间中）AWS 账户。

在创建跨两个 ID 映射工作流程之前 AWS 账户，您必须先完成[先决条件](#)。

创建 ID 映射工作流程后，您可以查看输出（ID 映射表）并将其用于分析。

以下主题将引导您完成一系列步骤，在两个步骤之间创建 ID 映射工作流程 AWS 账户：

主题

- [先决条件](#)
- [创建 ID 映射工作流程 \(基于规则 \)](#)
- [创建 ID 映射工作流程 \(提供者服务 \)](#)

先决条件

在创建跨两个 ID 映射工作流程之前 AWS 账户，您必须先执行以下操作：

- 完成[设置 AWS Entity Resolution 数据匹配服务](#)中所述的任务。
- [创建 ID 命名空间源](#)。
- [创建 ID 命名空间目标](#)。
- ARN如果您使用的是其他来源的 ID 命名空间源，请获取 ID 命名空间 AWS 账户。
- (仅限提供者服务) 在两者之间创建 ID 映射工作流程 AWS 账户 需要权限 LiveRamp 才能访问 S3 存储桶和 AWS Key Management Service (AWS KMS) 客户管理的密钥。

在创建跨两个 ID 映射工作流程之前 AWS 账户 使用 LiveRamp，添加以下权限策略，该策略 LiveRamp 允许访问 S3 存储桶和客户托管密钥。

```
{
  "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"
      }
    }
  ]
}
```

在前面的权限策略中，替换每个 `<user input placeholder>` 用你自己的信息。

`<KMSKeyARN>`

o ARN 为 AWS KMS 客户管理的密钥。

创建 ID 映射工作流程（基于规则）

完成[先决条件](#)后，您可以创建一个或多个 ID 映射工作流程，以使用匹配规则将第一方数据从源转换为目标。

在两者之间创建基于规则的 ID 映射工作流程 AWS 账户

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“工作流程”下，选择 ID 映射。
3. 在 ID 映射工作流程页面的右上角，选择创建 ID 映射工作流程。
4. 对于步骤 1：指定 ID 映射工作流程详细信息，请执行以下操作。
 - a. 输入 ID 映射工作流程名称和可选描述。

- b. 对于 ID 映射方法，请选择基于规则。
 - c. （可选）要为资源启用标签，请选择添加新标签，然后输入密钥和值对。
 - d. 选择下一步。
5. 对于“步骤 2：指定源和目标”，请执行以下操作。
 - a. 打开“高级选项”。

- b. 对于来源，选择匹配工作流程，然后从下拉列表中选择现有的匹配工作流程。
- c. 对于 Target，选择匹配工作流程，然后从下拉列表中选择现有的匹配工作流程。
- d. 对于规则参数，通过选择源还是目标可以在 ID 映射工作流程中提供规则来指定规则控件。

规则控件必须在源和目标之间兼容，才能在 ID 映射工作流程中使用。例如，如果源 ID 命名空间将规则限制在目标上，但目标 ID 命名空间将规则限制在源上，则会导致错误。

- e. 对于比较和匹配参数，请执行以下操作。
 - i. 通过根据您的目标选择一个选项来指定比较类型。

你的目标	推荐选项
查找存储在多个输入字段中的数据的任意匹配组合，无论数据位于相同还是不同的输入字段中。	多个输入字段
当存储在多个输入字段中的相似数据不应匹配时，请限制在单个输入字段内进行比较。	单一输入字段

- ii. 通过根据您的目标选择一个选项来指定“记录”匹配类型。

你的目标	推荐选项
创建 ID 映射工作流程时，将记录匹配类型限制为仅在源中为目标中的每条匹配记录存储一条匹配记录。	一个源对一个目标
创建 ID 映射工作流程时，将记录匹配类型限制为在源中存储目标中每条匹配记录的所有匹配记录。	同一个目标有多个来源

Note

您必须为源 ID 和目标 ID 命名空间指定兼容限制。

- f. 要指定服务访问权限，请选择一个选项并采取建议的操作。

Service access

AWS Entity Resolution requires permissions to read your data input from AWS Glue and write to S3 on your behalf. [View policy document](#)

Choose a method to authorize AWS Entity Resolution

- Create and use a new service role
Automatically create the role and add the necessary permissions policy.
- Use an existing service role

Service role name

entityresolution-id-mapping-workflow-20240117121045

51 of 64 characters. Use alphanumeric and '+=, @-_' characters. Don't include spaces. Name must be unique across all roles in the account.

- This data is encrypted with a KMS key
Specify the associated KMS key to enable AWS Entity Resolution to access each of your data inputs.

选项	推荐操作
创建并使用新的服务角色	<ul style="list-style-type: none"> • AWS Entity Resolution 数据匹配服务 使用此表所需的策略创建服务角色。 • 默认服务角色名称为 <code>entityresolution-id-mapping-workflow- -<timestamp></code> 。 • 您必须拥有创建角色并附加策略的权限。 • 如果您的输入数据已加密，请选择“此数据由KMS密钥加密”选项。然后，输入 AWS KMS 用于解密您的数据输入的密钥。

选项	推荐操作
使用现有服务角色	<p>1. 从下拉列表中选择一个现有服务角色名称。</p> <p>如果您有列出角色的权限，则会显示角色列表。</p> <p>如果您无权列出角色，则可以输入要使用的角色的 Amazon 资源名称 (ARN)。</p> <p>如果没有现有的服务角色，则使用现有服务角色选项不可用。</p> <p>2. 通过选择“在IAM外部查看”链接来查看服务角色。</p> <p>默认情况下，AWS Entity Resolution 数据匹配服务不会尝试更新现有角色策略以添加必要的权限。</p>

6. 选择下一步。
7. 对于步骤 3：指定数据输出位置（可选），请执行以下操作。
 - a. 对于数据输出目标，请执行以下操作。
 - i. 选择数据输出的 Amazon S3 位置。
 - ii. 对于加密，如果您选择自定义加密设置，则输入 AWS KMS 键ARN或选择“创建”AWS KMS 钥匙。
 - b. 查看LiveRamp 生成的输出。
 - c. 选择下一步。
8. 对于“步骤 4：查看并创建”，请执行以下操作。
 - a. 查看您在之前的步骤中所做的选择，并在必要时对其进行编辑。
 - b. 选择创建。

将出现一条消息，表明身份映射工作流程已创建。

创建 ID 映射工作流程后，就可以[运行身份映射工作流程](#)了。

创建 ID 映射工作流程 (提供者服务)

完成[先决条件](#)后，您可以使用 LiveRamp 提供者服务创建一个或多个 ID 映射工作流程。LiveRamp 使用维护的或派生的 RampIDs 将一组源 RampIDs 转换为另一个集合。

使用提供者服务创建 ID 映射工作流程

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“工作流程”下，选择 ID 映射。
3. 在 ID 映射工作流程页面的右上角，选择创建 ID 映射工作流程。
4. 对于步骤 1：指定 ID 映射工作流程详细信息，请执行以下操作。
 - a. 输入 ID 映射工作流程名称和可选描述。

The screenshot shows the AWS Entity Resolution console interface for creating an ID mapping workflow. The breadcrumb navigation at the top reads "AWS Entity Resolution > ID mapping workflows > Create ID mapping workflow". On the left, a vertical progress bar indicates four steps: Step 1 (Specify ID mapping workflow details), Step 2 (Specify source and target), Step 3 (Specify data output location), and Step 4 (Review and create). Step 1 is currently active. The main content area is titled "Specify ID mapping workflow details" and includes a sub-header "Provide details for your ID mapping workflow and choose an ID mapping method." Below this, there are two input fields: "Name" (ID mapping workflow name) and "Description - optional". The Name field has a placeholder "Enter name" and a character limit of "0 of 255 characters. Use alphanumeric, underscore (_), or hyphen (-) characters. Name must be unique across all ID mapping workflows in your account." The Description field has a placeholder "Enter description" and a character limit of "0 of 255 characters."

- b. 对于 ID 映射方法，请选择提供商服务。

AWS Entity Resolution 数据匹配服务 目前提供 LiveRamp 提供者服务作为 ID 映射方法。如果您订阅了 LiveRamp，则状态将显示为“已订阅”。有关如何订阅的更多信息 LiveRamp，请[参阅步骤 1：在上订阅提供商服务 AWS Data Exchange](#)。

ID mapping method [Info](#)

/LiveRamp

Currently we are only offering LiveRamp service as an ID mapping method.

Access to LiveRamp provider subscription

✔ Subscribed

ℹ To ensure a successful workflow run, your data input file format and normalization must be aligned with the provider service's guidelines. [Learn more](#) [↗](#)

ℹ Note

确保您的数据输入文件格式符合提供商服务的指南。有关输入文件格式指南 LiveRamp 的更多信息，请参阅 LiveRamp 文档网站 ADX 上的 [执行直通翻译](#)。

c. 要进行 LiveRamp 配置，请输入以下 LiveRamp 提供的值：

- 客户端 ID 管理器 ARN
- 客户机密管理器 ARN

LiveRamp configuration [Info](#)**Client ID manager ARN**

Enter the Client ID manager ARN provided by LiveRamp.

0 of 2,048 characters.

Client secret manager ARN

Enter the Client secret manager ARN provided by LiveRamp.

0 of 2,048 characters.

d. (可选) 要为资源启用标签，请选择添加新标签，然后输入密钥和值对。

e. 选择下一步。

5. 对于“步骤 2：指定源和目标”，请执行以下操作。

- 打开“高级选项”。
- 对于来源，选择 ID 命名空间。

AWS Entity Resolution > ID mapping workflows > Create ID mapping workflow

Step 1
Specify ID mapping workflow details

Step 2
Specify source and target

Step 3 - optional
Specify data output location

Step 4
Review and create

Specify source and target Info

Use a schema mapping or ID namespace to describe your input data depending on your ID mapping type.

Advanced options
Use advanced options if you are creating an ID mapping across AWS accounts and have created ID namespace resources to manage AWS account permissions.

Source Info

The source of the data in an ID mapping workflow.

Schema mapping
Use AWS Glue database, AWS Glue table, and schema mapping for ID mapping on your own AWS account.

ID namespace
Use an ID namespace to describe your source data for ID mapping across two AWS accounts.

ID namespace Info

Choose an AWS account associated with the ID namespace source. [Create ID namespace](#)

Your AWS account
 Another AWS account

Your ID namespaces

Select ID namespace ▼

- c. 对于 ID 命名空间，请确定 ID 命名空间所在的位置，然后采取建议的操作。

ID 命名空间的位置	推荐操作
你自己的 AWS 账户	<ol style="list-style-type: none"> 选择你的 AWS 账户。 从“您的 ID 命名空间”下拉列表中选择 ID 命名空间。
别人的 AWS 账户	<ol style="list-style-type: none"> 选择另一个 AWS 账户。 输入 ID 命名空间ARN。

- d. 对于目标，选择 ID 命名空间。

Target [Info](#)

Select how you want to provide the domain to which you want to translate your data using ID mapping.

Domain
Provide a specific target domain to which you want to translate the data to

ID namespace
Use an ID namespace to describe your target configuration for ID mapping across two AWS accounts.

ID namespace [Info](#)

Choose an AWS account associated with the ID namespace source. [Create ID namespace](#)

Your AWS account

Another AWS account

Your ID namespaces

- e. 要指定服务访问权限，请选择一个选项并采取建议的操作。

Service access

AWS Entity Resolution requires permissions to read your data input from AWS Glue and write to S3 on your behalf. [View policy document](#)

Choose a method to authorize AWS Entity Resolution

Create and use a new service role
Automatically create the role and add the necessary permissions policy.

Use an existing service role

Service role name

51 of 64 characters. Use alphanumeric and '+=, @-_' characters. Don't include spaces. Name must be unique across all roles in the account.

This data is encrypted with a KMS key
Specify the associated KMS key to enable AWS Entity Resolution to access each of your data inputs.

选项	推荐操作
创建并使用新的服务角色	<ul style="list-style-type: none"> • AWS Entity Resolution 数据匹配服务 使用此表所需的策略创建服务角色。 • 默认服务角色名称为 <code>entityresolution-id-mapping-workflow-<timestamp></code> 。 • 您必须拥有创建角色并附加策略的权限。 • 如果您的输入数据已加密，请选择“此数据由KMS密钥加密”选项。然后，输入 AWS KMS 用于解密您的数据输入的密钥。
使用现有服务角色	<ol style="list-style-type: none"> 1. 从下拉列表中选择现有服务角色名称。 <p>如果您有列出角色的权限，则会显示角色列表。</p> <p>如果您无权列出角色，则可以输入要使用的角色的 Amazon 资源名称 (ARN)。</p> <p>如果没有现有的服务角色，则使用现有服务角色选项不可用。</p> 2. 通过选择“在IAM外部查看”链接来查看服务角色。 <p>默认情况下，AWS Entity Resolution 数据匹配服务 不会尝试更新现有角色策略以添加必要的权限。</p>

6. 选择下一步。

7. 对于步骤 3：指定数据输出位置（可选），请执行以下操作。

a. 对于数据输出目标，请执行以下操作。

i. 选择数据输出的 Amazon S3 位置。

- ii. 对于加密，如果您选择自定义加密设置，则输入 AWS KMS 键ARN或选择“创建” AWS KMS 钥匙。
- b. 查看LiveRamp 生成的输出。
- c. 选择下一步。

Step 1 Specify ID mapping workflow details

Step 2 Specify source and target

Step 3 - optional Specify data output location

Step 4 Review and create

Specify data output location - optional Info

Choose your S3 location to write your data output.

Data output destination Info

Choose the Amazon S3 location for the data output.

Amazon S3 location

Encryption - optional Info

Your data is encrypted by default with a key that AWS owns and manages for you. To specify a different key, customize your encryption settings.

Customize encryption settings
Specify an AWS KMS key to customize your encryption settings.

LiveRamp generated output (2)
Additional information generated by LiveRamp.

Output field	Description
RAMPID	LiveRamp's universal identifier that is tied to devices in the LiveRamp Identity Graph
TRANSCODED_IDENTIFIER	LiveRamp's universal identifier that is tied to devices in the LiveRamp Identity Graph

8. 对于“步骤 4：查看并创建”，请执行以下操作。
 - a. 查看您在之前的步骤中所做的选择，并在必要时对其进行编辑。
 - b. 选择创建。

将出现一条消息，表明身份映射工作流程已创建。

创建 ID 映射工作流程后，就可以[运行身份映射工作流程](#)了。

运行 ID 映射工作流程

[为一个人创建 ID 映射工作流程之后 AWS 账户](#)或者在两者之间[创建 ID 映射工作流程 AWS 账户](#)，您可以运行 ID 映射工作流程。ID 映射工作流程输出一个 CSV 文件。

运行 ID 映射工作流程

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“工作流程”下，选择 ID 映射。
3. 选择 ID 映射工作流程。
4. 在 ID 映射工作流程详细信息页面的右上角，选择运行。
5. 在匹配的工作流程详细信息页面的指标选项卡上，在“上次作业指标”下查看以下内容：
 - Job ID
 - 工作流作业的完成时间
 - 匹配工作流作业的状态：已排队、进行中、已完成、失败
 - 已处理的记录数
 - 未处理的记录数
 - 输入记录的数量

在 Job History 下，您还可以查看之前运行的 ID 映射工作流程作业的作业指标。

6. 身份映射工作流程任务完成（状态为已完成）后，选择数据输出，然后选择您的 Amazon S3 位置以查看结果。

拿到 CSV 文件后，你可以 RAMPID 加入 TRANSCODED_ID。

使用新的输出目标运行 ID 映射工作流程

[为一个人创建 ID 映射工作流程之后 AWS 账户](#) 或者在两者之间 [创建 ID 映射工作流程 AWS 账户](#)，您可以选择不同的 S3 位置来写入数据输出。

使用新的输出目标运行 ID 映射工作流

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“工作流程”下，选择 ID 映射。
3. 选择 ID 映射工作流程。
4. 在右上角的 ID 映射工作流详细信息页面上，从“运行工作流”下拉列表中选择“使用新的输出目标运行”。

5. 对于数据输出目标，请执行以下操作。
 - a. 选择数据输出的 Amazon S3 位置。
 - b. 对于加密，如果您选择自定义加密设置，则输入 AWS KMS 键ARN或选择“创建”AWS KMS 钥匙。
6. 要指定服务访问权限，请选择一个选项并采取建议的操作。

选项	推荐操作
创建并使用新的服务角色	<ul style="list-style-type: none"> • AWS Entity Resolution 数据匹配服务 使用此表所需的策略创建服务角色。 • 默认服务角色名称为 <code>entityresolution-id-mapping-workflow- <timestamp></code> 。 • 您必须拥有创建角色并附加策略的权限。 • 如果您的输入数据已加密，请选择“此数据由KMS密钥加密”选项。然后，输入 AWS KMS 用于解密您的数据输入的密钥。
使用现有服务角色	<ol style="list-style-type: none"> 1. 从下拉列表中选择一个现有服务角色名称。 如果您有列出角色的权限，则会显示角色列表。 如果您无权列出角色，则可以输入要使用的角色的 Amazon 资源名称 (ARN)。 如果没有现有的服务角色，则使用现有服务角色选项不可用。 2. 通过选择“在IAM外部查看”链接来查看服务角色。 默认情况下，AWS Entity Resolution 数据匹配服务 不会尝试更新现有角色策略以添加必要的权限。

