



Informatica®
10.4.0

Analyst 工具指南

Informatica Analyst 工具指南

10.4.0

2019 年 12 月

© 版权所有 Informatica LLC 2009, 2020

本软件和文档仅根据包含使用与披露限制的单独许可协议提供。未事先征得 Informatica LLC 同意，不得以任何形式、通过任何手段（电子、影印、录制或其他手段）复制或传播本文档的任何部分。

美国政府权利交付给美国政府客户的程序、软件、数据库及相关文档和技术数据是指适用的联邦采购条例和政府机构特定补充条例中定义的"商业计算机软件"或"商业技术数据"。因此，使用、复制、披露、修改和改编应遵循适用的政府合同中规定的限制和许可条款、政府合同条款的适用范围以及 FAR 52.227-19 商用计算机软件许可中规定的额外权利。

Informatica 和 Informatica 标志是 Informatica LLC 在美国和世界其他许多司法管辖区的商标或注册商标。欲获得 Informatica 商标的最新列表，请访问 <https://www.informatica.com/trademarks.html>。其他公司和产品名称可能是其各自所有者的商业名称或商标。

本软件和/或文档中的若干部分受第三方版权约束。所需的第三方声明随产品一起提供。

本文档中的信息如有更改，恕不另行通知。如发现本文档中有什么问题，请通过以下电子邮件地址向我们报告：infa_documentation@informatica.com。

Informatica 产品根据对应协议的条款和条件进行担保。INFORMATICA 按"原样"提供本文档中的信息，无任何明示或暗示的担保，包括但不限于任何适销性和特定用途适用性担保，也没有任何非侵权担保或条件。

发布日期: 2020-03-02

目录

前言	6
Informatica 资源	6
Informatica Network	6
Informatica 知识库	6
Informatica 文档	6
Informatica 产品可用性矩阵	7
Informatica Velocity	7
Informatica Marketplace	7
Informatica 全球客户支持部门	7
第 1 章： Informatica Analyst 简介.....	8
Informatica Analyst 概览	8
Informatica Analyst 界面	8
Informatica Analyst 表头	9
Informatica Analyst 工作区	10
键盘快捷方式	11
Informatica Analyst 资产	11
登录到 Analyst 工具	12
第 2 章： 库工作区.....	13
库工作区概览	13
访问库工作区	13
库任务	14
创建用户定义的标记	14
分配和删除标记	15
第 3 章： 连接工作区.....	16
连接工作区概览	16
IBM DB2 Connection Properties	17
JDBC 连接属性	20
MS SQL Server 连接属性	22
ODBC 连接属性	26
Oracle 连接属性	27
Hive 连接属性	29
HDFS 连接属性	32
数据库连接中的标识符属性	33
常规标识符	34
分隔标识符	34
标识符属性	34
搜索数据库连接	36

创建数据库连接.....	36
编辑数据库连接.....	37
删除数据库连接.....	37
第 4 章： 作业状态工作区.....	38
作业状态工作区概览.....	38
访问作业状态工作区.....	38
作业属性.....	39
监视作业.....	40
第 5 章： 项目工作区.....	41
项目工作区概览.....	41
分配项目工作区.....	41
管理项目和文件夹.....	42
项目安全性.....	43
项目权限.....	43
分配项目的直接权限.....	43
查看项目的权限.....	44
第 6 章： 模型存储库.....	45
模型存储库概览.....	45
Informatica Analyst 资产.....	45
存储库资产锁定.....	46
资产锁定管理的规则和准则.....	46
使用受版本控制的对象进行基于团队的开发.....	47
受版本控制的资产管理.....	47
第 7 章： 数据对象.....	49
数据对象概览.....	49
平面文件数据对象.....	49
导入平面文件数据对象.....	50
平面文件选项.....	50
平面文件数据类型.....	50
日期时间数据类型.....	51
添加带分隔符的平面文件.....	52
添加固定宽度的平面文件.....	53
平面文件的规则和准则.....	53
表数据对象.....	54
添加表.....	54
表的规则和准则.....	54
同步数据对象.....	55
在 Informatica Analyst 中同步平面文件数据对象.....	55
在 Informatica Analyst 中同步关系数据对象.....	57

查看数据对象.	57
编辑数据对象.	58
第 8 章： 搜索.	59
搜索概览.	59
搜索结果.	59
搜索查询.	60
搜索属性.	60
附录 A： 配置 Web 浏览器.	61
配置 Web 浏览器.	61
索引.	62

前言

要了解如何在 Informatica Analyst (Analyst 工具) 中发现、定义和查看业务逻辑，请阅读《*Informatica® Analyst 工具指南*》。了解各种不同的工作区，您可以在这些工作区中管理资产或使用资产来支持业务运营。

Informatica 资源

Informatica 通过 Informatica Network 和其他在线门户为您提供一系列产品资源。使用这些资源，可以充分利用 Informatica 产品和解决方案，并向其他 Informatica 用户和主题专家学习。

Informatica Network

在 Informatica Network 中可以获得许多资源，包括 Informatica 知识库和 Informatica 全球客户支持。要进入 Informatica Network，请访问 <https://network.informatica.com>。

作为 Informatica Network 成员，您可以选择以下服务：

- 在知识库中搜索产品资源。
- 查看产品可用性信息。
- 创建并检查您的支持案例。
- 查找当地的 Informatica 用户组网络并与您的伙伴进行协作。

Informatica 知识库

使用 Informatica 知识库可查找产品资源，例如操作方法文章、最佳实践、视频教程以及常见问题的答案。

要搜索知识库，请访问 <https://search.informatica.com>。如果您对知识库有任何疑问、意见或建议，请与 Informatica 知识库团队联系，电子邮件地址为 KB_Feedback@informatica.com。

Informatica 文档

使用 Informatica 文档门户可浏览大量当前与最近产品版本的文档库。要浏览文档门户，请访问 <https://docs.informatica.com>。

如果您对产品文档有任何疑问、意见或建议，请与 Informatica 文档团队联系，电子邮件地址为 infa_documentation@informatica.com。

Informatica 产品可用性矩阵

产品可用性矩阵 (PAM) 指明了产品版本支持的操作系统版本、数据库以及数据源和目标的类型。您可以在以下网址中浏览 Informatica PAM:

<https://network.informatica.com/community/informatica-network/product-availability-matrices>。

Informatica Velocity

Informatica Velocity 是由 Informatica 专业服务根据数百个数据管理项目的实际经验所开发出来的，其中汇集了大量使用技巧和最佳实践。Informatica Velocity 代表了 Informatica 顾问的集体知识，这些顾问与世界各地的组织合作，共同计划、开发、部署和维护成功的数据管理解决方案。

您可以在以下网址中找到 Informatica Velocity 资源：<http://velocity.informatica.com>。如果您对 Informatica Velocity 有任何疑问、意见或建议，请通过 ips@informatica.com 与 Informatica 专业服务联系。

Informatica Marketplace

Informatica Marketplace 是一个论坛，该论坛中提供的解决方案可扩展和增强您的 Informatica 实施。利用 Informatica 开发人员和合作伙伴在 Marketplace 中提供的数以百计的解决方案，可提高您的工作效率并加快项目实施时间。您可以在以下网址中找到 Informatica Marketplace：<https://marketplace.informatica.com>。

Informatica 全球客户支持部门

您可以通过电话或 Informatica Network 与全球支持中心联系。

要查找您当地的 Informatica 全球客户支持部门电话号码，请访问 Informatica 网站，链接为：<https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>。

要在 Informatica Network 上查找在线支持资源，请访问 <https://network.informatica.com>，然后选择 eSupport 选项。

第 1 章

Informatica Analyst 简介

本章包括以下主题：

- [Informatica Analyst 概览, 8](#)
- [Informatica Analyst 界面, 8](#)
- [Informatica Analyst 资产, 11](#)
- [登录到 Analyst 工具, 12](#)

Informatica Analyst 概览

Informatica Analyst（Analyst 工具）是可用于多个 Informatica 产品的 Web 客户端工具，企业用户将其用于公司内的项目协作。例如，业务分析人员可以使用 Analyst 工具在组织内协作完成数据集成项目。

使用 Analyst 工具可发现、定义和查看组织中项目的业务逻辑。可以在 Analyst 工具中执行的任务取决于 Informatica 产品的许可证以及执行任务的特权。

根据您的组织具有的许可证，您可以使用 Analyst 工具来执行以下任务：

- 定义业务词汇表、术语和策略来维护组织中数据资产的标准化定义。
- 执行数据发现来查找数据源的内容、质量和结构，以及监视数据质量趋势。
- 定义数据集成逻辑以及对项目进行协作来加快项目交付。
- 定义和管理规则来验证数据对于业务策略的遵从性。
- 查看并解决数据质量问题以查找并修复组织中的数据质量问题。

分析服务会管理 Analyst 工具。Analyst 工具在模型存储库中存储项目、文件夹和数据对象。Analyst 工具会连接到模型存储库数据库以创建、更新和删除项目、文件夹以及数据对象。

创建、更新和删除项目、文件夹以及数据对象时，管理 Analyst 工具的分析服务必须与存储这些项目、文件夹和数据对象的模型存储库相关联。此外，使用 Developer tool 客户端打开 Analyst 工具中的对象时，分析服务必须与存储这些对象的模型存储库相关联。

Informatica Analyst 界面

使用 Analyst 工具的基于 Web 的界面对组织中的业务项目进行协作。

Analyst 工具界面具有表头和工作区。工作区是一个网页，在其中可根据通过 Analyst 工具中的选项卡访问的已许可功能执行任务。您还必须具有特权才能在工作区中执行任务。

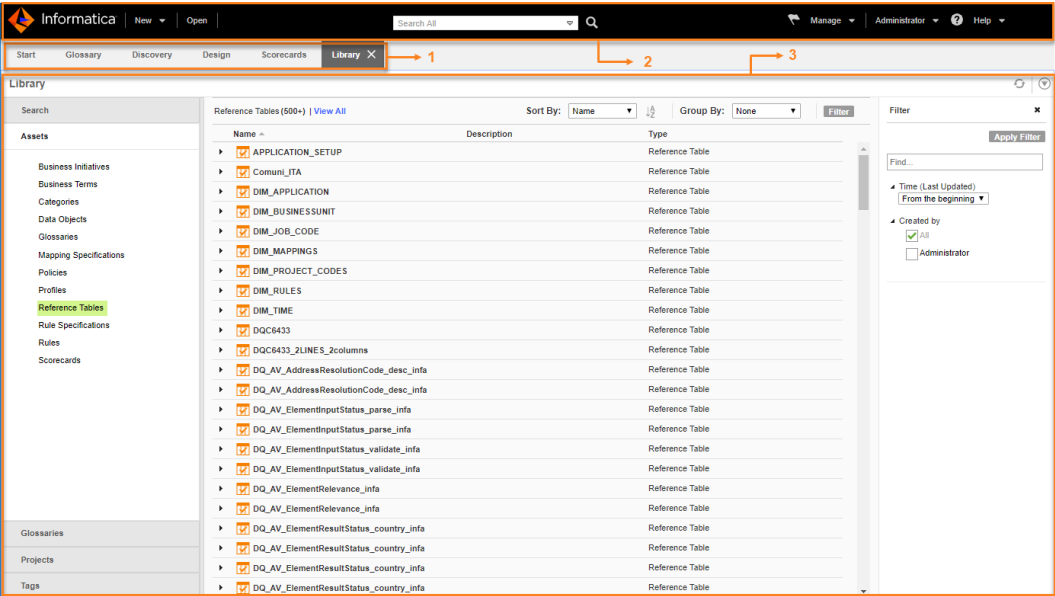
登录 Analyst 工具时，将显示**开始**工作区。您可以在 Analyst 工具界面中打开多个工作区。

例如，使用**发现**工作区分析源系统中数据和元数据的质量。您可以通过工作区选项卡或通过 Analyst 工具表头中的菜单访问某个工作区。

您可以使用某些工作区中的资产来执行任务，例如运行配置文件、创建业务规则或创建映射规范。资产是 Analyst 工具中某种类型的对象，可支持组织中的业务操作。

如果您具有相应许可证可使用业务词汇表，则可以查看业务词汇表资产的通知警告。在 Analyst 工具表头中查看通知警告。

下图显示了 Analyst 工具：



1. 工作区访问面板
2. 表头区域
3. 工作区

Informatica Analyst 表头

Analyst 工具表头显示在 Analyst 工具用户界面的顶部。

Analyst 工具具有以下表头项：

新建

在**词汇表、发现和设计**工作区中创建资产。

打开

打开**库**工作区。

通知警告

查看词汇表资产的通知。

管理

打开临时工作区和通知。您可以打开**连接、数据域、作业状态、项目**和**词汇表安全**工作区。

用户名

设置用户首选项以更改密码并注销 Analyst 工具。

帮助

访问当前工作区中的帮助。

Informatica Analyst 工作区

工作区是一个网页，可以根据许可证和特权进行访问。可以在工作区中执行任务。可以在一些工作区中管理资产或使用资产来执行任务。Analyst 工具具有永久工作区和临时工作区。

永久工作区始终可以通过工作区选项卡访问。您可以导航到其他工作区，但是无法关闭永久工作区。临时工作区可通过工作区选项卡访问。可以从 Analyst 工具表头或工作区中的访问面板打开临时工作区。当不需要某个临时工作区时，您可以从选项卡关闭该工作区。

永久工作区

Analyst 工具包含以下永久工作区：

启动

访问您有相应许可证可通过工作区上的访问面板访问的其他工作区。如果您有许可证来执行异常管理，则任务会显示在工作区的**我的任务**面板上。如果您必须以批准工作流程的审批者身份进行投票，则您挂起的任务将显示在**我的任务**面板中。

词汇表

定义和描述对您的组织非常重要的业务概念。您可以创建和管理业务术语、业务计划、类别、词汇表和策略。

发现

分析源系统中数据和元数据的质量。您可以创建和管理配置文件、平面文件数据对象和表数据对象。您可以查看和管理 Developer tool 对象，例如在模型存储库的项目中存储的 SAP 和大型机对象。

设计

帮助分析师和开发人员进行协作的设计业务逻辑。您可以创建和管理映射规范、引用表和规则规范。

结果卡

打开、编辑和运行通过配置文件结果创建的结果卡。您可以添加度量，在列上向下钻取，添加结果卡筛选器以及查看结果卡的趋势图表。

临时工作区

Analyst 工具包含以下临时工作区：

库

搜索模型存储库中的资产。您还可以在**库**工作区中查看元数据。打开某个资产时，该资产会在创建它的工作区中打开。

异常

查看和管理任务的异常记录数据。从**启动**工作区的**我的任务**面板中打开任务时，Analyst 工具会打开名为**异常**工作区的临时工作区。根据您要执行的任务类型查看重复记录群集或异常记录。查看有关您对任务中的记录所做更改的审计跟踪。

连接

创建和管理连接以导入关系数据对象、预览数据、运行配置文件以及运行映射规范。

数据域

创建、管理和删除数据域和数据域组。数据域介绍列数据的语义，例如社会保障号或电话号码。您可以将数据域分类在数据域组中，例如将社会保障号和电话号码分类在个人健康信息 (PHI) 数据域组中。

