

Information Management software

IBM InfoSphere QualityStage

分析、清理和管理高质量数据，实现更出色的业务成果



要点

- 分析数据, 确定数据质量水平现状和质量问题
- 对核心业务实体实施标准化、匹配和数据生存规则
- 根据组织的业务规则匹配客户、供应商、产品和位置数据, 在整个组织实现准确、一致的视图
- 在高度可扩展的并行平台上处理全局数据, 实现
- 最优性能
- 向关键企业计划提供可靠、高质量的数据, 同时实现成功的批处理架构和面向服务架构(SOA)

最充分利用企业的信息资产

企业需要从其操作系统中的海量信息中获取价值。能否清晰了解客户、产品、合作伙伴和提供商, 构成了决定能否发展业务并获得竞争成功的决定因素。没有有条理、标准化且准确的数据, 这种清晰理解将无法实现。而糟糕的数据质量会侵蚀和破坏关键业务计划, 比如信息治理、合规和主数据管理。

然而, 大部分企业都还未改进他们的流程、策略和基础架构来实现高数据质量。要解决这一需求, 企业需要更多地采用信息治理, 这是一种质量控制渠道, 能够使定义通用术语和管理、使用、改进及保护信息的流程更加严谨。有效的信息治理可支持跨组织协作和结构化的策略制定, 增强了公司数据的质量、可用性和完整性。

数据质量是信息治理的重要部分, 是IBM® InfoSphere™ Information Server平台上的一个核心领域, 有助于实现一致、准确、可信的信息交付。借助InfoSphere Information Server数据整合平台, IBM提供了广泛的数据质量功能, 从数据探查、标准化和匹配到主动数据质量监控。

企业在不同的时间点专注于数据质量的不同方面, InfoSphere Information Server提供了多项功能来解决其中每个接触点的数据质量需要。

IBM InfoSphere QualityStage™是InfoSphere Information Server数据整合平台的一部分, 它专注于清理数据: 它使企业能够创建和维护关键实体的准确视图, 包括客户、供应商、地点和产品。核心InfoSphere QualityStage功能包括数据分析、标准化、地址验证、概率匹配、数据生存和数据扩充。InfoSphere QualityStage可以部署在事务性、操作性或分析环境中, 可采用批量或实时部署形式。

IBM InfoSphere Information Analyzer也是InfoSphere Information Server平台的一部分, 它提供了另一组数据质量增强功能来帮助客户理解、分析和监控数据。借助整合的规则分析、异常管理和直观的用户界面, 客户可以保持高质量的数据来帮助实现业务目标。关于InfoSphere Information Analyzer的更多信息, 请访问:

ibm.com/software/data/integration/information-analyzer

InfoSphere QualityStage:一种实现数据质量收益的途径

InfoSphere QualityStage旨在提供高质量数据和帮助组织收获相关的收益, 包括:

- 改善的投资回报(ROI)
 - 实现企业资源规划(ERP)、客户关系管理(CRM)、数据仓库、商业智能、主数据管理和其他战略IT计划的时间、成本和风险更少
 - 整洁、整合的客户和内部视图, 支持交叉销售和向上销售工作
 - 改进的客户支持和服务, 拥有识别盈利能力最高客户的能力
 - 整合的提供商、部件和产品视图, 可实现更高效的分析、采购和库存管理
 - 与更广泛的InfoSphere Information Server数据整合平台紧密整合, 实现一种将数据质量作为数据整合的关键组件的整体方法
- 图1和2给出了InfoSphere QualityStage帮助标准化和转换数据的示例。

标准化部件							
输入文件:							
Operation Work Instruction WING ASSY DRILL 4 HOLE USE 5J868A HEXBOLT 1/4 INCH WING ASSEMBLY, USE 5J868-A HEX BOLT .25" - DRILL FOUR HOLES USE 4 5J868A BOLTS (HEX .25) - DRILL HOLES FOR EACH ON WING ASSEM RUDDER, TAP 6 WHOLES, SECURE W/KL2301 RIVETS (10 CM)							
结果文件:							
Assembly	Instruction	Qty	Type	Part	Size	Measure	SKU
WING	DRILL	4	HOLES	HEXBOLT	.25	INCH	5J868A
WING	DRILL	4	HOLES	HEXBOLT	.25	INCH	5J868A
WING	DRILL	4	HOLES	HEXBOLT	.25	INCH	5J868A
RUDDER	DRILL	6	HOLES	RIVET	10	CM	KL2301