7. 选择运行。

8. 在匹配的工作流程详细信息页面的指标选项卡上，在“上次作业指标”下查看以下内容：

- Job ID
- workflow作业的完成时间
- 匹配workflow作业的状态：已排队、进行中、已完成、失败
- 已处理的记录数
- 未处理的记录数
- 输入记录的数量

在 Job History 下，您还可以查看之前运行的 ID 映射 workflow 作业的作业指标。

9. 身份映射 workflow 任务完成（状态为已完成）后，选择数据输出，然后选择您的 Amazon S3 位置以查看结果。

拿到 CSV 文件后，你可以 RAMPID 加入 TRANSCODED_ID。

编辑 ID 映射 workflow

要编辑 ID 映射 workflow，请执行以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“workflow”下，选择 ID 映射。
3. 选择 ID 映射 workflow。
4. 在 ID 映射 workflow 详细信息页面的右上角，选择编辑。
5. 在“指定 ID 映射 workflow 详细信息”页面上，进行任何必要的更改，然后选择“下一步”。
6. 在“指定数据输出”页面上，进行必要的更改，然后选择“下一步”。
7. 在“查看并保存”页面上，进行必要的更改，然后选择“保存”。

删除 ID 映射 workflow

要删除 ID 映射 workflow，请执行以下操作：

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。

2. 在左侧导航窗格的“工作流程”下，选择 ID 映射。
3. 选择 ID 映射工作流程。
4. 在 ID 映射工作流程详细信息页面的右上角，选择删除。
5. 确认删除，然后选择删除。

为 ID 映射工作流程添加或更新资源策略

资源策略允许 ID 映射资源的创建者访问您的 ID 映射工作流程资源。

添加或更新资源策略

1. 登录 AWS Management Console 然后打开 [AWS Entity Resolution 数据匹配服务](#) 用你的控制台 AWS 账户，如果你还没有这样做的话。
2. 在左侧导航窗格的“工作流程”下，选择 ID 映射。
3. 选择 ID 映射工作流程。
4. 在 ID 映射工作流程详细信息页面上，选择权限选项卡。
5. 在资源策略部分中，选择编辑。
6. 在JSON编辑器中添加或更新策略。
7. 选择 Save changes (保存更改)。

与集成 AWS Entity Resolution 数据匹配服务 作为提供者

AWS Entity Resolution 数据匹配服务 第三方提供商集成可帮助客户保护消费者隐私并遵守数据主权法。第三方提供商（例如 LiveRamp 和 TransUnion）将消费者标识符转换为广告IDs，例如Ramp IDs 和Fabrck IDs。这些广告标识符通常用于广告和营销工具，以防止将消费者数据导出到非广告中AWS 托管系统。本节为提供者提供了与之集成的指南 AWS Entity Resolution 数据匹配服务 将消费者标识符编码或转码IDs为广告，以便在[基于提供商服务的](#)匹配工作流程中使用。

有关当前与之集成的提供商服务的更多信息 AWS Entity Resolution 数据匹配服务，请参阅 [创建基于提供商服务的匹配工作流程](#)。

主题

- [要求](#)
- [使用 AWS Entity Resolution 数据匹配服务 开放API规范](#)
- [测试提供商集成](#)

要求

在作为提供者服务集成之前 AWS Entity Resolution 数据匹配服务，完成以下要求。

主题

- [在上列出提供商服务 AWS Data Exchange](#)
- [确定你的属性](#)
- [申请 AWS Entity Resolution 数据匹配服务 开放API规范](#)

在上列出提供商服务 AWS Data Exchange

作为第三方提供商，您必须在 [AWSData Exchange \(ADX\)](#) 产品目录中发布您的产品。在您的产品上架后 AWS Data Exchange 产品目录，订阅者可以通过公开或私人报价订阅您的产品。

要在上列出提供者服务 AWS Data Exchange

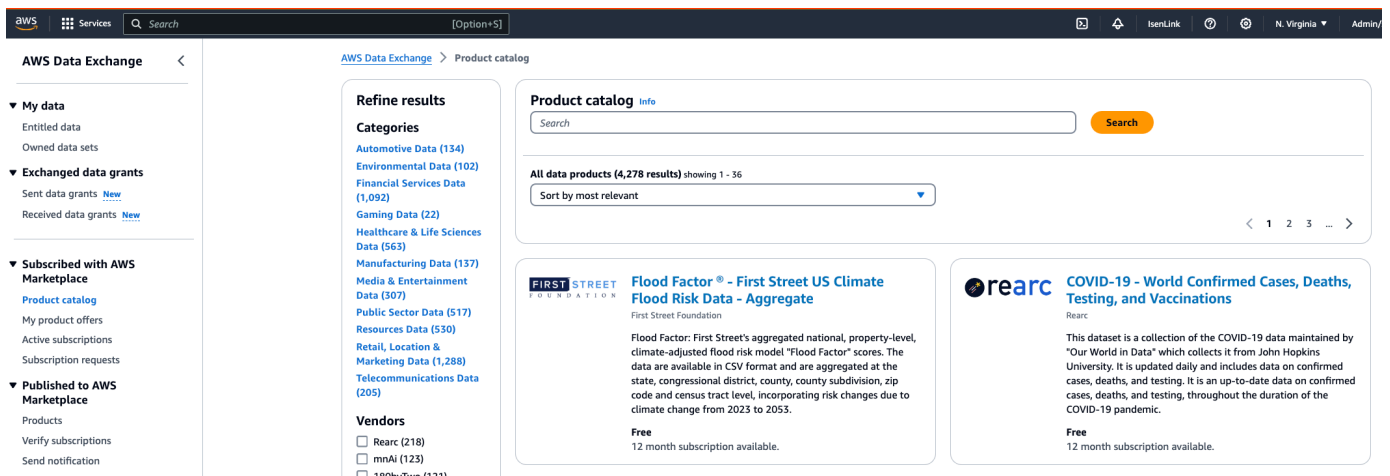
1. 如果您是新的数据产品提供商 AWS Data Exchange，完成标题为 [“作为提供者入门”](#) 一节中的步骤 AWS Data Exchange 用户指南。

2. 创建RESTAPI数据集并发布包含APIs在的新产品 AWS Data Exchange 按照标题为“[如何发布包含以下内容的产品](#)”一节APIs中的步骤进行操作 AWS Data Exchange 用户指南。您可以使用以下任一方法完成该过程 AWS Data Exchange 控制台或 AWS Command Line Interface.

如果您已将产品知名度设置为公开，则所有订阅者都可以公开报价。

如果您已将产品可见性设置为私有，请完成标题为“[创建自定义报价](#)”部分中的步骤 AWS Data Exchange 用户指南，视您的用例而定。

下图显示了中可用产品的示例 AWS Data Exchange 产品目录。



3. 产品上市后 AWS Data Exchange 产品目录，订阅者可以通过以下方式订阅产品。

- 订阅公共产品。
- 使用由[提供商服务发布的私有报价](#)（自定义报价）。
- 使用[自带订阅 \(BYOS\)](#) 优惠。

有关更多信息，请参阅[订阅和访问包含APIs的产品](#) AWS Data Exchange 用户指南。

确定你的属性

输入数据的@@ 属性是工作流程中要解析的实体的类型定义。属性的一些示例包括FirstName、LastNameEmail、或Custom String。

当你确定自己的属性时，你应该记下任何要求或指导方针。

Example 示例

以下是识别提供者属性的验证示例。

- “FirstName或LastName”属性是必需的。
- 如果该Email属性存在，则必须对其进行哈希处理。

作为提供者，您必须识别提供者服务产品中的属性，然后将这些属性传达给 AWS Entity Resolution 数据匹配服务。在继续操作之前，请 <aws-entity-resolution-bd##@amazon.com> 的业务开发团队进行进一步验证。

申请 AWS Entity Resolution 数据匹配服务 开放API规范

AWS Entity Resolution 数据匹配服务 有一个开放API规范，作为提供者，您可以将其用作包含集成 APIs 所涉及内容的握手。有关更多信息，请参阅 [使用 AWS Entity Resolution 数据匹配服务 开放API规范](#)。

要申请 Open API 定义，请联系 AWS Entity Resolution 数据匹配服务 <aws-entity-resolution-bd@amazon.com> 的业务发展团队。

使用 AWS Entity Resolution 数据匹配服务 开放API规范

开放API规范定义了与之相关的所有协议 AWS Entity Resolution 数据匹配服务。此规范是实现集成所必需的。

Open API 定义包含以下API操作：

- POST AssignIdentities
- POST CreateJob
- GET GetJob
- POST StartJob
- POST MapIdentities
- GET Schema

要申请 Open API 规范，请联系 AWS Entity Resolution 数据匹配服务 <aws-entity-resolution-bd@amazon.com> 的业务发展团队。

Open API 规范支持两种类型的集成，用于编码和转码消费者标识符、批处理和同步处理。获得 Open API 规范后，为您的用例实现处理集成类型。

主题

- [批处理集成](#)
- [同步处理集成](#)

批处理集成

批处理集成遵循异步设计模式。启动工作流程后 AWS Data Exchange，它通过提供商集成端点提交作业，然后工作流程通过定期轮询作业状态来等待任务完成。对于可能需要更长时间且提供商吞吐量较低的任务运行，则更适合使用此解决方案。提供者将以 Amazon S3 链接的形式获取数据集位置，他们可以自行处理该链接，并将结果写入预先确定的输出 S3 位置。

使用三个API定义启用批处理集成。AWS Entity Resolution 数据匹配服务 将调用提供者端点，该端点可通过以下方式获得 AWS Data Exchange 按以下顺序排列：

1. POST CreateJob：此API操作将工作信息提交给提供者进行处理。这些信息与任务类型有关；编码或转码、S3 位置、客户提供的架构以及所需的任何其他作业属性。

这将API返回 aJobId，而 Job 的状态将是以下状态之

一：PENDINGREADY、IN_PROGRESS、COMPLETE、或FAILED。

编码请求示例

```
POST /jobs
{
  "actionType": "ID_ASSIGNMENT",
  "s3SourceLocation": "string",
  "s3TargetLocation": "string",
  "jobProperties": {
    "assignmentJobProperties": {
      "fieldMappings": [
        {
          "name": "string",
          "type": "NAME"
        }
      ]
    }
  },
}
```

```
"customerSpecifiedJobProperties": {
  "property1": "string",
  "property2": "string"
},
"outputSourceConfiguration": {
  "KMSArn": "string"
}
}
```

示例响应

```
{
  "jobId": "string",
  "status": "PENDING"
}
```

2. POST StartJob : 这API让提供者知道要根据JobId提供的内容开始作业。这允许提供者执行CreateJob直到之前StartJob所需的任何验证。

这将API返回 a JobId、f Status or the Job statusMessage、和statusCode。

编码请求示例

```
POST/jobs/{jobId}
{
  "customerSpecifiedJobProperties": {
    "property1": "string",
    "property2": "string"
  }
}
```

示例响应

```
{
  "jobId": "string",
  "status": "PENDING",
  "statusMessage": "string",
  "statusCode": 200
}
```

3. GET GetJob: 这会API告知 AWS Entity Resolution 数据匹配服务 任务是否已完成或任何其他状态。

这将API返回 a JobId、f Status or the Job statusMessage、和statusCode。

编码请求示例

```
GET /jobs/{jobId}
```

示例响应

```
{
  "jobId": "string",
  "status": "PENDING",
  "statusMessage": "string",
  "statusCode": 200
}
```

APIs这些内容的完整定义载于 [AWS Entity Resolution 数据匹配服务 开放API规范](#)。

同步处理集成

对于具有近乎实时的响应时间、实时响应时间、更高吞吐量和更高吞吐量的提供商来说，同步处理解决方案更受青睐TPS。该 AWS Entity Resolution 数据匹配服务 workflow 对数据集进行分区，并行发出多个API请求。这些区域有：AWS Entity Resolution 数据匹配服务 然后，工作流程负责将结果写入所需的输出位置。

此过程是使用其中一个API定义启用的。AWS Entity Resolution 数据匹配服务 调用提供者端点，该端点可通过以下方式获得 AWS Data Exchange:

POST AssignIdentities : 这将使用source_id标识符向提供者API发送数据，并recordFields与该记录相关联。

这将API返回assignedRecords.

编码请求示例

```
POST /assignment
{
  "sourceRecords": [
    {
      "sourceId": "string",
```

```
    "recordFields": [
      {
        "name": "string",
        "type": "NAME",
        "value": "string"
      }
    ]
  }
]
```

示例响应

```
{
  "assignedRecords": [
    {
      "sourceRecord": {
        "sourceId": "string",
        "recordFields": [
          {
            "name": "string",
            "type": "NAME",
            "value": "string"
          }
        ]
      },
      "identity": any
    }
  ]
}
```

APIs这些内容的完整定义载于 [AWS Entity Resolution 数据匹配服务 开放API规范](#)。

视提供者选择哪种方法而定，AWS Entity Resolution 数据匹配服务 将为该提供者创建用于启动编码或转码的配置。此外，客户还可以使用APIs提供的配置来使用这些配置 AWS Entity Resolution 数据匹配服务。

此配置可使用 Amazon 资源名称 (ARN) 进行访问，该名称源自提供商服务所在的位置 AWS Data Exchange 是托管的，以及提供商服务的类型。AWS Entity Resolution 数据匹配服务 将其ARN称为providerServiceARN。

测试提供商集成

虽然 AWS Entity Resolution 数据匹配服务 主机数据匹配服务，提供商集成是 end-to-end 匹配工作流程的重要第三方组件。有几项测试可以 AWS Entity Resolution 数据匹配服务 已为提供商定义了集成失败时增加保护措施。这种方法为提供商提供了根据这些 end-to-end 测试案例监控其服务运行状况的机会。

提供者可以使用自己的测试账号和自己的数据来运行这些 end-to-end 测试用例 AWS Entity Resolution 数据匹配服务 软件开发套件 (SDK)。如果供应商有任何问题，AWS Entity Resolution 数据匹配服务 使用首选的上报路径来升级问题。此外，提供者需要对测试结果实施自己的监控。提供者需要分享他们的 AWS 账户 IDs 用来运行这些测试的 AWS Entity Resolution 数据匹配服务。

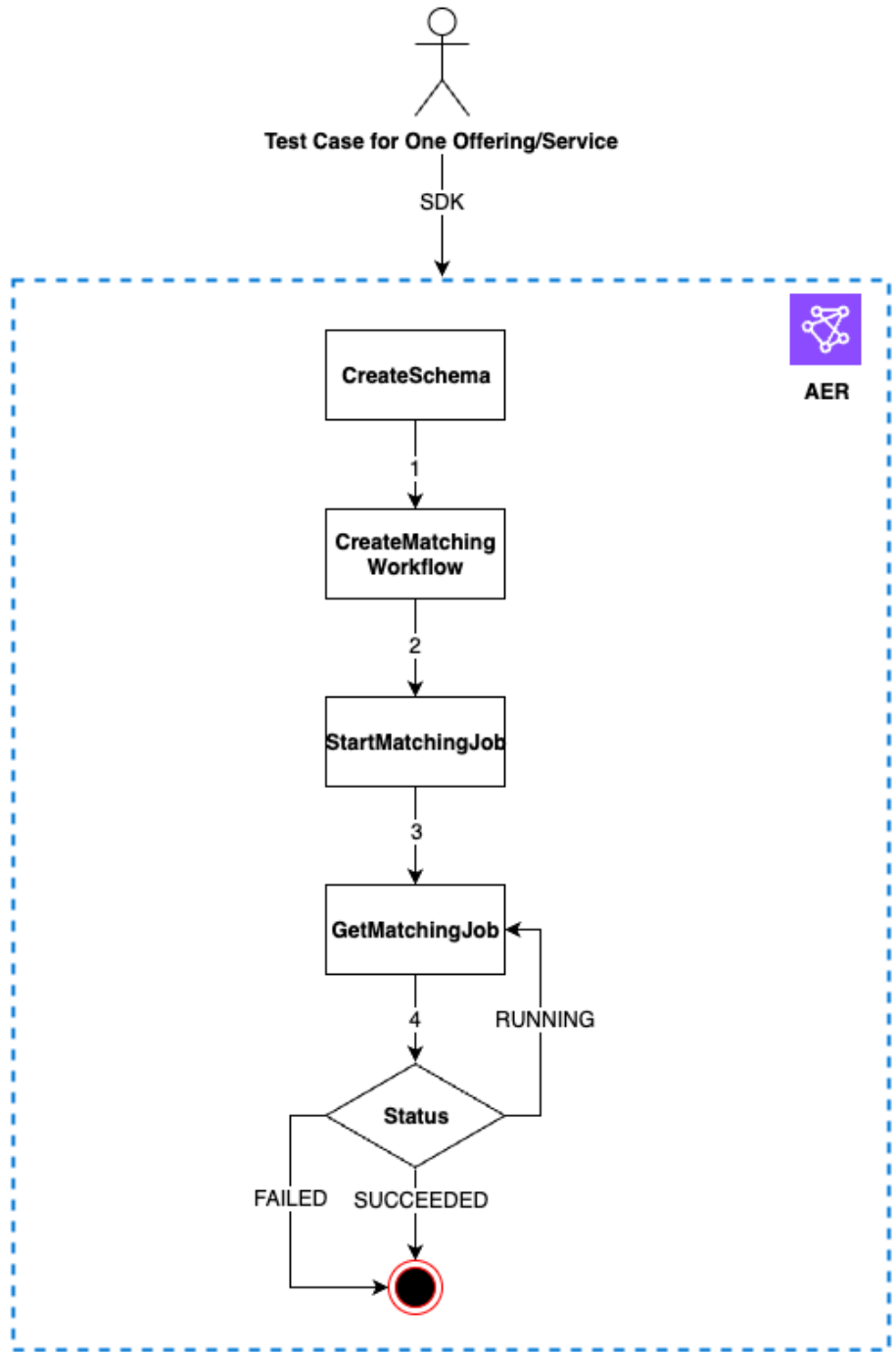
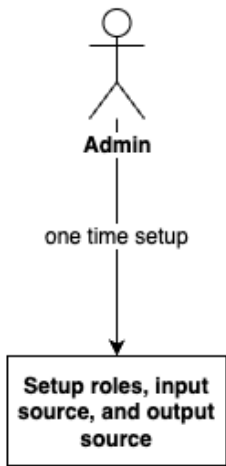
成功运行意味着提供商可以设置自己的数据，通过以下方式使用自己的服务 AWS Entity Resolution 数据匹配服务，任务状态将返回“已完成”，没有错误。这可以通过使用 APIs 提供的，以编程方式完成 AWS Entity Resolution 数据匹配服务。