作业状态

监视 Analyst 工具作业的状态，例如对所有数据对象的数据预览以及对配置文件的向下钻取操作。

项目

创建和管理文件夹和项目，并分配项目的权限。

词汇表安全

管理业务词汇表用户的权限、特权和角色。

设置

选择操作系统配置文件以运行映射、配置文件和结果卡。操作系统配置文件是一种安全类型，数据集成服务使用此安全类型来隔离运行时用户环境。当 Informatica 管理员将多个操作系统配置文件分配给用户时，**设置**工作区可用。要打开**设置**工作区，请转到 Analyst 工具的用户名表头项，然后单击**设置**。

键盘快捷方式

您可以使用键盘快捷方式浏览并使用 Analyst 工具界面。

对象的导航顺序是从上到下、从左到右。

您可以对键盘快捷方式执行以下任务：

浏览各元素，然后选择一个元素。

按 Tab。

浏览工作区中的 portlet 和窗格。

按 Alt+P。

关闭临时工作区。

按 Ctrl+Shift+X。

Informatica Analyst 资产

您可以在一些工作区中管理资产。资产是一种用于在企业内支持业务操作的对象。

例如，配置文件是分析师可以创建用于发现数据源的内容、质量和结构的资产。

您可以创建以下类型的资产：

词汇表资产

在**词汇表**工作区中创建词汇表资产。您可以创建以下类型的词汇表资产：

- 业务术语。使用业务语言为组织中的企业用户定义相关概念的词语或短语。
- 业务计划。用于对词汇表资产集合进行批量更改的业务决策。
- 类别。业务术语和策略的描述性分类。
- 词汇表。类别、业务术语和策略的集合。
- 策略。管理与业务术语相关的业务实践的业务目的、进程或协议。

发现资产

在**发现**工作区中创建发现资产。您可以创建以下类型的发现资产：

- 配置文件。在单个数据源或多个数据源中发现列特征和数据域的单个数据对象配置文件或企业发现配置文件。
- 平面文件数据对象。基于平面文件的数据表示。
- 表数据对象。基于关系表的数据表示。

设计资产

在**设计**工作区中创建设计资产。您可以创建以下类型的设计资产：

- 映射规范。描述数据从源到目标的移动及转换的模板。
- 引用表。包含一组数据值的标准版本和替代版本的一种表。
- 规则规范。代表业务规则逻辑的一种对象。

结果卡资产

在**结果卡**工作区中打开结果卡资产。结果卡是配置文件中质量度量的图形表示形式。

登录到 Analyst 工具

使用 Analyst 工具 URL 可登录到 Analyst 工具。登录 Analyst 工具时，您需要指定 Informatica 登录名、密码及本地域或 LDAP 安全域。

1. 启动 Microsoft Internet Explorer 或 Google Chrome 浏览器。
2. 在“地址”字段中，输入 Analyst 工具的 URL：
`http[s]://<fully qualified host name>:<port number>/analyst/`
3. 如果域使用 LDAP 或本地身份验证，请在登录页面输入登录名和密码。
4. 选择**本地**或特定安全域的名称。

如果 Informatica 域使用 LDAP 或 Kerberos 身份验证，将显示“安全域”字段。如果您不知道用户帐户所属的安全域，请联系 Informatica 域管理员。

5. 单击**登录**。

Analyst 工具会在**启动**工作区上打开。

第 2 章

库工作区

本章包括以下主题：

- [库工作区概览, 13](#)
- [访问库工作区, 13](#)
- [库任务, 14](#)

库工作区概览

使用库工作区可浏览、搜索或筛选您有特权访问的资产集合。您可以浏览和查看资产、词汇表、项目和标记。

开始通过库导航器浏览模型存储库内容。库导航器包含的部分取决于您可能需要查询模型存储库内容的方式。例如，如果您要查找与一个词汇表关联的所有词汇表资产，则可以开始从词汇表部分进行浏览。

您可以通过库导航器按以下方式查找模型存储库内容：

- 按资产
- 按项目
- 按词汇表
- 按标记

打开一个部分时，您可以选择资产的类型，并查看资产列表中列出的资产。可以按资产属性对资产列表进行排序或分组，以便组织资产。不能按说明对资产列表进行排序。

使用库导航器执行发现搜索。发现搜索可查找资产及其与组织中其他资产的关系。您可以添加筛选器（包括搜索筛选器）以精简资产列表。可以按搜索结果中的资产说明进行排序。

您可以从资产列表打开资产。单击资产时，资产会在其对应的工作区中打开。可以编辑资产、查看历史记录、添加注释以及查看相关的资产。

访问库工作区

访问库工作区可查看和管理您有特权查看或管理的资产集合。

- 从 Analyst 工具表头中，单击**打开**。
库工作区将打开。

库任务

您可以管理有特权访问和执行库任务的资产集合。

可以执行以下库任务：

执行发现搜索。

发现搜索可查找资产及其与组织中其他资产的关系。例如，您需要查找所有客户信息在财务组织中的存放位置。执行发现搜索可针对客户搜索字符串查找符合搜索条件的数据对象。发现搜索结果包括与您在 Analyst 工具中搜索的数据对象相关联的资产。相关联的资产包括根据数据对象运行的配置文件、关联的结果卡以及业务术语。

有关详细信息，请参阅《Data Discovery 指南》。

查看资产。

从库导航器中的某个部分查找模型存储库内容时，Analyst 工具会在资产列表中显示列出的资产。例如，选择**数据对象**时，Analyst 工具会显示您有特权访问的数据对象的列表。

从库导航器中，单击**资产**部分并选择一个资产。您可以查看属于资产列表的一系列资产。

查看词汇表。

查看您有特权访问的词汇表。从**词汇表**部分中选择词汇表时，Analyst 工具会在资产列表中列出词汇表的内容，例如业务术语、类别或策略。

从库导航器中，单击**词汇表**部分并选择一个词汇表。可以在资产列表中查看词汇表资产。

查看项目。

查看项目和文件夹及其内容。选择项目或文件夹时，Analyst 工具会在资产列表中列出项目或文件夹的内容。

从库导航器中，单击**项目**部分并选择项目或文件夹。您可以在**资产**面板中查看项目或文件夹内容。

查看、添加或删除标记。

按系统定义的标记查看标记的业务术语，或按用户定义的标记查看资产。系统定义的标记根据业务术语的用法对其进行分组。您可以通过**标记**部分创建标记。您可以通过**项目**部分分配标记或从资产中删除标记。

搜索资产。

通过搜索字符串搜索资产或应用筛选器来搜索资产。输入筛选器属性可筛选搜索结果。

从**筛选器**面板中，在搜索框中输入搜索字符串或添加筛选器属性以筛选搜索结果。

创建用户定义的标记

创建用户定义的标记以根据业务使用情况对资产分组。

1. 从库导航器的**标记**部分中，右键单击**用户定义**并选择**新建标记**。
将显示**新建标记**对话框。
2. 输入名称和可选说明。
3. 单击**确定**。

分配和删除标记

为资产分配标记以根据业务使用情况对资产分组。您还可以在不再需要某个标记时，将其从资产中删除。

1. 从库导航器的**项目**部分中，选择某个项目。
2. 从资产列表中，右键单击某个资产并选择**标记**。
将显示**标记**对话框。
3. 选择添加或删除标记。
 - 要添加标记，请在**新建标记**面板中输入用户定义的标记名称，然后单击**添加**。
 - 要删除标记，请从**标记**面板中选择一个标记，然后单击**删除**。
4. 单击**确定**。

第 3 章

连接工作区

本章包括以下主题：

- [连接工作区概览, 16](#)
- [IBM DB2 Connection Properties, 17](#)
- [JDBC 连接属性, 20](#)
- [MS SQL Server 连接属性, 22](#)
- [ODBC 连接属性, 26](#)
- [Oracle 连接属性, 27](#)
- [Hive 连接属性, 29](#)
- [HDFS 连接属性, 32](#)
- [数据库连接中的标识符属性, 33](#)
- [搜索数据库连接, 36](#)
- [创建数据库连接, 36](#)
- [编辑数据库连接, 37](#)
- [删除数据库连接, 37](#)

连接工作区概览

使用**连接**工作区可查看、创建和管理连接。一个连接就是一个存储库对象，可用于在域配置存储库中定义一个连接。

创建连接以导入数据对象、预览数据及运行配置文件或映射。Analyst 工具在您导入数据对象时使用连接。数据集成服务在您预览数据、运行配置文件或运行映射时使用该连接。

Analyst 工具会将连接存储在域配置存储库中。您在 Analyst 工具中创建的任何连接都可在 Developer tool 或 Administrator 工具中使用。

您可以在**连接**工作区中执行以下任务：

- 搜索连接。
- 创建连接。
- 测试连接。
- 编辑连接。
- 删除连接。

在 Analyst 工具中可以创建以下类型的连接：

- IBM DB2
- Microsoft SQL Server
- ODBC
- Oracle
- Hive
- Hadoop 文件系统

您可以从 IBM DB2/zOS 连接浏览和导入表。但是，您必须在 Administrator 工具或 Developer tool 中创建 IBM DB2/zOS 连接。

IBM DB2 Connection Properties

Use an IBM DB2 connection to access IBM DB2. An IBM DB2 connection is a relational database connection. You can create and manage an IBM DB2 connection in the Administrator tool, the Developer tool, or the Analyst tool.

注意：连接属性的顺序可能会因查看属性所用的工具而异。

The following table describes DB2 connection properties:

Property	Description
Database Type	The database type.
Name	Name of the connection. The name is not case sensitive and must be unique within the domain. The name cannot exceed 128 characters, contain spaces, or contain the following special characters: ~ ` ! \$ % ^ & * () - + = { [] } \ : ; " ' < , > . ? /
ID	String that the Data Integration Service uses to identify the connection. The ID is not case sensitive. It must be 255 characters or less and must be unique in the domain. You cannot change this property after you create the connection. Default value is the connection name.
Description	The description of the connection. The description cannot exceed 765 characters.
User Name	The database user name.
Password	The password for the database user name.
Pass-through security enabled	Enables pass-through security for the connection. When you enable pass-through security for a connection, the domain uses the client user name and password to log into the corresponding database, instead of the credentials defined in the connection object.
Connection String for data access	The DB2 connection URL used to access metadata from the database. dbname Where dbname is the alias configured in the DB2 client.

Property	Description
Metadata Access Properties: Connection String	<p>Use the following metadata connection string URL:</p> <p>jdbc:informatica:db2://<host name>:<port>;DatabaseName=<database name></p> <p>When you import a table, by default, all tables are displayed under the default schema name. To view tables under a specific schema instead of the default schema, you can specify the schema name from which you want to import the table. Include the ischename parameter in the URL to specify the schema name. For example, use the following syntax to import a table from a specific schema:</p> <p>jdbc:informatica:db2://<host name>:<port>;DatabaseName=<database name>;ischename=<schema_name></p> <p>To search for a table in multiple schemas and import it, you can specify multiple schema names in the ischename parameter. The schema name is case sensitive. You cannot use special characters when you specify multiple schema names. Use the pipe () character to separate multiple schema names. For example, use the following syntax to search for a table in three schemas and import it:</p> <p>jdbc:informatica:db2://<host name>:<port>;DatabaseName=<database name>;ischename=<schema_name1> <schema_name2> <schema_name3></p> <p>When you specify multiple schema names, you must clear the Show Default Schema Only option to view the tables under the specified schema names.</p>

Property	Description
AdvancedJDBCSecurityOptions	<p>Database parameters for metadata access to a secure database. Informatica treats the value of the AdvancedJDBCSecurityOptions field as sensitive data and stores the parameter string encrypted.</p> <p>To connect to a secure database, include the following parameters:</p> <ul style="list-style-type: none"> - EncryptionMethod. Required. Indicates whether data is encrypted when transmitted over the network. This parameter must be set to SSL. - ValidateServerCertificate. Optional. Indicates whether Informatica validates the certificate that is sent by the database server. If this parameter is set to True, Informatica validates the certificate that is sent by the database server. If you specify the HostNameInCertificate parameter, Informatica also validates the host name in the certificate. If this parameter is set to false, Informatica does not validate the certificate that is sent by the database server. Informatica ignores any truststore information that you specify. - HostNameInCertificate. Optional. Host name of the machine that hosts the secure database. If you specify a host name, Informatica validates the host name included in the connection string against the host name in the SSL certificate. - cryptoProtocolVersion. Optional. If you enable TLS for the IBM DB2 instance, set the cryptoProtocolVersion parameter as follows: cryptoProtocolVersion=TLSv<version number>. For example, cryptoProtocolVersion=TLSv1.2 注意: The version number must be the same as the TLS version you configured for the server. - TrustStore. Required. Path and file name of the truststore file. 注意: If you configure SSL or TLS and specify only the file name, you must copy the truststore file to the Informatica installation directory. To test the connection and import metadata, copy the truststore file to the following directory: <Informatica client installation directory>/clients/DeveloperClient To run the mapping, copy the truststore file to the following directory: <Informatica server installation directory>/tomcat/bin 注意: If you configure SSL or TLS and specify only the file name, you must copy the truststore file to the following directory to test the connection: <Informatica server installation directory>/tomcat/bin - TrustStorePassword. Required. Password for the truststore file for the secure database. 注意: Informatica appends the secure JDBC parameters to the connection string. If you include the secure JDBC parameters directly to the connection string, do not enter any parameters in the AdvancedJDBCSecurityOptions field.
Data Access Properties: Connection String	<p>The connection string used to access data from the database. For IBM DB2 this is <database name></p>
Code Page	<p>The code page used to read from a source database or to write to a target database or file.</p>
Environment SQL	<p>SQL commands to set the database environment when you connect to the database. The Data Integration Service runs the connection environment SQL each time it connects to the database.</p>

Property	Description
Transaction SQL	SQL commands to set the database environment when you connect to the database. The Data Integration Service runs the transaction environment SQL at the beginning of each transaction.
Retry Period	This property is reserved for future use.
Tablespace	The tablespace name of the database.
SQL Identifier Character	<p>数据库用于在 SQL 查询中为分隔标识符加上引号的字符类型。可用的字符取决于数据库类型。</p> <p>如果数据库使用常规标识符，选择（无）。数据集成服务生成 SQL 查询时，该服务不会将带分隔符的字符放置在任何标识符周围。</p> <p>如果数据库使用分隔标识符，则选择一个字符。数据集成服务生成 SQL 查询时，该服务将在此字符内为分隔标识符加上引号。</p>
Support Mixed-case Identifiers	<p>如果数据库使用区分大小写的标识符，则启用。启用后，数据集成服务在为 SQL 标识符字符属性选择的字符内的所有标识符加上引号。</p> <p>当 SQL 标识符字符属性设置为 none 时，支持混合大小写标识符属性将禁用。</p>
ODBC Provider	<p>ODBC. The type of database to which ODBC connects. For pushdown optimization, specify the database type to enable the Data Integration Service to generate native database SQL. The options are:</p> <ul style="list-style-type: none"> - Other - Sybase - Microsoft_SQL_Server <p>Default is Other.</p>

JDBC 连接属性

您可以使用 JDBC 连接访问数据库中的表。您可以在 Administrator 工具、Developer tool 或 Analyst 工具中创建和管理 JDBC 连接。

