



主数据管理（MDM）

应用指南

主数据管理（MDM）应用指南

随着业务发展以及监管的需要，企业对主数据的实时性、准确性、一致性有了更高的要求，主数据管理 (Master Data Management，MDM) 也应运而生，它是指一组约束和方法用来保证一个企业内主题域和系统内相关数据和跨主题域和系统的相关数据的实时性、含义和质量。本次的技术手册对主数据管理（MDM）进行了一个比较系统的介绍，从基本定义概念，到实施主数据管理的最佳实践，我们将带您领略更多 MDM 的知识。

主数据管理基础知识

简单而言，主数据管理（MDM）是公司权衡和重用常见、和准确业务数据的一种全面的方法。不管你是否考虑实现新的 MDM 系统还是评估可能的工具和软件来帮助你目前的数据管理和数据整合主动性，了解基础的知识都是十分必要的。

- ❖ 主数据管理详解（一）
- ❖ 主数据管理详解（二）
- ❖ 主数据管理详解（三）
- ❖ 最常见的 13 种主数据管理（MDM）词汇和定义

主数据管理最佳实践

在考虑像主数据管理（MDM，Master Data Management）这样的学科时，寻找已经经历过或做过主数据管理的过来人咨询是唯一自然的途径。因此在本部分中，我们将请相关人员讲述主数据管理的最佳实践。

- ❖ 主数据管理（MDM）的七个最佳实践（上）
- ❖ 主数据管理（MDM）的七个最佳实践（下）

SQL Server 2008 R2: 主数据服务

Master Data Services, 也称为 MDS, 是 SQL Server 2008 R2 增加的 关键商业智能特性 之一。Master Data Services 的基本目标是为企业信息提供单个权威来源。然后这个信息可以被其它应用和数据使用, 这样您的环境中每一个应用便都是使用相同信息的同一份权威副本。

- ❖ SQL Server 2008 R2 最新功能: 主数据服务
- ❖ Master Data Services 的潜在问题
- ❖ 主数据服务的相关业务案例
- ❖ Master Data Services 存在可提升的空间
- ❖ Master Data Services 入门: 配置 MDS
- ❖ Master Data Services 入门: 创建数据模型
- ❖ Master Data Services 进阶指导: 加载数据
- ❖ Master Data Services 进阶指导: 处理和验证数据

MDM 专家指导

利用现有的数据质量工具, 你也许可以进行数据标准化的工作, 但是数据质量工具不可能满足所有 MDM 项目的需求。那么在进行 MDM 项目时, 应该注意哪些问题呢? 专家的指导可以起到很大的帮助。

- ❖ 是否可以使用数据质量工具执行 MDM 项目
- ❖ 在 MDM hub 与交易数据库之间保持参照完整性
- ❖ 数据仓库的未来: 开源与 MDM

主数据管理详解（一）

主数据是指在整个企业范围内各个系统(操作/事务型应用系统以及分析型系统)间要共享的数据，比如，可以是与客户(customers)， 供应商(suppliers)， 帐户(accounts)以及组织单位(organizational units)相关的数据。主数据通常需要在整个企业范围内保持一致性(consistent)、完整性(complete)、可控性 (controlled)，为了达成这一目标，就需要进行主数据管理(Master Data Management ，MDM)。需要注意的是，主数据不是企业内所有的业务数据，只是有必要在各个系统间共享的数据才是主数据，比如大部分的交易数据、帐单数据等都不是主数据，而像描述核心业务实体的数据，而像客户、供应商、帐户、组织单位、员工、合作伙伴、位置信息等都是主数据。主数据是企业内能够跨业务重复使用的高价值的数据。这些主数据在进行主数据管理之前经常存在于多个异构或同构的系统中。

主数据管理(Master Data Management ，MDM)是指一组约束和方法用来保证一个企业内主题域和系统内相关数据和跨主题域和系统的相关数据的实时性、含义和质量。这是从深层次来说来说明主动主数据管理(MDM)的深度和复杂性，简单的说，主数据管理(MDM)保证你的系统协调和重用通用、正确的业务数据(主数据)。通常，我们会把主数据管理作为应用流程的补充，通过从各个操作/事务型应用以及分析型应用中分离出主要的信息，使其成为一个集中的、独立于企业中各种其他应用核心资源，从而使得企业的核心信息得以重用并确保各个操作/事务型应用以及分析型应用间的核心数据的一致性。通过主数据管理，改变企业数据利用的现状，从而更好地为企业信息集成做好铺垫。

主数据管理(MDM)可以帮助我们创建并维护整个企业内主数据的单一视图(Single View)，保证单一视图的准确性、一致性以及完整性，从而提供数据质量，统一商业实体的定义，简化改进商业流程并提供业务的响应速度。从变化的频率来看，主数据和日常交易数据不一样，变化相对缓慢，另外，主数据由于跨各个系统，所以对数据的一致性、实时性以及版本控制要求很高。

主数据管理其实在很早之前就一直存在，只不过现在随着业务发展以及监管的需要，对主数据的实时性、准确性、一致性有了更高的要求，才被业界广泛接受，各个厂商相应的推出了一系列的主数据管理集成与基础套件以及特定领域的解决方案。近年来最明显的变化是，客户在以前的时候经常问的问题是：“主数据管理是什么？”，而现在客户经常问的问题演变成了：“我们的业务的确存在一些问题，主数据管理正好可以解决这个问题，我们怎么开始？”。与以前相比，客户对主数据管理(MDM)的认识有了巨大的进步，并开

始尝试用主数据管理 (MDM) 解决他们在整个企业范围内进行跨业务、跨主题域时遇到的各种挑战和问题：比如税务行业，税务局在按纳税人在一些分析统计时，就发现关于纳税人的基本信息分布在核心征收管理系统、发票管理系统、个人所得税系统、增值税管理系统等多达几十个系统中，使得统计分析变得困难起来，在比如在医疗设备公司，由于没有按照供应商进行产品层次的分类，各个产品的描述也很不一样，使得产品目录的维护十分困难。随着业务的发展，对各行各业来说，生成并维护一个统一的主数据系统变的十分迫切和必要，特别是对一些跨国公司，如何在不同的地区(各个国家和地区) 的业务系统之间维护关于客户、产品目录、供应商等信息的单一视图更是重要。

需要注意的是，主数据 (Master Data) 和元数据 (Meta Data) 是两个完全不同的概念。元数据是指表示数据的相关信息，比如数据定义等，而主数据是指实例数据，比如产品目录信息等。比如，某省地税开发了一套征收管理软件，以市为单位部署了 17 套，每套征收管理软件中的元数据都是一样的，但是主数据还是需要进行管理的。主数据管理和传统数据仓库解决方案不是一个概念，数据仓库会将各个业务系统的数据集中在一起在进行业务的分析，而主数据管理系统不会把所有数据都管理起来，只是把需要在各个系统间共享的主数据进行采集和发布。相对于传统数据仓库解决方案的单向集成，主数据管理正注重将主数据的变化同步发布到各个关联的业务系统中(主数据管理数据是双向的)。

主数据管理问题存在的根源

对于大多数的企业都存在主数据管理的问题，个人以为这是由于业务发展的渐进性以及 IT 技术发展的渐进性造成的，正是由于这种渐进性，各大企业的业务系统从经历了从无到有，从简单到复杂，从而形成了一个又一个的业务竖井。从根本上来说，不可能只使用一个业务系统就能覆盖企业的所有业务，即便对一些国际大型的公司提供的套件来说也是一个不可能完成的任务(即便对套件来说，经常也存在一个跨国企业在不同的国家或地区部署多个实例的现象，也就是没有集中部署该套件，而是在很多地方分散部署了该套件)。对企业来说，业务系统的构建更多是以项目为中心，从下而上的构建系统，而不是至上而下的构建系统，必然缺乏整个企业范围内的统一规划，从而使得一些需要在各个业务中共享的数据(主数据)被分散到了各个业务系统进行分别管理。分散管理的主数据由于没有不具备一致性、准确性、完整性，使得各个企业普遍存在着产品管理不力、供应商管理不力、订单管理不力等现象。解决这一问题的根本方法就是引入主数据管理 (MDM)，主数据不光指需要共享的数据，更包含需要共享的业务规则和策略。