例如，提供商可以根据其服务设置其 S3 存储桶、输入源、角色、架构和工作流程。这些设置完成后，提供商可以每天运行一次包含 200 条记录的工作流程来测试其服务。在这种方法中，提供商使用自己的选择 SDK 并对通过以下方式提供的服务进行 end-to-end 测试 AWS Data Exchange 使用他们的测试账户。供应商应针对其每项产品或服务进行这些测试。

Note

提供者需要提供 AWS Entity Resolution 数据匹配服务 这 AWS 账户 ID (他们 `accountId`) 用来运行这些工作流程进行测试。此外，提供商需要监控这些测试并确保它们通过，这意味着提供商需要在失败时启用通知并相应地解决问题。

下图显示了一个典型 end-to-end 的工作流程测试用例。



测试提供商集成

1. (一次性设置) 设置资源 AWS Entity Resolution 数据匹配服务 按照中的步骤操作 [设置 AWS Entity Resolution 数据匹配服务](#)。

完成一次性设置过程后，应准备好角色、数据和数据源。现在，您已准备好使用以下任一方法来测试提供商集成 AWS Entity Resolution 数据匹配服务 控制台或APIs。

2. 使用以下任一方法测试提供商集成 AWS Entity Resolution 数据匹配服务 APIs或控制台。

API

要测试提供商集成，请使用 AWS Entity Resolution 数据匹配服务 APIs

1. 使用创建架构映射[CreateSchemaMapping API](#)。有关支持的编程语言的完整列表，请参阅的“[另请参阅](#)”部分[CreateSchemaMapping API](#)。

架构映射是您告知的过程 AWS Entity Resolution 数据匹配服务 如何解释您的数据以进行匹配。您可以定义输入数据表的架构，希望AWS实体解析将其读入匹配的工作流程。

创建架构映射时，必须为AWS实体解析读取的每行输入数据指定并分配一个[唯一标识符](#)。例如：Primary_key、Row_ID、Record_ID。

Example 为包含id和的数据源创建架构映射 **email**

以下是包含id和的数据源的架构映射示例email：

```
[
  {
    "fieldName": "id",
    "type": "UNIQUE_ID"
  },
  {
    "fieldName": "email",
    "type": "EMAIL_ADDRESS"
  }
]
```

Example 为包含id和email使用 Java 的数据源创建架构映射 SDK

以下是包含id并email使用 Java 的数据源的架构映射示例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. 使用创建匹配的工作流程 [CreateMatchingWorkflow API](#)。有关支持的编程语言的完整列表，请参阅的“[另请参阅](#)”部分 [CreateMatchingWorkflow API](#)。

Example 使用 Java 创建匹配的工作流程 SDK

以下是使用 Java 的匹配工作流程的示例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()
    )

```

设置匹配的工作流程后，您可以运行工作流程。

3. 使用运行匹配的工作流程 [StartMatchingJob API](#)。要运行匹配的工作流程，您必须使用 `CreateMatchingWorkflow` 端点创建了匹配的工作流程。

有关支持的编程语言的完整列表，请参阅的“[另请参阅](#)”部分 [StartMatchingJob API](#)。

Example 使用 Java 运行匹配的工作流程 SDK

以下是使用 Java 运行匹配工作流程的示例SDK：

```

EntityResolutionClient.startMatchingJob(StartMatchingJobRequest.builder()
    .workflowName(<name-of-workflow-from-step3>)
    .build()
)

```

4. 使用监控工作流程的状态 [GetMatchingJob API](#)。

这将API返回与任务关联的状态、指标和错误（如果有）。

Example 使用 Java 监控匹配的工作流程 SDK

以下是使用 Java 监控匹配工作流作业的示例SDK：

```

EntityResolutionClient.getMatchingJob(GetMatchingJobRequest.builder()
    .workflowName(<name-of-workflow-from-step3>)
    .jobId(jobId-from-startMatchingJob)
    .build()
)

```

如果工作流程成功完成，则 end-to-end 测试已完成。

Console

要测试提供商集成，请使用 AWS Entity Resolution 数据匹配服务 Console

1. 按照中的步骤创建架构映射[创建架构映射](#)。

架构映射是您告知的过程 AWS Entity Resolution 数据匹配服务 如何解释您的数据以进行匹配。您可以根据需要定义输入数据表的架构 AWS Entity Resolution 数据匹配服务 以了解匹配的工作流程。

创建架构映射时，必须为每行输入数据指定并分配一个[唯一标识符](#) AWS Entity Resolution 数据匹配服务 读。例如：Primary_key、Row_ID、Record_ID。

Example 包含id和的数据源的架构映射 **email**

以下是包含id和的数据源的架构映射示例email：

```
[
  {
    "fieldName": "id",
    "type": "UNIQUE_ID"
  },
  {
    "fieldName": "email",
    "type": "EMAIL_ADDRESS"
  }
]
```

2. 按照中的步骤创建并运行匹配的工作流程[创建基于提供商服务的匹配工作流程](#)。

创建匹配工作流是您为指定要匹配的输入数据以及如何执行匹配而设置的过程。在基于提供商的工作流程中，如果账户通过订阅了提供商服务 AWS Data Exchange，您可以将您的已知标识符与您的首选提供商进行匹配。根据您用于执行端到端测试的提供商和服务，您可以相应地配置匹配的工作流程。

这些区域有：AWS Entity Resolution 数据匹配服务 控制台将创建和运行的操作组合在一个按钮中。选择“创建并运行”后，将显示一条消息，表明匹配的工作流程已创建且作业已启动。

3. 在“匹配工作流程”页面上监控工作流程的状态。

如果工作流成功完成（Job 状态为“已完成”），则 end-to-end 测试已完成。

在匹配工作流程详细信息页面的指标选项卡上，您可以在“上次任务指标”下查看以下内容：

- 作业 ID。
- 匹配 workflow 作业的状态：已排队、进行中、已完成、失败
- workflow 作业的完成时间。
- 已处理的记录数。
- 未处理的记录数。
- IDs 生成的唯一匹配项。
- 输入记录的数量。

您还可以查看任务历史记录下先前运行过的匹配 workflow 作业的作业指标。

AWS Entity Resolution 数据匹配服务 中的安全性

AWS 十分重视云安全性。为了满足对安全性最敏感的组织的需求，我们打造了具有超高安全性的数据中心和网络架构。作为 AWS 客户，您也将从这些数据中心和网络架构受益。

安全性是 AWS 和您的共同责任。[shared responsibility model](#) (责任共担模式) (责任共担模式) (责任共担模式) 将其描述为云的安全性和云中的安全性：

- 云的安全性 – AWS 负责保护在 AWS Cloud 中运行 AWS 服务的基础设施。AWS 还向您提供可安全使用的服务。第三方审核员定期测试和验证我们的安全性的有效性，作为 [AWS 合规性计划](#) 的一部分。要了解适用于 AWS Entity Resolution 数据匹配服务的合规性计划，请参阅[合规性计划范围内的 AWS 服务](#)。
- 云中的安全性 – 您的责任由您使用的 AWS 服务 决定。您还需要对其它因素负责，包括您的数据的敏感性、您公司的要求以及适用的法律法规。

此文档将帮助您了解如何在使用 AWS Entity Resolution 数据匹配服务 时应用责任共担模型 以下主题说明如何配置 AWS Entity Resolution 数据匹配服务 以实现您的安全性和合规性目标。您还将了解如何使用其他 AWS 服务以帮助您监控和保护 AWS Entity Resolution 数据匹配服务 资源。

主题

- [中的数据保护 AWS Entity Resolution 数据匹配服务](#)
- [的身份和访问管理 AWS Entity Resolution 数据匹配服务](#)
- [合规性验证 AWS Entity Resolution 数据匹配服务](#)
- [韧性在 AWS Entity Resolution 数据匹配服务](#)

中的数据保护 AWS Entity Resolution 数据匹配服务

这些区域有：AWS [分担责任模型](#)适用于以下领域的的数据保护 AWS Entity Resolution 数据匹配服务。如本模型所述，AWS 负责保护运行所有内容的全球基础设施 AWS Cloud。您有责任保持对托管在此基础架构上的内容的控制。您还负责以下各项的安全配置和管理任务 AWS 服务 你用的。有关数据隐私的更多信息，请参阅[数据隐私FAQ](#)。有关欧洲数据保护的信息，请参阅 [AWS 责任共担模型和 GDPR](#)博客文章 AWS 安全博客。

出于数据保护的目，我们建议您进行保护 AWS 账户 凭据并使用设置个人用户 AWS IAM Identity Center 或者 AWS Identity and Access Management (IAM)。这样，每个用户只获得履行其工作职责所需的权限。我们还建议您通过以下方式保护数据：

- 对每个账户使用多重身份验证 (MFA)。
- 使用SSL/TLS与之通信 AWS 资源的费用。我们需要 TLS 1.2，建议使用 TLS 1.3。
- 使用API进行设置和用户活动记录 AWS CloudTrail。有关使用 CloudTrail 轨迹捕获的信息 AWS 活动，请参阅[使用中的 CloudTrail 轨迹](#) AWS CloudTrail 用户指南。
- 使用 AWS 加密解决方案，以及其中的所有默认安全控件 AWS 服务。
- 使用高级托管安全服务（例如 Amazon Macie），它有助于发现和保护存储在 Amazon S3 中的敏感数据。
- 如果您在访问时需要 FIPS 140-3 经过验证的加密模块 AWS 通过命令行界面或API，使用FIPS端点。有关可用FIPS端点的更多信息，请参阅[联邦信息处理标准 \(FIPS\) 140-3](#)。

我们强烈建议您切勿将机密信息或敏感信息（如您客户的电子邮件地址）放入标签或自由格式文本字段（如名称字段）。这包括你何时使用 AWS Entity Resolution 数据匹配服务 或其他 AWS 服务 使用控制台，API，AWS CLI，或 AWS SDKs。在用于名称的标签或自由格式文本字段中输入的任何数据都可能会用于计费或诊断日志。如果您URL向外部服务器提供，我们强烈建议您不要在中包含凭据信息，URL以验证您对该服务器的请求。

静态数据加密 AWS Entity Resolution 数据匹配服务

AWS Entity Resolution 数据匹配服务 默认提供加密以保护敏感的静态客户数据 AWS 拥有的加密密钥。

AWS拥有的密钥 — AWS Entity Resolution 数据匹配服务 默认使用这些密钥来自动加密个人身份数据。您无法查看、管理或使用 AWS 拥有的密钥，或者审核其使用情况。但是，您无需采取任何措施来保护加密数据的密钥。有关更多信息，请参阅中的[AWS自有密钥](#) AWS Key Management Service 开发者指南。

默认情况下，静态数据加密有助于降低保护敏感数据的操作开销和复杂性。同时，您可以使用它来构建符合严格加密合规性和监管要求的安全应用程序。

或者，您也可以在创建匹配的工作流程资源时提供用于加密的客户托管KMS密钥。

客户管理的密钥 — AWS Entity Resolution 数据匹配服务 支持使用由您创建、拥有和管理的对称客户托管KMS密钥来加密您的敏感数据。由于您可以完全控制这层加密，因此可以执行以下任务：

- 制定和维护关键策略
- 制定和维护IAM政策及补助金

- 启用和禁用密钥策略
- 轮换加密材料
- 添加标签
- 创建密钥别名
- 安排密钥删除

有关更多信息，请参阅《[客户托管密钥](#)》AWS Key Management Service 开发者指南。

有关 AWS KMS，请参阅[什么是AWS密钥管理服务？](#)

密钥管理

操作方法 AWS Entity Resolution 数据匹配服务 将补助金用于 AWS KMS

AWS Entity Resolution 数据匹配服务 需要获得[授权](#)才能使用您的客户托管密钥。当您创建使用客户托管密钥加密的匹配工作流程时，AWS Entity Resolution 数据匹配服务 通过向发送[CreateGrant](#)请求来代表您创建补助金 AWS KMS。补助金 AWS KMS 是用来给 AWS Entity Resolution 数据匹配服务 访问客户账户中的KMS密钥。AWS Entity Resolution 数据匹配服务 需要获得授权才能使用您的客户托管密钥进行以下内部操作：

- 将[GenerateDataKey](#)请求发送到 AWS KMS 生成由您的客户托管密钥加密的数据密钥。
- 将[解密请求](#)发送到 AWS KMS 解密加密的数据密钥，以便它们可用于加密您的数据。

您可以随时撤销授予访问权限，或删除服务对客户托管密钥的访问权限。如果你这样做，AWS Entity Resolution 数据匹配服务 将无法访问由客户托管密钥加密的任何数据，这会影响依赖该数据的操作。例如，如果您通过授权删除了对密钥的服务访问权限，并尝试为使用客户密钥加密的匹配工作流程启动任务，则该操作将返回AccessDeniedException错误。

创建客户托管的密钥

您可以使用以下方法创建对称的客户托管密钥 AWS Management Console，或者 AWS KMS APIs.

创建对称的客户托管密钥

AWS Entity Resolution 数据匹配服务 支持使用[对称加密KMS密钥进行加密](#)。按照中[创建对称客户托管密钥](#)的步骤进行操作 AWS Key Management Service 开发者指南。

关键政策声明

密钥策略控制对客户托管密钥的访问。每个客户托管式密钥必须只有一个密钥策略，其中包含确定谁可以使用密钥以及如何使用密钥的声明。创建客户托管式密钥时，可以指定密钥策略。有关更多信息，请参阅中的[管理客户托管密钥的访问权限](#) AWS Key Management Service 开发者指南。

将您的客户托管密钥与您的 AWS Entity Resolution 数据匹配服务 资源，密钥策略中必须允许以下API操作：

- [kms:DescribeKey](#)— 提供诸如密钥ARN、创建日期（以及删除日期，如果适用）、密钥状态以及密钥材料的来源和到期日期（如果有）等信息。它包括可帮助您区分不同类型的KMS密钥的字段（例如KeySpec）。它还显示密钥的使用情况（加密、签名或生成和验证MACs）以及KMS密钥支持的算法。AWS Entity Resolution 数据匹配服务 验证是SYMMETRIC_DEFAULT和KeySpecKeyUsage是ENCRYPT_DECRYPT。
- [kms:CreateGrant](#) – 添加客户托管密钥授权。授予对指定KMS密钥的控制访问权限，从而允许访问[授予操作](#) AWS Entity Resolution 数据匹配服务 要求。有关[使用拨款](#)的更多信息，请参阅 AWS Key Management Service 开发者指南。

这允许 AWS Entity Resolution 数据匹配服务 执行以下操作：

- 调用 GenerateDataKey 生成加密的数据密钥并将其存储，因为数据密钥不会立即用于加密。
- 调用 Decrypt 使用存储的加密数据密钥访问加密数据。
- 设置停用主体，以允许服务 RetireGrant。

以下是您可以为其添加的政策声明示例 AWS Entity Resolution 数据匹配服务:

```
{
  "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"
    }
  }
}
```

```
}
}
```

用户的权限

当您将KMS密钥配置为默认加密密钥时，默认KMS密钥策略允许任何有权访问所需KMS操作的用户使用此密KMS钥来加密或解密资源。要使用客户托管KMS密钥加密，您必须授予用户调用以下操作的权限：

- kms:CreateGrant
- kms:Decrypt
- kms:DescribeKey
- kms:GenerateDataKey

在[CreateMatchingWorkflow](#)请求期间，AWS Entity Resolution 数据匹配服务 将向发送 a [DescribeKey](#)和 a [CreateGrant](#)请求 AWS KMS 代表你。这将要求使用客户托管KMS密钥CreateMatchingWorkflow提出请求的IAM实体拥有密KMS钥策略的kms:DescribeKey权限。

在 and [CreateIdMappingWorkflowStartIdMappingJob](#)请求期间，AWS Entity Resolution 数据匹配服务 将向发送 a [DescribeKey](#)和 a [CreateGrant](#)请求 AWS KMS 代表你。这将要求使用客户托管KMS密钥进行CreateIdMappingWorkflow和StartIdMappingJob请求的IAM实体拥有密KMS钥策略的kms:DescribeKey权限。提供商将能够访问客户管理的密钥来解密其中的数据 AWS Entity Resolution 数据匹配服务 Amazon S3 桶。

以下是您可以添加的策略声明示例，供提供商解密中的数据 AWS Entity Resolution 数据匹配服务 Amazon S3 存储桶：