注意：连接属性的顺序可能会因查看属性所用的工具而异。

下表介绍了 JDBC 连接属性：

属性	说明
数据库类型	数据库类型。
名称	<p>连接的名称。名称不区分大小写，但在域中必须唯一。名称不能超出 128 个字符，也不能包含空格或以下特殊字符：</p> <p>~ ` ! \$ % ^ & * () - + = { [] } \ : ; " ' < , > . ? /</p>
ID	<p>数据集成服务用来标识连接的字符串。ID 不区分大小写。不得超过 255 个字符，并且在域中必须唯一。不能在创建连接后更改此属性。默认值是连接名称。</p>
说明	连接的说明。说明不得超过 765 个字符。

属性	说明
用户名	数据库用户名。
密码	数据库用户名的密码。
JDBC 驱动程序类名称	<p>JDBC 驱动程序类的名称。</p> <p>以下列表提供了可以为适用的数据库类型输入的驱动程序类名称：</p> <ul style="list-style-type: none"> - Oracle 的 DataDirect JDBC 驱动程序类名称： com.informatica.jdbc.oracle.OracleDriver - IBM DB2 的 DataDirect JDBC 驱动程序类名称： com.informatica.jdbc.db2.DB2Driver - Microsoft SQL Server 的 DataDirect JDBC 驱动程序类名称： com.informatica.jdbc.sqlserver.SQLServerDriver - Sybase ASE 的 DataDirect JDBC 驱动程序类名称： com.informatica.jdbc.sybase.SybaseDriver - Informix 的 DataDirect JDBC 驱动程序类名称： com.informatica.jdbc.informix.InformixDriver - MySQL 的 DataDirect JDBC 驱动程序类名称： com.informatica.jdbc.mysql.MySQLDriver - Databricks Delta Lake 的 JDBC 驱动程序：已从 Databricks 下载的驱动程序名称。有关驱动程序的信息，请参见《<i>Data Engineering Integration 指南</i>》的“Databricks 集成准备工作”一章中有关配置存储访问的主题。 <p>有关要将哪一驱动程序类与特定数据库使用的详细信息，请参阅供应商文档。</p>
连接字符串	<p>用于连接到数据库的连接字符串。使用以下连接字符串：</p> <p>jdbc:<subprotocol>:<subname></p> <p>有关可用于特定驱动程序连接字符串的详细信息，请参阅供应商文档。</p>
环境 SQL	<p>可选。输入在连接到数据库时用于设置数据库环境的 SQL 命令。数据集成服务在每次连接到数据库时执行连接环境 SQL。</p> <p>注意: 如果启用 Sqoop，Sqoop 将忽略此属性。</p>
事务 SQL	<p>可选。输入在连接到数据库时用于设置数据库环境的 SQL 命令。数据集成服务在每个事务开始时执行事务环境 SQL。</p> <p>注意: 如果启用 Sqoop，Sqoop 将忽略此属性。</p>
SQL 标识符字符	<p>数据库用于在 SQL 查询中为分隔标识符加上引号的字符类型。可用的字符取决于数据库类型。</p> <p>如果数据库使用常规标识符，选择（无）。数据集成服务生成 SQL 查询时，该服务不会将带分隔符的字符放置在任何标识符周围。</p> <p>如果数据库使用分隔标识符，则选择一个字符。数据集成服务生成 SQL 查询时，该服务将在此字符内为分隔标识符加上引号。</p> <p>注意: 如果启用 Sqoop，Sqoop 将忽略此属性。</p>
支持混合大小写标识符	<p>如果数据库使用区分大小写的标识符，则启用。启用后，数据集成服务在为 SQL 标识符字符 属性选择的字符内的所有标识符加上引号。</p> <p>当 SQL 标识符字符 属性设置为 none 时，支持混合大小写标识符 属性将禁用。</p> <p>注意: 如果启用 Sqoop，那么当您生成并执行 DDL 脚本以便在运行时创建或替换目标时，Sqoop 会遵守此属性。在所有其他情形下，Sqoop 都会忽略此属性。</p>

属性	说明
使用 Sqoop 连接器	<p>为使用 JDBC 连接的数据对象启用 Sqoop 连接。数据集成服务在 Hadoop 运行时环境中通过 Sqoop 运行映射。</p> <p>您可以为基于 JDBC 数据库的关系数据对象、自定义数据对象和逻辑数据对象配置 Sqoop 连接。</p> <p>选择 Sqoop v1.x 可启用 Sqoop 连接。</p> <p>默认值为无。</p>
Sqoop 参数	<p>输入 Sqoop 必须用来连接到数据库的参数。请用空格分隔多个参数。</p> <p>要在 Blaze 引擎上使用适用于 Sqoop 的 Teradata Connector for Hadoop (TDCH) 专用连接器来运行映射，必须在 Sqoop 参数中定义 TDCH 连接工厂类。连接工厂类根据要使用的 TDCH Sqoop 连接器而异。</p> <ul style="list-style-type: none"> - 要使用由 Teradata 提供技术支持的 Cloudera 连接器，请配置以下 Sqoop 参数： -Dsqoop.connection.factories=com.cloudera.connector.teradata.TeradataManagerFactory - 要使用 Hortonworks Connector for Teradata（由 Teradata Connector for Hadoop 提供技术支持），请配置以下 Sqoop 参数： -Dsqoop.connection.factories=org.apache.sqoop.teradata.TeradataManagerFactory <p>要在 Spark 引擎上运行映射，您无需在 Sqoop 参数中定义 TDCH 连接工厂类。默认情况下，数据集成服务会调用由 Teradata 提供技术支持的 Cloudera 连接器和 Hortonworks Connector for Teradata（由 Teradata Connector for Hadoop 提供技术支持）。</p> <p>注意: 要使用通用 JDBC 连接器代替专用 Cloudera 或 Hortonworks 连接器来运行映射，必须在 JDBC 连接中定义 <code>--driver</code> 和 <code>--connection-manager</code> Sqoop 参数。如果在映射的读取或写入转换中定义了 <code>--driver</code> 和 <code>--connection-manager</code> 参数，Sqoop 将忽略这些参数。</p> <p>如果未输入 Sqoop 参数，则数据集成服务会基于 JDBC 连接属性构建 Sqoop 命令。</p>

MS SQL Server 连接属性

使用 Microsoft SQL Server 连接可访问 Microsoft SQL Server。Microsoft SQL Server 连接是指与 Microsoft SQL Server 关系数据库的连接。您可以在 Administrator 工具或 Developer tool 中创建和管理 Microsoft SQL Server 连接。

注意: 连接属性的顺序可能会因查看属性所用的工具而异。

下表介绍了 MS SQL Server 连接属性：

属性	说明
数据库类型	数据库类型。
名称	<p>连接的名称。名称不区分大小写，但在域中必须唯一。名称不能超出 128 个字符，也不能包含空格或以下特殊字符：</p> <p>~ ` ! \$ % ^ & * () - + = { [] } \ : ; " ' < , > . ? /</p>
ID	<p>数据集成服务用来标识连接的字符串。ID 不区分大小写。不得超过 255 个字符，并且在域中必须唯一。不能在创建连接后更改此属性。默认值是连接名称。</p>
说明	连接的说明。说明不得超过 765 个字符。

属性	说明
使用受信任连接	<p>启用应用程序服务以使用 Windows 身份验证访问数据库。启动应用程序服务的用户名必须是有效的 Windows 用户并拥有访问数据库的权限。默认情况下，此选项被清除。</p> <p>注意: Windows 和 NTLM 身份验证未经可用于 Linux 上托管的 Microsoft SQL Server 2017 版本的认证。</p>
用户名	数据库用户名。如果 Microsoft SQL Server 使用 NTLMv1 或 NTLMv2 身份验证，则为必填。
密码	数据库用户名的密码。如果 Microsoft SQL Server 使用 NTLMv1 或 NTLMv2 身份验证，则为必填。
已启用传递安全	为该连接启用传递安全。启用连接的传递安全时，域使用客户端用户名和密码登录相应的数据库，而不是连接对象中定义的凭据。
元数据访问属性：连接字符串	<p>用于从数据库访问元数据的连接字符串。</p> <p>使用以下连接字符串：</p> <p>jdbc:informatica:sqlserver://<host name>:<port>;DatabaseName=<database name></p> <p>要使用 NTLM 身份验证测试连接，请在连接字符串中包括以下参数：</p> <ul style="list-style-type: none"> - 身份验证方法。要使用的 NTLM 身份验证版本。 <p>注意: UNIX 支持 NTLMv1 和 NTLMv2，但不支持 NTLM。</p> <ul style="list-style-type: none"> - 域。SQL server 所属的域。 <p>以下示例显示 SQL Server（在名为 Informatica.com 的 NT 域中使用 NTLMv2 身份验证）的连接字符串：</p> <p>jdbc:informatica:sqlserver://host01:1433;DatabaseName=SQL1;AuthenticationMethod=ntlm2java;Domain=Informatica.com</p> <p>如果使用 NTLM 身份验证连接，则可以在 MS SQL Server 连接属性中启用使用受信任连接选项。如果使用 NTLMv1 或 NTLMv2 身份验证连接，则必须在连接属性中提供用户名和密码。</p>

属性	说明
AdvancedJDBCSecurityOptions	<p>数据库参数，用于访问安全数据库的元数据。Informatica 将 AdvancedJDBCSecurityOptions 字段的值视为敏感数据，并以加密的方式存储该参数字符串。</p> <p>要连接到安全数据库，请包括以下参数：</p> <ul style="list-style-type: none"> - EncryptionMethod。必需。指示数据在通过网络传送时是否进行了加密。此参数必须设置为 SSL。 - ValidateServerCertificate。可选。指示 Informatica 是否验证由数据库服务器发送的证书。 如果该参数设置为 True，则 Informatica 将验证由数据库服务器发送的证书。如果指定 HostNameInCertificate 参数，Informatica 还会验证证书中的主机名。 如果该参数设置为 false，则 Informatica 不验证由数据库服务器发送的证书。Informatica 将忽略您指定的任何信任库信息。 - HostNameInCertificate。可选。托管安全数据库的计算机的主机名。如果指定主机名，Informatica 将根据 SSL 证书中的主机名来验证连接字符串中包含的主机名。 - cryptoProtocolVersion。可选。如果您为 Microsoft SQL Server 实例启用 TLS，请按如下方式设置 cryptoProtocolVersion 参数： cryptoProtocolVersion=TLSv<版本号>。 例如，cryptoProtocolVersion=TLSv1.2 注意: 版本号必须与您为服务器配置的 TLS 版本相同。 - TrustStore。必需。信任库文件的路径和文件名。 注意: 如果您配置 SSL 或 TLS 且只指定文件名，则必须将信任库文件复制到 Informatica 安装目录。要测试连接并导入元数据，请将信任库文件复制到以下目录： <Informatica 客户端安装目录>/clients/DeveloperClient 要运行映射，请将信任库文件复制到以下目录： <Informatica 服务器安装目录>/tomcat/bin 注意: 如果您配置 SSL 或 TLS 且只指定文件名，则必须将信任库文件复制到以下目录以测试连接：<Informatica 服务器安装目录>/tomcat/bin - TrustStorePassword。必需。安全数据库的信任库文件的密码。 不适用于 ODBC。 注意: Informatica 会将安全的 JDBC 参数附加到该连接字符串。如果将安全的 JDBC 参数直接附加到该连接字符串，那么请勿在 AdvancedJDBCSecurityOptions 字段中输入任何参数。
数据访问属性：提供程序类型	<p>要用来连接到 Microsoft SQL Server 数据库的连接提供程序。</p> <p>可以选择以下提供程序类型：</p> <ul style="list-style-type: none"> - ODBC - Oldeb（已弃用） <p>默认类型为 ODBC。</p> <p>注意: 尽管 Microsoft SQL Server 连接用户界面将 OLEDB 提供程序类型显示为已弃用，但 Informatica 仍会为 OLEDB 提供程序类型提供支持。有关 OLEDB 提供程序类型支持语句的详细信息，请参阅以下知识库文章：KB 522895。</p>
使用 DSN	<p>使数据集成服务能够将数据源名称用于连接。</p> <p>如果选择“使用 DSN”选项，数据集成服务将从 DSN 检索数据库名称和服务器名称。</p> <p>如果未选择“使用 DSN”选项，则您必须提供数据库名称和服务器名称。</p>

属性	说明
连接字符串	<p>如果未启用 DSN 模式，请使用以下连接字符串：</p> <p><server name>@<database name></p> <p>如果已启用 DSN 模式，请使用以下连接字符串：</p> <p><DSN Name></p>
代码页	用于从源数据库读取或写入目标数据库或文件的代码页。
域名	域的名称。
数据包大小	用于传输数据的包大小。用于优化 Microsoft SQL Server 的本地驱动程序。
所有者名称	<p>架构所有者的名称。</p> <p>注意: 通过动态映射或生成并执行 DDL 选项生成表 DDL 时，DDL 元数据不包含架构名称属性和所有者名称属性。</p>
架构名称	<p>数据库中架构的名称。如果架构名称与数据库用户名不同，则必须为剖析仓库指定架构名称。如果架构名称与数据库用户名不同，并且您配置了用户管理的缓存表，您必须指定数据对象缓存数据库的架构名称。</p> <p>注意: 通过动态映射或生成并执行 DDL 选项生成表 DDL 时，DDL 元数据不包含架构名称属性和所有者名称属性。</p>
环境 SQL	连接到数据库时，用于设置数据库环境的 SQL 命令。数据集成服务在每次连接到数据库时运行连接环境 SQL。
事务 SQL	连接到数据库时，用于设置数据库环境的 SQL 命令。数据集成服务在每个事务开始时运行事务环境 SQL。
重试周期	此属性保留供将来使用。
SQL 标识符字符	<p>数据库用于在 SQL 查询中为分隔标识符加上引号的字符类型。可用的字符取决于数据库类型。</p> <p>如果数据库使用常规标识符，选择（无）。数据集成服务生成 SQL 查询时，该服务不会将带分隔符的字符放置在任何标识符周围。</p> <p>如果数据库使用分隔标识符，则选择一个字符。数据集成服务生成 SQL 查询时，该服务将在此字符内为分隔标识符加上引号。</p>
支持混合大小写标识符	<p>如果数据库使用区分大小写的标识符，则启用。启用后，数据集成服务为在为 SQL 标识符字符属性选择的字符内的所有标识符加上引号。</p> <p>当 SQL 标识符字符属性设置为 none 时，支持混合大小写标识符属性将禁用。</p>
ODBC 提供程序	<p>ODBC。ODBC 连接的数据库的类型。要实现下推优化，请指定数据库类型，以使数据集成服务能够生成本地数据库 SQL。选项如下：</p> <ul style="list-style-type: none"> - 其他 - Sybase - Microsoft_SQL_Server <p>默认值为“其他”。</p>

ODBC 连接属性

使用 ODBC 连接可访问 ODBC 数据。ODBC 连接属于关系数据库连接。您可以在 Administrator 工具、Developer tool 或 Analyst 工具中创建和管理 ODBC 连接。