图1. 示例产品部件标准化

经典转换:帐户到客户						
帐户视图						
Source	Legacy Key	Name	Address	Phone	Birth Date	Cust-ID
Life	70328574	John Smith Jr.	10 Main St Boston MA 02110	781-259-9945	02/05/1940	
Home	80328575	Mr. John Smith	10 Main St Unit 10 Boston MA 02111	617-259-9000		
Auto	90238495	J. Smyth	Main St Bostan Mass 02110	781-295-9945	02/05/1941	
客户视图 链接相关记录来创建交叉引用ID						
Source	Legacy Key	Name	Address	Phone	Birth Date	Cust-ID
Life	70328574	John Smith Jr.	10 Main St Boston MA 02110	781-259-9945	02/05/1940	0001
Home	80328575	Mr. John Smith	10 Main St Unit 10 Boston MA 02111	617-259-9000	0001	
Auto	90238495	J. Smyth	Main St Bostan Mass 02110	781-295-9945	02/05/1941	0002
客户个人资料 使用来自所有来源的最佳信息创建客户个人资料						
Source	Legacy Key	Name	Address	Phone	Birth Date	Cust-ID
CP		Mr. John Smith Jr.	10 Main St Unit 10 Boston MA 02111	617-259-9000	02/05/1940	0001
CP		J. Smyth	Main St Bostan Mass 02110	781-295-9945	02/05/1941	0002

图2. 数据转换示例

企业必须确保战略系统提供整个企业的业务用户可以信赖的准确、全面的信息。通过其易于使用、可自定义的用户界面, InfoSphere QualityStage帮助业务用户获得对国际名称和地址, 电话号码、生日、电子邮件地址及其他描述性评论字段等相关数据的控制能力。InfoSphere QualityStage使用高度准确的概率匹配算法来匹配数据元素, 发现它们之间的关系(企业中和Internet环境中), 以及实现批量和实时处理。

从分散的源数据到有关核心业务实体的高质量信息

通过执行字符级分析, InfoSphere QualityStage有助于揭示不规则和被埋藏的数据, 然后转换它以用于数据库加载或事务处理。首先, 来自不同来源的数据标准化为固定字段, 业务驱动的规则向输入数据分配正确的语义含义, 以方便匹配。

接下来, InfoSphere QualityStage的强大匹配功能可检测数据中的重复和关系(无论是不规则、不一致还是丢失数据值)。InfoSphere QualityStage独有的统计匹配引擎两个或更多数据值

集合引用相同业务实体的概率进行评估, 从而提供极其准确的匹配结果。这些功能在一个整合的设计环境中通过转换技术来提供, 有助于将数据质量嵌入到关键信息整合流程中。

一旦确认了匹配, InfoSphere QualityStage就构建链接关键字, 使用户可以完成事务或加载具有真正实体完整性的目标系统, 并将相关数据作为信息而查看。通过在初始加载、系统更新和实时数据输入期间使用InfoSphere QualityStage的数据质量增强功能, 公司能够访问任何个体或业务实体及其在整个企业中的关系的准确、一致且统一的视图。这种强大的匹配和数据清理发生在一个可扩展的并行处理框架中, 提供了专为已扩展企业的需求量身订造的一流性能。

InfoSphere QualityStage特性

- 易于使用、整合且直观的点击式用户界面, 用于指定自动化的数据质量流程: 数据分析、标准化、匹配和生存
- 增强的匹配设计器工具实现了更轻松的设置和更高的灵活性
- 全球地址清理、验证、认证(针对具体的位置)和地理定位

- 标准化和匹配报告, 获得了对您的数据质量流程的更深入洞察, 改进了部署质量
- 额外的标准化规则用于设置拉丁美洲、荷兰和印度, 以及传统的中文和日本假名等覆盖范围
- 产品数据的规则集加速
- SOA, 用于为实时部署创建数据质量服务
- 数据质量服务
- 强大、准确的匹配, 以概率匹配技术和易于设置和维护的全面的模糊匹配功能为基础
- 严格、科学的匹配判断, 以及轻松的审计和验证
- 有效的运行时和系统资源使用及高度可伸缩性
- 与其他InfoSphere Information Server功能全面整合, 包括共享的元数据, 数据监控、探查和转换

统一平台中的高品质数据

作为InfoSphere Information Server平台的一部分, InfoSphere QualityStage在一个完整的信息整合平台上下文中提供了重要的数据质量功能。它利用统一的安装、部署和源控制来实现快速启动, 使用统一的数据质量和转换功能(结合IBM InfoSphere DataStage[®])帮助减少整合项目的开发时间, 帮助确保所交付数据的质量。