```
{
  "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"
    }
}
}]
}

```

替换每个 `<user input placeholder>` 用你自己的信息。

`<KMSKeyARN>` AWS KMS Amazon 资源名称。

同样，调用匹配工作流程中提供的客户托管KMS密钥的[StartMatchingJobAPI](#)必须具备 `kms:Decrypt` 和 `kms:GenerateDataKey` 权限的IAM实体。

有关在[策略中指定权限](#)的更多信息，请参阅 AWS Key Management Service 开发者指南。

有关对[密钥访问进行故障排除](#)的更多信息，请参阅 AWS Key Management Service 开发者指南。

为指定客户管理的密钥 AWS Entity Resolution 数据匹配服务

您可以指定客户托管密钥作为以下资源的第二层加密：

[匹配工作流程](#)-创建匹配的工作流资源时，可以通过输入 `a` 来指定数据密钥 `KMSArn`，其中 AWS Entity Resolution 数据匹配服务 用于加密资源存储的可识别个人数据。

`KMSArn`— 输入密钥ARN，这是[密钥标识符](#) AWS KMS 客户管理的密钥。

如果您要在两个资源之间创建或运行 ID 映射工作流程，则可以将客户托管密钥指定为以下资源的第二层加密 AWS 账户：

[ID 映射 workflow](#) 或 [启动 ID 映射 workflow](#) — 创建 ID 映射 workflow 资源或启动 ID 映射 workflow 作业时，您可以通过输入 `a` 来指定数据密钥 `KMSArn`，其中 AWS Entity Resolution 数据匹配服务 用于加密资源存储的可识别个人数据。

`KMSArn`— 输入密钥ARN，这是[密钥标识符](#) AWS KMS 客户管理的密钥。

监控您的加密密钥 AWS Entity Resolution 数据匹配服务 服务

当你使用 AWS KMS 客户管理的密钥与你的 AWS Entity Resolution 数据匹配服务 服务资源，您可以使用[AWS CloudTrail](#)或 [Amazon CloudWatch Logs](#) 来跟踪以下请求 AWS Entity Resolution 数据匹配服务 发送到 AWS KMS。

以下示例是 AWS CloudTrail 用于 CreateGrant、GenerateDataKeyDecrypt、和监视 DescribeKey 的事件 AWS KMS 由调用的操作 AWS Entity Resolution 数据匹配服务 要访问由您的客户托管密钥加密的数据，请执行以下操作：

主题

- [CreateGrant](#)
- [DescribeKey](#)
- [GenerateDataKey](#)
- [Decrypt](#)

CreateGrant

当你使用 AWS KMS 客户管理的密钥用于加密您的匹配工作流程资源，AWS Entity Resolution 数据匹配服务 代表您发送访问您的KMS密钥的CreateGrant请求 AWS 账户。 那笔补助金 AWS Entity Resolution 数据匹配服务 创建特定于与之关联的资源 AWS KMS 客户管理的密钥。此外，AWS Entity Resolution 数据匹配服务 当您删除资源时，使用RetireGrant操作来移除授权。

以下示例事件记录了 CreateGrant 操作：

```
{
  "eventVersion": "1.08",
  "userIdentity": {
    "type": "AssumedRole",
    "principalId": "AROAIQDTESTANDEXAMPLE:Sampleuser01",
    "arn": "arn:aws:sts::111122223333:assumed-role/Admin/Sampleuser01",
    "accountId": "111122223333",
    "accessKeyId": "AKIAIOSFODNN7EXAMPLE3",
    "sessionContext": {
      "sessionIssuer": {
        "type": "Role",
        "principalId": "AROAIQDTESTANDEXAMPLE: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 数据匹配服务 使用该DescribeKey操作来验证是否 AWS KMS 与您的匹配资源关联的客户托管密钥存在于账户和区域中。

以下示例事件记录了该DescribeKey操作。

```
{
  "eventVersion": "1.08",
  "userIdentity": {
    "type": "AssumedRole",
    "principalId": "AROAIQDTESTANDEXAMPLE:Sampleuser01",
    "arn": "arn:aws:sts::111122223333:assumed-role/Admin/Sampleuser01",
    "accountId": "111122223333",
    "accessKeyId": "AKIAIOSFODNN7EXAMPLE3",
    "sessionContext": {
      "sessionIssuer": {
        "type": "Role",
        "principalId": "AROAIQDTESTANDEXAMPLE: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

当您启用 AWS KMS 您的匹配工作流程资源的客户管理密钥，AWS Entity Resolution 数据匹配服务通过亚马逊简单存储服务 (Amazon S3) Simple Service 向 GenerateDataKey AWS KMS 这指定了 AWS KMS 客户管理的资源密钥。

以下示例事件记录了该GenerateDataKey操作。

```

{
  "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

当你启用 AWS KMS 您的匹配工作流程资源的客户管理密钥，AWS Entity Resolution 数据匹配服务通过亚马逊简单存储服务 (Amazon S3) Simple Service 向 Decrypt AWS KMS 这指定了 AWS KMS 客户管理的资源密钥。

以下示例事件记录了该Decrypt操作。

```

{
  "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"
}
```

注意事项

AWS Entity Resolution 数据匹配服务 不支持使用新的客户托管KMS密钥更新匹配的工作流程。在这种情况下，您可以使用客户托管KMS密钥创建新的工作流程。

了解更多

以下资源提供有关静态数据加密的更多信息。

有关[AWS密钥管理服务基本概念](#)的更多信息，请参阅 AWS Key Management Service 开发者指南。

有关[AWS密钥管理服务安全最佳实践的更多信息](#)，请参阅 AWS Key Management Service 开发者指南。

访问 AWS Entity Resolution 数据匹配服务 使用接口端点 (AWS PrivateLink)

您可以使用 ... AWS PrivateLink 在您的VPC和之间创建私人连接 AWS Entity Resolution 数据匹配服务。你可以访问 AWS Entity Resolution 数据匹配服务 就好像在你的里面一样VPC，无需使用互联网网关、NAT设备、VPN连接或 AWS Direct Connect 连接。你中的实例VPC不需要公有 IP 地址即可访问 AWS Entity Resolution 数据匹配服务。

您可以通过创建接口端点来建立此私有连接，该端点由 AWS PrivateLink。我们在您为接口终端节点启用的每个子网中创建一个终端节点网络接口。这些是请求者管理的网络接口，可用作发往的流量的入口点 AWS Entity Resolution 数据匹配服务。

有关更多信息，请参阅[访问权限 AWS 服务 通过 AWS PrivateLink](#)中的AWS PrivateLink 指南。

的注意事项 AWS Entity Resolution 数据匹配服务

在为设置接口终端节点之前 AWS Entity Resolution 数据匹配服务，请查看 [《注意事项》](#) 中的 AWS PrivateLink 指南。

AWS Entity Resolution 数据匹配服务 支持通过接口端点调用其所有API操作。

VPC支持终端节点策略 AWS Entity Resolution 数据匹配服务。默认情况下，完全访问权限 AWS Entity Resolution 数据匹配服务 允许通过接口端点。或者，您可以将安全组与终端节点网络接口关联以控制流量 AWS Entity Resolution 数据匹配服务 通过接口端点。

为创建接口终端节点 AWS Entity Resolution 数据匹配服务

您可以为创建接口终端节点 AWS Entity Resolution 数据匹配服务 使用 Amazon VPC 控制台或 AWS Command Line Interface (AWS CLI)。有关更多信息，请参阅中的[创建接口终端节点 AWS PrivateLink 指南](#)。

为创建接口终端节点 AWS Entity Resolution 数据匹配服务 使用以下服务名称：

```
com.amazonaws.region.entityresolution
```

如果您DNS为接口终端节点启用私有功能，则可以向发出API请求 AWS Entity Resolution 数据匹配服务 使用其默认的区域DNS名称。例如，entityresolution.us-east-1.amazonaws.com。

为 VPC 端点创建端点策略

终端节点策略是您可以附加到接口终端节点的IAM资源。默认终端节点策略允许完全访问 AWS Entity Resolution 数据匹配服务 通过接口端点。控制允许的访问权限 AWS Entity Resolution 数据匹配服务 从您的VPC，将自定义终端节点策略附加到接口终端节点。

端点策略指定以下信息：

- 可以执行操作的委托人 (AWS 账户、IAM用户和IAM角色)。
- 可执行的操作。
- 可对其执行操作的资源。

有关更多信息，请参阅中的[使用终端节点策略控制对服务的访问权限](#) AWS PrivateLink 指南。

示例：的VPC终端节点策略 AWS Entity Resolution 数据匹配服务 actions

以下是自定义端点策略的示例。当您将此策略附加到您的接口终端节点时，它会授予对所列内容的访问权限 AWS Entity Resolution 数据匹配服务 所有委托人对所有资源执行操作。

```
{
  "Statement": [
    {
      "Principal": "*",
      "Effect": "Allow",
      "Action": [
        "entityresolution:CreateMatchingWorkflow",
        "entityresolution:StartMatchingJob",
        "entityresolution:GetMatchingJob"
      ],
      "Resource": "*"
    }
  ]
}
```

的身份和访问管理 AWS Entity Resolution 数据匹配服务

AWS Identity and Access Management (IAM) 是一个 AWS 服务 可帮助管理员安全地控制对的访问权限 AWS 资源的费用。IAM管理员控制谁可以进行身份验证（登录）和授权（拥有权限）才能使用 AWS Entity Resolution 数据匹配服务 资源的费用。IAM是一个 AWS 服务 无需额外付费即可使用。

Note

AWS Entity Resolution 数据匹配服务 支持跨账户政策。有关更多信息，请参阅《IAM用户指南》IAM中的[跨账户资源访问权限](#)。

主题

- [受众](#)
- [使用身份进行身份验证](#)
- [使用策略管理访问](#)
- [操作方法 AWS Entity Resolution 数据匹配服务 配合使用 IAM](#)
- [基于身份的策略示例 AWS Entity Resolution 数据匹配服务](#)
- [AWS 的托管策略 AWS Entity Resolution 数据匹配服务](#)

- [故障排除 AWS Entity Resolution 数据匹配服务 身份和访问权限](#)

受众

你怎么用 AWS Identity and Access Management (IAM) 会有所不同，具体取决于你所做的工作 AWS Entity Resolution 数据匹配服务。

服务用户-如果您使用 AWS Entity Resolution 数据匹配服务 服务来完成您的工作，然后您的管理员会为您提供所需的凭证和权限。当你用得越多 AWS Entity Resolution 数据匹配服务 功能才能完成工作，您可能需要其他权限。了解如何管理访问权限有助于您向管理员请求适合的权限。如果您无法访问中的功能 AWS Entity Resolution 数据匹配服务，请参阅 [故障排除 AWS Entity Resolution 数据匹配服务 身份和访问权限](#)。

服务管理员-如果你负责 AWS Entity Resolution 数据匹配服务 您公司的资源，您可能拥有完全访问权限 AWS Entity Resolution 数据匹配服务。你的工作是确定哪个 AWS Entity Resolution 数据匹配服务 您的服务用户应访问的功能和资源。然后，您必须向IAM管理员提交更改服务用户权限的请求。查看此页面上的信息以了解的基本概念IAM。要详细了解贵公司如何IAM使用 AWS Entity Resolution 数据匹配服务，请参阅 [操作方法 AWS Entity Resolution 数据匹配服务 配合使用 IAM](#)。

IAM管理员-如果您是IAM管理员，则可能需要详细了解如何编写策略来管理访问权限 AWS Entity Resolution 数据匹配服务。查看示例 AWS Entity Resolution 数据匹配服务 您可以在中使用的基于身份的策略IAM，请参阅 [基于身份的策略示例 AWS Entity Resolution 数据匹配服务](#)

使用身份进行身份验证

身份验证是您登录的方式 AWS 使用您的身份凭证。您必须经过身份验证 (登录到 AWS) 作为 AWS 账户根用户、以IAM用户身份或担任IAM角色。

您可以登录 AWS 使用通过身份源提供的凭证作为联合身份。AWS IAM Identity Center (IAM身份中心) 用户、贵公司的单点登录身份验证以及您的 Google 或 Facebook 凭据就是联合身份的示例。当您以联合身份登录时，您的管理员之前使用IAM角色设置了联合身份。当你访问时 AWS 通过使用联合，您可以间接担任角色。

根据您的用户类型，您可以登录 AWS Management Console 或者 AWS 访问门户。有关登录的更多信息 AWS，请参阅[如何登录您的 AWS 账户](#)中的AWS 登录 用户指南。

如果你访问 AWS 以编程方式，AWS 提供了一个软件开发套件 (SDK) 和一个命令行界面 (CLI)，用于使用您的凭证对您的请求进行加密签名。如果你不使用 AWS 工具，你必须自己签署请求。有关使用推荐的方法自行签署请求的更多信息，请参阅[签名](#)AWS APIIAM用户指南中的请求。

无论使用何种身份验证方法，您可能需要提供其他安全信息。例如，AWS 建议您使用多重身份验证 (MFA) 来提高帐户的安全性。要了解更多信息，请参阅中的[多重身份验证](#) AWS IAM Identity Center 用户指南和[使用多因素身份验证 \(MFA\) 在 AWS](#) (在 IAM 用户指南中)。

AWS 账户 根用户

当你创建 AWS 账户，您从一个登录身份开始，该身份可以完全访问所有人 AWS 服务 以及账户中的资源。这个身份叫做 AWS 账户 root 用户，使用您创建账户时使用的电子邮件地址和密码登录即可访问。强烈建议您不要使用根用户执行日常任务。保护好根用户凭证，并使用这些凭证来执行仅根用户可以执行的任务。有关需要您以 root 用户身份登录的任务的完整列表，请参阅《用户指南》中的“[需要根用户凭据的IAM任务](#)”。

联合身份

作为最佳实践，要求人类用户 (包括需要管理员访问权限的用户) 使用与身份提供商的联合身份进行访问 AWS 服务 通过使用临时证书。

联合身份是企业用户目录中的用户、Web 身份提供商、AWS Directory Service、身份中心目录或任何访问的用户 AWS 服务 通过使用通过身份源提供的凭证。当联合身份访问时 AWS 账户，他们扮演角色，角色提供临时证书。

对于集中访问管理，我们建议您使用 AWS IAM Identity Center。您可以在 Ident IAM ity Center 中创建用户和群组，也可以连接并同步到您自己的身份源中的一组用户和群组，以便在您的所有用户和群组中使用 AWS 账户 和应用程序。有关IAM身份中心的信息，请参阅[什么是IAM身份中心？](#) 在 AWS IAM Identity Center 用户指南。

IAM 用户和组

[IAM用户](#)是你内部的身份 AWS 账户 对个人或应用程序具有特定权限。在可能的情况下，我们建议使用临时证书，而不是创建拥有密码和访问密钥等长期凭证的IAM用户。但是，如果您有需要IAM用户长期凭证的特定用例，我们建议您轮换访问密钥。有关更多信息，请参阅《IAM用户指南》中的[定期轮换需要长期凭证的用例的访问密钥](#)。

[IAM群组](#)是指定IAM用户集合的身份。您不能使用组的身身份登录。您可以使用组来一次性为多个用户指定权限。如果有大量用户，使用组可以更轻松地管理用户权限。例如，您可以拥有一个名为的群组，IAMAdmins并授予该群组管理IAM资源的权限。

用户与角色不同。用户唯一地与某个人员或应用程序关联，而角色旨在让需要它的任何人代入。用户具有永久的长期凭证，而角色提供临时凭证。要了解更多信息，请参阅《[IAM用户指南](#)》中的[何时创建IAM用户 \(而不是角色\)](#)。

IAM角色

[IAM角色](#)是你内在的身份 AWS 账户 具有特定权限的。它与IAM用户类似，但与特定人员无关。你可以暂时扮IAM演一个角色 AWS Management Console 通过[切换角色](#)。你可以通过调用来扮演角色 AWS CLI 或者 AWS API操作或使用自定义URL。有关使用角色的方法的更多信息，请参阅《IAM用户指南》中的[使用IAM角色](#)。