注意: 连接属性的顺序可能会因查看属性所用的工具而异。

下表介绍了 ODBC 连接属性：

属性	说明
数据库类型	数据库类型。
名称	连接的名称。名称不区分大小写，但在域中必须唯一。名称不能超出 128 个字符，也不能包含空格或以下特殊字符： ~ ` ! \$ % ^ & * () - + = { [] } \ : ; " ' < , > . ? /
ID	数据集成服务用来标识连接的字符串。ID 不区分大小写。不得超过 255 个字符，并且在域中必须唯一。不能在创建连接后更改此属性。默认值是连接名称。
说明	连接的说明。说明不得超过 765 个字符。
用户名	数据库用户名。
密码	数据库用户名的密码。
已启用传递安全	为该连接启用传递安全。启用连接的传递安全时，域使用客户端用户名和密码登录相应的数据库，而不是连接对象中定义的凭据。
数据访问属性：连接字符串	用于访问数据库中元数据的 ODBC 连接 URL。 <data source name>
代码页	用于从源数据库读取或写入目标数据库或文件的代码页。
环境 SQL	连接到数据库时，用于设置数据库环境的 SQL 命令。数据集成服务在每次连接到数据库时运行连接环境 SQL。
事务 SQL	连接到数据库时，用于设置数据库环境的 SQL 命令。数据集成服务在每个事务开始时运行事务环境 SQL。
重试周期	此属性保留供将来使用。
SQL 标识符字符	数据库用于在 SQL 查询中为分隔标识符加上引号的字符类型。可用的字符取决于数据库类型。 如果数据库使用常规标识符，选择（无）。数据集成服务生成 SQL 查询时，该服务不会将带分隔符的字符放置在任何标识符周围。 如果数据库使用分隔标识符，则选择一个字符。数据集成服务生成 SQL 查询时，该服务将在此字符内为分隔标识符加上引号。

属性	说明
支持混合大小写标识符	如果数据库使用区分大小写的标识符，则启用。启用后，数据集成服务为在为 SQL 标识符字符属性选择的字符内的所有标识符加上引号。 当 SQL 标识符字符属性设置为 none 时，支持混合大小写标识符属性将禁用。
ODBC 提供程序	ODBC 连接的数据库的类型。要实现下推优化，请指定数据库类型，以使数据集成服务能够生成本地数据库 SQL。选项如下： - 其他 - Sybase - Microsoft_SQL_Server - Snowflake 默认值为“其他”。

注意：当数据集成服务在 UNIX 或 Linux 上运行时，使用 ODBC 连接来连接到 Microsoft SQL Server。当数据集成服务在 Windows 上运行时，使用本地连接来连接到 Microsoft SQL Server。

Oracle 连接属性

使用 Oracle 连接可连接到 Oracle 数据库。Oracle 连接属于关系连接类型。您可以在 Administrator 工具、Developer tool 或 Analyst 工具中创建和管理 Oracle 连接。

注意：连接属性的顺序可能会因查看属性所用的工具而异。

下表介绍了 Oracle 连接属性：

属性	说明
数据库类型	数据库类型。
名称	连接的名称。名称不区分大小写，但在域中必须唯一。名称不能超出 128 个字符，也不能包含空格或以下特殊字符： ~ ` ! \$ % ^ & * () - + = { [] } \ : ; " ' < , > . ? /
ID	数据集成服务用来标识连接的字符串。ID 不区分大小写。不得超过 255 个字符，并且在域中必须唯一。不能在创建连接后更改此属性。默认值是连接名称。
说明	连接的说明。说明不得超过 765 个字符。
用户名	数据库用户名。
密码	数据库用户名的密码。
已启用传递安全	为该连接启用传递安全。启用连接的传递安全时，域使用客户端用户名和密码登录相应的数据库，而不是连接对象中定义的凭据。
元数据访问属性：连接字符串	用于从数据库访问元数据的连接字符串。 使用以下连接字符串： <code>jdbc:informatica:oracle://<host_name>:<port>;SID=<database name></code>

属性	说明
AdvancedJDBCSecurityOptions	<p>数据库参数，用于访问安全数据库的元数据。Informatica 将 AdvancedJDBCSecurityOptions 字段的值视为敏感数据，并以加密的方式存储该参数字符串。</p> <p>要连接到安全数据库，请包括以下参数：</p> <ul style="list-style-type: none"> - EncryptionMethod。必需。指示数据在通过网络传送时是否进行了加密。此参数必须设置为 SSL。 - ValidateServerCertificate。可选。指示 Informatica 是否验证由数据库服务器发送的证书。 如果该参数设置为 True，则 Informatica 将验证由数据库服务器发送的证书。如果指定 HostNameInCertificate 参数，Informatica 还会验证证书中的主机名。 如果该参数设置为 false，则 Informatica 不验证由数据库服务器发送的证书。Informatica 将忽略您指定的任何信任库信息。 - HostNameInCertificate。可选。托管安全数据库的计算机的主机名。如果指定主机名，Informatica 将根据 SSL 证书中的主机名来验证连接字符串中包含的主机名。 - cryptoProtocolVersion。可选。如果您为 Oracle 实例启用 TLS，请按如下方式设置 cryptoProtocolVersion 参数： cryptoProtocolVersion=TLSv<版本号>。 例如，cryptoProtocolVersion=TLSv1.2 注意：版本号必须与您为服务器配置的 TLS 版本相同。 - TrustStore。必需。信任库文件的路径和文件名。 注意：如果您配置 SSL 或 TLS 且只指定文件名，则必须将信任库文件复制到 Informatica 安装目录。要测试连接并导入元数据，请将信任库文件复制到以下目录： <Informatica 客户端安装目录>/clients/DeveloperClient 要运行映射，请将信任库文件复制到以下目录： <Informatica 服务器安装目录>/tomcat/bin 注意：如果您配置 SSL 或 TLS 且只指定文件名，则必须将信任库文件复制到以下目录以测试连接：<Informatica 服务器安装目录>/tomcat/bin - TrustStorePassword。必需。安全数据库的信任库文件的密码。 - KeyStore。必需。密钥库文件的路径和文件名。 - KeyStorePassword。必需。安全数据库的密钥库文件的密码。 <p>注意：Informatica 会将安全的 JDBC 参数附加到该连接字符串。如果将安全的 JDBC 参数直接附加到该连接字符串，那么请勿在 AdvancedJDBCSecurityOptions 字段中输入任何参数。</p>
数据访问属性：连接字符串	<p>使用以下连接字符串：</p> <p><database name>.world</p>
代码页	用于从源数据库读取或写入目标数据库或文件的代码页。
环境 SQL	连接到数据库时，用于设置数据库环境的 SQL 命令。数据集成服务在每次连接到数据库时运行连接环境 SQL。
事务 SQL	连接到数据库时，用于设置数据库环境的 SQL 命令。数据集成服务在每个事务开始时运行事务环境 SQL。
重试周期	此属性保留供将来使用。

属性	说明
启用并行模式	在批量模式下将数据加载到表中时启用并行处理。默认情况下，此选项被清除。
SQL 标识符字符	<p>数据库用于在 SQL 查询中为分隔标识符加上引号的字符类型。可用的字符取决于数据库类型。</p> <p>如果数据库使用常规标识符，选择（无）。数据集成服务生成 SQL 查询时，该服务不会将带分隔符的字符放置在任何标识符周围。</p> <p>如果数据库使用分隔标识符，则选择一个字符。数据集成服务生成 SQL 查询时，该服务将在此字符内为分隔标识符加上引号。</p>
支持混合大小写标识符	<p>如果数据库使用区分大小写的标识符，则启用。启用后，数据集成服务为在为 SQL 标识符字符属性选择的字符内的所有标识符加上引号。</p> <p>当 SQL 标识符字符属性设置为 none 时，支持混合大小写标识符属性将禁用。</p>

Hive 连接属性

使用 Hive 连接可访问 Hive 数据。Hive 连接属于数据库类型连接。您可以在 Administrator 工具、Analyst 工具或 Developer tool 中创建和管理 Hive 连接。Hive 连接属性区分大小写，另有说明除外。

注意：连接属性的顺序可能会因查看属性所用的工具而异。

下表介绍了 Hive 连接属性：

属性	说明
名称	<p>连接的名称。名称不区分大小写，但在域中必须唯一。可以在创建连接后更改此属性。名称不能超出 128 个字符，也不能包含空格或以下特殊字符：</p> <p>~ ` ! \$ % ^ & * () - + = { [] } \ : ; " ' < , > . ? /</p>
ID	<p>数据集成服务用来标识连接的字符串。ID 不区分大小写。不得超过 255 个字符，并且在域中必须唯一。不能在创建连接后更改此属性。默认值是连接名称。</p>
说明	连接的说明。说明不得超过 4000 个字符。
位置	要在其中创建连接的域。对 Analyst 工具不适用。
类型	连接类型。选择 Hive。

属性	说明
LDAP 用户名	<p>数据集成服务模拟的用户的 LDAP 用户名，用于在 Hadoop 群集中运行映射。该用户名取决于在元数据连接字符串或数据访问连接字符串中为本机环境指定的 JDBC 连接字符串。</p> <p>如果 Hadoop 群集使用 Kerberos 身份验证，则 JDBC 连接字符串的主体名称必须与用户名相同。否则，用户名将取决于 JDBC 驱动程序的行为。使用 Hive JDBC 驱动程序时，您可以通过多种方法指定用户名，而且用户名可以成为 JDBC URL 的一部分。</p> <p>如果 Hadoop 群集不使用 Kerberos 身份验证，则用户名将取决于 JDBC 驱动程序的行为。</p> <p>如果不指定用户名，则 Hadoop 群集会根据以下条件对作业进行身份验证：</p> <ul style="list-style-type: none"> - Hadoop 群集不使用 Kerberos 身份验证。它根据运行数据集成服务的计算机的操作系统配置文件用户名对作业进行身份验证。 - Hadoop 群集 Kerberos 身份验证。它根据数据集成服务的 SPN 对作业进行身份验证。将忽略 LDAP 用户名。
密码	LDAP 用户名的密码。
环境 SQL	<p>用于设置 Hadoop 环境的 SQL 命令。在本机环境类型中，数据集成服务会在每次创建与 Hive 元存储的连接时执行环境 SQL。如果使用 Hive 连接在 Hadoop 群集上运行配置文件，则数据集成服务会在每个 Hive 会话开始时执行环境 SQL。</p> <p>在两种连接模式下使用环境 SQL 时适用以下规则和准则：</p> <ul style="list-style-type: none"> - 使用环境 SQL 指定 Hive 查询。 - 使用环境 SQL 设置 Hive 用户定义的函数的类路径，然后使用环境 SQL 或 PreSQL 指定 Hive 用户定义的函数。不能在数据对象属性中使用 PreSQL 来指定类路径。如果使用 Hive 用户定义函数，则必须将 .jar 文件复制到以下目录中： <p><Informatica 安装目录>/services/shared/hadoop/<Hadoop 发行版名称>/extras/hive-auxjars</p> <ul style="list-style-type: none"> - 可以使用环境 SQL 来定义要在 PreSQL 命令或自定义查询中使用的 Hadoop 或 Hive 参数。 - 如果为环境 SQL 属性使用多个值，请确保这些值之间没有空格。
SQL 标识符字符	<p>用于标识特殊字符和预留的 SQL 关键字（例如 WHERE）的字符类型。数据集成服务会将选定字符放置在特殊字符和预留的 SQL 关键字的两旁。数据集成服务还会将该字符用于支持混合大小写标识符属性。</p>

用于访问作为源或目标的 Hive 的属性

下表介绍了配置为访问作为源或目标的 Hive 的连接属性：

属性	说明
JDBC 驱动程序类名称	Hive JDBC 驱动程序类的名称。如果将此选项留空，则 Developer tool 将使用随发行版一起提供的默认 Apache Hive JDBC 驱动程序。如果默认的 Apache Hive JDBC 驱动程序并不符合您的要求，可以通过指定驱动程序类名来将 Apache Hive JDBC 驱动程序替代为第三方 Hive JDBC 驱动程序。
元数据连接字符串	<p>用于访问 Hadoop 服务器中的元数据的 JDBC 连接 URI。</p> <p>可以使用 PowerExchange for Hive 与 HiveServer 服务或 HiveServer2 服务进行通信。要连接到 HiveServer，请使用以下格式指定连接字符串：</p> <pre>jdbc:hive2://<hostname>:<port>/<db></pre> <p>其中</p> <ul style="list-style-type: none">- <hostname> 表示运行 HiveServer2 的计算机的名称或 IP 地址。- <port> 表示 HiveServer2 用来执行侦听的端口号。- <db> 表示要连接到的数据库的名称。如果未提供数据库名称，数据集成服务会使用默认的数据库详细信息。 <p>要连接到 HiveServer2，请使用 Apache Hive 为该特定 Hadoop 发行版实施的连接字符串格式。有关 Apache Hive 连接字符串格式的详细信息，请参见 Apache Hive 文档。</p> <p>对于用户模拟，必须将 hive.server2.proxy.user=<xyz> 添加到 JDBC 连接 URI 中。如果不配置用户模拟，则会使用当前用户的凭据来连接到 HiveServer2。</p> <p>如果 Hadoop 群集使用 SSL 或 TLS 身份验证，您必须将 ssl=true 添加到 JDBC 连接 URI。例如：jdbc:hive2://<主机名>:<端口>/<db>;ssl=true</p> <p>如果对 SSL 或 TLS 身份验证使用自签名证书，请确保该证书文件在客户端计算机和数据集成服务计算机上可用。有关详细信息，请参阅 <i>Data Engineering Integration 指南</i>。</p>
绕过 Hive JDBC 服务器	<p>JDBC 驱动程序模式。要使用 JDBC 驱动程序嵌入模式，请选中此复选框。</p> <p>要使用 JDBC 嵌入模式，请执行以下任务：</p> <ul style="list-style-type: none">- 验证 Hive 客户端与 Informatica 服务是否安装在同一台计算机上。- 将 Hive 连接属性配置为在 Hadoop 群集上运行映射。 <p>如果选择非嵌入模式，则必须配置数据访问连接字符串。</p> <p>Informatica 建议使用 JDBC 嵌入模式。</p>
查找粒度授权	<p>如果选择此选项以遵守 Hive 源中的细粒度授权，映射将遵守：</p> <ul style="list-style-type: none">- 行列级限制。适用于启用了 Sentry 或 Ranger 安全模式的 Hadoop 群集。- 数据屏蔽规则。应用于 Dynamic Data Masking 为包含敏感数据的列设置的屏蔽规则。 <p>如果没有选择此选项，Blaze 和 Spark 引擎会忽略限制和屏蔽规则，并且结果将包括限制或敏感数据。</p>

属性	说明
数据访问连接字符串	<p>用于访问 Hadoop 数据存储中的数据的连接字符串。要连接到 HiveServer，请使用以下格式指定非嵌入 JDBC 模式连接字符串：</p> <p>jdbc:hive2://<hostname>:<port>/<db></p> <p>其中</p> <ul style="list-style-type: none"> - <hostname> 表示运行 HiveServer2 的计算机的名称或 IP 地址。 - <port> 表示 HiveServer2 用来执行侦听的端口号。 - <db> 表示要连接到的数据库。如果未提供数据库名称，数据集成服务会使用默认的数据库详细信息。 <p>要连接到 HiveServer2，请使用 Apache Hive 为该特定 Hadoop 发行版实施的连接字符串格式。有关 Apache Hive 连接字符串格式的详细信息，请参见 Apache Hive 文档。</p> <p>对于用户模拟，必须将 hive.server2.proxy.user=<xyz> 添加到 JDBC 连接 URI 中。如果不配置用户模拟，则会使用当前用户的凭据来连接到 HiveServer2。</p> <p>如果 Hadoop 群集使用 SSL 或 TLS 身份验证，您必须将 ssl=true 添加到 JDBC 连接 URI。例如：jdbc:hive2://<主机名>:<端口>/<db>;ssl=true</p> <p>如果对 SSL 或 TLS 身份验证使用自签名证书，请确保该证书文件在客户端计算机和数据集成服务计算机上可用。有关详细信息，请参阅 <i>Data Engineering Integration 指南</i>。</p>
HDFS 上的 Hive 暂存目录	<p>Hive 暂存表的 HDFS 目录。必须为 Hadoop 模拟用户和映射模拟用户授予执行权限。</p> <p>在本地环境中将数据写入到 Hive 目标时，此选项适用且必需。</p>
Hive 暂存数据库名称	<p>Hive 暂存表的命名空间。</p> <p>Hive 暂存数据库名称自动从数据访问连接字符串更新。如果要替代默认名称，您需要在 Hive 连接中配置 Hive 暂存数据库名称。</p> <p>在本地环境中运行映射以将数据写入到 Hive 目标中时，此选项适用。</p> <p>如果在 Blaze 或 Spark 引擎上运行映射，则无需在 Hive 连接中配置 Hive 暂存数据库名称。数据集成服务使用在 Hadoop 连接中配置的值。</p>