(作者：程永 / 王雪梅 来源：互联网)

原文标题：主数据管理详解（一）

链接: http://www.searchdatabase.com.cn/showcontent_17542.htm

主数据管理详解（二）

主数据管理 (MDM) 的成熟度

根据主数据管理实施的复杂程度，参照 Jill Dyché, Evan Levy 的观点大体可以把主数据管理可以分为五个层次，从低到高反映了主数据管理 (MDM) 的不同成熟度。下面我们简单介绍一下这五个层次：

Level 0：没有实施任何主数据管理 (MDM)

在 Level 0 的情况下，意味着企业的各个应用之间没有任何的数据共享，整个企业没有数据定义元素存在。比如，一个公司销售很多产品，对这些产品的生产和销售由多个独立的系统来处理，各个系统独立处理产品数据并拥有自己独立的产品列表，各个系统之间不共享产品数据。在 Level 0，每个独立的应用负责管理和维护自己的关键数据（比如产品列表、客户信息等），各个系统间不共享这些信息，这些数据是不连通的。

Level 1：提供列表

不管公司大还是小，列表管理是我们常用的一种方式。在公司内部，会通过手工的方式维护一个逻辑或物理的列表。当各个异构的系统 and 用户需要某些数据的时候，就可以索取该列表了。对于这个列表的维护，包括数据添加、删除、更新以及冲突处理，都是由各个部门的工作人员通过一系列的讨论和会议进行处理的。业务规则 (Business Rules) 是用来反映价值的一致性，当业务规则发生改变或者出现类似的情况时，这样高度手工管理的流程容易发生错误。由于列表管理是通过手工管理的，其列表维护的质量取决于谁参加了变更管理流程，一旦某人缺席，将会影响列表的维护。

MDM Level 1 比 MDM Level 0 的不同就是，各个部门虽然还是独立维护各自的关键数据，但会通过列表管理维护一个松散的主数据列表，能够向其他各个部门提供其需要的数据。在 MDM Level 1 中，数据变更决定以及数据变更操作都是由人决定的，因此，只有人完成数据变更决定后才会变更数据。在实际情况中，虽然数据变更流程有严格的规定，但是由于缺乏集中的、基于规则的数据管理，当数据量比较大时，数据维护的成本会变的很高，效率也会很低。当主数据，比如客户信息、产品目录信息等数量比较少时，列表管理的方式是可行的，但是当产品目录或客户列表出现爆炸式增长以后，列表管理的变更流程将变得困难起来。MDM Level 1 依赖于人的协作。如果产品经理需要更新过后的产品价格列表，那需要联系 ERP 系统所有者，让其发送邮件给她。在企业范围内实现客户或产品

列表就如同维护不同部门之间人们的关系一样。如果客户或产品存在层次或分组，列表将很难提供，并且通常在 Level 1 因为过于复杂难以被管理。

Level 2：同等访问(通过接口的方式，各个系统与主数据主机之间直接互联)

MDM Level 2 与 MDM Level 1 相比，引入了对主数据的(自动)管理。通过建立数据标准，定义对存储在中央知识库(Central Repository)中详细数据的访问和共享，为各个系统间共享使用数据提供了严密的支持。中央知识库(Central Repository)通常会被称为“主数据主机(Master Data Host)”。这个知识库可以是一个数据库或者一个应用系统，通过在线的方式支持数据的访问和共享。

创建、读取、更新和删除 (CRUD) 是处理基本功能的典型编程术语。即便在 MDM 中，CRUD 处理也是基本功能。你的数据库如果仅仅支持 CRUD 处理并不意味着你实现了 MDM。MDM Level 2 引入了“同等访问”(peer-based access)，也就是说一个应用可以调用另一个应用来更新或刷新需要的数据。当 CRUD 处理规则定义完成后，MDM Level 2 需要客户或“同等”应用格式化请求(和数据)，以便和 MDM 知识库保持一致。MDM 知识库提供集中的数据存储和供应(provisioning)。在这个阶段，规则管理、数据质量和变更管理必须要在企业范围内作为附加功能定制构建。

比如，一个数据库或一个打包应用(比如一个销售自动化系统)对外部应用提供数据访问功能。当一个外部应用(比如呼叫中心应用)需要增加一个客户，这个外部应用将提交一个事务，请求数据所有者增加一个客户条目。主数据主机(Master Data Host)将增加数据并告知外部应用。CRUD 处理方式比纸上办公有了很大提高，其是基于会话的数据管理。在 MDM Level 1，数据变更是基于手工的方式。在 MDM Level 2，数据变更是自动完成的一通过由具体技术实现的标准流程，允许多应用系统修改数据。MDM Level 2 可以支持不同的应用使用和变更单一、共享的数据知识库。MDM Level 2 需要每个同等应用理解基本的业务规则以便访问主列表、与主列表进行交互。因此，每个同等应用必须正确恰当地创建、增加、更新和删除数据。授权应用有责任坚持数据管理原则和约束。

Level 3：集中总线处理

与 MDM Level 2 相比，MDM Level 3 打破了各个独立应用的组织边界，使用各个系统都能接受的数据标准统一建立和维护主数据(MDM Level 2 的主数据主机上存储的数据还是按照各个系统分开存储的，没有真正的整合在一起)。

集中处理意味着为 MDM 构建了一个通用的、基于目标构建的平台。大多数公司发现 MDM 正在挑战他们现有的 IT 架构：他们拥有太多的独立平台处理主数据。MDM Level 3

集中数据访问、控制跨不同应用和系统使用数据。这极大的降低了应用数据访问的复杂性，大大简化了面向数据规则的管理，使 MDM 比一个分散环境具有更多的功能和特点。企业主数据面临一致性的挑战。数据在不同的地方存在，数据所代表的含义也是不同的，数据的规则各个系统之间也是不一样的。集中 MDM 处理-通过一个公共的平台作为一个总线 (HUB)-说明一个共识，从多个系统整合主题域数据，意味着使用集中、标准化的方法转换异构操作数据，不管其在源系统中是什么样子，都会被整合起来。在 MDM Level 3，公司对主题域内容采用集中管理方式。这意味着应用系统，作为消费者或使用主数据，拥有一个共识就是数据是主题数据内容的映像，打破了各个独立应用的组织边界。MDM Level 3 支持分布主参考数据的存在。

MDM 的核心之一就是保证所有系统都能接受数据表示的唯一公认方法。这有点类似于语言翻译，通过其他语言的翻译，英语已经称为一个全球性的语言。在 MDM Level 3，一个公司可以让任意两个系统共享数据和说对方的语言。MDM Level 3 还降低了等同访问的复杂性。“消费”应用不再需要支持系统定位和操作逻辑。任何与源系统数据相关的分布式细节都会被 MDM 总线集中处理。在 MDM Level 3 自动数据标准意味着：建立目标数据值表示和通过必要的步骤提供精确的主数据值捕获。在所有的分类中从 MDM Level 3 开始第一次支持一致性的企业数据视图。数据质量规则在这里进行数据清洗和错误纠正。

Level 4：业务规则和政策支持

一旦数据从多个数据源整合在一起，主题域视图超越单独的应用并表现为一个企业视图，你将获得事实的单一版本。当事实的单一版本已经能够提供出来时，来自业务主管和执行人员的必然反应经常是：“证明它”。MDM Level 4 可以保证主数据反映一个公司业务规则和流程，并证实其正确性。MDM Level 4 通过引入主数据来支持规则，并对 MDM 总线以及其它外部系统进行完整性检查。由于多数公司相对比较复杂，影响业务数据访问和操作的规则以及策略 (rules and policies) 相对也比较复杂。假定任何一个单一系统可以包含并管理与主参考数据相关的各种类型的规则是不切实际的。因此，如果一个 MDM 总线真正打算提供企业范围内数据的精确性，工作流和流程整合的支持是必不可少的。