IAM具有临时证书的角色在以下情况下很有用：

- 联合用户访问 – 要向联合身份分配权限，请创建角色并为角色定义权限。当联合身份进行身份验证时，该身份将与角色相关联并被授予由此角色定义的权限。有关用于联合身份验证的角色的信息，请参阅《IAM用户指南》中的[为第三方身份提供商创建角色](#)。如果您使用 IAM Identity Center，则需要配置权限集。为了控制您的身份在进行身份验证后可以访问的内容，IAM Identity Center 会将权限集关联到中的IAM角色。有关权限集的信息，请参阅中的[权限集](#) AWS IAM Identity Center 用户指南。
- 临时IAM用户权限-IAM 用户或角色可以代入一个IAM角色，为特定任务临时获得不同的权限。
- 跨账户访问-您可以使用IAM角色允许其他账户中的某人（受信任的委托人）访问您账户中的资源。角色是授予跨账户访问权限的主要方式。但是，有些 AWS 服务，您可以将策略直接附加到资源（而不是使用角色作为代理）。要了解角色和基于资源的跨账户访问策略之间的区别，请参阅IAM用户指南[IAM中的跨账户资源访问权限](#)。
- 跨服务访问 — 一些 AWS 服务 使用其他功能 AWS 服务。例如，当您在服务中拨打电话时，该服务通常会在 Amazon 中运行应用程序EC2或在 Amazon S3 中存储对象。服务可能会使用发出调用的主体的权限、使用服务角色或使用服务相关角色来执行此操作。
- 转发访问会话 (FAS)-当您使用IAM用户或角色在中执行操作时 AWS，你被视为校长。使用某些服务时，您可能会执行一个操作，然后此操作在其他服务中启动另一个操作。FAS使用委托人的权限调用 AWS 服务，再加上请求的 AWS 服务 向下游服务发出请求。FAS只有当服务收到需要与其他服务进行交互的请求时，才会发出请求 AWS 服务 或需要完成的资源。在这种情况下，您必须具有执行这两个操作的权限。有关提出FAS请求时的政策详情，请参阅[转发访问会话](#)。
- 服务角色-服务[IAM角色](#)是服务代替您执行操作的角色。IAM管理员可以在内部创建、修改和删除服务角色IAM。有关更多信息，请参阅[创建角色以向某人委派权限 AWS 服务](#)（在 IAM 用户指南中）。
- 服务相关角色-服务相关角色是一种与服务相关联的服务角色 AWS 服务。该服务可以代替您执行操作。服务相关角色显示在您的 AWS 账户 并归该服务所有。IAM管理员可以查看但不能编辑服务相关角色的权限。
- 在 Amazon 上运行的应用程序 EC2 — 您可以使用IAM角色来管理在EC2实例上运行并制作的应用程序的临时证书 AWS CLI 或者 AWS API请求。这比在EC2实例中存储访问密钥更可取。要分配

AWS 在 EC2 实例中扮演角色并使其可供其所有应用程序使用，则可以创建附加到该实例的实例配置文件。实例配置文件包含该角色，并允许在 EC2 实例上运行的程序获得临时证书。有关更多信息，请参阅 IAM 用户指南中的 [使用 IAM 角色向在 Amazon EC2 实例上运行的应用程序授予权限](#)。

要了解是使用 IAM 角色还是使用 IAM 用户，请参阅 [《用户指南》中的何时创建 IAM 角色（而不是 IAM 用户）](#)。

使用策略管理访问

您可以控制访问权限 AWS 通过创建策略并将其附加到 AWS 身份或资源。策略是中的一个对象 AWS 当与身份或资源关联时，它定义了他们的权限。AWS 在委托人（用户、root 用户或角色会话）发出请求时评估这些策略。策略中的权限确定是允许还是拒绝请求。大多数策略存储在 AWS 作为 JSON 文件。有关 JSON 策略文档结构和内容的更多信息，请参阅 [《IAM 用户指南》中的 JSON 策略概述](#)。

管理员可以使用 AWS JSON 用于指定谁有权访问什么的策略。也就是说，哪个主体可以对什么资源执行操作，以及在什么条件下执行。

默认情况下，用户和角色没有权限。要授予用户对其所需资源执行操作的权限，IAM 管理员可以创建 IAM 策略。然后，管理员可以将 IAM 策略添加到角色中，用户可以代入角色。

IAM 无论您使用何种方法执行操作，策略都会定义该操作的权限。例如，假设您有一个允许 `iam:GetRole` 操作的策略。拥有该策略的用户可以从中获取角色信息 AWS Management Console，AWS CLI，或者 AWS API。

基于身份的策略

基于身份的策略是可以附加到身份（例如 IAM 用户、用户组或角色）的 JSON 权限策略文档。这些策略控制用户和角色可在何种条件下对哪些资源执行哪些操作。要了解如何创建基于身份的策略，请参阅 IAM 用户指南中的 [创建 IAM 策略](#)。

基于身份的策略可以进一步归类为内联策略或托管策略。内联策略直接嵌入单个用户、组或角色中。托管策略是独立的策略，您可以将其附加到您的多个用户、群组和角色 AWS 账户。托管策略包括 AWS 托管策略和客户托管策略。要了解如何在托管策略或内联策略之间进行选择，请参阅 [《IAM 用户指南》中的在托管策略和内联策略之间进行选择](#)。

基于资源的策略

基于资源的 JSON 策略是您附加到资源的策略文档。基于资源的策略的示例包括 IAM 角色信任策略和 Amazon S3 存储桶策略。在支持基于资源的策略的服务中，服务管理员可以使用它们来控制对特定资

源的访问。对于在其中附加策略的资源，策略定义指定主体可以对该资源执行哪些操作以及在什么条件下执行。您必须在基于资源的策略中[指定主体](#)。委托人可以包括账户、用户、角色、联合用户或 AWS 服务。

基于资源的策略是位于该服务中的内联策略。你不能用 AWS 基于资源的策略 IAM 中的托管策略。

访问控制列表 (ACLs)

访问控制列表 (ACLs) 控制哪些委托人 (账户成员、用户或角色) 有权访问资源。ACLs 与基于资源的策略类似，尽管它们不使用 JSON 策略文档格式。

亚马逊 S3，AWS WAF，Amazon VPC 就是支持的服务示例 ACLs。要了解更多信息 ACLs，请参阅《亚马逊简单存储服务开发者指南》中的[访问控制列表 \(ACL\) 概述](#)。

其他策略类型

AWS 支持其他不太常见的策略类型。这些策略类型可以设置更常用的策略类型向您授予的最大权限。

- 权限边界-权限边界是一项高级功能，您可以在其中设置基于身份的策略可以向 IAM 实体 (IAM 用户或角色) 授予的最大权限。您可为实体设置权限边界。这些结果权限是实体基于身份的策略及其权限边界的交集。在 Principal 中指定用户或角色的基于资源的策略不受权限边界限制。任一项策略中的显式拒绝将覆盖允许。有关权限边界的更多信息，请参阅《IAM 用户指南》中的[IAM 实体的权限边界](#)。
- 服务控制策略 (SCPs)-SCPs 是指定组织或组织单位 (OU) 的最大权限的 JSON 策略 AWS Organizations。AWS Organizations 是一项用于对多个进行分组和集中管理的 AWS 账户 您的企业拥有的。如果您启用组织中的所有功能，则可以将服务控制策略 (SCPs) 应用于您的任何或所有帐户。SCP 限制了成员账户中实体的权限，包括每个 AWS 账户根用户。有关 Organizations 和的更多信息 SCPs，请参阅《》中的[服务控制策略](#) AWS Organizations 用户指南。
- 会话策略 – 会话策略是当您以编程方式为角色或联合用户创建临时会话时作为参数传递的高级策略。结果会话的权限是用户或角色的基于身份的策略和会话策略的交集。权限也可以来自基于资源的策略。任一项策略中的显式拒绝将覆盖允许。有关更多信息，请参阅《IAM 用户指南》中的[会话策略](#)。

多个策略类型

当多个类型的策略应用于一个请求时，生成的权限更加复杂和难以理解。要了解如何做 AWS 决定在涉及多种策略类型时是否允许请求，请参阅《IAM 用户指南》中的[策略评估逻辑](#)。

操作方法 AWS Entity Resolution 数据匹配服务 配合使用 IAM

在使用管理IAM访问权限之前 AWS Entity Resolution 数据匹配服务，了解哪些IAM功能可供使用 AWS Entity Resolution 数据匹配服务。

IAM你可以使用的功能 AWS Entity Resolution 数据匹配服务

IAM功能	AWS Entity Resolution 数据匹配服务 支持
基于身份的策略	是
基于资源的策略	是
策略操作	是
策略资源	是
策略条件键	是
ACLs	不支持
ABAC (策略中的标签)	部分
临时凭证	是
转发访问会话 (FAS)	是
服务角色	是
服务相关角色	否

要从更高层次的角度了解如何 AWS Entity Resolution 数据匹配服务 和其他 AWS 服务适用于大多数 IAM功能，请参阅 [AWS IAM](#)在《IAM用户指南》中使用的服务。

基于身份的策略 AWS Entity Resolution 数据匹配服务

支持基于身份的策略：是

基于身份的策略是可以附加到身份（例如IAM用户、用户组或角色）的JSON权限策略文档。这些策略控制用户和角色可在何种条件下对哪些资源执行哪些操作。要了解如何创建基于身份的策略，请参阅IAM用户指南中的[创建IAM策略](#)。

使用IAM基于身份的策略，您可以指定允许或拒绝的操作和资源，以及允许或拒绝操作的条件。您无法在基于身份的策略中指定主体，因为它适用于其附加的用户或角色。要了解可以在JSON策略中使用的所有元素，请参阅IAM用户指南中的[IAMJSON策略元素参考](#)。

基于身份的策略示例 AWS Entity Resolution 数据匹配服务

查看以下示例 AWS Entity Resolution 数据匹配服务 基于身份的策略，请参阅。[基于身份的策略示例 AWS Entity Resolution 数据匹配服务](#)

内部基于资源的政策 AWS Entity Resolution 数据匹配服务

支持基于资源的策略：是

基于资源的JSON策略是您附加到资源的策略文档。基于资源的策略的示例包括IAM角色信任策略和Amazon S3 存储桶策略。在支持基于资源的策略的服务中，服务管理员可以使用它们来控制对特定资源的访问。对于在其中附加策略的资源，策略定义指定主体可以对该资源执行哪些操作以及在什么条件下执行。您必须在基于资源的策略中[指定主体](#)。委托人可以包括账户、用户、角色、联合用户或AWS服务。

要启用跨账户访问权限，您可以将整个账户或另一个账户中的IAM实体指定为基于资源的策略中的委托人。将跨账户主体添加到基于资源的策略只是建立信任关系工作的一半而已。当委托人和资源处于不同状态时AWS账户，可信账户中的IAM管理员还必须向委托人实体（用户或角色）授予访问资源的权限。他们通过将基于身份的策略附加到实体以授予权限。但是，如果基于资源的策略向同一个账户中的主体授予访问权限，则不需要额外的基于身份的策略。有关更多信息，请参阅《IAM用户指南》IAM中的[跨账户资源访问权限](#)。

的政策行动 AWS Entity Resolution 数据匹配服务

支持策略操作：是

管理员可以使用AWS JSON用于指定谁有权访问什么的策略。也就是说，哪个主体可以对什么资源执行操作，以及在什么条件下执行。

JSON策略Action元素描述了可用于在策略中允许或拒绝访问的操作。策略操作通常与关联的同名AWS API操作。也有一些例外，例如没有匹配API操作的仅限权限的操作。还有一些操作需要在策略中执行多个操作。这些附加操作称为相关操作。

在策略中包含操作以授予执行关联操作的权限。

要查看清单 [AWS Entity Resolution 数据匹配服务 操作](#)，请参阅[操作定义者 AWS Entity Resolution 数据匹配服务](#)在《服务授权参考》中。

中的政策行动 [AWS Entity Resolution 数据匹配服务](#) 在操作前使用以下前缀：

```
entityresolution
```

要在单个语句中指定多项操作，请使用逗号将它们隔开。

```
"Action": [  
  "entityresolution:action1",  
  "entityresolution:action2"  
]
```

查看以下示例 [AWS Entity Resolution 数据匹配服务 基于身份的策略](#)，请参阅。[基于身份的策略示例 AWS Entity Resolution 数据匹配服务](#)

的政策资源 [AWS Entity Resolution 数据匹配服务](#)

支持策略资源：是

管理员可以使用 AWS JSON用于指定谁有权访问什么的策略。也就是说，哪个主体 可以对什么资源执行操作，以及在什么条件下执行。

ResourceJSON策略元素指定要应用操作的一个或多个对象。语句必须包含 Resource 或 NotResource 元素。最佳做法是，使用资源的 [Amazon 资源名称 \(ARN\)](#) 来指定资源。对于支持特定资源类型（称为资源级权限）的操作，您可以执行此操作。

对于不支持资源级权限的操作（如列出操作），请使用通配符 (*) 指示语句应用于所有资源。

```
"Resource": "*"
```

要查看清单 [AWS Entity Resolution 数据匹配服务 资源类型及其类型ARNs](#)，参见[资源定义者 AWS Entity Resolution 数据匹配服务](#)在《服务授权参考》中。要了解您可以使用哪些操作来指定每ARN种资源，请参阅[操作定义者 AWS Entity Resolution 数据匹配服务](#)。

查看以下示例 [AWS Entity Resolution 数据匹配服务 基于身份的策略](#)，请参阅。[基于身份的策略示例](#)
[AWS Entity Resolution 数据匹配服务](#)

的策略条件密钥 AWS Entity Resolution 数据匹配服务

支持特定于服务的策略条件键：是

管理员可以使用 AWS JSON用于指定谁有权访问什么的策略。也就是说，哪个主体可以对什么资源执行操作，以及在什么条件下执行。

在 Condition 元素 (或 Condition 块) 中，可以指定语句生效的条件。Condition 元素是可选的。您可以创建使用[条件运算符](#) (例如，等于或小于) 的条件表达式，以使策略中的条件与请求中的值相匹配。

如果您在一条语句中指定多个Condition元素，或者在单个Condition元素中指定多个键，AWS 使用逻辑AND运算对其进行评估。如果您为单个条件键指定多个值，AWS 使用逻辑OR运算评估条件。在授予语句的权限之前必须满足所有的条件。

在指定条件时，您也可以使用占位符变量。例如，只有在资源上标有IAM用户的用户名时，您才能向IAM用户授予访问该资源的权限。有关更多信息，请参阅《IAM用户指南》中的[IAM策略元素：变量和标签](#)。

AWS 支持全局条件密钥和特定于服务的条件键。要查看全部 AWS 全局条件键，请参见 [AWS 《IAM用户指南》](#) 中的全局条件上下文密钥。

要查看清单 [AWS Entity Resolution 数据匹配服务 条件键](#)，参见[条件键 AWS Entity Resolution 数据匹配服务](#)在《服务授权参考》中。要了解可以使用条件键的操作和资源，请参阅[操作定义者 AWS Entity Resolution 数据匹配服务](#)。

查看以下示例 [AWS Entity Resolution 数据匹配服务 基于身份的策略](#)，请参阅。[基于身份的策略示例](#)
[AWS Entity Resolution 数据匹配服务](#)

ACLs在 AWS Entity Resolution 数据匹配服务

支持ACLs：否

访问控制列表 (ACLs) 控制哪些委托人 (账户成员、用户或角色) 有权访问资源。ACLs与基于资源的策略类似，尽管它们不使用JSON策略文档格式。

ABAC与 AWS Entity Resolution 数据匹配服务

支持ABAC (策略中的标签)：部分

基于属性的访问控制 (ABAC) 是一种基于属性定义权限的授权策略。在 AWS 中，这些属性称为标签。您可以将标签附加到 IAM 实体（用户或角色）和许多实体 AWS 资源上。为实体和资源添加标签是第一步。ABAC 然后，您可以设计 ABAC 策略，允许在委托人的标签与他们尝试访问的资源上的标签匹配时进行操作。

ABAC 在快速增长的环境中很有用，也有助于解决策略管理变得繁琐的情况。

要基于标签控制访问，您需要使用 `aws:ResourceTag/key-name`、`aws:RequestTag/key-name` 或 `aws:TagKeys` 条件键在策略的 [条件元素](#) 中提供标签信息。

如果某个服务对于每种资源类型都支持所有这三个条件键，则对于该服务，该值为是。如果某个服务仅对于部分资源类型支持所有这三个条件键，则该值为部分。

有关的更多信息 ABAC，请参阅 [什么是 ABAC？](#) 在《IAM 用户指南》中。要查看包含设置步骤的教程 ABAC，请参阅 IAM 用户指南中的 [使用基于属性的访问控制 \(ABAC\)](#)。

将临时证书与 AWS Entity Resolution 数据匹配服务

支持临时凭证：是

一段时间 AWS 服务使用临时证书登录时不起作用。欲了解更多信息，包括哪个 AWS 服务使用临时证书，请参阅 [AWS 服务可以 IAM](#) 在《IAM 用户指南》中使用。

如果您登录，则使用的是临时证书 AWS Management Console 使用除用户名和密码之外的任何方法。例如，当您访问时 AWS 使用贵公司的单点登录 (SSO) 链接，该过程会自动创建临时证书。当您以用户身份登录控制台，然后切换角色时，您还会自动创建临时凭证。有关切换角色的更多信息，请参阅《IAM 用户指南》中的 [切换到角色（控制台）](#)。

您可以使用手动创建临时证书 AWS CLI 或者 AWS API。然后，您可以使用这些临时证书进行访问 AWS。AWS 建议您动态生成临时证书，而不是使用长期访问密钥。有关更多信息，请参阅 [中的临时安全证书 IAM](#)。

转发访问会话 AWS Entity Resolution 数据匹配服务

支持转发访问会话 (FAS)：是

当您使用 IAM 用户或角色在中执行操作时 AWS，您被视为校长。使用某些服务时，您可能会执行一个操作，然后此操作在其他服务中启动另一个操作。FAS 使用委托人的权限调用 AWS 服务，再加上请求的 AWS 服务 向下游服务发出请求。FAS 只有当服务收到需要与其他服务进行交互的请求时，才会发出请求 AWS 服务 或需要完成的资源。在这种情况下，您必须具有执行这两个操作的权限。有关提出 FAS 请求时的政策详情，请参阅 [转发访问会话](#)。

的服务角色 AWS Entity Resolution 数据匹配服务

支持服务角色：是

服务IAM角色是服务代替您执行操作的角色。IAM管理员可以在内部创建、修改和删除服务角色IAM。有关更多信息，请参阅[创建角色以向某人委派权限 AWS 服务](#)（在 IAM 用户指南中）。