整个InfoSphere Information Server平台中有效的共享数据有助于简化在整个整合范围内对元数据的收集和管理。来自InfoSphere Information Analyzer的元数据可在InfoSphere QualityStage内共享和利用, 实现卓越的协作。这种整合水平可带来重大的收益, 包括对信息一致性的更高自信, 以及在整个InfoSphere Information Server上执行影响分析的能力。

数据质量和信息治理

信息治理可增强公司数据的质量、可用性和完整性, 促进跨组织的协作和结构化的决策制定。只要一致地应用, 它可帮助平衡派系孤岛与组织利益, 直接影响任何业务的4个最重要的目标: 增长收入, 降低成本、减少风险和提高对其数据的自信。此外, 信息治理使组织能够以端到端系统的形式监控其信息供应链, 帮助确保信息得到了一致定义和良好理解, 可靠且具有高质量, 在其整个生命周期得

到了管理, 以及在任何地方都受到了保护。

InfoSphere Information Server, InfoSphere QualityStage和InfoSphere Information Analyzer为企业提供了制定和启用信息治理策略所需的数据质量功能。

信息治理论坛

较之以前, 数据保护和管理是一个更加普遍的业务问题。为了帮助企业更好地理解新兴的信息治理领域, IBM于2004年11月为关注信息治理问题的首席数据官以及安全、风险、合规和隐私官员创建了一个领导者论坛。

从那时起, IBM信息治理委员会就在稳步发展, 现已包含近55家领先企业、大学和IBM业务合作伙伴, 包括大型金融机构、电信企业、零售商和政府机构。该委员会设计了一个框架来帮助业务人员理解信息治理的核心和支持学科, 以及推动力。它还制定了一个成熟度模型来帮助评估企业内的信息治理。为了拓宽该委员会所涉及的范围, IBM启动了一个在线社区来鼓励企业参与, 使用众包技术进一步增强该成熟度模型和整体信息治理。

关于IBM信息治理委员会的更多信息, 请访问:

www.infogovcommunity.com

InfoSphere Information Server带来了价值

企业面临着持续的信息挑战: 它在何处? 我在需要它时如何以我需要的形式获取它? 我能否信任它? 我如何控制它? 这些阻碍会继续加剧, 如果业务人员无法保证他们能够访问权威、一致、及时且完整的信息。

InfoSphere Information Server是一个市场领先的数据整合平台, 能够帮助企业从分布在各种系统的复杂、异构信息中发现更多价值。它支持组织整合不同的数据, 并在需要的时间和地点、在相关环境中有序地为特定人员、应用程序和过程交付可信的信息。它可帮助业务和IT人员协作理解任何来源范围的任何类型数据的含义、结构和内容。它提供了突破性的效率和性能, 能够在整个企业中一致且安全地清理、转换和迁移此信息, 所以它可以全新的方式访问和使用来促进创新, 提高操作效率和帮助降低风险。



更多信息

要了解InfoSphere QualityStage的更多信息, 包括详细的硬件和软件系统需求, 请联系您的IBM营销代表或IBM业务合作伙伴, 或者访问:ibm.com/software/data/infosphere/qualitystage

关于IBM数据质量解决方案的更多信息, 请访问:

ibm.com/software/data/integration/capabilities/cleanse.html

要了解InfoSphere Information Server或其他IBM信息整合解决方案的更多信息, 请联系您的IBM营销代表或IBM业务合作伙伴, 或者访问:ibm.com/software/data/integration

© 版权所有IBM Corporation 2011

IBM Software Group
Route 100
Somers, NY 10589
U.S.A.

在中国印刷
2011年8月
保留所有权利。

IBM、IBM徽标、ibm.com、InfoSphere和QualityStage是国际商业机器公司在美国和/或其他国家/地区的商标或注册商标。如果这些商标和其他IBM商标在本文中第一次出现时标注了商标符号(®或TM), 均代表在本文出版之际, 它们是IBM在美国注册的商标或约定俗成的商标。此类商标也可能是其他国家/地区的注册商标或普通法规定的商标。IBM商标的最新列表可在网络上获得, 请访问:ibm.com/legal/copytrade.shtml的“Copyright and trademark information”部分。

其他公司、产品或服务名称可能是其他公司的商标或服务标记。

本出版物中对IBM产品或服务的引用, 不代表它们可用于所有IBM运营的国家/地区。关于IBM未来方向或打算的声明仅代表IBM的发展目标, 如有变更, 恕不另行通知。



可回收, 请回收再利用