举例来说，在一个 HMO 内，需要多个应用来支持一个病人的护理。一个单一的访问 (visit) 可能包括入院、房间和床位分配、监控设备、化验、身体检查以及其他程序等。一旦一个病人准备离开医院，出院流程需要确保和这个病人相关的所有活动、资源都被结清。MDM 技术在召集多个应用系统一起保证病人辨识方面是十分有效的，处理是正确的。虽然病人辨识很重要，业务规则整合同样重要。临床系统依靠一系列的业务流程和数据规则来辨别所有显著的病人详细资料。这包括返回所有基于房间的资源 (监护设备、床位等) 以得到有用的详细目录，当病人要出院时分解其所有的费用。MDM 保证当 John Smith 出院

时，正确的房间和设备放入到该 John Smith 的详细目录中，而不是其他的 John Smith(正在另一个楼层做身体治疗)。

MDM 系统必须不仅支持基于规则的整合，还要能够整合外部的工作流。这些规则可能包括通过总线与临床系统交互或等待另一个系统或者人(有权限做出改变的人)审批。通过一个 MDM 总线，规则定义可以不仅局限在逻辑上，还可以依赖于其他系统的输入。当然，协调和审计数据意味着可以回退其他系统(或业务流程)来保证数据变化经过严格的审批，这样错误可以被发现并且事务在需要的时候可以被回滚。MDM Level 4 提出对规则和策略扩展性的支持。通过总线以一个灵活可持续的方式支持任何面向业务的规则集合这很重要。

比如，如果一个商店经理更新一个产品的价格，总线系统需要能够和一个可信系统(比如，商品管理系统)进行协商以便使规则生效。详细规则将支持另一个系统中存在产品价格的变更—总线需要能够理解能够处理和批准变更的权限系统或方法。这些规则可能涉及到复杂性或隐私限制，禁止它们直接在总线上存在。在 MDM Level 4，一个企业可以支持一套步骤或任务，在一个特殊的创建、读取、更新和删除任务被允许之前这些步骤或任务必须遵守。工作流自动化经常用来支持发生在总线上的事件或活动的授权。但是变更管理远远不仅仅是工作流：它可以包括基于逻辑的流程和基于人的决策。变更管理的存在可以支持动态业务，允许变更。举例说明，在 911 之前，任何人都可以在美国国内的航空公司运载货物。没有规定以外的其他某种形式的鉴定和付款方式。911 之后，美国联邦航空协会 (FAA) 指导建立了一个更加全面的规定，指示一个人是否被允许运载货物。在这个特殊的例子中，要求各个系统都部署 FAA 对托运人的要求是不现实的。部署一个规则管理系统，为所有的系统(包括 MDM 总线)集中托运人批准规则，更加容易实现(也更现实)。集中数据定义和标准化在 MDM Level 2 就已经引入，与 MDM Level 4 的集中规则管理相比，相对简单。业务流程越复杂、业务流程越多，对总线的需求就越多，以便对针对共同数据的跨职能、异构规则进行更好的支持。重要的是 MDM Level 4 支持集中规则管理，但是规则本身和相关的处理是可以分开的。换句话说，MDM 总线需要保证规则是集中应用的，即便这个规则是在总线外居住的。

Level 5：企业数据集中

在 MDM Level 5，总线和相关的主数据被集成到独立的应用中。主数据和应用数据之间没有明显的分隔。他们是一体的。当主数据记录详细资料被修改后，所有应用的相关数据元素都将被更新。这意味着所有的消费应用和源系统访问的是相同的数据实例。这本质上是一个闭环的 MDM：所有的应用系统通过统一管理的主数据集成在一起。在这个级别，所有在系统看起来都是事实的同一个版本。操作应用系统和 MDM 内容是同步的，所以当变更发生时，操作应用系统都将更新。在那些熟悉的 MDM 架构风格中，持久总线架构，当一

个总线更新所有的操作应用系统将体现这种变更，形成改变的直接操作视图。在注册环境中，当数据数据更新时，总线将通过 Web 服务连接相关系统应用事务更新。因此，MDM Level 5 提供一个集成的，同步的架构，当一个有权限的系统更新一个数据值时，公司内所有的系统将反映这个变更。系统更新完数据值后不要单选其他系统中相应值的更新：MDM 将使这种更新变的透明。

从 MDM Level 4 到 MDM Level 5 意味着 MDM 功能性不是在一个应用内被特殊设计或编码的。这还意味着主数据传播和供应不需要源系统专门的开发或支持。所有的应用清楚的知道他们并不拥有或控制主数据。他们仅仅使用数据来支持他们自己的功能和流程。由于 MDM 总线和支持的 IT 基础架构，所有的应用可以访问主参考数据。一个公司在完成 MDM Level 5 后将使他们所有的应用连在一起—既包括操作的也包括分析的一所有访问主数据是透明的。举例说明，当一个客户更新她的状态—不要管注册该变更的系统—数据变更将被广播到所有的应用平台(因此一致起来)。MDM Level 5 是把数据概念作为一种 service 来实现。MDM Level 5 保证了一个一致的主数据主题域企业映像。定义“客户”和其他应用接受客户主数据业务规则变化实际上是一回事。MDM Level 5 移走了主数据的最后一个障碍：统一采用数据定义、授权使用和变更传播。

(作者：程永 / 王雪梅 来源：互联网)

原文标题：主数据管理详解（二）

链接：http://www.searchdatabase.com.cn/showcontent_17543.htm

主数据管理详解（三）

如何构建一个主数据管理(MDM)的解决方案

在开始构建主数据管理(MDM)解决方案之前，首先需要明确我们当前的数据管理现状是什么样子的，而我们的目标是什么，具体可以参照上一小节：主数据管理(MDM)的成熟度。

第二步，需要确定我们的每个主数据域的范围(这也是前期需求分析的一部分)。常见的主题域有：

Party：可以反映任何合法的实体，无论是个体还是组织。

Product：既包括物理存在的货物，也可以是任何服务。

Account：包括期限和条件，以及相关的各种关系。

Location：既可以独立存在，也常常与其他主数据域共存。

第三步，进行数据管理系统的设计，在设计时要注意以下几点：

数据采集和发布是否实时，最小的响应时间是多少。

数据转换规则能否让客户定制，而不是硬编码。

如果根据数据质量标准清理主数据域中的主数据。

权限控制。

主数据的历史版本控制以及变更监控控制(当主数据变化时，要能记录该变化，另外还要对主数据形成层次并记录其不同的版本值)。

第四步，开发部署测试。

(作者: 程永 / 王雪梅 来源: 互联网)

原文标题: 主数据管理详解 (三)

链接: http://www.searchdatabase.com.cn/showcontent_17544.htm

最常见的 13 种主数据管理（MDM）词汇和定义

简单而言，主数据管理（MDM）是公司权衡和重用常见、和准确业务数据的一种全面的方法。然而，它同时也是一种不只有助于组织做表面工夫的快速改进数据管理的策略。MDM 系统是为维护公司数据主版本而设计的，它可以通过将其影射到业务规则上来让组织自由处理数据，这种方法比以前的数据管理方法更加的主动。公司之所以关注 MDM 的其中一个主要的原因是为了从不同的资源和应用上协调功能。

不管你是否考虑实现新的 MDM 系统还是评估可能的工具和软件来帮助你目前的数据管理和数据整合主动性，了解目前市场的趋势，包括最新的 MDM 定义和产业词汇都是至关重要的。

为了便于你学习（熟悉）MDM，我们编排了市场上最常见的 13 种词汇，包括数据管理、产品信息管理（PIM）、企业主索引（EMPI）、MDM 集线器等等。

如果没有**原数据**，那么将不存在 MDM，依 Forrester's Rob Karel 看来。他说道任何激活 MDM 的，包括定义、业务规则、政策、数据关系和质量优点都是通过使用原数据类维护的。“换言之，原数据是主数据真正值得信任的唯一方式，”他说道。