Warning

更改服务角色的权限可能会中断 AWS Entity Resolution 数据匹配服务 功能。仅在以下情况下编辑服务角色 AWS Entity Resolution 数据匹配服务 提供了执行此操作的指导。

的服务相关角色 AWS Entity Resolution 数据匹配服务

支持服务相关角色：否

服务相关角色是一种与服务相关联的服务角色 AWS 服务。该服务可以代替您执行操作。服务相关角色显示在您的 AWS 账户 并归该服务所有。IAM管理员可以查看但不能编辑服务相关角色的权限。

有关创建或管理服务相关角色的详细信息，请参阅 [AWS 与之配合使用的服务IAM](#)。在表中查找服务相关角色列中包含 Yes 的表。选择是链接以查看该服务的服务相关角色文档。

基于身份的策略示例 AWS Entity Resolution 数据匹配服务

默认情况下，用户和角色无权创建或修改 AWS Entity Resolution 数据匹配服务 资源的费用。他们也无法使用来执行任务 AWS Management Console, AWS Command Line Interface (AWS CLI), 或 AWS API。要授予用户对其所需资源执行操作的权限，IAM管理员可以创建IAM策略。然后，管理员可以将IAM策略添加到角色中，用户可以代入角色。

要了解如何使用这些示例策略文档创建IAM基于身份的JSON策略，请参阅IAM用户指南中的[创建IAM策略](#)。

有关由定义的操作和资源类型的详细信息 AWS Entity Resolution 数据匹配服务，包括每种资源类型的格式，请参见[操作、资源和条件键 ARNs AWS Entity Resolution 数据匹配服务](#)在《服务授权参考》中。

主题

- [策略最佳实践](#)

- [使用 AWS Entity Resolution 数据匹配服务 Console](#)
- [允许用户查看他们自己的权限](#)

策略最佳实践

基于身份的策略决定了某人是否可以创建、访问或删除 AWS Entity Resolution 数据匹配服务 您账户中的资源。这些操作可能会使您付出代价 AWS 账户。创建或编辑基于身份的策略时，请遵循以下准则和建议：

- 开始使用 AWS 托管策略并转向最低权限权限 — 要开始向用户和工作负载授予权限，请使用 AWS 为许多常见用例授予权限的托管策略。它们在你的 AWS 账户。我们建议您通过定义来进一步减少权限 AWS 特定于您的用例的客户托管政策。有关更多信息，请参阅 [AWS 托管策略](#) 或 [AWS 《IAM 用户指南》](#) 中工作职能的托管策略。
- 应用最低权限权限-使用IAM策略设置权限时，仅授予执行任务所需的权限。为此，您可以定义在特定条件下可以对特定资源执行的操作，也称为最低权限许可。有关使用应用权限IAM的更多信息，请参阅《IAM用户指南》IAM[中的策略和权限](#)。
- 使用IAM策略中的条件进一步限制访问权限-您可以在策略中添加条件以限制对操作和资源的访问权限。例如，您可以编写一个策略条件来指定所有请求都必须使用发送SSL。如果通过特定条件使用服务操作，则也可以使用条件来授予对服务操作的访问权限 AWS 服务之外的压缩算法（例如 AWS CloudFormation。有关更多信息，请参阅《IAM用户指南》中的[IAMJSON策略元素：条件](#)。
- 使用 A IAM ccess Analyzer 验证您的IAM策略以确保权限的安全性和功能性 — A IAM ccess Analyzer 会验证新的和现有的策略，以便策略符合IAM策略语言 (JSON) 和IAM最佳实践。IAMAccess Analyzer 提供了 100 多项策略检查和可行的建议，可帮助您制定安全和实用的策略。有关更多信息，请参阅《IAM用户指南》中的 [IAMAccess Analyzer 策略验证](#)。
- 需要多因素身份验证 (MFA)-如果您的场景需要IAM用户或 root 用户 AWS 账户，请打开MFA以提高安全性。要要求MFA何时调用API操作，请在策略中添加MFA条件。有关更多信息，请参阅《IAM用户指南》中的[配置MFA受保护的API访问权限](#)。

有关最佳做法的更多信息IAM，请参阅《IAM用户指南》IAM[中的安全最佳实践](#)。

使用 AWS Entity Resolution 数据匹配服务 Console

要访问 AWS Entity Resolution 数据匹配服务 控制台，您必须拥有一组最低权限。这些权限必须允许您列出和查看有关以下内容的详细信息 AWS Entity Resolution 数据匹配服务 您账户中的资源 AWS 账户。如果您创建的基于身份的策略比所需的最低权限更严格，则控制台将无法按预期运行，适用于使用该策略的实体（用户或角色）。

您无需为仅拨打控制台的用户设置最低控制台权限 AWS CLI 或者 AWS API。相反，只允许访问与他们尝试执行的API操作相匹配的操作。

为确保用户和角色仍然可以使用 AWS Entity Resolution 数据匹配服务 控制台，还要连接 AWS Entity Resolution 数据匹配服务 *ConsoleAccess* 或 *ReadOnly* AWS 实体的托管策略。有关更多信息，请参阅《[用户指南](#)》中的[向IAM用户添加权限](#)。

允许用户查看他们自己的权限

此示例说明如何创建允许IAM用户查看附加到其用户身份的内联和托管策略的策略。此策略包括通过控制台或以编程方式使用控制台完成此操作的权限 AWS CLI 或者 AWS API。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "ViewOwnUserInfo",
      "Effect": "Allow",
      "Action": [
        "iam:GetUserPolicy",
        "iam:ListGroupForUser",
        "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 的托管策略 AWS Entity Resolution 数据匹配服务

AWS 托管策略是由创建和管理的独立策略 AWS。AWS 托管策略旨在为许多常见用例提供权限，以便您可以开始为用户、组和角色分配权限。

请记住，AWS 托管策略可能不会为您的特定用例授予最低权限权限，因为它们可供所有 AWS 客户使用。我们建议通过定义特定于您的使用场景的[客户管理型策略](#)来进一步减少权限。

您无法更改 AWS 托管策略中定义的权限。如果 AWS 更新 AWS 托管策略中定义的权限，则更新会影响该策略所关联的所有委托人身份（用户、组和角色）。AWS 最有可能在启动新的 API 或现有服务可以使用新 AWS 服务的 API 操作时更新 AWS 托管策略。

有关更多信息，请参阅《IAM 用户指南》中的[AWS 托管策略](#)。

AWS 托管策略：AWSEntityResolutionConsoleFullAccess

您可以将 AWSEntityResolutionConsoleFullAccess 策略附加到 IAM 身份。

此策略授予对 AWS Entity Resolution 数据匹配服务 端点和资源的完全访问权限。

此策略还允许对相关内容（AWS 服务 例如 S3、AWS Glue、Tagging 等）进行某些读取，AWS KMS 以便控制台可以显示选择并使用所选选项来执行实体解析操作。某些资源的范围已缩小到包含服务名称entityresolution。

由于 AWS Entity Resolution 数据匹配服务 依赖传递的角色对相关 AWS 资源执行操作，因此该策略还授予选择和传递所需角色的权限。

权限详细信息

该策略包含以下权限。

- EntityResolutionAccess— 允许委托人完全访问 AWS Entity Resolution 数据匹配服务 端点和资源。
- GlueSourcesConsoleDisplay— 授予将 AWS Glue 表列为数据源选项和导入数据源的表架构的访问权限，以获得用户体验。
- S3BucketsConsoleDisplay— 授予将所有 S3 存储桶列为数据源选项的权限。
- S3SourcesConsoleDisplay— 授予将 S3 存储桶显示为数据源选项的权限。

- `TaggingConsoleDisplay`— 授予读取标记键和值的权限。
- `KMSConsoleDisplay`— 授予描述密钥和列出别名的访问权限 `AWS Key Management Service` ，以解密和加密数据源。
- `ListRolesToPickForPassing`— 授予列出所有角色的权限，以便用户可以选择要传递的角色。
- `PassRoleToEntityResolutionService`— 授予将缩小范围的角色传递给 `AWS Entity Resolution 数据匹配服务` 服务的访问权限。
- `ManageEventBridgeRules`— 授予创建、更新和删除用于获取 S3 通知的 `Amazon EventBridge` 规则的权限。
- `ADXReadAccess`— 授予访问权限 `AWS Data Exchange` 以验证客户是否有权利或订阅。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "EntityResolutionAccess",
      "Effect": "Allow",
      "Action": [
        "entityresolution:*"
      ],
      "Resource": "*"
    },
    {
      "Sid": "GlueSourcesConsoleDisplay",
      "Effect": "Allow",
      "Action": [
        "glue:GetSchema",
        "glue:SearchTables",
        "glue:GetSchemaByDefinition",
        "glue:GetSchemaVersion",
        "glue:GetSchemaVersionsDiff",
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:GetTable",
        "glue:GetTables",
        "glue:GetTableVersion",
        "glue:GetTableVersions"
      ],
      "Resource": "*"
    },
    {
```

```
    "Sid": "S3BucketsConsoleDisplay",
    "Effect": "Allow",
    "Action": [
        "s3:ListAllMyBuckets"
    ],
    "Resource": "*"
},
{
    "Sid": "S3SourcesConsoleDisplay",
    "Effect": "Allow",
    "Action": [
        "s3:ListBucket",
        "s3:GetBucketLocation",
        "s3:ListBucketVersions",
        "s3:GetBucketVersioning"
    ],
    "Resource": "*"
},
{
    "Sid": "TaggingConsoleDisplay",
    "Effect": "Allow",
    "Action": [
        "tag:GetTagKeys",
        "tag:GetTagValues"
    ],
    "Resource": "*"
},
{
    "Sid": "KMSConsoleDisplay",
    "Effect": "Allow",
    "Action": [
        "kms:DescribeKey",
        "kms:ListAliases"
    ],
    "Resource": "*"
},
{
    "Sid": "ListRolesToPickRoleForPassing",
    "Effect": "Allow",
    "Action": [
        "iam:ListRoles"
    ],
    "Resource": "*"
},
}
```

```
{
  "Sid": "PassRoleToEntityResolutionService",
  "Effect": "Allow",
  "Action": [
    "iam:PassRole"
  ],
  "Resource": "arn:aws:iam::*:role/*entityresolution*",
  "Condition": {
    "StringEquals": {
      "iam:PassedToService": [
        "entityresolution.amazonaws.com"
      ]
    }
  }
},
{
  "Sid": "ManageEventBridgeRules",
  "Effect": "Allow",
  "Action": [
    "events:PutRule",
    "events>DeleteRule",
    "events:PutTargets",
  ],
  "Resource": [
    "arn:aws:events::*:rule/entity-resolution-automatic*"
  ]
},
{
  "Sid": "ADXReadAccess",
  "Effect": "Allow",
  "Action": [
    "dataexchange:GetDataSet"
  ],
  "Resource": "*"
},
]
```

AWS 托管策略：AWSEntityResolutionConsoleReadOnlyAccess

您可以将 AWSEntityResolutionConsoleReadOnlyAccess 附加到 IAM 实体。

此策略授予对 AWS Entity Resolution 数据匹配服务 端点和资源的只读访问权限。

权限详细信息

该策略包含以下权限。

- **EntityResolutionRead**— 允许委托人对 AWS Entity Resolution 数据匹配服务 端点和资源进行只读访问。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "EntityResolutionRead",
      "Effect": "Allow",
      "Action": [
        "entityresolution:Get*",
        "entityresolution:List*"
      ],
      "Resource": "*"
    }
  ]
}
```

AWS Entity Resolution 数据匹配服务AWS 托管策略的更新

查看 [AWS Entity Resolution 数据匹配服务](#) 自该服务开始跟踪这些更改以来 AWS 托管策略更新的详细信息。要获得有关此页面变更的自动提醒，请订阅“[AWS Entity Resolution 数据匹配服务 文档历史记录](#)”页面上的 RSS feed。

更改	描述	日期
AWSEntityResolutionConsoleFullAccess - 对现有策略的更新	添加ADXReadAccess ManageEventBridgeRules 并启用匹配工作流程中的提供者服务选项。	2023 年 10 月 16 日
AWS Entity Resolution 数据匹配服务 开始跟踪更改	AWS Entity Resolution 数据匹配服务 开始跟踪其 AWS 托管策略的更改。	2023 年 8 月 18 日

故障排除 AWS Entity Resolution 数据匹配服务 身份和访问权限

使用以下信息来帮助您诊断和修复在使用时可能遇到的常见问题 AWS Entity Resolution 数据匹配服务和IAM。

主题

- [我无权在以下位置执行操作 AWS Entity Resolution 数据匹配服务](#)
- [我无权执行 iam : PassRole](#)
- [我想允许我以外的人进入 AWS 账户 访问我的 AWS Entity Resolution 数据匹配服务 资源](#)

我无权在以下位置执行操作 AWS Entity Resolution 数据匹配服务

如果 AWS Management Console 告诉您您无权执行某项操作，那么您必须联系管理员寻求帮助。管理员是指提供用户名和密码的人员。

当mateojacksonIAM用户尝试使用控制台查看虚构`my-example-widget`资源的详细信息但没有虚构权限时，就会出现以下示例错误。entityresolution:`GetWidget`

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
entityresolution:GetWidget on resource: my-example-widget
```

在这种情况下，Mateo 请求他的管理员更新其策略，以允许他使用 entityresolution:`GetWidget` 操作访问 `my-example-widget` 资源。

我无权执行 iam : PassRole

如果您收到一条错误消息，说您无权执行iam:PassRole操作，则必须更新您的策略以允许您将角色传递给 AWS Entity Resolution 数据匹配服务。

一段时间 AWS 服务 允许您将现有角色传递给该服务，而不是创建新的服务角色或服务相关角色。为此，您必须具有将角色传递到服务的权限。

当名为的IAM用户marymajor尝试使用控制台在中执行操作时，会出现以下示例错误 AWS Entity Resolution 数据匹配服务。但是，该操作要求服务拥有由服务角色授予的权限。Mary 不具有将角色传递到服务的权限。

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
iam:PassRole
```

在这种情况下，必须更新 Mary 的策略以允许她执行 `iam:PassRole` 操作。

如果您需要帮助，请联系您的 AWS 管理员。您的管理员是提供登录凭证的人。

我想允许我以外的人进入 AWS 账户 访问我的 AWS Entity Resolution 数据匹配服务 资源

您可以创建一个角色，以便其他账户中的用户或您组织外的人员可以使用该角色来访问您的资源。您可以指定谁值得信赖，可以担任角色。对于支持基于资源的策略或访问控制列表 (ACLs) 的服务，您可以使用这些策略向人们授予访问您的资源的权限。

要了解更多信息，请参阅以下内容：

- 要了解是否 AWS Entity Resolution 数据匹配服务 支持这些功能，请参阅[操作方法 AWS Entity Resolution 数据匹配服务 配合使用 IAM](#)。
- 了解如何提供对您的资源的访问权限 AWS 账户 您拥有的，请参阅[向其他IAM用户提供访问权限 AWS 账户 您在《IAM用户指南》中拥有的](#)。
- 了解如何向第三方提供对您的资源的访问权限 AWS 账户，请参阅[提供访问权限 AWS 账户 IAM用户指南中归第三方所有](#)。
- 要了解如何通过联合身份验证提供访问权限，请参阅《用户指南》中的[向经过外部身份验证的用户提供访问权限 \(联合身份验证\)](#)。IAM
- 要了解使用角色和基于资源的策略进行跨账户访问的区别，请参阅IAM用户指南[IAM中的跨账户资源访问权限](#)。

合规性验证 AWS Entity Resolution 数据匹配服务

要了解是否属于特定合规计划的范围，请参阅AWS 服务“[按合规计划划分的范围](#)”，然后选择您感兴趣的合规计划。AWS 服务 有关一般信息，请参阅[AWS 合规计划AWS](#)。

您可以使用下载第三方审计报告 AWS Artifact。有关更多信息，请参阅中的“[下载报告](#)”中的“[AWS Artifact](#)”。

您在使用 AWS 服务 时的合规责任取决于您的数据的敏感性、贵公司的合规目标以及适用的法律和法规。AWS 提供了以下资源来帮助实现合规性：

- [安全与合规性快速入门指南](#) — 这些部署指南讨论了架构注意事项，并提供了在这些基础上 AWS 部署以安全性和合规性为重点的基准环境的步骤。

- 在 [Amazon Web Services 上进行HIPAA安全与合规架构](#) — 本白皮书描述了各公司如何使用 AWS 来创建HIPAA符合条件的应用程序。

Note

并非所有 AWS 服务 人都有HIPAA资格。有关更多信息，请参阅《[HIPAA符合条件的服务参考](#)》。

- [AWS 合规资源AWS](#) — 此工作簿和指南集可能适用于您所在的行业和所在地区。
- [AWS 客户合规指南](#) — 从合规角度了解责任共担模式。这些指南总结了保护的最佳实践，AWS 服务 并将指南映射到跨多个框架（包括美国国家标准与技术研究院 (NIST)、支付卡行业安全标准委员会 (PCI) 和国际标准化组织 (ISO)) 的安全控制。
- [使用AWS Config 开发人员指南中的规则评估资源](#) — 该 AWS Config 服务评估您的资源配置在多大程度上符合内部实践、行业指导方针和法规。
- [AWS Security Hub](#)— 这 AWS 服务 提供了您内部安全状态的全面视图 AWS。Security Hub 通过安全控件评估您的 AWS 资源并检查其是否符合安全行业标准和最佳实践。有关受支持服务及控件的列表，请参阅 [Security Hub 控件参考](#)。
- [Amazon GuardDuty](#) — 它通过监控您的 AWS 账户环境中是否存在可疑和恶意活动，来 AWS 服务 检测您的工作负载、容器和数据面临的潜在威胁。GuardDuty 可以帮助您满足各种合规性要求 PCIDSS，例如满足某些合规性框架规定的入侵检测要求。
- [AWS Audit Manager](#)— 这 AWS 服务 可以帮助您持续审计 AWS 使用情况，从而简化风险管理以及对法规和行业标准的合规性。