HDFS 连接属性

使用 Hadoop 文件系统 (HDFS) 连接可访问 Hadoop 群集中的数据。HDFS 连接属于文件系统类型连接。您可以在 Administrator 工具、Analyst 工具或 Developer 工具中创建和管理 HDFS 连接。HDFS 连接属性区分大小写，另有说明除外。

注意：连接属性的顺序可能会因查看属性所用的工具而异。

下表介绍了 HDFS 连接属性：

属性	说明
名称	<p>连接的名称。名称不区分大小写，但在域中必须唯一。名称不能超出 128 个字符，也不能包含空格或以下特殊字符：</p> <p>~ ` ! \$ % ^ & * () - + = { [] } \ : ; " ' < , > . ? /</p>
ID	<p>数据集成服务用来标识连接的字符串。ID 不区分大小写。不得超过 255 个字符，并且在域中必须唯一。不能在创建连接后更改此属性。默认值是连接名称。</p>
说明	<p>连接的说明。说明不得超过 765 个字符。</p>

属性	说明
位置	要在其中创建连接的域。对 Analyst 工具不适用。
类型	连接类型。默认为 Hadoop 文件系统。
用户名	用于访问 HDFS 的用户名。
NameNode URI	<p>用于访问存储系统的 URI。</p> <p>可以在群集配置的 <code>core-site.xml</code> 配置集中找到 <code>fs.defaultFS</code> 的值。</p> <p>注意: 如果在导入群集配置时创建连接，则会默认填充 NameNode URI 属性，并在每次刷新群集配置时更新该属性。如果您手动设置此属性或替代属性值，则刷新操作不会更新此属性。</p>

访问多个存储类型

在连接参数中使用 NameNode URI 属性来连接到多个存储类型。下表介绍了存储类型以及存储类型的 NameNode URI 格式：

存储	NameNode URI 格式
HDFS	<p><code>hdfs://<namenode>:<port></code></p> <p>其中：</p> <ul style="list-style-type: none"> - <名称节点> 表示名称节点的主机名或 IP 地址。 - <端口> 表示名称节点用于侦听远程过程调用 (RPC) 的端口。 <p><code>hdfs://<nameservice></code>，对于 NameNode 高可用性。</p>
MapR-FS	<code>maprfs:///</code>
HDInsight 中的 WASB	<p><code>wasb://<container_name>@<account_name>.blob.core.windows.net/<path></code></p> <p>其中：</p> <ul style="list-style-type: none"> - <container_name> 标识特定的 Azure 存储 Blob 容器。 <p>注意: <container_name> 为可选。</p> <ul style="list-style-type: none"> - <account_name> 标识 Azure 存储 Blob 对象。 <p>示例：</p> <p><code>wasb://infabdmoffering1storage.blob.core.windows.net/infabdmoffering1cluster/mr-history</code></p>
HDInsight 中的 ADLS	<code>adl://home</code>

从 Azure HDInsight 群集创建群集配置时，该群集配置将使用 ADLS 或 WASB 作为主存储。不能以 ADLS 或 WASB 作为辅助存储来创建群集配置。可以在 HDFS 连接中编辑 NameNode URI 属性以连接到本地 HDFS 位置。

数据库连接中的标识符属性

在创建大部分关系数据库连接时，您必须配置数据库标识符属性。标识符属性决定了当服务生成 SQL 查询来访问数据库时，数据集成服务是否将标识符括在分隔字符内。

数据库标识符是一个数据库对象名称。表、视图、列、索引、触发器、过程、约束和规则可以具有标识符。您使用标识符在 SQL 查询中引用对象。数据库可以具有常规标识符或必须括在分隔字符内的分隔标识符。

常规标识符

常规标识符合标识符的格式规则。常规标识符在 SQL 查询中使用时不需要分隔字符。

例如，以下 SQL 语句使用常规标识符 *MYTABLE* 和 *MYCOLUMN*：

```
SELECT * FROM MYTABLE
WHERE MYCOLUMN = 10
```

分隔标识符

分隔标识符必须括在分隔字符内，因为它们不符合标识符的格式规则。

数据库可以使用以下类型的分隔标识符：

使用预留关键字的标识符

如果标识符使用预留关键字，必须在 SQL 查询中将标识符括在分隔字符内。例如，以下 SQL 语句访问一个名为 *ORDER* 的表：

```
SELECT * FROM "ORDER"
WHERE MYCOLUMN = 10
```

使用特殊字符的标识符

如果标识符使用特殊字符，必须在 SQL 查询中将标识符括在分隔字符内。例如，以下 SQL 语句访问一个名为 *MYTABLE\$@* 的表：

```
SELECT * FROM "MYTABLE$@"
WHERE MYCOLUMN = 10
```

区分大小写的标识符

默认情况下，IBM DB2、Microsoft SQL Server 和 Oracle 数据库中的标识符不区分大小写。数据库对象名称以大写字母存储，但是 SQL 查询可使用任意大小写字母引用它们。例如，以下 SQL 语句访问名为 *MYTABLE* 的表：

```
SELECT * FROM mytable
SELECT * FROM MyTable
SELECT * FROM MYTABLE
```

要使用区分大小写的标识符，必须在 SQL 查询中将标识符括在分隔字符内。例如，以下 SQL 语句访问一个名为 *MyTable* 的表：

```
SELECT * FROM "MyTable"
WHERE MYCOLUMN = 10
```

标识符属性

在创建大部分数据库连接时，您必须配置数据库标识符属性。所配置的标识符属性取决于数据库是使用常规标识符、在标识符中使用关键字或特殊字符，还是使用区分大小写的标识符。

在数据库连接中配置以下标识符属性：

SQL 标识符字符

数据库用于在 SQL 查询中为分隔标识符加上引号的字符类型。可用的字符取决于数据库类型。

如果数据库使用常规标识符，选择（无）。数据集成服务生成 SQL 查询时，该服务不会将带分隔符的字符放在任何标识符周围。

如果数据库使用分隔标识符，则选择一个字符。数据集成服务生成 SQL 查询时，该服务将在此字符内为分隔标识符加上引号。

支持混合大小写标识符

如果数据库使用区分大小写的标识符，则启用。启用后，数据集成服务为在为 **SQL 标识符字符** 属性选择的字符内的所有标识符加上引号。

在 Informatica 客户端工具中，必须使用正确的大小写引用标识符。例如，创建数据库连接时，必须使用正确的大小写输入数据库用户名。

当 **SQL 标识符字符** 属性设置为 none 时，**支持混合大小写标识符** 属性将禁用。

示例：数据库使用常规标识符

在此示例中，数据库使用常规标识符。没有标识符含有预留关键字或特殊字符。数据库使用不区分大小写的标识符。

在数据库连接中，将 **SQL 标识符字符** 属性设置为“(无)”。当 **SQL 标识符字符** 设置为无时，**支持混合大小写标识符** 属性将被禁用。

数据集成服务生成 SQL 查询时，服务不会在任何标识符周围放置分隔字符。

示例：数据库在标识符中使用关键字或特殊字符

在此示例中，数据库在某些标识符中使用关键字或特殊字符。数据库使用不区分大小写的标识符。

在数据库连接中，按如下所示配置标识符属性：

1. 将 **SQL 标识符字符** 属性设置为数据库为分隔标识符使用的字符。

此示例将该属性设置为 `"""` (引号)。

2. 清除 **支持混合大小写标识符** 属性。

当数据集成服务生成 SQL 查询时，服务会在使用预留关键字或特殊字符的标识符周围放置选定的字符。例如，数据集成服务可生成以下查询：

```
SELECT * FROM "MYTABLE$@" /* identifier with special characters enclosed within delimited
                           character */
WHERE MYCOLUMN = 10 /* regular identifier not enclosed within delimited character */
```

示例：数据库使用区分大小写的标识符

在此示例中，数据库使用区分大小写的标识符。数据库可能会在某些标识符中使用关键字或特殊字符，也可能不会如此。

在数据库连接中，按如下所示配置标识符属性：

1. 将 **SQL 标识符字符** 属性设置为数据库为分隔标识符使用的字符。

此示例将该属性设置为 `"""` (引号)。

2. 选择 **支持混合大小写标识符** 属性。

当数据集成服务生成 SQL 查询时，服务会在所有标识符周围放置选定的字符。例如，数据集成服务可生成以下查询：

```
SELECT * FROM "MyTable" /* case-sensitive identifier enclosed within delimited character */
WHERE "MYCOLUMN" = 10 /* regular identifier enclosed within delimited character */
```

搜索数据库连接

您可以搜索数据库连接。Analyst 工具会在具有搜索字符串的列表中突出显示第一个数据库连接。选择连接后，您可以测试连接是否成功。

1. 单击**查找**图标。
“查找”文本字段会显示在连接列表上方。
2. 输入搜索字符串。
Analyst 工具会在包含搜索字符串的列表中突出显示第一个连接名称。
从列表中选择连接，然后单击**测试**图标以测试连接。

创建数据库连接

可以在 Analyst 工具中创建数据库连接。选择一个简单连接可包含基本数据库属性。选择一个高级连接可包含其他数据库特定的属性。

1. 单击**新建**可打开**新建连接**对话框。
2. 输入以下信息：

选项	说明
名称	连接的名称。名称不区分大小写，但在域中必须唯一。名称不能超出 128 个字符，也不能包含空格或以下特殊字符： ~ ` ! \$ % ^ & * () - + = { [] } \ : ; " ' < , > . ? /
ID	数据集成服务用来标识连接的字符串。ID 不区分大小写。不得超过 255 个字符，并且在域中必须唯一。不能在创建连接后更改此属性。默认值是连接名称。
说明	连接的说明（可选）。

3. 选择数据库类型。
其他字段会根据您选择的数据库类型进行显示。
4. 选择一个简单连接或高级连接。
 - 要选择简单连接，请选择**简单连接**并指定连接属性。
 - 要选择高级连接，请选择**高级连接**并指定其他数据库连接属性。
5. 单击**确定**。
Analyst 工具会测试连接并显示测试状态。

编辑数据库连接

编辑连接以对连接的属性进行更改。不能更改连接的 ID。

1. 选择一个连接，然后单击**编辑**。
此时将显示**编辑连接**对话框。
2. 进行所需的更改并单击**确定**。
Analyst 工具会验证连接。
3. 单击**确定**，然后单击**关闭**。

删除数据库连接

您可以删除数据库连接。您必须具有数据库连接的写入权限才能删除连接。

1. 选择连接，然后单击**删除**图标。
2. 单击**关闭**。

第 4 章

作业状态工作区

本章包括以下主题：

- [作业状态工作区概览, 38](#)
- [访问作业状态工作区, 38](#)
- [作业属性, 39](#)
- [监视作业, 40](#)

作业状态工作区概览

使用**作业状态**工作区监视临时作业（例如配置文件、结果卡和映射规范作业）的状态。临时作业是用户从 Developer tool 或 Analyst 工具中运行的作业。

您可以监视临时作业（例如资产的数据预览以及对配置文件的向下钻取操作）的状态。例如，如果 Analyst 工具无法执行数据预览，您可能需要查看映射规范的数据预览作业的状态。您可以按作业类型进行筛选以精简数据预览作业的结果。

默认情况下，您可以监视您运行的作业。如果您有相应的特权，则也可以查看其他用户运行的作业。

选择一个作业时，可以查看该作业的日志，查看该作业的上下文或取消该作业。您也可以在作业面板中查看该作业的属性 and 消息。

注意: 如果 Analyst 工具使用 HTTPS 安全协议而 Administrator 工具使用 HTTP 安全协议，则您可能无法查看作业状态。请联系管理员为两个工具均配置 HTTPS 安全协议。

访问作业状态工作区

访问“作业状态”工作区可查看和监视作业。

- 从**管理菜单**中，选择**作业状态**。
此时将显示**作业状态**工作区。

作业属性

您可以查看每个作业的属性，例如作业状态、开始作业的用户以及作业持续时间。

您可以查看以下作业属性：

名称

作业的名称。

类型

作业的类型。您可以按特定作业类型进行筛选以查看作业的状态。选择**自定义**可按多个作业类型进行筛选。您可以选择以下选项：

- 预览
- 映射
- 引用表
- 企业发现配置文件
- 配置文件
- 结果卡

省/自治区/直辖市

作业的状态。您可以按特定作业状态进行筛选以查看作业的进度。选择**自定义**可按多个作业状态进行筛选。您可以选择查看以下状态：

- 正在运行。分析服务正在运行作业。
- 已完成。分析服务已成功完成作业。
- 失败。分析服务在处理作业时遇到致命错误。
- 已中止。分析服务已中止作业。
- 已取消。您选择取消正在运行的作业。
- 已排队。分析服务已排队作业以进行处理。
- 未知。分析服务无法确定作业的状态。

作业 ID

作业的唯一标识符。

启动者

启动作业的用户名。

开始时间

作业的开始时间。您可以按特定开始时间进行筛选。选择**自定义**可按日期和时间范围进行筛选。您可以为开始时间选择查看以下选项之一：

- 最后 30 分钟
- 过去 4 小时
- 过去 1 天
- 过去 1 周

已用时间

作业持续运行的时间。选择**自定义**可按日期和时间范围进行筛选。

结束时间

作业结束的时间。您可以按特定结束时间进行筛选。选择**自定义**可按日期和时间范围进行筛选。您可以为结束时间选择以下选项：

- 最后 30 分钟
- 过去 4 小时
- 过去 1 天
- 过去 1 周

用户安全域

用户名的安全域。安全域可以是本地域、LDAP 域或 Kerberos 域。

监视作业

您可以监视与数据预览或配置文件向下钻取相关的作业的状态。

在监视作业时，您可以执行以下任务：

搜索作业。

按作业状态属性或通过搜索筛选器搜索作业。应用搜索筛选器后，您可以清除筛选器。

要按工作状态属性进行搜索，请在搜索字段中输入作业状态属性。

要通过应用筛选器进行搜索，请单击作业状态属性上的筛选器菜单。（可选）为**开始时间**和**已用时间**属性输入自定义筛选器。

要清除搜索筛选器，请单击**重置筛选器**图标。

查看作业的上下文。

在与选定作业几乎相同的时间内开始的其他作业的上下文中查看作业。

要查看作业的上下文，请从“操作”菜单中选择**查看上下文**。Analyst 工具会显示在与选定作业几乎相同的时间内开始的作业列表。

刷新作业列表。

要刷新作业列表，请从“操作”菜单中选择**刷新**。

请求有关新作业的通知。

要请求有关新作业的通知，请从“操作”菜单中选择**新建作业通知**。

取消作业。

可以取消正在运行的作业。您可能希望取消挂起的作业或所需完成时间过长的作业。

要取消作业，请从“操作”菜单中单击**取消选定作业**。

查看作业日志事件。

您可以查看选定作业的日志事件。事件严重性值包括信息、错误、警告、跟踪、调试、严重错误。默认值为信息。

要查看某个作业的日志事件，请从“操作”菜单中单击**查看选定对象的日志**。Analyst 工具会创建包含日志的文本文件。您可以打开或下载文本文件以查看日志。

第 5 章

项目工作区

本章包括以下主题：

- [项目工作区概览, 41](#)
- [分配项目工作区, 41](#)
- [管理项目和文件夹, 42](#)
- [项目安全性, 43](#)