产品信息管理（PIM） 在一个唯一的位置上管理组织的产品数据，因此产品数据是一致的，不管系统或者应用是否需要它。Gartner 定义 PIM 为“通过产品主数据的语义调和来支持不同数据来源的全局产品信息认证、连接和同步软件产品；创建和管理一个关于记录的中央数据系统；（向所有 stakeholders）提供单个产品视图；通过监控和校正技术支持数据质量和规范。”

Enterprise master patient index（EMPI），是针对于医疗行业 CDI 的一种形式。一些医疗组织正在使用 EMPI 来合并系统之间的病人标识并将这些数据充实到更优分析的数据仓库中。EMPI 目录一般包括病人人口统计和病人所使用的器械，并且它们通常能够根据 Gillogley Services 确定来自其它器械的信息是否应该与已经存在的目录信息相匹配。

数据管治（Data Governance） 指的是企业使用的数据的有效性、可用性、一致性和安全性的总体管理。一个强大的数据管治程序将包括一个中央实体或委员会来创建管治规则、一套流程和一个完成这些流程的规划。与数据管理不同的是，数据管理是数据管治主体策略的一次战术执行。

客户数据整合 (CDI) 是整合和管理不同来源的客户信息的过程。客户信息可能包括合同细节、客户评估数据和通过交互收集的信息 (如, 直接销售)。如果实施恰当, CDI 可以保证组织中所有相关部门能够持续访问最新和最完整的客户信息。因此 CDI 是客户关系管理 (CRM) 的一个至关重要的元素。

MDM 中心是一个数据库及管理数据库中主数据的软件并保持它与使用主数据的事务系统同步。关于 MDM 中心的最常见问题是“MDM 中心是否能够同时处理客户和产品数据?”。

根据 MDM 专家 Jill Dyché 的观点, 不同类型的 **MDM 体系结构**都有特定的优点和缺点。MDM 体系结构风格包括持久化、注册、混合、共存和事务。

根据 MDM 专家意见, 诸如“**协作的**”和“**分析的**”MDM 的词汇都只会混淆 MDM 的功能。开发 MDM 的动因是主数据是如何存储、访问和传播到企业中的。而且, 如果数据用法给 MDM 增加了复杂性时, 第二个动因就是组织使用和管理它的主数据的方式。

在实现 MDM 系统后, 专家建议不要太担心 **MDM 投资回报 (ROI)**。相反, 重要的是你的实现是如何解决公司的数据管理问题, 以及 MDM 系统是否有助于节省业务规划的开支。

MDM 项目可能是一项复杂的工作, 但如果有合适的专业人员可以使项目变得很顺利。主要的 MDM 人员包括以下 **MDM 角色**: 高级管理、业务客户、应用负责人、信息架构师、数据管治和数据质量人员、元数据分析师、系统开发人员和运营人员。

企业组织管理 (Enterprise hierarchy management) 是公司更新、建模和维护信息层次 (如, 客户和/或产品信息) 的地方。以前, 当业务发生变化, 如收购或重组, 有关的组织层次必须在不同的分析系统中更新以使报告精确。然而, 新产品提供了一个独立部分来完成所有必需的修改, 这些修改被发布到多个商业智能 (Business Intelligence, BI) 或企业绩效管理 (Corporate Performance Management, CPM) 软件系统。

开发一套 **MDM 度量**来测量和演示 MDM 对于一个公司的影响是 MDM 项目成功的关键。然而, 根据 Gartner Inc. 分析师 John Radcliffe 的观点, 这些度量应该关注于业务, 而不是 IT。比如, 显示 MDM 有助于增加客户数据 10%准确性的度量是不太可能影响到管理的; 然而, 显示客户保有和/或越区销售率增加的度量将会产生很大的影响, 因为它们是业务成功的关键。

数据职能中心, 或者称为数据管理职能中心, 是一个组织内的一个小组 (一般是一组以数据为中心和有数据处理能力的专家), 它为 MDM 制定策略和流程, 并帮助在公司内部实施这些 MDM 规则。

(作者: *Justin Aucoin* 译者: 曾少宁 来源: TT 中国)

原文标题: 最常见的 13 种主数据管理 (MDM) 词汇和定义

链接: http://www.searchdatabase.com.cn/showcontent_24416.htm

主数据管理（MDM）的七个最佳实践（上）

在考虑像主数据管理（MDM，Master Data Management）这样的学科时，寻找已经经历过或做过主数据管理的过来人咨询是唯一自然的途径。

但是主数据管理可能最佳实践仍处于起步阶段仍然显现出来了，因为让有经验的机构去讲他们主数据管理的经验不是一件容易的事。Kalido 公司是一家马萨诸塞州伯灵顿的主数据管理技术供应商，该公司也认为让客户对新闻界讲这方面的话题是很困难的。

所有成功的主数据管理项目都保密，因此对企业寻找最佳实践不太容易没有好处，这也是 Kalido 为什么发起一次客户审计和主数据管理最佳实践研究（由加州圣马特奥分析公司 Ventana 研究院研究）的部分原因。其研究员调查了五个 Kalido 匿名客户的最佳最终实践来推导结论。为了主数据管理的成功，Ventana 研究所，一个有经验的顾问和一家欧洲电信制造商最终提供了一些最佳实践（和最差实践）。

1. 让业务参与进来（或者在管理下进行）。

“主数据管理必须由企业需求驱动，否则它可能变成为一个只是必须与所有其他数据库同步的另一个数据库”，David Loshin 说。他是马里兰州银泉 Knowledge Integrity 公司的总裁，该公司提供主数据管理策略开发服务咨询以及从事企业级策略研究工作。

同样，Ventana 研究发现，推动主数据管理进程的应该是业务人员，而不是 IT 人员。Ventana 发现，要想获得主数据管理成功，从 C 级主管到高级管理者到业务终端用户的广泛支持是至关重要的。我们常常很难激励组织向着主数据管理干巴巴的前景迈进，但是从长远来看，早点开始企业范围内的支持是很重要的，用户们说。如果公司的重要目标通过一个稳定的商业案例绑定了该项目，那证明它的利益和激励士气的任务应该就简单了。

2. 允许使用充足的时间评估和计划。

Kalido 用户告诉 Ventana 研究员们说，要计划至少为期三个月的评估，与相关客户交谈，用真实的公司数据做例子做一个价值验证项目。用户们说，不要低估了时间和开发基础数据模型的过程。

“这事比人们想象的更复杂——而且它需要及早开始，并且计划使用真实数据”，David Waddington 说，他是 Ventana 的一位副总裁兼研究主管，一直在从事这项研究。

Ventana 发现，IT 公司的合作是一个令人关切的领域，因为一些公司曾经经历过由于许可和访问权限导致的项目延迟。

(作者: Hannah Smalltree 译者: 冯昀晖 来源: TT 中国)

原文标题：主数据管理（MDM）的七个最佳实践（上）

链接：http://www.searchdatabase.com.cn/showcontent_24196.htm

主数据管理（MDM）的七个最佳实践（下）

3. 定一个大目标，但采取小步骤。

用户们告诉 Ventana，要考虑基本目标，但是会限制战略部署的视野。他们建议，一旦主数据管理在一个地方好用了，再一步一步地扩展。他们说，业务流程常常是缓解因素，而不是技术，所以在过程中及早得到终端用户的输入时很重要的。

“如果你只对保持客户数据一致性感兴趣，那就一定要反对‘怎样我才能长期管理我全部的主数据？’这一大前提，这一点是很重要的”，Waddington 解释到。“那样你最后就不会落到不得不把整个一大堆解决方案连接到一起的困境中。”

4. 考虑潜在的性能问题

Loshin 警告说，性能像一只 800 磅重的猩猩悄悄潜伏在主数据管理的探讨中。

不同的架构能导致不同的性能损失。例如，如果公司为主数据管理采用主 hub 方式，记录创建流经唯一一点，它就能变成一个瓶颈。而且，随着应用越来越依赖主数据管理，工作流，系统优先级和运行顺序会变成要马上考虑的紧急问题。Loshin 说，公司怎样解决潜在的性能问题有多种选择，因为它本质上与他们独特的架构有关系。