AWS Entity Resolution 数据匹配服务 合规性最佳实践

本节提供使用时的最佳实践和合规性建议 AWS Entity Resolution 数据匹配服务。

支付卡行业数据安全标准 (PCIDSS)

AWS Entity Resolution 数据匹配服务 支持商家或服务提供商处理、存储和传输信用卡数据，并且已被验证符合支付卡行业 (PCI) 数据安全标准 (DSS)。有关更多信息 PCIDSS，包括如何索取 Compliance Packag AWS PCI e 的副本，请参阅[PCIDSS级别 1](#)。

系统和组织控制 (SOC)

AWS Entity Resolution 数据匹配服务 符合系统和组织控制 (SOC) 措施，包括 SOC 1、SOC 2 和 SOC 3。SOC报告是独立的第三方检查报告，用于展示如何 AWS 实现关键合规控制和目标。这些审计可确

保采取适当的安全措施和程序，防止出现风险，影响客户和公司数据的安全性、机密性和可用性。这些第三方审计的结果可在[AWS SOC合规网站上](#)查看，您可以在其中查看已发布的报告，以获取有关支持 AWS 运营和合规的控制措施的更多信息。

韧性在 AWS Entity Resolution 数据匹配服务

AWS 全球基础设施是围绕 AWS 区域 可用区构建的。AWS 区域 提供多个物理隔离和隔离的可用区，这些可用区通过低延迟、高吞吐量和高度冗余的网络连接。利用可用区，您可以设计和操作在可用区之间无中断地自动实现失效转移的应用程序和数据库。与传统的单个或多个数据中心基础设施相比，可用区具有更高的可用性、容错性和可扩展性。

有关 AWS 区域 和可用区的更多信息，请参阅[AWS 全球基础设施](#)。

除了 AWS 全球基础架构外，还 AWS Entity Resolution 数据匹配服务 提供多项功能来帮助支持您的数据弹性和备份需求。

监控 AWS Entity Resolution 数据匹配服务

监控是维护和其他 AWS 解决方案的可靠性、可用性和性能的重要组成部分。AWS Entity Resolution 数据匹配服务 AWS 提供以下监控工具 AWS Entity Resolution 数据匹配服务，供您监视、报告问题并在适当时自动采取措施：

- AWS CloudTrail 捕获由您或代表您发起的 API 调用和相关事件，AWS 账户 并将日志文件传输到您指定的 Amazon S3 存储桶。您可以识别哪些用户和账户拨打了电话 AWS、发出呼叫的源 IP 地址以及呼叫发生的时间。有关更多信息，请参阅 [AWS CloudTrail 用户指南](#)。

主题

- [使用记录 AWS Entity Resolution 数据匹配服务 API 调用 AWS CloudTrail](#)

使用记录 AWS Entity Resolution 数据匹配服务 API 调用 AWS CloudTrail

AWS Entity Resolution 数据匹配服务 与 AWS CloudTrail 一项服务集成，该服务提供用户、角色或 AWS 服务在中执行的操作的记录 AWS Entity Resolution 数据匹配服务。CloudTrail 将所有 API 调用捕获 AWS Entity Resolution 数据匹配服务 为事件。捕获的调用包括来自 AWS Entity Resolution 数据匹配服务 控制台的调用和对 AWS Entity Resolution 数据匹配服务 API 操作的代码调用。如果您创建了跟踪，则可以允许将 CloudTrail 事件持续传输到 Amazon S3 存储桶，包括的事件 AWS Entity Resolution 数据匹配服务。如果您未配置跟踪，您仍然可以在 CloudTrail 控制台的“事件历史记录”中查看最新的事件。使用收集的信息 CloudTrail，您可以确定向哪个请求发出 AWS Entity Resolution 数据匹配服务、发出请求的 IP 地址、谁发出了请求、何时发出请求以及其他详细信息。

要了解更多信息 CloudTrail，请参阅 [《AWS CloudTrail 用户指南》](#)。

AWS Entity Resolution 数据匹配服务 信息在 CloudTrail

CloudTrail 在您创建账户 AWS 账户 时已在您的账户上启用。当活动发生在中时 AWS Entity Resolution 数据匹配服务，该活动会与其他 AWS 服务 CloudTrail 事件一起记录在事件历史记录中。您可以在 AWS 账户中查看、搜索和下载最新事件。有关更多信息，请参阅 [使用事件历史记录查看 CloudTrail 事件](#)。

要持续记录您的 AWS 账户事件（包括的事件）AWS Entity Resolution 数据匹配服务，请创建跟踪。跟踪允许 CloudTrail 将日志文件传输到 Amazon S3 存储桶。预设情况下，在控制台中创建跟踪记

录时，此跟踪记录应用于所有 AWS 区域。跟踪记录 AWS 分区中所有区域的事件，并将日志文件传送到您指定的 Amazon S3 存储桶。此外，您可以配置其他 AWS 服务，以进一步分析和处理 CloudTrail 日志中收集的事件数据。有关更多信息，请参阅以下内容：

- [创建跟踪记录概述](#)
- [CloudTrail 支持的服务和集成](#)
- [配置 Amazon SNS 通知 CloudTrail](#)
- [接收来自多个区域的 CloudTrail 日志文件](#)和[接收来自多个账户的 CloudTrail 日志文件](#)

所有 AWS Entity Resolution 数据匹配服务 操作均由《API 参考》记录 CloudTrail 并记录在《[AWS Entity Resolution 数据匹配服务 API 参考](#)》中。

每个事件或日记账条目都包含有关生成请求的人员信息。身份信息有助于您确定以下内容：

- 请求是使用根证书还是 AWS Identity and Access Management (IAM) 用户凭证发出。
- 请求是使用角色还是联合用户的临时安全凭证发出的。
- 请求是否由其他 AWS 服务发出。

有关更多信息，请参阅[CloudTrail 用户身份元素](#)。

了解 AWS Entity Resolution 数据匹配服务 日志文件条目

跟踪是一种配置，允许将事件作为日志文件传输到您指定的 Amazon S3 存储桶。CloudTrail 日志文件包含一个或多个日志条目。事件代表来自任何来源的单个请求，包括有关请求的操作、操作的日期和时间、请求参数等的信息。CloudTrail 日志文件不是公共 API 调用的有序堆栈跟踪，因此它们不会按任何特定的顺序出现。

使用创建AWS实体解析资源 AWS CloudFormation

AWS Entity Resolution 与一项服务集成，该服务可帮助您对 AWS 资源进行建模和设置，从而减少创建和管理资源和基础架构所花费的时间。您可以创建一个描述所需的所有 AWS 资源（例如 `AWS::EntityResolution::MatchingWorkflow`，`AWS::EntityResolution::SchemaMapping`，`AWS::EntityResolution::IdMappingWorkflow`，`AWS::EntityResolution::IdNamespace` 和 `AWS::EntityResolution::PolicyStatement`）的模板，然后为 AWS CloudFormation 预置和配置这些资源。

使用时 AWS CloudFormation，您可以重复使用模板来一致且重复地设置AWS实体解析资源。只需描述一次您的资源，然后在多个 AWS 账户 区域中一遍又一遍地配置相同的资源。

AWS实体分辨率和 AWS CloudFormation 模板

要为AWS实体解析和相关服务配置和配置资源，必须了解[AWS CloudFormation 模板](#)。模板是JSON或格式化的文本文件YAML。这些模板描述了您要在 AWS CloudFormation 堆栈中配置的资源。如果您不熟悉JSON或YAML，可以使用 AWS CloudFormation Designer 来帮助您开始使用 AWS CloudFormation 模板。有关更多信息，请参阅《AWS CloudFormation 用户指南》中的[什么是 AWS CloudFormation Designer ?](#)。

AWS实体解析支持创建 `AWS::EntityResolution::MatchingWorkflow`，`AWS::EntityResolution::SchemaMapping`，`AWS::EntityResolution::IdMappingWorkflow`，`AWS::EntityResolution::IdNamespace` 和输入 `AWS::EntityResolution::PolicyStatement` 入 AWS CloudFormation。有关更多信息，包括JSON和的示例和YAML模板 `AWS::EntityResolution::PolicyStatement`，请参阅AWS CloudFormation 用户指南中的[AWS实体解析资源类型参考](#)。 `AWS::EntityResolution::MatchingWorkflow`，`AWS::EntityResolution::SchemaMapping`，`AWS::EntityResolution::IdMappingWorkflow`，`AWS::EntityResolution::IdNamespace`

可用模板如下：

- 匹配工作流程

创建一个`MatchingWorkflow`对象，用于存储要运行的数据处理作业的配置。

有关更多信息，请参阅以下主题：

《AWS CloudFormation 用户指南》中的 [AWS::EntityResolution::MatchingWorkflow](#)

[CreateMatchingWorkflow](#)在AWS Entity Resolution 数据匹配服务 API参考文献中

- 架构映射

创建架构映射，用于定义输入客户记录表的架构。

有关更多信息，请参阅以下主题：

《AWS CloudFormation 用户指南》中的 [AWS::EntityResolution::SchemaMapping](#)

[CreateSchemaMapping](#)在AWS Entity Resolution 数据匹配服务 API参考文献中

- 身份映射工作流程

创建一个IdMappingWorkflow对象，用于存储要运行的数据处理作业的配置。

有关更多信息，请参阅以下主题：

《AWS CloudFormation 用户指南》中的 [AWS::EntityResolution::IdMappingWorkflow](#)

[CreateIdMappingWorkflow](#)在AWS Entity Resolution 数据匹配服务 API参考文献中

- ID 命名空间

创建一个IdNamespace对象，该对象存储解释数据集及其使用方法的元数据。

有关更多信息，请参阅以下主题：

《AWS CloudFormation 用户指南》中的 [AWS::EntityResolution::IdNamespace](#)

[CreateIdNamespace](#)在AWS Entity Resolution 数据匹配服务 API参考文献中

- PolicyStatement

创建一个 PolicyStatement 对象。

有关更多信息，请参阅以下主题：

《AWS CloudFormation 用户指南》中的 [AWS::EntityResolution::PolicyStatement](#)

[AddPolicyStatement](#)在AWS Entity Resolution 数据匹配服务 API参考文献中

了解更多关于 AWS CloudFormation

要了解更多信息 AWS CloudFormation，请参阅以下资源：

了解更多信息 AWS CloudFormation

- [AWS CloudFormation](#)
- [AWS CloudFormation 用户指南](#)
- [AWS CloudFormation API参考](#)
- [AWS CloudFormation 命令行界面用户指南](#)

的配额 AWS Entity Resolution 数据匹配服务

您的每个配额 AWS 账户 都有默认配额，以前称为限制 AWS 服务。除非另有说明，否则，每个配额都特定于区域。您可以申请增加某些配额，但无法增加其他配额。

要查看的配额 AWS Entity Resolution 数据匹配服务，请打开 [Service Quotas 控制台](#)。在导航窗格中，选择 AWS 服务，然后选择 AWS Entity Resolution 数据匹配服务。

要请求提高配额，请参阅《Service Quotas 用户指南》中的[请求提高配额](#)。如果配额在服务配额中尚不可用，请使用[提高限制表格](#)。

您的 AWS 账户 配额与以下有关 AWS Entity Resolution 数据匹配服务。

名称	默认值	可调整	描述
并发 ID 映射作业	1	否	当前 AWS 区域可以同时处理的 ID 映射作业的最大数量。
并发匹配作业	1	否	当前 AWS 区域可以同时处理的最大匹配任务数。
并发提供商服务匹配作业	1	否	当前 AWS 区域可以同时处理的提供者服务匹配任务的最大数量。
数据输入	20	否	要在匹配流程中使用此列表。每个输入对应于 AWS Glue 输入数据表中的一列，其中包含列名和 AWS Entity Resolution 数据匹配服务 用于匹配目的的其他信息。输入必须包含一个唯一 ID 以及至少一个额外的输入字段。
数据输出	750	否	这是对象列表，每个 OutputAttribute 对象都有“名称”和“哈希”字段。这些对象中的每一个都表示要包含在 AWS Glue 输出表中的一列，以及您是否要对该列中的值进行哈希处理。
数据架构	25	否	数据架构输入字段的最大数量。

名称	默认值	可调整	描述
ID 映射工作流程	10	是	当前您可以在此处创建的 ID 映射工作流程的最 AWS 账户 大数量 AWS 区域。
ID 命名空间	10	是	当前您可以在此 AWS 账户 处创建的最大 ID 命名空间数。AWS 区域
比赛 ID	500	否	每个工作负载可合并到一个 matchID 下的最大记录数。
匹配规则	15	否	对于基于规则的匹配，这是生成匹配记录集所应用的规则编号。这是将包含在输出中的匹配工作流程元数据的一部分。
匹配工作流程	10	是	匹配流程的最大数量。
每个流程的规则数量	15	否	每个匹配流程的最大规则数。
GetMatchId API 请求的速率	50	是	每秒 GetCustomerID API 请求的最大数量。
架构映射	50	是	您可以在当前 AWS 区域的此账户中创建的最大架构映射数。
每个规则集都有唯一的匹配密钥	15	否	每个规则集的唯一匹配密钥的最大数量。匹配键指示 AWS Entity Resolution 数据匹配服务 哪些输入字段应被视为相似数据，哪些应被视为不同的数据。这有助于 AWS Entity Resolution 数据匹配服务 自动配置基于规则的匹配规则，并比较存储在不同输入字段中的相似数据。

API 节流配额

资源	默认值	描述
GetMatchId 请求速率	50 TPS	每秒 GetMatchId API 调用的最大次数。

《AWS Entity Resolution 数据匹配服务 用户指南》的文档历史记录

下表描述了文档版本 AWS Entity Resolution 数据匹配服务。

要获得有关本文档更新的通知，您可以订阅该订阅 RSS Feed。要订阅RSS更新，您必须为正在使用的浏览器启用RSS插件。

变更	说明	日期
提供商集成	仅文档更新。客户可以学习如何作为提供商服务与集成 AWS Entity Resolution 数据匹配服务。	2024 年 8 月 8 日
身份映射工作流程-更新	现在，客户可以使用匹配规则在 ID 映射工作流程中转换第三方数据。	2024 年 7 月 23 日
匹配工作流程-更新	客户现在可以从基于规则或基于机器学习的匹配工作流程中删除记录，以帮助遵守数据管理法规。	2024 年 4 月 8 日
身份映射工作流程-更新	客户现在可以跨多个使用身份映射工作流程 AWS 账户。	2024 年 4 月 2 日
AWS CloudFormation 资源-新的和更新的资源	AWSEntity Resolution 添加了以下资源：AWS::EntityResolution::IdNamespace AWS::EntityResolution::PolicyStatement 和，并更新了以下资源：AWS::EntityResolution::IdMappingWorkflow。	2024 年 4 月 2 日

查找比赛编号	现在，客户可以为已处理的基于规则的工作流程找到相应的 Match ID 和关联规则。	2024 年 3 月 25 日
匹配工作流程-更新	AWS Entity Resolution 数据匹配服务 现在支持在 PII 基于 LiveRamp 提供者服务的匹配工作流程中进行基于 RAMPID 分配。	2024 年 2 月 12 日
AWS PrivateLink	AWS Entity Resolution 数据匹配服务 现在支持额外的数据安全性 AWS PrivateLink，可帮助客户私密访问托管在上的服务 AWS。	2023 年 10 月 20 日
AWS CloudFormation 资源-新的和更新的资源	AWS Entity Resolution 数据匹配服务 添加了以下资源：AWS::EntityResolution:IdMappingWorkflow 并更新了以下资源：AWS::EntityResolution::MatchingWorkflow 和 AWS::EntityResolution::Schemamapping。	2023 年 10 月 19 日
更新现有策略	AWS Entity Resolution Console Full Access 托管策略中添加了以下新权限：ADXReadAccess 和 ManageEventBridgeRules。	2023 年 10 月 16 日
架构映射-更新	客户现在可以编辑和更新现有的数据架构。	2023 年 10 月 16 日

匹配工作流程-更新	客户现在可以选择首选的数据提供商服务来帮助匹配和关联他们的数据。	2023 年 10 月 16 日
身份映射工作流程	客户可以使用此新工作流程来指定 ID 映射详细信息、选择所需的 ID 映射方法以及指定数据输入和输出字段。	2023 年 10 月 16 日
AWS CloudFormation 整合	AWS Entity Resolution 数据匹配服务 现在与集成 AWS CloudFormation。	2023 年 8 月 24 日
AWS 托管策略更新-新策略	AWS Entity Resolution 数据匹配服务 添加了两个新的托管策略。	2023 年 8 月 18 日
初始版本	《AWS Entity Resolution 数据匹配服务 用户指南》的初始版本	2023 年 7 月 26 日

AWS Entity Resolution 数据匹配服务 词汇表

亚马逊资源名称 (ARN)