项目工作区概览

使用**项目**工作区管理项目和文件夹，并为项目和文件夹分配权限。项目和文件夹会显示在**项目**面板中。

项目是用于存储文件夹和存储库内容的顶级容器。还可以通过 Analyst 工具在项目中存储资产。使用项目可组织和管理文件夹和资产。

使用文件夹可组织项目内容。创建文件夹可根据业务需求分组资产。您可以在项目或在其他文件夹中创建文件夹。创建项目或文件夹时，Analyst 工具会在模型存储库中存储项目或文件夹。

例如，您需要对按国家中的地区构建的多个系统评估数据质量。创建名为“East”和“West”的项目，使其分别对应于东部和西部地区的数据。在“East”和“West”项目中创建名为“Customers”和“Accounts”的文件夹以组织这些项目中的数据。您可以在“Customers”和“Accounts”文件夹中导入诸如表对象和平面文件对象等资产。

分配项目工作区

访问**项目**工作区可管理项目和文件夹。

- 从**管理**菜单中，选择**项目**。
将显示**项目**工作区。

管理项目和文件夹

您可以在**项目**工作区中执行任务来管理项目和文件夹。

您可以对项目或文件夹执行以下任务：

创建项目或文件夹。

在 Analyst 工具中创建项目以存储数据对象和资产。可以在项目中创建文件夹。

从**项目**面板的“操作”菜单中，单击**新建 > 项目**或单击**新建 > 文件夹**，然后输入项目或文件夹名称或者可选的说明。

复制项目或文件夹。

在项目中复制项目或文件夹可使用相同的内容执行不同的任务。例如，复制包含客户地址表的“Customers”项目以对“Customer_Accounts”项目使用相同的表。

选择您要复制的项目或文件夹。不能在具有相同名称的另一项目中复制项目。不能将项目中的文件夹复制到另一项目中的另一文件夹。复制项目不会复制项目上的用户权限。默认情况下，项目的所有者会获得复制项目上的所有权限。

从**项目**面板的“操作”菜单中，单击**复制**，然后输入项目或文件夹名称或者可选的说明。

重命名项目或文件夹。

创建项目或文件夹后对其重命名以更改项目或文件夹名称来适合特定业务用途或命名约定。选择您要重命名的项目或文件夹。

从**项目**面板的“操作”菜单中，单击**编辑**并输入另一项目或文件夹名称。

编辑项目或文件夹说明。

创建项目或文件夹后，编辑其说明。选择您要编辑的项目或文件夹。

从**项目**面板的“操作”菜单中，单击**编辑**并输入项目或文件夹说明。

删除项目或文件夹。

不再需要某个项目或文件夹时可将其删除。选择您要删除的项目或文件夹。在删除项目或文件夹之前，确认其相关内容未在其他项目或文件夹中使用。

从**项目**面板的“操作”菜单中，单击**删除**。

刷新项目或文件夹。

刷新项目或文件夹的内容可查看最新的内容和项目权限。

从**项目**面板的“操作”菜单中，单击**刷新**。

移动文件夹。

移动项目中另一文件夹内的文件夹，从而将项目内容组织成文件夹层次结构。不能将文件夹移至其在项目中的一个子文件夹。请选择要移动的文件夹。

从**项目**面板的“操作”菜单中，单击**移动**。

查看或分配项目上的权限。

查看或分配用户或组的项目权限。选择您分配或查看权限的项目。

在**直接权限**面板中查看项目上的权限。

从**编辑权限**对话框分配项目上的权限。

项目安全性

在 Analyst 工具中管理项目上的权限可控制对项目的访问。您可以将用户添加到项目，并为项目上的用户分配权限。

即使用户具有执行某些操作的特权，该用户还可能需要对某个特定资产执行操作的权限。

默认情况下，创建项目后，您便是此项目的所有者。所有者拥有所有权限，对此您无法更改。所有者可以为用户分配权限。

您可以为用户或组分配以下权限：

读取

用户或组可以打开、预览、导出、验证和部署项目中的所有资产。用户或组还可以查看项目详细信息。

写入

用户或组具有对项目中的所有资产的读取权限。此外，用户或组可以编辑项目中的所有资产、编辑项目详细信息以及删除项目中的所有资产。

授予

用户或组具有对项目中的所有资产的读取权限。此外，用户或组还可以为其他用户或组分配权限。

项目权限

为用户或组分配项目权限。项目权限决定用户或组是否可以查看资产、编辑资产或为其他用户或组分配权限。权限可以是直接权限、继承权限或有效权限。

直接权限是直接分配给用户或组的权限。用户和组拥有对对象的权限后，如果还拥有相应的特权，则可以对该对象执行管理任务。您可以编辑直接权限。

继承权限是用户继承的权限。用户具有项目的权限时，他们会继承项目中所有文件夹和数据对象的权限。组具有项目的权限时，属于该组的所有子组 and 用户都会继承项目的权限。例如，某个项目具有名为“Customers”的文件夹，其中包含多个文件夹。如果分配对该项目的组权限，则属于该组的所有子组 and 用户都会继承对“Customers”文件夹及该文件夹中所有文件夹的权限。

有效权限是某个用户或组的所有权限的超集。这些权限包括直接权限和继承权限。

分配到模型存储库服务管理员角色的用户会继承模型存储库服务中所有项目的全部权限。分配到组的用户会继承组权限。

分配项目的直接权限

您可以将用户添加到项目，分配项目的直接权限以施加限制、提供访问权限或管理项目中的资产。

1. 选择要分配直接权限的项目。
2. 单击**编辑权限**图标。
此时将显示**编辑权限**对话框。
3. 从**用户和组**面板中选择用户和/或组。
4. （可选）单击**添加用户和组**图标可将用户和组添加到项目。
此时将显示**添加组和用户**对话框。
5. 选择要为其分配权限的用户和组。
6. 单击**下一步**。
7. 选择用户和组权限。

8. 单击**保存**。
9. （可选）选择按名称、安全域或者用户或组的类型来筛选用户和组列表。
 - 要按名称筛选，请在**名称**字段上方输入名称或字符串。
 - 要按安全域筛选，请单击**安全域**字段上方的筛选器菜单。
 - 要按类型筛选，请单击**类型**字段上方的筛选器菜单并选择用户或组。
10. 选择或清除**权限**面板中的**读取**、**写入**和**授予**权限。
11. 单击**确定**。

查看项目的权限

查看项目的权限时，您可以查看有效权限的来源。权限详细信息显示分配给用户或组的直接权限、分配给父组的直接权限，以及从父对象继承的权限。

1. 选择要查看其权限的项目。
2. 单击**有效权限**图标。
此时将显示**有效权限**对话框。
3. 查看用户和组的有效权限。您查看的权限包括直接权限和继承权限。
4. （可选）选择按名称、安全域或者用户或组的类型来筛选用户和组列表。
 - 要按名称筛选，请在**名称**字段上方输入名称或字符串。
 - 要按安全域筛选，请单击**安全域**字段上方的筛选器菜单。
 - 要按类型筛选，请单击**类型**字段上方的筛选器菜单并选择用户或组。
5. 单击**关闭**。

第 6 章

模型存储库

本章包括以下主题：

- [模型存储库概览, 45](#)
- [Informatica Analyst 资产, 45](#)
- [存储库资产锁定, 46](#)
- [使用受版本控制的对象进行基于团队的开发, 47](#)

模型存储库概览

模型存储库是一种关系数据库，用于存储项目和文件夹的元数据。

每次打开 Analyst 工具时，将连接到模型存储库来访问项目和文件夹。

当您编辑资产时，模型存储库会锁定资产，以便您进行独占编辑。管理员也可以将模型存储库与第三方版本控制系统进行集成。利用版本控制系统集成，您可以签出和签入资产。

Informatica Analyst 资产

您可以在一些工作区中管理资产。资产是一种用于在企业内支持业务操作的对象。

例如，配置文件是分析师可以创建用于发现数据源的内容、质量和结构的资产。

您可以创建以下类型的资产：

词汇表资产

在**词汇表**工作区中创建词汇表资产。您可以创建以下类型的词汇表资产：

- 业务术语。使用业务语言为组织中的企业用户定义相关概念的词语或短语。
- 业务计划。用于对词汇表资产集合进行批量更改的业务决策。
- 类别。业务术语和策略的描述性分类。
- 词汇表。类别、业务术语和策略的集合。
- 策略。管理与业务术语相关的业务实践的业务目的、进程或协议。

发现资产

在**发现**工作区中创建发现资产。您可以创建以下类型的发现资产：

- 配置文件。在单个数据源或多个数据源中发现列特征和数据域的单个数据对象配置文件或企业发现配置文件。
- 平面文件数据对象。基于平面文件的数据表示。
- 表数据对象。基于关系表的数据表示。

设计资产

在**设计**工作区中创建设计资产。您可以创建以下类型的设计资产：

- 映射规范。描述数据从源到目标的移动及转换的模板。
- 引用表。包含一组数据值的标准版本和替代版本的一种表。
- 规则规范。代表业务规则逻辑的一种对象。

结果卡资产

在**结果卡**工作区中打开结果卡资产。结果卡是配置文件中质量度量的图形表示形式。

存储库资产锁定

模型存储库会锁定资产来防止用户覆盖所做的工作。模型存储库可以锁定 Analyst 工具在库工作区中显示的除项目和文件夹以外的任何资产。

当您开始在 Analyst 工具中编辑资产时，模型存储库会锁定资产，使其他用户无法保存对该资产的修改。当您保存资产时，您会保持锁定。当您关闭资产时，模型存储库会解除资产锁定。

如果您打开其他用户已锁定的资产，Analyst 工具会通知您该资产已被另一用户锁定。对象可能在 Analyst 工具或 Developer tool 中处于锁定状态。您可以选择在只读模式下审阅资产，也可以用另一名称保存资产。

如果 Analyst 工具关闭，模型存储库会保持资产锁定。当您再次连接到模型存储库时，可以继续编辑之前已锁定的资产。要编辑已被另一用户锁定的资产，请联系该用户或管理员。

每个已锁定的资产的**属性**视图会显示其锁定时间和日期，以及锁定所有者的用户 ID。

资产锁定管理的规则和准则

管理资产锁定时，请注意以下规则和准则：

- 当您打开资产时，模型存储库并不会锁定资产。只有当您开始编辑资产后，模型存储库才会锁定资产。例如，当您在映射规范的可编辑字段中插入光标或重命名资产时，模型存储库才会锁定映射规范。
- 您可以使用多种客户端工具开发资产：例如，您可以在一台计算机上编辑资产，然后在另一台计算机上打开该资产并继续编辑。当回到第一台计算机时，必须关闭资产然后重新打开它，以重新获得锁定。具有管理特权的用户解锁您之前已打开的资产时，也适用同样的原则。
- 管理员可以撤消您对已锁定资产的写入权限，或将锁定权限重新分配给其他用户。在这种情况下，您无法编辑或保存资产，但可以使用其他名称保存资产。

使用受版本控制的对象进行基于团队的开发

Team-based development is the integration of the Model repository with a third-party version control system. The version control system saves multiple versions of assets and assigns each version a version number. You can check assets out and in and undo the checkout of assets.

模型存储库可以保护资产不受开发团队其他成员的覆盖。如果您打开已被另一用户签出的资产，将收到一条通知，指示已签出该资产的用户。您可以在只读模式下打开已签出的资产，也可以使用不同的名称保存它。

使用**我的签出资产**视图管理已签出的资产。例如，您可能想要撤消签出，以删除对资产所做的更改。

当与版本控制系统的连接处于活动状态时，模型存储库具有每个资产的最新版本。

如果模型存储库失去与版本控制系统的连接，将保持已签出资产的状态。当与版本控制系统的连接中断时，您可以继续打开、编辑、保存和关闭资产。模型存储库会跟踪并保持资产状态。

连接恢复后，您可以继续执行与版本控制系统相关的操作，如签入资产或撤消资产的签出。如果您在连接中断期间打开并编辑了某个资产，模型存储库会将该资产签出给您。

受版本控制的资产管理

模型存储库与版本控制系统集成时，您可以管理资产的版本。例如，您可以签出和签入资产、撤消签出和查看已签出的资产。

您可以执行以下操作：

签出资产。

当您签出资产时，资产将保持签出状态，直到您将其签入或撤消签出。您可以在**我的签出资产**视图中查看已签出的资产。要签出资产，在“对象库”中右键单击资产并选择“签出”。

撤消资产的签出。

当您撤消签出时，会签入资产，但不会保存任何更改，也不会使版本号或版本历史记录发生递增。在签出资产后对该资产所做的所有更改都将丢失。要撤消签出，可以使用**我的签出资产**视图。

签入资产。

当您签入资产时，版本控制系统会更新版本历史记录并递增版本号。您可以添加最大 4 KB 的签入注释。要签入资产，使用**我的签出资产**视图或对象的右键单击菜单。

删除资产。

对于受版本控制的资产，您必须先将其签出，然后才能将其删除。如果资产在执行删除操作时未签出，模型存储库会将它签出给您，并将其标记为待删除。要完成删除操作，必须签入资产。

当您删除受版本控制的资产时，版本控制系统会删除所有版本。

要删除资产，可以使用**我的签出资产**视图。

我的签出资产视图

我的签出资产视图列出您已签出的所有资产。

您可以在**我的签出资产**视图中执行以下操作：

- 撤消资产的签出。
- 签入资产。
- 删除资产。

要访问该视图，在**我的签出资产**视图中右键单击资产，然后选择一项操作。

删除资产

当您删除受版本控制的资产时，必须将资产标记为待删除，然后再将其签入。

1. 在**库导航器**视图或**我的签出资产**视图中右键单击资产，然后选择**删除**。
2. 在**我的签出资产**视图中选择资产，然后选择**签入**。

资产将从模型存储库中删除。

第 7 章

数据对象

本章包括以下主题：

- [数据对象概览, 49](#)
- [平面文件数据对象, 49](#)
- [表数据对象, 54](#)
- [同步数据对象, 55](#)
- [查看数据对象, 57](#)
- [编辑数据对象, 58](#)

数据对象概览

数据对象表示您要从中提取元数据的源。您可以将平面文件和表导入为数据对象以分析数据的结构。

平面文件数据对象和表数据对象是发现资产，可以用作在组织中进行协作式项目的起点。您可以通过将数据对象导入 Analyst 工具来添加它们。您可以为表数据对象和平面文件数据对象表示的源数据创建配置文件。运行配置文件时，Analyst 工具会连接到该数据库表或平面文件。然后，您可以使用表数据对象和平面文件数据对象来执行任务，例如数据分析或数据集成任务。