5. 建立数据管理策略和流程。

专家们一致认为，应该允许在人力和过程变更管理方面花时间和投钱，不要低估了工作量。瑞典电信设备制造商爱立信公司了解到，根据高级项目经理 Roderick Hall 的观点，数据管理政策可能是相当难满足的。

在部署 SAP 主数据管理很早以前，斯德哥尔摩的公司建立了一个主数据小组来管理关键数据资产。它是一个“共享服务”小组，同时对 IT 和商业提供服务。该组刚成立时是财务部门的一部分，但是随着主数据管理受到公司范围的关注，它的职能也发生了变化，Hall 说。他们的工作并不总是很容易。

Hall 说，尽管一些部门（比如财务部门）看到了集中主数据管理的价值，其但他群体都不太乐意放弃自己管理数据的权利。

“人们已经放弃了自己管理数据的自由，把他们自己的主数据放到斯德哥摩尔（在那里主数据小组是本地的）一些不露面的组织。为了得到人们对这一事实的认可，他们真的打了一场硬仗”，Hall 说。

6. 仔细计划部署。

Ventana 发现，主数据管理相对仍然比较新，所以对业务和技术人员的培训比以往更重要。Waddington 说，使用未参加培训的或者是准培训的系统集成人员和

尝试外包，对主数据管理用户来说，会引起大问题，甚至项目延期。

7. 考虑过渡计划。

那么，有了铺开项目的前景，该项目会对关键流程和系统造成一定影响——没有无缘无故的担心。Loshin 推荐公司应该计划一个主数据管理的过渡策略，允许静态和动态的数据同步。

“试图在不影响日常操作的前提下调整潜在的基础设施，可能与不中断交通的情况下在公路上补洞一样具有挑战”。Loshin 说。

(作者: Hannah Smalltree 译者: 冯昀晖 来源: TT 中国)

原文标题：主数据管理（MDM）的七个最佳实践（下）

链接：http://www.searchdatabase.com.cn/showcontent_24198.htm

SQL Server 2008 R2 最新功能：主数据服务

如果说 Microsoft 在开发 SQL Server 2008 R2 时有特别关注的方面，那就是商业智能。Microsoft 的新最数据库管理系统(DBMS)产品将在 2010 年上半年发布，它增加了支持 BI 的改进，如 Excel 的 PowerPivot，之前的代码称号为 Gemini。

R2 是在 SQL Server 2008 发布后不到两年时间内定下来的，它包含了自己的 BI 特性共享以及 SQL Server Analysis Services、Integration Services 和 Reporting Services 的改进。

下表概括地显示了 SQL Server 2008 R2 版本中将发布的其它一些特性：

PowerPivot （代码前称为 Gemini）	Excel 和 SharePoint 的插件 PowerPivot 可以管理自助 BI。
Parallel DataWarehouse （代码前称为 Madison）	新版本的 SQL Server 2008 R2 的设计改进了多服务器上数据仓库的可扩展性。
Reporting Services 增强	SQL Server Reporting Services (SSRS) 的增强包括报表部分、共享数据集和 SharePoint 整合

尽管出现了这个新技术，但是大多数组织仍然在生产环境中使用 SQL Server 2005 —— 并且大多都愿意继续这样做。

那么如何才能说服公司进行升级呢?很可能，只有实质性的商业利益才能促使公司投入宝贵的时间和资源进行迁移 —— 而这些都是小调整和改进所不能实现的。

然而，Microsoft 通过 SQL Server 2008 R2 向它的 BI 平台添加了一个新特性：Master Data Services，这个特性一定能够吸引到这些还在犹豫的公司进行升级。

虽然 MDS 对于 SQL Server 而言还是新东西，但是它是一个起源于 Stratature 产品的成熟技术，而 Stratature 是 Microsoft 在 2007 年收购的一家公司。

Stratature 技术当时被认为是很强的，因为它已经在 SQL Server 上开发了。它也填补了 Microsoft 相对于其它企业供应商在基础架构上的空白，如 IBM、Oracle 和 SAP。

(作者: BRENDAN COURNOYER 译者: 曾少宁 来源: TT 中国)

原文标题: SQL Server 2008 R2 最新功能: 主数据服务

链接: http://www.searchdatabase.com.cn/showcontent_33855.htm

Master Data Services 的潜在问题

Master Data Services —— 和通常的主数据管理 (MDM) —— 的主要目的是一致性。对于运行多个数据库的较大型企业而言，这是一个极大的挑战，因为这些数据库的相同属性上都存在细微差别。

Quest Software 负责 SQL Server 解决方案的主管 Kevin Kline 说：“当你想要将所有数据放在一起并生成报表，组织通常会发现他们会遇到真正的数据质量问题。当你想要实现一个主数据策略的时候；它是一种版本事实，或者我们所谓的‘引用数据’。”

存在多种“版本的事实”往往就是大多数数据质量问题的根源。举一个简单的例子，有两个数据库，它们分别都装有客户数据。其中第一个数据库可能记录着 John Smith 在 Hewlett-Packard 工作；而另一个数据库则记录他在 HP 工作。这样，同一个公司就存在有两种“版本”的记录。

对于报表来说，清除类似的不一致性会使数据库人员增加不少的时间和压力。Microsoft 开发 Master Data Services 的目标是提供一种更简单的、更划算的方法来保持数据在多个系统中的一致。

“MDS 做得很好的是它能让你创建你的记录数据，” Kline 说：“它会给你一个版本的事实，当你拥有许多数据库时，这将是很有用的。例如，你就不会记录一个客户对应 8 个不同公司名称。”

Herain Oberoi 是 Microsoft SQL Server 业务小组的产品经理，他认为 MDS 是公司 BI 平台的一个重要的辅助。Oberoi 说当客户在创建一个 BI 解决方案时，他们并不愿意去猜测他们的主数据引用在什么位置，而且他补充道，MDS 设计拥有多种解决该类问题的方法。

“有一个关于运行的主数据概念，它是保持系统同步的，” Oberoi 说：“…… 然后还有一个分析的主数据概念，它并不只是保持业务系统和传统产品线的同步，它还保持这两个系统与我的数据仓库同步。”

(作者: BRENDAN COURNOYER 译者: 曾少宁 来源: TT 中国)

原文标题: Master Data Services 的潜在问题

链接: http://www.searchdatabase.com.cn/showcontent_33860.htm

主数据服务的相关业务案例

Microsoft 将 Master Data Services 看作是目前 BI 特性的一个自然辅助，而 Kline 也同意这个观点。“它能够与 Microsoft BI 平台的其余部分很好地协作，”他说，但是要注意 MDS 并不适合于那些商业智能还没有成为第一优先级的组织。

“这可能是一种小公司并不感兴趣的东西，但是一旦你创建了 BI 策略，你就会开始看到所有地方都会需要它。它在 SQL Server Analysis Service 出现之后的几年里产生是很有意义的，因为许多人都将它作为第一个 BI 工具使用，”Kline 说。

建立了 BI 策略的较大型公司对于 MDS 会有更大的兴趣，这对于 SQL Server 2008 R2 而言也是一样适用的。虽然主数据管理解决方案并不是什么新鲜事，但是实现这样一个项目一般是高成本和费时的过程。

Kline 说类似于 MDS 的服务对于运行着 SQL Server 的组织来说是非常划算的。“Microsoft 在其它 BI 工具市场上也处于相同的地位，这原先是一个高端市场，而现在公司将对它进行简化和降低价格。”

他指出 MDS 的另一个省钱的特性源自于它的过程自动化功能。“每次你雇人完成这些工作时，总是需要一定的资金投入。”他解释说。“MDS 一个很大的功能是过程改进和过程自动化。你不再需要雇佣一个顾问，事情就会变得更简单，你只需要自己去完成。”

MDS 的其它好处包含改进的数据清理操作和过程管理功能。而后者是用来帮助指导数据库人员正确地清理他们的数据的。

(作者: BRENDAN COURNOYER 译者: 曾少宁 来源: TT 中国)