AWS 资源的唯一标识符。ARNs 当您需要在所有内容中明确指定资源时，例如在 AWS Entity Resolution 数据匹配服务 策略 AWS Entity Resolution 数据匹配服务、Amazon Relational Database Service (AmazonRDS) 标签和 API 调用中，则需要使用此选项。

自动处理

匹配 workflow 作业的处理节奏选项，当您的数据输入发生变化时，它可以自动运行。

此选项仅适用于[基于规则的匹配](#)。

默认情况下，匹配 workflow 作业的处理节奏设置为“[手动](#)”，这样便可以按需运行。您可以将自动处理设置为在数据输入发生变化时自动运行匹配的工作流程作业。这样可以保留匹配的工作流程输出 up-to-date。

AWS KMS key ARN

这是用于静态加密的 AWS KMS Amazon 资源名称 (ARN)。如果未提供，则系统将使用 AWS Entity Resolution 数据匹配服务 托管 KMS 密钥。

明文

未受加密保护的数据。

置信度 (ConfidenceLevel)

对于 ML 匹配，这是 ML 识别匹配的记录集 AWS Entity Resolution 数据匹配服务 时所应用的置信水平。这是将包含在输出中的[匹配 workflow 元数据](#)的一部分。

解密

将加密数据转换回其原始形式的过程。只有获得密钥，才能进行解密。

加密

使用称为密钥的机密值将数据编码成看似随机的形式的过程。如果无法访问密钥，就无法确定原始明文。

组名

组名引用整个输入字段组，可以帮助您将已解析的数据分组在一起以进行匹配。

例如，如果有三个输入字段：**first_name**、**middle_name**、**last_name**、和 **suffix**，则可以通过在组名中输入匹配和输出来将它们分组在一起。**full_name**

哈希

哈希意味着应用一种加密算法，该算法会生成不可逆且唯一的固定大小的字符串，称为哈希。AWS Entity Resolution 数据匹配服务 使用安全哈希算法 256 位 (SHA256) 哈希协议，并将输出 32 字节的字符串。在中 AWS Entity Resolution 数据匹配服务，您可以选择是否对输出中的数据值进行哈希处理。

哈希协议 (Hashing Protocol)

AWS Entity Resolution 数据匹配服务 使用安全哈希算法 256 位 (SHA256) 哈希协议，并将输出 32 字节的字符串。这是将包含在输出中的 [匹配工作流程元数据](#) 的一部分。

身份映射方法

您希望如何执行 ID 映射。

有两种 ID 映射方法：

- 基于规则-在 ID 映射工作流程中使用匹配规则将第一方数据从源转换为目标的方法。
- 提供者服务-在 ID 映射工作流程中，使用提供者服务将第三方编码的数据从源转换为目标的方法。

AWS Entity Resolution 数据匹配服务 目前支持 LiveRamp 作为基于提供商服务的身份映射方法。您必须订阅 LiveRamp 直通 AWS Data Exchange 才能使用此方法。有关更多信息，请参阅 [步骤 1：在上订阅提供商服务 AWS Data Exchange](#)。

身份映射工作流程

一种数据处理作业，它根据指定的 ID 映射方法将输入数据源中的数据映射到输入数据目标。它生成一个 ID 映射表。此工作流程要求您指定 [ID 映射方法](#) 以及要从源转换为目标的输入数据。

您可以将 ID 映射工作流程设置为自己运行 AWS 账户 或跨两个运行 AWS 账户。

ID 命名空间

中的一种资源 AWS Entity Resolution 数据匹配服务 ，其中包含解释多个数据集 AWS 账户 以及如何在 [ID 映射工作流程](#) 中使用这些数据集的元数据。

ID 命名空间有两种类型：SOURCE 和 TARGET。SOURCE 包含将在 ID 映射工作流程中处理的源数据的配置。TARGET 包含所有源都将解析为目标数据的配置。要定义要跨两个集合解析的输入数据 AWS 账户，请创建一个 ID 命名空间源和一个 ID 命名空间目标，以将数据从一个集合 (SOURCE) 转换为另一个集合 (TARGET)。

在您和其他成员创建 ID 命名空间并运行 ID 映射工作流程后，您可以加入协作，在 AWS Clean Rooms ID 映射表上运行多表联接，并分析数据。

有关更多信息，请参阅 [用户指南。AWS Clean Rooms](#)

输入字段

输入字段对应于 AWS Glue 输入数据表中的列名。

输入源 ARN (InputSourceARN)

为 AWS Glue 表输入生成的 Amazon 资源名称 (ARN)。这是将包含在输出中的 [匹配工作流程元数据](#) 的一部分。

输入类型

输入数据的类型。您可以从预先配置的值列表中进行选择，例如姓名、地址、电话号码或电子邮件地址。输入类型告诉您呈现的是 AWS Entity Resolution 数据匹配服务 哪种数据，从而可以对其进行正确分类和标准化。

基于机器学习的匹配

基于机器学习的匹配（机器学习匹配）可在您的数据中查找可能不完整或可能看起来不完全相同的匹配项。机器学习匹配是一个预设过程，它将尝试匹配您输入的所有数据的记录。机器学习匹配返回每个[匹配数据集的匹配 ID](#)和[置信度](#)。

手动处理

匹配 workflow 作业的处理节奏选项，使其能够按需运行。

此选项是默认设置的，可用于[基于规则的匹配和基于机器学习的匹配](#)。

多对多匹配

Many-to-many matching 会比较相似数据的多个实例。已分配相同匹配键的输入字段中的值将相互匹配，无论它们位于同一个输入字段还是不同的输入字段中。

例如，您可能有多个电话号码输入字段，例如 mobile_phone 和 home_phone，它们具有相同的匹配键“Phone”。使用 many-to-many matching 将 mobile_phone 输入字段中的数据与输入字段中的数据以及 mobile_phone 输入字段中的 home_phone 数据进行比较。

匹配规则通过（或）运算评估具有相同匹配键的多个输入字段中的数据，而 one-to-many 匹配则比较多个输入字段中的值。这意味着，如果两条记录之间有任何组合 mobile_phone 或 home_phone 匹配，“电话”匹配键将返回匹配项。对于匹配键“Phone”来查找匹配项，Record One mobile_phone = Record Two mobile_phone 或 Record One mobile_phone = Record Two home_phone 或 Record One home_phone = Record Two home_phone 或 Record One home_phone = Record Two mobile_phone。

匹配 ID (matchID)

对于基于规则的匹配和 ML 匹配，这是由每个匹配的记录集生成 AWS Entity Resolution 数据匹配服务并应用于的 ID。这是将包含在输出中的[匹配 workflow 元数据](#)的一部分。

匹配密钥 (MatchKey)

Match key 指示将 AWS Entity Resolution 数据匹配服务 哪些输入字段视为相似数据，哪些输入字段应视为不同的数据。这有助于 AWS Entity Resolution 数据匹配服务 自动配置基于规则的匹配规则，并比较存储在不同输入字段中的相似数据。

如果您想将数据中的多种类型的电话号码信息（例如mobile_phone输入字段和home_phone输入字段）进行比较，则可以为它们提供匹配键“Phone”。然后，可以将基于规则的匹配配置为在所有输入字段中使用“或”语句与“电话”匹配键比较数据（参见“[匹配工作流程](#)”部分中的“[一对一匹配](#)”和“[多对多匹配](#)”定义）。

如果您希望基于规则的匹配完全分开考虑不同类型的电话号码信息，则可以创建更具体的匹配键，例如“Mobile_Phone”和“Home_Phone”。然后，在设置匹配工作流程时，您可以指定在基于规则的匹配中如何使用每个电话匹配键。

如果没有 MatchKey 为特定的输入字段指定“否”，则该字段不能用于匹配，但可以执行匹配工作流程，并且可以根据需要输出。

匹配密钥名称

分配给匹配密钥的名称。

匹配规则 (MatchRule)

对于基于规则的匹配，这是生成匹配记录集所应用的规则编号。这是将包含在输出中的[匹配工作流程元数据](#)的一部分。

匹配

组合和比较来自不同输入字段、表或数据库的数据，并根据满足某些匹配条件（例如，通过匹配规则或模型）确定其中哪个相似或“匹配”的过程。

匹配工作流程

您为指定要匹配的输入数据而设置的过程以及应如何执行匹配。

匹配的工作流程描述

对您可能选择输入的匹配工作流程的可选描述。如果您创建多个工作流程，描述可以帮助您区分匹配的工作流程。

匹配工作流程名称

您指定的匹配工作流程的名称。

Note

匹配的工作流程名称必须是唯一的。它们的名称不能相同，否则将返回错误。

匹配工作流程元数据

在匹配的工作流作业 AWS Entity Resolution 数据匹配服务 期间生成和输出的信息。输出时需要此信息。

标准化 (ApplyNormalization)

选择是否按照架构中的定义对输入数据进行标准化。标准化通过删除多余的空格和特殊字符并标准化为小写格式来标准化数据。

例如，如果输入字段的输入类型为PHONE_NUMBER，并且输入表中的值格式为(123) 456-7890，则AWS Entity Resolution 数据匹配服务 会将值标准化为1234567890。

以下各节描述了标准化规则。

主题

- [名称](#)
- [Email](#)
- [Phone](#)
- [地址](#)
- [经过哈希处理](#)
- [来源_ID](#)

名称

- TRIM= 修剪前导和尾随的空格
- LOWERCASE= 小写所有字母字符
- CONVERT_ACCENT = 将带有重音符号的字母转换为普通字母
- REMOVE_ALL_NON_ALPHA = 移除所有非字母字符 [a-zA-Z]

Email

- TRIM= 修剪前导和尾随的空格
- LOWERCASE= 小写所有字母字符
- CONVERT_ACCENT = 将带有重音符号的字母转换为普通字母
- REMOVE__ ALL NON EMAIL _ CHARS = 移除所有 non-alpha-numeric 字符 [a-zA-Z0-9] 和 [.@-]

Phone

- TRIM= 修剪前导和尾随的空格
- REMOVE_ ALL NON _ NUMERIC = 移除所有非数字字符 [0-9]
- REMOVE_ ALL LEADING _ ZEROES = 移除所有前导零

地址

- TRIM= 修剪前导和尾随的空格
- LOWERCASE= 小写所有字母字符
- CONVERT_ACCENT = 将带有重音符号的字母转换为普通字母
- REMOVE_ ALL NON _ ALPHA = 移除所有非字母字符 [a-zA-Z]
- RENAME_ WORDS 使用 ADDRESS _ RENAME WORD _ MAP = 将地址字符串中的单词替换为 [ADDRESS__ RENAME WORD__](#) 中的单词 MAP
- RENAME_ DELIMITERS 使用 ADDRESS _ RENAME DELIMITER _ MAP = 将地址字符串中的分隔符替换为 [ADDRESS__ RENAME _ DELIMITER](#) 中的字符串 MAP
- RENAME_ DIRECTIONS 使用 ADDRESS _ RENAME DIRECTION _ MAP = 将地址字符串中的分隔符替换为 [ADDRESS__ RENAME _ DIRECTION](#) 中的字符串 MAP
- RENAME_ NUMBERS 使用 ADDRESS _ RENAME NUMBER _ MAP = 将地址字符串中的数字替换为 [ADDRESS__ RENAME NUMBER__](#) 中的字符串 MAP
- RENAME_ SPECIAL _ CHARS 使用 ADDRESS __ RENAME _ SPECIAL CHAR _ MAP = 将地址字符串中的特殊字符替换为 [ADDRESS__ __ RENAME SPECIAL CHAR__](#) 中的字符串 MAP

ADDRESS_RENAME_WORD_MAP

这些是标准化地址字符串时将重命名的单词。

```
"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"
```

ADDRESS_RENAME_DELIMITER_MAP

这些是标准化地址字符串时将重命名的分隔符。

```
"," : " ",
"." : " ",
"[" : " ",
"]" : " ",
"/" : " ",
"-" : " ",
"#" : " number "
```

ADDRESS_RENAME_DIRECTION_MAP

这些是在标准化地址字符串时将重命名的方向标识符。

```
"east": "e",  
"north": "n",  
"south": "s",  
"west": "w",  
"northeast": "ne",  
"northwest": "nw",  
"southeast": "se",  
"southwest": "sw"
```

ADDRESS_RENAME_NUMBER_MAP

这些是在标准化地址字符串时将重命名的数字字符串。

```
"número": "number",  
"numero": "number",  
"no": "number",  
"núm": "number",  
"num": "number"
```

ADDRESS_RENAME_SPECIAL_CHAR_MAP

这些是在标准化地址字符串时将被重命名的特殊字符串。

```
"ß": "ss",  
"ä": "ae",  
"ö": "oe",  
"ü": "ue",  
"ø": "o",  
"æ": "ae"
```

经过哈希处理

- TRIM= 修剪前导和尾随的空格

来源_ID

- TRIM= 修剪前导和尾随的空格

一对一匹配

One-to-one 匹配比较相似数据的单个实例。具有相同匹配键和相同输入字段中的值的输入字段将相互匹配。

例如，您可能有多个电话号码输入字段，例如mobile_phone和home_phone，它们具有相同的匹配键“Phone”。使用 one-to-one 匹配将输入字段中的数据与mobile_phone输入字段中的数据进行比较，并将mobile_phone输入字段中的数据与home_phone输入字段中的home_phone数据进行比较。mobile_phone输入字段中的数据不会与home_phone输入字段中的数据进行比较。

匹配规则通过 (或) 运算评估具有相同匹配键的多个输入字段中的数据，而 one-to-many 匹配则比较单个输入字段中的值。这意味着，如果mobile_phone或在两条记录之间匹home_phone配，“电话”匹配键将返回匹配项。要使用匹配键“电话”来查找匹配项，Record One mobile_phone = Record Two mobile_phone或者Record One home_phone = Record Two home_phone。

匹配规则通过 (和) 运算评估具有不同匹配键的输入字段中的数据。如果您希望基于规则的匹配完全分开考虑不同类型的电话号码信息，则可以创建更具体的匹配键，例如“mobile_phone”和“home_phone”。如果您想在规则中同时使用两个匹配键来查找匹配项，Record One mobile_phone = Record Two mobile_phoneANDRecord One home_phone = Record Two home_phone。

输出

对象列表，每个OutputAttribute对象都有“名称”和“哈希”字段。这些对象中的每一个都表示要包含在AWS Glue 输出表中的一列，以及您是否要对该列中的值进行哈希处理。

outputs3Path

AWS Entity Resolution 数据匹配服务 将输出表写入的 S3 目标。

OutputSourceConfig

对象列表，每个 OutputSource 对象都有字段 outputs3Path、ApplyNormalization和 Output t。

基于提供商服务的匹配

基于提供商服务的匹配过程旨在将您的记录与首选数据服务提供商和许可数据集进行匹配、关联和增强。您必须通过 AWS Data Exchange 提供商服务订阅才能使用此匹配技术。

AWS Entity Resolution 数据匹配服务 目前已与以下数据服务提供商集成：

- LiveRamp
- TransUnion
- UID2.0

基于规则的匹配

基于规则的匹配是旨在查找精确匹配项的过程。基于规则的匹配是一组分层的瀑布匹配规则 AWS Entity Resolution 数据匹配服务，由您根据输入的数据建议并完全由您配置。规则条件中提供的所有匹配键都必须完全匹配，才能将比较的数据声明为匹配项并输出关联的元数据。基于规则的匹配会为每个[匹配的数据集返回一个匹配 ID](#) 和一个规则编号。

我们建议定义能够唯一标识实体的规则。对规则进行排序，先找到更精确的匹配项。

例如，假设你有两个规则，规则 1 和规则 2。

这些规则具有以下匹配密钥：

- 规则 1 包括全名和地址
- 规则 2 包括全名、地址和电话

由于规则 1 首先运行，因此规则 2 不会找到任何匹配项，因为规则 1 本来可以找到所有匹配项。

要查找按电话区分的匹配项，请重新排列规则，如下所示：

- 规则 2 包括全名、地址和电话
- 规则 1 包括全名和地址

架构

该术语用于定义一组数据的组织和连接方式的结构或布局。

架构描述

您可以选择输入的架构的可选描述。如果您创建了多个架构映射，则描述可以帮助您区分架构映射。

架构名称

架构的名称。

Note

架构名称必须是唯一的。它们的名称不能相同，否则将返回错误。

架构映射

中的架构映射 [AWS Entity Resolution 数据匹配服务](#) 是告诉您 [AWS Entity Resolution 数据匹配服务](#) 如何解释数据以进行匹配的过程。您可以定义要 [AWS Entity Resolution 数据匹配服务](#) 读入匹配工作流程的输入数据表的架构。

架构映射 ARN

为 [架构映射](#) 生成的 Amazon 资源名称 (ARN)。

唯一标识

您指定的唯一标识符，必须将其分配给 [AWS Entity Resolution 数据匹配服务](#) 读取的每一行输入数据。

Example

例如，**Primary_key**、**Row_ID** 或 **Record_ID**。

“唯一 ID” 列为必填字段。

唯一 ID 必须是单个表中的唯一标识符。

在不同的表中，唯一 ID 可以有重复的值。

运行 [匹配工作流程](#) 时，如果唯一 ID 满足以下条件，则该记录将被拒绝：

- 未指定
- 在同一个表中不是唯一的
- 在不同源的属性名称方面存在重叠。

- 超过 38 个字符 (仅限基于规则的匹配工作流程)

本文属于机器翻译版本。若本译文内容与英语原文存在差异，则一律以英文原文为准。