

Informatica PowerCenter Data Cleanse and Match 选项

标准化、验证并更正企业名称和地址数据

Dave Ramos

商务智能和分析部门

LinkShare

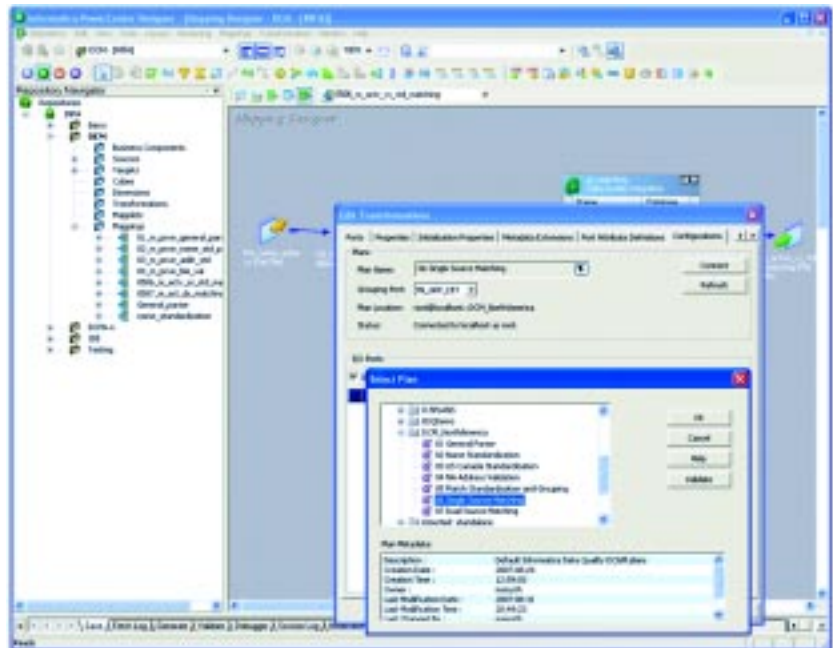
“令我们印象深刻的是，Informatica 能够持续提供产品线上的各种功能，满足甚至预测企业乃至技术的需求。版本 8.5 继承了该传统。此版本有太多方面表明这是一个愿意倾听客户意见，然后尽力提供更多服务的公司。”

PowerCenter® Data Cleanse and Match 选项，为需要简单处理数据集成项目中的名称和地址数据质量问题的企业提供了一个理想起点。为了将数据质量方案主动扩展，以便涵盖所有数据类型和企业数据，Informatica 为所有 Data Cleanse and Match 选项客户提供了升级至 Informatica Data Quality 的简单途径。

此选项具有强大的算法：数据匹配功能以确定各数据记录间的关系，以便进行重复数据删除或基于组的处理。通过利用 Informatica Data Quality 的数据分析和增强功能以及 PowerCenter 的数据集成功能，该选项可使您轻松将数据质量流程添加至 PowerCenter 转换。它使为地址验证、标准化和重复数据删除设计的预建计划来提高您邮政地址数据的质量。

优势

- 提高了名称和地址数据的准确性，从而建立了对数据质量的信心
- 提高了开发人员的工作效率，减少了清洗和转换数据所需的工作量
- 通过消除数据重复保持整个企业内数据的一致性和完整性



Data Cleanse and Match 选项使 IT 企业能够轻松将用于名称和地址标准化、验证及重复数据删除的数据质量流程嵌入 PowerCenter 转换。

Data Cleanse and Match 选项的关键功能

数据清洗和解析

- 标准化、验证、增强和更正客户联系信息（名称和地址）以及公司或第三方数据。
- 隔离数据并将结构应用到每个数据元素。
- 帮助将数据转换成其它格式、结构和内容以创建所需的输出记录。
- 使用拉丁文和非拉丁文字符集，例如俄语、波兰语、汉语和日语。
- 帮助比较地址信息和美国及加拿大邮政服务目录。

数据匹配

- 确定各数据记录间的关系，以便进行重复数据删除或基于组的处理。
- 并行处理多组业务规则。

在整个数据集成平台中全面集成

- 利用 PowerCenter 体系结构的性能和可扩展性，在清洗和匹配数据时自动利用平台的所有复杂并行能力和网格功能。
- 充分利用 PowerCenter 的元数据管理功能以提供数据血缘报告和审计。
- 使用企业范围数据分析、清洗、匹配和监控功能，提供升级至 Informatica Data Quality 的简单途径。

Data Cleanse and Match 选项的优势

提高了名称和地址数据的准确性， 从而建立了对数据质量的信心

该选项具有丰富的解析功能，使客户文件中的数据元素得以识别并分成单独的字段。Data Cleanse and Match 选项提高了名称和地址数据的准确性，使最终用户对其数据质量充满信心，同时也降低了邮寄成本。

提高了开发人员的工作效率， 减少了清洗和转换数据所需的工作

包括 CASS 认证的地址验证和增强计划的预建数据质量计划减少了清洗和转换数据所需的开发工作，提高了开发人员的工作效率并且降低了培训要求。开发人员可直接利用 PowerCenter 工具栏中的数据清洗和转换工具。将数据清洗业务规则写入数据集成平台并从中读取。开发人员可以访问可重复使用的数据清洗中的特定转换，用于在一次操作中集成和清洗数据。

通过消除数据重复保持整个企业内数据的一致性和完整性

与数据质量欠佳相关的主要问题之一是类似记录间的数据重复问题。该选项使 IT 团队能够识别和删除重复数据，有助于确保数据在整个企业内的一致性和准确性。Data Cleanse and Match 选项使用强大的“模糊逻辑”和特定的业务定义标准来搜索类似记录的现有记录。匹配标准和业务规则消除了有关两个记录是否指向同一个人、客户或家庭的所有疑问。这些功能简化了执行此重复数据删除过程的逻辑的创建。

关于 Informatica

Informatica 是全球领先的企业数据集成和数据质量软件服务提供商。借助 Informatica，企业可通过集成其内部的所有信息资产获得更大的商业价值。拥有全球愈 3,600 家客户，他们依赖 Informatica 降低成本并加速解决任何复杂程度和规模的数据集成需求。



www.informatica.com.cn

北京办事处

地址：北京市朝阳区建国门外大街乙 12 号
LG 双子座大厦东塔 19 层 1906 室
邮编：100022
电话：86-10-5879 3366
传真：86-10-5879 3130

上海办事处

地址：上海市浦东世纪大道 201 号
渣打银行大厦 5 楼
邮编：200120
电话：86-21-6182 6825
传真：86-21-6182 6755

广州办事处

地址：广州市天河区体育东路 138 号
金利来数码网络大厦 3 层 321 室
邮编：510620
电话：86-20-2886 0646
传真：86-20-3878 1801