原文标题: 主数据服务的相关业务案例

链接: http://www.searchdatabase.com.cn/showcontent_33862.htm

Master Data Services 存在可提升的空间

虽然它来源于成熟的技术，但是 SQL Server 2008 R2 的 Master Data Services 还是第一个版本。因此，相对于其它竞争产品，这个技术还有一些问题需要解决，这一点也不奇怪，这很可能给 Microsoft 的合作伙伴带来参与 MDS 改进的机会。

“由于 SQL Server 产品确实是很普通，所以它将不会马上为你提供相同的内置工具集，” SQL Server MVP 和 BI 架构师 John Welch 说：“可能在接下来几年内，我们可以看到供应商开始提供扩展的插件来解决这些问题。有待改进的一个重要领域可能涉及到识别相同客户的不同版本的逻辑方面。”

Welch 说虽然 MDS 的许多功能都是用来清理各种不同的属性，但是它可能在识别实际上不同的相似客户时还不够成熟。例如，区分名为 Jon Smith 和 John Smith 的客户——他们有相似的姓名但拼写不同。

“有一些现成的产品更擅长于进行这些类型的处理，而且能够找到相匹配的数据，然后指出这个是相同的，而另一个不是，” Welch 说。“我们在 MDS 的第一轮使用中将不会看到它很有效地完成这样的处理。”

在这个领域有另一个供应商走在前面，它就是 Zoomix，Microsoft 在 2008 年收购的一个数据质量软件公司。虽然 Zoomix 的数据质量软件将不会成为 SQL Server 2008 R2 Master Data Services 的一部分，但是 Welch 说它将很可能在 MDS 的下一个版本中发挥作用。

尽管如此，SQL Server 2008 R2 仍然会给正在寻求扩展商业智能策略以保证更高水平的数据质量的公司带来很大的变化——特别是如果这些公司已经在使用 SQL Server 的 BI 工具。

“记住，这是一个你不愿意在第一个 BI 项目就使用的工具，” Kline 说。“当你说，‘现在一切都运行良好，但是我们发现我们的数据有些混乱，那么我们应该怎么清理它呢？’这就是使用它的时候了。实际上，它是你完成 BI 项目之后的跟进工具。”

(作者: BRENDAN COURNOYER 译者: 曾少宁 来源: TT 中国)

原文标题: Master Data Services 存在可提升的空间

链接: http://www.searchdatabase.com.cn/showcontent_33864.htm

Master Data Services 入门：配置 MDS

Master Data Services, 也称为 MDS, 是 SQL Server 2008 R2 增加的 [关键商业智能特性](#) 之一。Master Data Services 的基本目标是为企业信息提供单个权威来源。然后这个信息可以被其它应用和数据使用, 这样您的环境中每一个应用便都是使用相同信息的同一份权威副本。

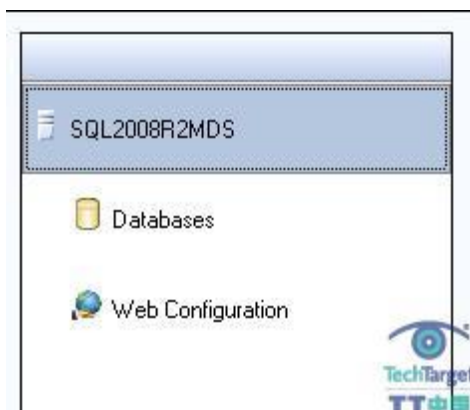
您需要先安装 Master Data Services 才能使用它。安装要求使用 64 位服务器, 因为 MDS 安装程序只支持 64 位。您还需要 Windows PowerShell 1.0 或以上版本。

要启动安装过程, 您需要打开 SQL Server 2008 R2 DVD, 然后进入 MasterDataServices\x64\1033_ENU 文件夹。您可以运行这个文件夹中的安装程序 MasterDataServices.msi。MDS 软件的安装是非常简单的, 唯一您需要做的是选择合适的安装路径。

配置 Master Data Services

在安装好 MDS 服务后, 您可以从 Start Menu 的 Start→All Programs→Microsoft SQL Server 2008 R2 November CTP→Master Data Services→Configuration Manager 打开 Master Data Services Configuration Manager。然后您可以配置 MDS 库, 并指定从哪里获取主数据。

要配置 Master Data Services, 可以打开 MDS Configuration Manager, 然后选择如图 1 所示的菜单中的 Databases 选项。然后单击右边的 Create Database 按钮。这样就可以打开配置装载 MDS 库元数据的数据库的向导。



在配置 MDS 库时，您可以选择一个本地或远程的数据库实例。在选择数据库实例后，您还需要指定用来管理 MDS 数据库的服务和管理员帐号。如果您已经知道您想要用来装载 MDS 库的数据是哪一个，那么就单击 Select Database 按钮。

当 MDS 库配置好后，您就可以创建 Web 配置。启动一个基于 Web 的配置界面，然后您可以在界面上根据 Master Data Services 的位置来配置数据管理位置。

Master Data Services 网站

在配置好数据库后，您就可以配置 Internet Information Services (IIS) 网站了。您需要提供 IP 地址、TCP 端口号、主机头和网站内用于访问 Master Data Services 的路径。此外，您还需要 创建一个应用池 并指定这个应用池运行所使用的用户名和密码，如图 2 所示。

The screenshot shows the 'Create Web Site' dialog box. It has two main sections: 'Web Site' and 'Application Pool'. In the 'Web Site' section, the 'Web site name' is 'MDS', the 'Protocol' is 'http', the 'IP address' is 'All Unassigned', and the 'Port' is '80'. The 'Host header' field is empty. In the 'Application Pool' section, the 'Name' is 'MDS Application Pool', the 'User name' is 'CORP\SQLServer', and the 'Password' and 'Confirm password' fields are filled with masked characters. The dialog has 'OK' and 'Cancel' buttons at the bottom.

注意：在为本文配置 Master Data Services 时，我发现应用池运行的帐号必须是 MDS 服务器上的本地管理员。如果这个帐号没有本地管理员权限，那么 MDS 网站上不会出现所有的选项。如果您使用的是没有管理员权限的帐号来安装应用池，那么只需要将帐号添加到本地管理员组，然后重启服务器就可以。

(作者: Denny Cherry 译者: 曾少宁 来源: TT 中国)

原文标题: Master Data Services 入门: 配置 MDS

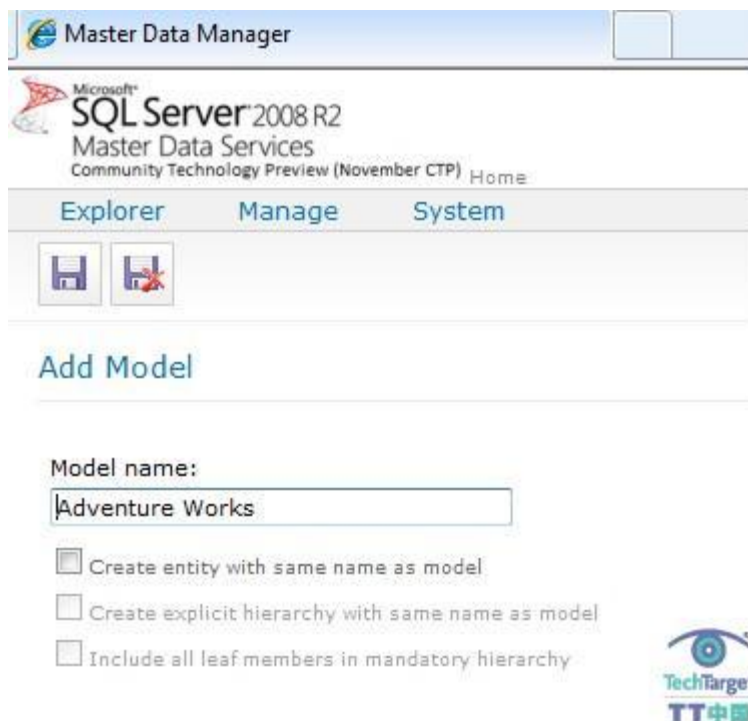
链接: http://www.searchdatabase.com.cn/showcontent_33421.htm