导入数据对象时，您必须访问源以提取元数据。通过 Analyst 工具中可用的连接对象访问关系源。通过网络路径访问平面文件源。

在**发现**工作区中创建平面文件数据对象和表数据对象。使用工作区悬停菜单或使用**新建资产**面板创建数据对象。您还可以从 Analyst 工具表头的**新建**菜单中创建数据对象。将数据对象添加到项目或文件夹后，可以在**库**工作区的**项目**面板中查看数据对象。

平面文件数据对象

平面文件数据对象包含平面文件的元数据。使用平面文件数据对象作为协作项目的起点。添加平面文件数据对象时，Analyst 工具会连接到网络路径位置或上传源平面文件的位置以提取元数据。

要添加平面文件数据对象，您必须选择平面文件、配置文件选项以及配置列数据类型。添加平面文件数据对象后，可以预览其属性和列数据。

可以添加平面文件数据对象作为固定宽度或带分隔符的对象。添加平面文件数据对象作为固定宽度的对象时，可以按照固定宽度的分列线格式化数据。添加平面文件数据对象作为带分隔符的对象时，可以使用逗号等分隔符作为分列线来格式化数据。

如果源平面文件更改，还可以同步对平面文件数据对象的更改以获得更新的元数据。

导入平面文件数据对象

您可以通过将平面文件导入项目或文件夹中，在 Analyst 工具中添加平面文件数据对象。导入平面文件数据对象时，可以选择从本地计算机上传平面文件或选择一个网络路径。如果平面文件的大小超过 10 MB，则选择网络路径可导入该平面文件数据对象。

从本地计算机上传平面文件时，Analyst 工具会将平面文件的一个副本上传到 Informatica 服务安装目录（可供 Analyst 工具访问）中的平面文件缓存目录。联系管理员以配置 Analyst 工具用于网络路径的平面文件缓存。选择网络路径时，您可以指定本地计算机上平面文件的位置。

如果修改源平面文件，则可以同步对平面文件数据对象的更改。

导入平面文件数据对象时，Analyst 工具会根据前 10,000 行推断平面文件字段的数值或字符串数据类型。

平面文件选项

导入平面文件数据对象时，可以在**添加平面文件**向导中配置每个列的平面文件选项。您配置的选项会确定向导如何从源平面文件读取数据。

可以在**添加平面文件**向导中配置一下平面文件选项：

代码页

平面文件对象中的数据代码页。选择与平面文件对象中数据的代码页相匹配的代码页。

分隔符

用于分隔数据列的字符。使用“其他”字段输入其他分隔符。可打印的字符必须与转义符和引号字符（如果选择）不同。您可以输入以下非打印多字节字符：\1、\01 或 \001。

文本限定符

定义文本字符串边界的引号字符。选择“无引号”、“单引号”或“双引号”。如果选择引号字符，向导将忽略引号对内的分隔符。

列名称

用于从第一行导入列名称的选项。如果列名称显示在第一行中，请选择该选项。向导在预览中使用第一行中的数据作为列名称。

如果第一行包含数值字符，向导将使用 COLUMNx 作为默认列名称。如果第一行包含特殊字符，向导会将特殊字符转换为下划线并在列名称中使用有效字符。向导会跳过列名称中的以下特殊字符： " . + - = ~ ` ! % ^ & * () [] { } ' \ " ; : , < > \ \ | \t \r \n。默认情况下未启用。

值

选择从某一行开始进行值导入。当向导导入文件时，指示在预览中向导开始读取的行号。

平面文件数据类型

在**添加平面文件**向导中为每个列中的数据配置数据类型。配置的数据类型将确定向导如何导入源平面文件的数据。

配置以下数据类型：

- 长整型。可以在**数值格式**窗口中指定格式。可以使用默认值，或者指定其他数值格式并选择将其作为默认数值格式。
- 日期时间。可以在**日期时间格式**窗口中指定格式。可以使用默认值，或者指定其他日期时间格式并选择将其作为默认日期时间格式。

- 双精度。可以在**数值格式**窗口中指定格式。可以使用默认值，或者指定其他数值格式并选择将其作为默认数值格式。
- 整型。可以在**数值格式**窗口中指定格式。可以使用默认值，或者指定其他数值格式并选择将其作为默认数值格式。
- Nstring。可以指定精度值。不能指定格式。
- 数字。可以指定精度值和小数位数值。可以在**数值格式**窗口中指定格式。可以使用默认值，或者指定其他数值格式并选择将其作为默认数值格式。
- 字符串。可以指定精度值。不能指定格式。

日期时间数据类型

配置日期时间数据类型时，可以在**日期时间格式**窗口中指定格式。可以使用默认值，或者指定其他日期时间格式并选择将其作为默认日期时间格式。

可以指定以下日期时间格式字符串作为日期的一部分：

AM、a.m.、PM、p.m.

子午线指示器。使用任何这些格式字符串指定 AM 和 PM 小时。AM 和 PM 与 a.m. 和 p.m. 返回相同的值。

DAY

日期的名称，包括多达九个字符。DAY 格式字符串不区分大小写。

DD

每月的某一日。

DDD

每年的某一日，包括闰年。

DY

由三个字符组成的缩写日期名称。DY 格式字符串不区分大小写。

HH、HH12

一天中的时段。

HH24

一天中的时段，从 0 到 23，其中 0 表示 12AM。

J

修改的儒略日。

MI

分钟数，从 0 到 59。

MM

月

MONTH

月份的名称，包括多达九个字符。不区分大小写。

MON

由三个字符组成的缩写月份名称。不区分大小写。

MS

毫秒数，从 0 到 999。

NS

纳秒数，从 0 到 999999999。

RR

四位数年份。当源字符串包括两位数年份时使用。

SS

秒数，从 0 到 59。

SSSSS

自午夜算起的秒数。

CN

微秒数，从 0 到 999999。

Y

当前年份，年份的最后一位替换为字符串值。

YY

当前年份，年份的最后两位替换为字符串值。

YYY

当前年份，年份的最后三位替换为字符串值。

YYYY

四位数年份。在传递两位数年份时不要使用此格式字符串。换用 RR 或 YY 格式字符串。

添加带分隔符的平面文件

将平面文件数据对象导入项目或文件夹时，可以设置分隔符来格式化数据。您可以更改列属性来匹配数据预览。

1. 在**新建表头**上，单击**平面文件数据对象**。
将显示**添加平面文件**向导。
2. 选择浏览某个位置或输入网络路径以导入平面文件。
 - 要浏览某个位置，请选择**浏览和上载**并单击**选择文件**以从您的计算机可访问的目录中选择平面文件。
 - 要输入网络路径，请选择**输入网络路径**并配置文件的路径和文件名。
3. 单击**下一步**。
4. 接受默认的**带分隔符**选项。
5. 单击**下一步**。
6. 配置平面文件选项，并预览平面文件数据。
注意：选择与文件中数据的代码页相匹配的代码页。
7. （可选）单击**预览**面板中的**刷新**图标以更新对平面文件数据的预览更改。
8. 单击**下一步**。
9. （可选）更改**列属性**。
10. 单击**下一步**。
11. 在您要添加平面文件的**文件夹**面板中配置名称、可选说明和位置。
平面文件面板显示项目或文件夹中包含的平面文件。
12. 单击**完成**。
Analyst 工具会在**属性**选项卡上显示平面文件的属性。在**数据预览**选项卡上查看平面文件的数据预览。

添加固定宽度的平面文件

将固定长度的平面文件导入项目或文件夹时，可以设置分列线来格式化数据。

1. 在**新建表头**上，单击**平面文件数据对象**。
将显示**添加平面文件**向导。
2. 选择浏览某个位置或输入网络路径以导入平面文件。
 - 要浏览某个位置，请选择**浏览和上传**并单击**选择文件**以从您的计算机可访问的目录中选择平面文件。
 - 要输入网络路径，请选择**输入网络路径**并配置文件的路径和文件名。
3. 单击**下一步**。
4. 选择**固定宽度**。
5. 单击**下一步**。
6. 配置平面文件选项，并预览平面文件数据。
注意：选择与文件中数据的代码页相匹配的代码页。
7. （可选）单击**预览**面板中的**刷新**图标以更新对平面文件数据的预览更改。
8. 选择设置、删除、移动或编辑分列线。
 - 要设置分列线，请在**预览**面板中单击。
 - 要删除分列线，请双击分列线。
 - 要移动分列线，请拖动它们。
 - 要编辑分列线，请单击**编辑分列线**图标并使用**编辑分列线**对话框修改分列线。
9. 单击**下一步**。
10. （可选）更改**列属性**。
11. 单击**下一步**。
12. 在您要添加平面文件的**文件夹**面板中配置名称、可选说明和位置。
平面文件面板显示项目或文件夹中包含的平面文件。
13. 单击**完成**。

Analyst 工具会在**属性**选项卡上显示平面文件的属性。在**数据预览**选项卡上查看平面文件的数据预览。

平面文件的规则和准则

在处理平面文件时考虑以下规则和准则：

将小文件上传到 Informatica 服务安装目录。

将多达 10 MB 的文件上传到运行 Analyst 工具的计算机上的 Informatica 服务安装目录。Analyst 工具会访问此位置以提取不会频繁更改的平面文件元数据。使用大小高达 10 MB 的文件时，Analyst 工具会访问 Informatica 服务安装目录中的文件副本。如果修改原始文件，则需要再次上传文件。

将大文件上传到网络路径位置。

使 Analyst 工具可连接到网络路径位置以处理大于 10 MB 的平面文件。Analyst 工具会访问此位置以提取频繁更改的平面文件元数据。网络路径位置应是 Analyst 工具可访问的共享目录或文件系统。使用大小超过 10 MB 的文件时，Analyst 工具可以通过网络路径连接到平面文件。如果修改原始平面文件，则必须在 Analyst 工具中刷新平面文件。刷新大型平面文件的元数据会比较耗时。

不会导入空数据行。

在导入平面文件时，Analyst 工具不会导入第一个数据行上方的空行、空中间行和最后一个数据行之后的空行。

刷新数据预览。

预览之后，可以更改**添加平面文件**向导在导入文件时开始读取的行号。该行号与预览对应。如果选择从第一行开始导入列名，则刷新预览以更新预览数据的行号。

表数据对象

表数据对象包含 Analyst 工具中关系数据库源的元数据。使用表数据对象分析源数据。添加表数据对象时，Analyst 工具使用数据库连接来连接源数据库，从而提取元数据。

您可以通过将表导入项目或文件夹，在 Analyst 工具中添加表数据对象。在导入表数据对象之前，选择或创建数据库连接并选择要添加的数据库表。可以通过连接添加多个表作为数据对象。在导入表数据对象时还可以搜索表或表架构。

使用**新建表**向导可将表数据对象添加到项目或文件夹。使用连接工作区可创建数据库连接，以在将源表作为表数据对象导入时连接到源表。

添加表

使用**新建表**向导可将表数据对象添加到项目。添加要针对其分析源数据的表数据对象。要添加表数据对象，请选择连接、选择架构和表，然后添加表数据对象。

1. 在**新建表**头上，单击**表数据对象**。
显示**新建表**向导。
2. 选择一个连接。
3. 单击**下一步**。
4. （可选）取消选择**仅显示默认架构**可显示与选定连接关联的所有架构。
5. 选择要添加的表。
6. （可选）选择按表名称和/或架构名称搜索表。
 - 要按表名称搜索表，请在**表**搜索框中输入表名称，然后单击**查找**图标按表名称进行搜索。单击**清除**图标可按名称显示所有表。
 - 要按架构名称搜索表，请在**架构**搜索框中输入表架构名称，然后单击**查找**图标按表架构名称进行搜索。单击**清除**图标可按名称显示所有架构。
 - 要按表名称和架构名称搜索表，请在**表**搜索框中输入表名称并在**架构**搜索框中输入架构名称，然后单击**查找**图标按表名称和架构名称显示所有表。单击**清除**图标可按名称显示所有架构。
7. （可选）在**属性**选项卡上查看表的属性和列元数据。
8. （可选）单击**数据预览**选项卡查看表的列和数据。
9. 单击**下一步**。
10. 在**文件夹**面板中选择要添加表的项目或文件夹。
表面板会显示项目或文件夹中存在的表。
11. 单击**完成**。

表的规则和准则

使用表时，请考虑以下规则和指导原则：

- 预览表的数据时，Analyst 工具默认显示前 100 行。Analyst 工具可能不会在宽表中显示所有数据列。

- Analyst 工具可以导入具有超过 30 个列的宽表以剖析数据。导入宽表时，Analyst 工具不会在数据预览中显示所有列。Analyst 工具将在数据预览中显示前 30 列。但是，您可以包含宽表和平面文件中的所有列以进行剖析。
- 您可以导入具有小写和混合大小写字符的表和列。
- 您可以导入表名称或列名称中具有特殊字符的表。导入表名称或列名称中具有特殊字符的表时，Analyst 工具会将表名称或列名称中的特殊字符转换为下划线字符。您可以在表名称或列名称中使用以下特殊字符：
" \$. + - = ~ ` ! % ^ & * () [] { } ' \ " ; : / ? , < > \ \ | \t \r \n
- 您可以将具有 Microsoft SQL92 或 Microsoft SQL99 保留字（例如“concat”）的表和列导入 Analyst 工具中。
- 您可以在 Analyst 工具中使用 ODBC 连接导入 Microsoft SQL Server、MySQL、Teradata 和 Sybase 表。ODBC 连接需要用户名和密码。
- 使用 Microsoft SQL Server 连接访问 Microsoft SQL Server 数据库中的表时，Analyst 工具不显示表的同义词。
- 预览 Oracle、IBM DB2、IBM DB2 for zOS、IBM DB2/iOS、Microsoft SQL Server 和 ODBC 数据库的关系表数据时，如果表、视图、架构、同义词和列名称包含混合大小写或小写字符，则 Analyst 工具无法显示预览。要预览区分大小写的数据库中的表数据，请在 Developer 工具或 Administrator 工具中，针对 Oracle、IBM DB2、IBM DB2 for zOS、IBM DB2/iOS、Microsoft SQL Server 和 ODBC 数据库的连接，将**支持混合大小写标识符**属性设置为 true。
- 您可以在将表导入 Analyst 工具之后，预览源数据表的注释。要查看源表注释，请在用于从数据库访问元数据的 JDBC 连接 URL 中使用其他参数。在数据库连接属性中的**元数据访问字符串**选项中，使用 CatalogOptions=1 或 CatalogOptions=3。例如，对于 Oracle 数据库连接使用以下 JDBC 连接 URL：

Oracle: jdbc:informatica:oracle:// <host_name>:<port>;SID=<database name>;CatalogOptions=1