Master Data Services 入门：创建数据模型

确定主数据

在完成 MDS 数据库的 Web 配置后，您就可以开始创建 数据模型 了。这些数据模型包括您的主数据将使用的数据引用。您将需要定义模型、实体、属性、业务规则和数据建模时派生的层次。

在连接到 Master Data Services 管理工具后，单击 System Administration 选项。这样您就可以创建一个装载数据的模型。单击顶部的 Manage 选项，选择 Model。然后单击绿色的 + 号添加如图 3 所示的一个新模型。当模型创建后，您可以选择使用一个与模型同名的实体，或者在之后手动创建一个不同名称的实体。在完成时单击保存按钮（带有小磁盘图标而没有红色 X 的那个按钮）。



在创建模型后，回到 Explorer 视图，选择刚刚创建的模型。然后与这个模型相关的实体将显示在右边。单击图标编辑这些实体。注意您也可以通过单击 Manage 下拉菜单，

选择 Entities，然后单击绿色的 + 号，打开这个页面。输入一个实体名称，然后选择您是否需要启动显式分层和收集(撰写本文时我选择了不需要)。然后单击 Save 图标。

在创建实体后，您可能还希望添加一个属性组，也称为“叶组”。要添加属性组，可以单击 Manage 下拉菜单，然后选择 Attribute Groups。选择您希望编辑的模型和实体，然后单击 Leaf Groups。然后就会出现绿色的 + 号，单击这个 + 号您就可以创建一个叶组。

要添加叶属性，您需要单击顶部的 Explorer 菜单返回主模型树结构。从模型树中选择带有所需要的属性树的模型，然后单击带有铅笔图标的按钮。您就可以看到叶属性，您可以添加更多的叶到叶组中。要添加叶属性，您可以单击绿色的 + 号。这样您可以选择属性名称、数据类型和数据最大大小。

在创建属性后，再一次单击顶部的 Manage 下拉菜单，然后选择 Attribute Group。选择您想要编辑的模型和实体(这里是 Adventure Works 和 Employee)。单击您想要编辑的叶组旁边的 + 号，然后选择 Attributes。选择您想要包含在组中的属性，单击 Save 按钮。你也可以修改已分配的属性的顺序。

使用 MDS 保证主数据安全

在创建主数据模型后，您可以通过使用 Microsoft Active Directory 用户和组来保证它的安全。这是通过选择 Master Data Services 首页的 User and Group Permissions 选项实现的。然后您可以单击绿色的 + 号将另一个用户添加到 MDS。这是通过直接输入用户或组名实现的。

在您保存新的用户或组后，您就可以从用户或组列表选择单击铅笔图标来编辑该用户或组。您可以给我们刚创建的每一个对象分配权限。

您可能觉得很奇怪，Master Data Services 在维护主数据时确实是很复杂的。但是，现在您已经注意到了本文还没有介绍如何实际地在 MDS 中加载数据。我们将在后面的文章进行介绍。

(作者: Denny Cherry 译者: 曾少宁 来源: TT 中国)

原文标题: Master Data Services 入门: 创建数据模型

链接: http://www.searchdatabase.com.cn/showcontent_33422.htm

Master Data Services 进阶指导：加载数据

本文的示例将使用名为 Adventure Works 的一个 Master Data Services 模型，它有唯一一个实体 Employee。在这个实体中有三个属性，分别是 LastName、FirstName 和 EmployeeId。我们将使用 AdventureWorks 2008 数据库所提供的示例数据，将这些数据加载到 Master Data Services 中。

加载数据是通过将数据表内容插入到 Master Data Services 知识库的一些数据表实现的。这些数据表包括：

- mdm.tblStgMember - 筹划叶成员，统一的成员和集合。
- mdm.tblStgMemberAttribute - 更新属性值。
- mdm.tblStgRelationship - 用以在显式层次中移动成员以及将成员添加到集合中。

我们将使用 T-SQL 来加载数据到这些表中。然而，在一个更大型的数据导入过程中，您可以使用 SQL Server Integration Services (SSIS) 或其它 ETL 工具 来导入数据。

加载数据

首先我们需要将记录插入到 mdm.tblStgMember 表中。这将告诉导入过程哪些实体是用来加载数据的。下面是示例代码：

```
INSERT INTO mdm.tblStgMember
(ModelName, HierarchyName, EntityName, MemberType_ID, MemberName, MemberCode)
VALUES
('Adventure Works', NULL, 'FirstName', 1, '', 1),
('Adventure Works', NULL, 'LastName', 1, '', 2),
('Adventure Works', NULL, 'EmployeeId', 1, '', 3)
```

在这些 T-SQL 代码运行后，查询 mdm.tblStgMember 表，注意记录以 Status_ID 为 0 插入，如图 1。如果我们查看关于 mdm.tblStgMember 表的 MSDN 文档，我们可以看到 Status_ID 字段的各种值所表达的意思。

Results Messages											
	ID	Batch_ID	UserName	ModelName	HierarchyName	EntityName	MemberType_ID	MemberName	MemberCode	Status_ID	ErrorCode
1	1	NULL	NULL	Adventure Works	NULL	FirstName	1		1	0	
2	2	NULL	NULL	Adventure Works	NULL	LastName	1		2	0	
3	3	NULL	NULL	Adventure Works	NULL	EmployeeId	1		3	0	

mdm.tblStgMemberAttribute 表是我们需要加载数据的第二个表。这能够告诉导入过程我们希望导入的数据的值。当查看这个表时，我们会注意到它看起来像一个名称-值配对的表。这样，我们可以导入两个员工的值到 mdm.tblStgMemberAttribute 表。当加载数据到 mdm.tblStgMemberAttribute 和 mdm.tblStgMember 表的查询时，我们会发现 MemberCode 的值似乎是随意赋值的，而且应该在加载批处理中是唯一的。

```

INSERT INTO mdm.tblStgMemberAttribute
(ModelName, EntityName, MemberType_ID, MemberCode, AttributeName, AttributeValue)
VALUES
('Adventure Works', 'Employee', 1, 1, 'FirstName', 'Bill'),
('Adventure Works', 'Employee', 1, 2, 'LastName', 'Gates'),
('Adventure Works', 'Employee', 1, 3, 'EmployeeId', '1')
INSERT INTO mdm.tblStgMemberAttribute
(ModelName, EntityName, MemberType_ID, MemberCode, AttributeName, AttributeValue)
VALUES
('Adventure Works', 'Employee', 1, 1, 'FirstName', 'Steve'),
('Adventure Works', 'Employee', 1, 2, 'LastName', 'Balmer'),
('Adventure Works', 'Employee', 1, 3, 'EmployeeId', '2')

```

如果我们查询 mdm.tblStgMemberAttribute，我们可以看到只包含在这个表中的数据，如图 2 所示。您可以在关于 mdm.tblStgMemberAttribute 的 MSDN 文档上查看不同的 Status_ID 值。

Results Messages											
	ID	Batch_ID	UserName	ModelName	EntityName	MemberType_ID	MemberCode	AttributeName	AttributeValue	Status_ID	ErrorCode
1	1	NULL	NULL	Adventure Works	Employee	1	1	FirstName	Bill	0	
2	2	NULL	NULL	Adventure Works	Employee	1	2	LastName	Gates	0	
3	3	NULL	NULL	Adventure Works	Employee	1	3	EmployeeId	1	0	
4	4	NULL	NULL	Adventure Works	Employee	1	1	FirstName	Steve	0	
5	5	NULL	NULL	Adventure Works	Employee	1	2	LastName	Balmer	0	
6	6	NULL	NULL	Adventure Works	Employee	1	3	EmployeeId	2	0	

如果您在您的主数据中使用层次，您将需要加载数据到 `mdm.tblStgRelationship` 表中。因为本文使用的示例并没有使用层次，所以请参考 MDS Team 关于这个问题的 博客，以及关于 `mdm.tblStgRelationship` 表的 MSDN 文档。