同步数据对象

将平面文件或表数据对象的更改与其外部数据源同步。如果外部平面文件或表数据源发生更改，则可以在 Analyst 工具中同步对平面文件或表数据对象所做的更改。

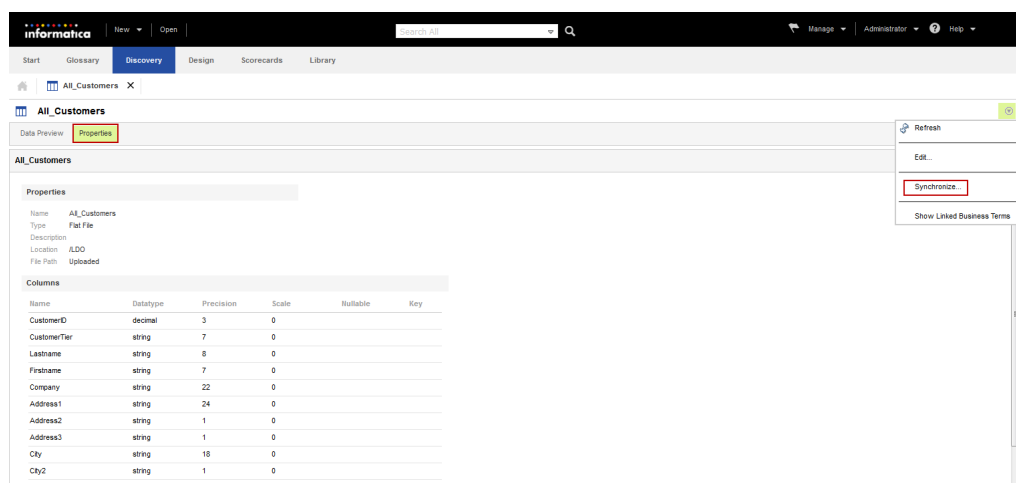
您可以通过**发现**工作区将对数据对象所做的更改与其外部数据源进行同步。还可以在**库**工作区中打开一个数据对象，然后右键单击数据对象以将其与外部数据源同步。

在 Informatica Analyst 中同步平面文件数据对象

可以将外部平面文件数据源的更改与其在 Analyst 工具中的数据对象同步。使用**同步平面文件**向导同步数据对象。

1. 打开**库**工作区。
2. 在**项目**部分中，选择项目中的平面文件数据对象。
Analyst 工具将在**属性**选项卡中显示表属性。
3. 从“操作”菜单中，单击**同步**。

下图显示了“属性”选项卡和“操作”菜单中的“同步”选项：

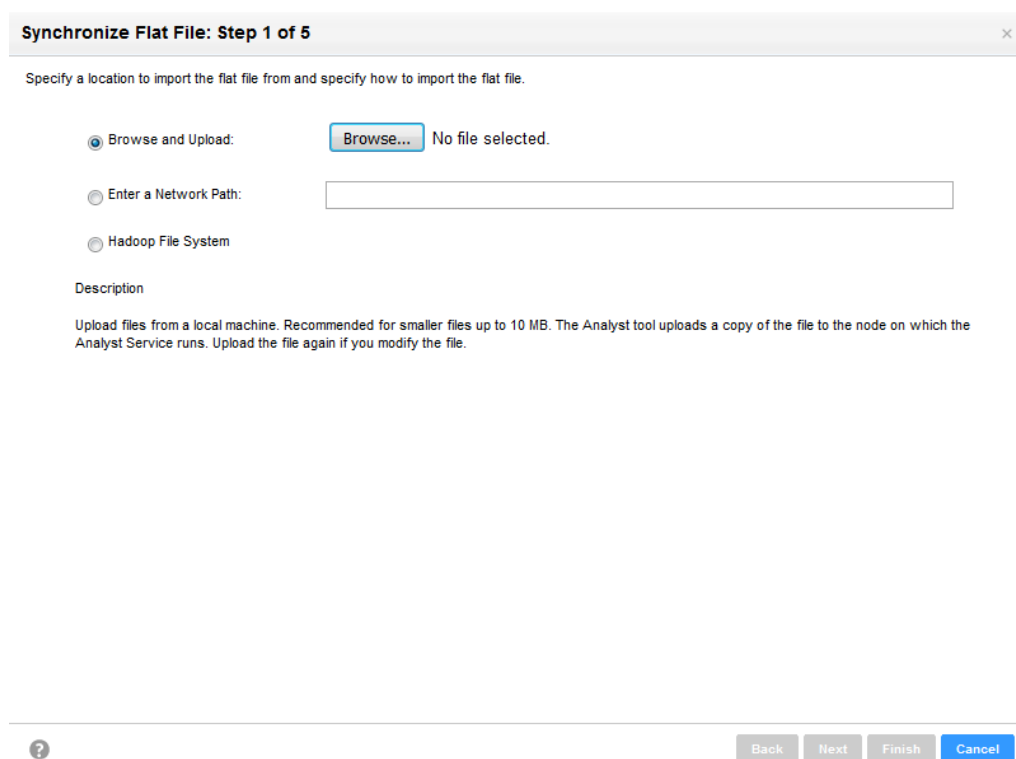


将显示同步平面文件向导。

4. 选择浏览某个位置或输入网络路径以导入平面文件。

- 要浏览位置，请单击**选择文件**以从计算机可访问的目录中选择平面文件。
- 要输入网络路径，请选择**输入网络路径**并配置文件路径和文件名。

下图显示了“同步平面文件”向导：



5. 单击**下一步**。

6. 选择导入带分隔符或固定宽度的平面文件。

- 要导入带分隔符的平面文件，请接受**带分隔符**选项。
- 要导入固定宽度的平面文件，请选择**固定宽度**选项。

7. 单击**下一步**。

8. 为带分隔符或固定宽度的平面文件配置平面文件选项。
9. 单击**下一步**。
10. (可选) 更改列属性。
11. 单击**下一步**。
12. 接受默认名称或为平面文件输入其他名称。
13. (可选) 输入说明。
14. 单击**完成**。
此时将显示一条同步消息，提示您确认操作。
15. 单击**是**可同步平面文件。
将显示一条消息指明同步已完成。要查看元数据更改的详细信息，请单击**显示详细信息**。
16. 单击**确定**。

在 Informatica Analyst 中同步关系数据对象

您可以将外部关系数据源的更改与其表数据对象进行同步。外部数据源更改包括添加、更改和删除源列和规则列。

1. 打开**库**工作区。
2. 在**项目**部分中，选择项目中的表数据对象。
Analyst 工具将在**属性**选项卡中显示表属性。
3. 从“操作”菜单中，单击**同步**。
此时将显示一条消息，提示您确认操作。
4. 要完成同步过程，请单击**是**。
将显示同步状态消息。
5. 将显示一条消息指明同步已完成。
要查看元数据更改的详细信息，请单击**显示详细信息**。
6. 单击**确定**。

查看数据对象

您可以查看项目或文件夹中每个数据对象的属性。您可以在选项卡中打开数据对象以预览数据。您可以预览数据对象的内容和对象类型以查看数据的结构并分析数据质量结果。

1. 打开**库**工作区，然后从**项目**或**资产**部分中查找数据对象。
Analyst 工具会在资产列表中显示数据对象。
2. 选择数据对象。
Analyst 工具会显示平面文件数据对象或表数据对象的属性。
注意: Analyst 工具使用以下格式显示系统参数的文件路径：\$\$[Parameter Name]/[Path]。例如，“\$\$SourceDir/ff_dept.txt”。
3. 单击**数据预览**选项卡。
Analyst 工具会在**数据预览**面板中显示平面文件数据对象或表数据对象的数据预览。

编辑数据对象

您可以在查看表和平面文件期间编辑表和平面文件的名称和说明属性。

1. 打开**库**工作区，然后从**项目**或**资产**部分中查找数据对象。
Analyst 工具会在资产列表中显示数据对象。
2. 选择数据对象。
Analyst 工具会显示平面文件数据对象或表数据对象的属性。
3. 单击**数据预览**选项卡。
Analyst 工具会在**数据预览**面板中显示平面文件数据对象或表数据对象的数据预览。
4. 从“操作”菜单中单击**编辑**可编辑数据对象。
此时将显示**编辑**对话框。
5. 输入名称和可选说明。
(可选) 对于表数据对象，输入所有者名称。
6. 单击**确定**。

第 8 章

搜索

本章包括以下主题：

- [搜索概览, 59](#)
- [搜索结果, 59](#)

搜索概览

您可以搜索诸如数据对象、映射规范和配置文件等资产。使用 Analyst 工具表头中的搜索框可执行搜索。您可以将搜索结果限制为某个工作区或您有特权访问的所有工作区。

必须启用搜索服务才能在 Analyst 工具中执行搜索。

从 Analyst 工具表头执行搜索时，您所在的工作区的底部会显示一个搜索面板。如果按工作区进行搜索，则搜索面板的名称会显示为**搜索 <工作区名称>**，或者如果搜索所有工作区，则显示为**搜索全部**。您可以关闭搜索面板。

您可以在搜索面板的**搜索框**中输入另一搜索查询。Analyst 工具会显示找到的结果数并列出搜索结果。

您可以从**筛选器**面板应用搜索筛选器以精简搜索结果。还可以优化搜索查询以使用关键字匹配、通配符和运算符。

搜索结果

执行搜索时，Analyst 工具会在搜索面板中显示搜索结果数并列出搜索结果。搜索面板显示在您所在的工作区的底部。

搜索结果包括资产、相关资产、业务术语和策略。结果还可以包括来自剖析仓库的列配置文件结果和域发现结果。

您可以应用筛选器以精简搜索结果。从搜索面板中的**筛选器**面板为搜索结果应用筛选器。在将筛选器应用到搜索结果时，可以配置筛选器属性。可以隐藏**筛选器**面板或再次将其打开。

可以在搜索结果中按资产属性对资产进行排序或分组。可以从搜索结果中选择资产并在其工作区中打开资产。

提示: 如果搜索不返回结果，可能是您没有权限查看工作区中的项目。请咨询您的管理员以确认您是否具有工作区中项目的“读取”特权。

搜索查询

使用关键字匹配、通配符或运算符优化搜索查询。

您可以在搜索查询中使用以下字符：

关键字

在搜索中使用完全关键字匹配。将搜索查询用引号 (" ") 括起来以搜索完全关键字匹配。Analyst 工具会返回名称与关键字完全匹配的资产。

通配符

在搜索中使用 * 和 ? 通配符。使用通配符可在搜索中定义一个或多个字符。在搜索中使用通配符作为后缀或中缀。

. 代表字符。例如，在搜索 customer 时，Analyst 工具会返回 customer、customer_name 和 CustomerID。将 * 与至少一个字符配合使用。不能在搜索查询的开头使用 *。

?. 代表单个字符。例如，在搜索 Customer? 时，Analyst 工具可返回 Customer1、Customer2 和 CustomerA。

运算符

在搜索中使用 + 或 Space 运算符。

+. 包括搜索字词。例如，要包括销售和数据，请使用以下查询：+sales +data.

Space. 包括搜索字词中的任何一个。例如 sales data。

搜索属性

您可以应用筛选器在搜索结果中搜索资产。如果不需要添加筛选器，则可以隐藏**筛选器**面板。业务词汇表用户可以指定其他资产状态筛选器。

您可以使用以下筛选器属性：

搜索

在**筛选器**面板的**查找**框中输入搜索字符串。

位置

资产在业务词汇表或存储库中的位置。

时间(上次更新)

上次更新资产的时间。可以选择以下时间：

- 从头开始
- 过去 1 小时
- 过去 24 小时
- 过去一周
- 过去一个月
- 过去一年

创建者

至少创建了资产列表中的一项资产的用户的名称。选择**全部**可选择所有用户。

附录 A

配置 Web 浏览器

本附录包括以下主题：

- [配置 Web 浏览器 , 61](#)

配置 Web 浏览器

您可以使用 Microsoft Internet Explorer 或 Google Chrome 在 Informatica 平台中启动 Analyst 工具。

要使用 Analyst 工具，请在浏览器中配置以下选项：

脚本和 ActiveX

在 Microsoft Internet Explorer 上启用以下控制项：

- 活动脚本
- 允许程式剪贴板访问
- 运行 ActiveX 控件和插件
- 标记为脚本安全的脚本 ActiveX 控件

要配置控件，单击 **工具 > Internet 选项 > 安全 > 自定义级别**。

受信任站点

配置浏览器以允许访问 Analyst 工具。在 Microsoft Internet Explorer 中，将 Analyst 工具 URL 添加到受信任站点列表。在 Google Chrome 中，将 Analyst 工具主机名添加到受信任站点的白名单。

索引

B

- 版本控制 [47](#)
- 标记
 - 创建 [14](#)
 - 分配 [15](#)
 - 删除 [15](#)
- 标识符
 - 分隔 [34](#)
 - 正常字体 [34](#)
- 表数据对象
 - 添加 [54](#)
 - 同步 [57](#)
- 表头
 - Informatica Analyst [9](#)

C

- 常规标识符
 - 数据库连接 [34](#)

F

- 访问库工作区
 - 作业状态 [13](#)
- 分隔标识符
 - 数据库连接 [34](#)

G

- 工作区
 - 词汇表安全工作区 [10](#)
 - 词汇表工作区 [10](#)
 - 发现工作区 [10](#)
 - Informatica Analyst [10](#)
 - 结果卡工作区 [10](#)
 - 库 [13](#)
 - 连接工作区 [16](#)
 - 启动工作区 [10](#)
 - 设计 [10](#)
 - 数据域工作区 [10](#)
 - 项目工作区 [41](#)
 - 异常项工作区 [10](#)
 - 作业状态工作区 [38](#)

H

- HDFS 连接
 - 属性 [32](#)
- Hive 连接
 - 属性 [29](#)

I

- IBM DB2 connections
 - properties [17](#)
- Informatica Analyst
 - 表头 [9](#)
 - 工作区 [10](#)
 - 界面 [8](#)
 - 资产 [11, 45](#)
- Informatica Analyst 界面
 - 登录 [12](#)

J

- JDBC 连接
 - 属性 [20](#)
- 基于团队的开发 [47](#)
- 界面
 - Informatica Analyst [8](#)

K

- 库
 - 工作区 [13](#)
 - 库任务 [14](#)

L

- 连接
 - 数据库标识符属性 [33](#)
 - 数据库连接 [36](#)
- 连接工作区
 - 编辑 [37](#)
 - 创建 [36](#)
 - 删除 [37](#)
 - 搜索 [36](#)

M

- MS SQL Server 连接
 - 属性 [22](#)
- 模型存储库
 - 基于团队的开发 [47](#)
 - 签出和签入资产 [47](#)
 - 受版本控制 [47](#)
 - 无版本控制 [47](#)
 - 说明 [45](#)

O

ODBC 连接
属性 [26](#)
Oracle 连接
属性 [27](#)

P

平面文件数据对象
导入 [50](#)
平面文件选项 [50](#)
日期时间数据类型 [51](#)
数据类型 [50](#)
带分隔符 [52](#)
固定宽度 [53](#)
同步 [55](#)

Q

签出和签入资产 [47](#)

S

数据对象
表数据对象 [54](#)
平面文件数据对象 [49](#)
资产 [49](#)
编辑 [58](#)
查看 [57](#)

数据库连接

标识符属性 [33](#)

搜索

筛选器属性 [60](#)

搜索结果 [59](#)

搜索语法 [60](#)

X

项目工作区

访问 [41](#)

管理项目 [42](#)

权限 [43](#)

项目权限 [43](#)

项目权限

有效权限 [44](#)

直接权限 [43](#)

Z

作业状态

访问库工作区 [13](#)

监视 [40](#)

属性 [39](#)

作业状态工作区

访问 [38](#)

资产

Informatica Analyst [11](#), [45](#)

搜索 [59](#)