(作者: Denny Cherry 译者: 曾少宁 来源: TT 中国)

原文标题: Master Data Services 进阶指导: 加载数据

链接: http://www.searchdatabase.com.cn/showcontent_33908.htm

Master Data Services 进阶指导：处理和验证数据

在筹划数据后，我们需要通过运行 `mdm.udpStagingSweep` 存储过程来提醒 Master Data Services 开始处理数据了。这个存储过程有三个参数：筹划这个批处理的用户的 `UserId`，所处理模型的 `VersionId`，以及一个告诉 MDS 数据应该马上处理还是经过一定时间间隔后再处理的值。指定的用户必须是一个进入 Master Data Services 的 Web 应用的模型管理员。下面的代码显示的是如何马上运行数据处理过程。

```
DECLARE @ModelName NVARCHAR(50) = 'Adventure Works'
DECLARE @UserName NVARCHAR(50) = 'CORP\dcherry'
DECLARE @UserId INT
DECLARE @VersionId INT
SET @UserId = (SELECT ID
FROM mdm.tblUser u
WHERE u.UserName = @UserName)
SET @VersionId = (SELECT MAX(ID)
FROM mdm.viw_SYSTEM_SCHEMA_VERSION
WHERE Model_Name = @ModelName)

EXEC mdm.udpStagingSweep @UserId=@UserId, @VersionId=@VersionId, @Process=1
```

如果要在指定时间间隔后处理数据，那么要将 `@Processparameter` 的值从 1 改为 0。

要验证用于处理的批处理是否正确，需要查询 `themdm.tblStgBatch` 表。`Status_ID` 字段值为 2 表示您的批处理过程失败了，而 `Status_ID` 字段值为 1 表示成功。

您可以通过查看 `mdm.tblStgMember` 和 `andmdm.tblStgMemberAttribute` 表的 `Status_ID` 和 `ErrorCode` 字段查看处理后的每一行记录的状态。关于 MDS Staging 表可能发生的错误代码的一个完整列表可以在 [KnightReign 博客](#) 和这个 [MSDN 论坛话题](#) 中找到。

所有的处理错误也将会被记录到运行您的 MDS 数据库的服务器上的 `ERRORLOG` 文件中。这些错误将会显示为一个激活的服务中间人过程的输出消息，如图 3 所示。

```
Selected row details:
Date          4/3/2010 8:10:45 PM
Log           SQL Server (Current - 4/3/2010 7:52:00 PM)
Source        spid23s

Message
The activated proc "[mdm].[udpSecurityMemberQueue.Activate]" running on queue
'mds.mdm.microsoft/mdm/queue/securitymembertimer' output the following: 'Creating member security items for Entity_ID: 11 and
Version_ID: 5 finished at: 20:10:45'
```

验证数据

在处理数据后，您可以选择使用 `themdm.udpValidateModel` 存储过程验证这些数据。这个存储过程有四个参数：`UserId`、`ModelId`、`VersionId` 和 `StatusId`。完成处理的 `Status` 值是 1。下面的示例代码将会验证最近处理的一批数据。

```
DECLARE @ModelName NVARCHAR(50) = 'Adventure Works'
        DECLARE @ModelId INT
        DECLARE @UserName NVARCHAR(50) = 'CORP\dcherry'
        DECLARE @UserId INT
        DECLARE @VersionId INT
        SET @UserId = (SELECT ID
                        FROM mdm.tblUser u
                        WHERE u.UserName = @UserName)
        SET @ModelId = (SELECT Model_ID
                        FROM mdm.viw_SYSTEM_SCHEMA_VERSION
                        WHERE Model_Name = @ModelName)
        SET @VersionId = (SELECT MAX(ID)
                        FROM mdm.viw_SYSTEM_SCHEMA_VERSION
                        WHERE Model_ID = @ModelId)
EXECUTE mdm.udpValidateModel @UserId, @ModelId, @VersionId, 1
```

(作者: Denny Cherry 译者: 曾少宁 来源: TT 中国)

原文标题: Master Data Services 进阶指导: 处理和验证数据

链接: http://www.searchdatabase.com.cn/showcontent_33911.htm

是否可以使用数据质量工具执行 MDM 项目

问：目前我们有一个 MDM 的项目需要执行，请问能否使用我们现有的数据质量工具来进行呢？还是我们要在购买一套独立的解决方案？

答：利用现有的数据质量工具，你也许可以进行数据标准化的工作，但是数据质量工具不可能满足所有 MDM 项目的需求。比如，你的数据质量工具是否可以实现数据匹配的功能？或者追踪数据血统？利用数据质量工具如何来更新数据源？当然还有商业规则、访问安全等等的问题都不是数据质量工具可以解决的。人们在进行 MDM 项目时，往往会发现所需要的功能并不是以“打包”的方式存在与他们企业内部的，因此要么去自己建立，要么就去买一整套解决方案比较合适。

(作者: Jill Dyché 译者: 孙瑞 来源: TT 中国)

原文标题: 是否可以使用数据质量工具执行 MDM 项目

链接: http://www.searchdatabase.com.cn/showcontent_24695.htm

在 MDM hub 与交易数据库之间保持参照完整性

问：请问如何在一个主数据管理中心 (MDM hub) 与一个交易数据库之间确保其参照完整性，使得不用进行数据拷贝？

答：参照完整性可以用来连接两个领域的相关数据。MDM hub 事实上就是一个特定的 subject area，因此你不支持不同 subject area 之间的链接。

MDM 独立地控制每一个 subject area (对于多域 MDM 同样适用)，这样就导致每一个 subject area 都有各自的参照列表。而整合数据以确保参照完整性是由系统来完成的 (数据仓库、数据集市)。事实上，当数据被加载到 MDM hub 中时，相关的数据细节会被维护在源系统键的 hub 中。

(作者: Jill Dyché 译者: 孙瑞 来源: TT 中国)

原文标题: 在 MDM hub 与交易数据库之间保持参照完整性

链接: http://www.searchdatabase.com.cn/showcontent_31340.htm

数据仓库的未来：开源与 MDM

开源数据仓库

开源已经达到了临界，开源领域已有足够的产品构成一个完整的数据仓库架构，正如前端，中间层和后台一样简单。展望未来，开源数据仓库软件产品在功能方面将得到发展、加强和丰富。对于后台， Infobright 和 KickFire 使用的是 MySQL。而 ParAccel， Netezza 和 GreenPlum 公司使用的是某个版本的 PostgreSQL。DatAllegro(现在是微软的一部分)使用的是 Open Ingres。(“版本”是这里的关键词，因为在某些情况下，厂商已经加强了代码，并使之成为专利。)该架构使 Pentaho 和 Talend 的数据集成和在中间层的 ETL 更完善，前端的 JasperSoft 和 Pentaho 在信息获取和传递功能更完美。因此，开源是一个重要的创新，将打破传统软件的开发，市场和价格。

主数据管理和数据仓库

主数据管理(MDM)作为驱动数据仓库的一个趋势，已经迎来了自己的时代。有时候我们可以把数据仓库比作尾巴，而主数据管理比作狗，但有时情况也会正好相反。从数据仓库的角度来看，主数据的维度与基本事实(例如客户购买的产品，时间和地点)是有区别的。“事实”是最基本的星型模式的交叉点，它通常是一个有关销售，运输和相关业务指标的量化数据点。从主数据的角度来看，客户、产品、市场和公司内部其他方面的内容是他们自身的处理过程，他不能减少商业智能反而提醒构建了它。关于主数据管理的好处是能了解客户的业务、产品和市场，因此可以涉及到至少在一个较高的水平的技术。因此，它可以使用主数据管理在业务上构建一座桥梁，并利用各种接触合作伙伴的机会来增值。通常，主数据被清理和被明晰的管理支配，以便它能够促进数据仓库的结构化和成功运作。

(作者: Lou Agosta 译者: 张峰 来源: TT 中国)

原文标题：数据仓库的未来：开源与 MDM

链接：http://www.searchdatabase.com.cn/showcontent_31197.htm