

元数据管理及其在数据仓库中的应用初探

王孝良

中国人民大学 北京 100085

摘要: 随着当前信息化建设的不断发展和进步,企业在发展的过程中不断加大对IT系统的研发和应用力度。IT系统研发运行之后,有效的提高了企业的经营效率和管理效率,对于企业的发展起到了一定的推动作用。除了给企业带来一定的积极影响之外,企业在进行信息化建设的过程中也面临一些问题。尤其是在数据管理方面,部分企业的管理模式相对分散数据结构不一致,导致数据孤岛情况出现。基于此,企业应该开展元数据管理工作,真正将元数据管理应用于数据仓库建设发展中全面推动企业的信息化建设和发展。

关键词: 元数据管理; 数据仓库; 应用初探

Metadata management and its application in data warehouse

Xiaoliang Wang

Renmin University of China ,Beijing 100085

Abstract: with the continuous development and progress of the current information construction, enterprises continue to increase the R & D and application of IT systems in the process of development. After the R & D and operation of IT system, it effectively improves the business efficiency and management efficiency of the enterprise, and plays a certain role in promoting the development of the enterprise. In addition to bringing a certain positive impact to enterprises, enterprises also face some problems in the process of information construction. Especially in the aspect of data management, the data management mode of some enterprises is relatively decentralized and the data structure is inconsistent, resulting in data islands. Based on this, enterprises should carry out metadata management, truly apply metadata management to the construction and development of data warehouse, and comprehensively promote the information construction and development of enterprises.

Keywords: metadata management; Data warehouse; Preliminary study on Application

前言:

企业管理者要明确企业统一原数据管理系统和数据仓库系统的基本概念,了解两大系统之间的关系,分析元数据管理的主要内容,以及对数据仓库系统运行所产生的辅助作用。根据企业当前发展的情况进行分析,将两大系统进行系统的整合之后可以全面推动企业的发展。

1 意义

元数据管理主要是指通过一些系统的数据和技术对现有的数据内容进行管理,要明确具体的流程程序和相关的工具。充分发挥原数据管理的技术作用,将原数据

集中在一起,能够方便用户进行及时的访问和沟通。元数据管理的主要目标是能够实现企业数据的一致性,而且明确数据之间的关系,真正实现数据信息的共享。企业进行元数据信息化管理建设的过程中能够有效的提供企业完整的数据现状,也可以全面增强信息的共享程度。对数据库进行调整,将数据库和相关的资料内容分散到各个应用系统中。不断对业务流程和相关的文件进行分析,对现有的原数据模式进行集中的管理,真正为企业提供系统的管理模式,让相关人员能够对现有的数据信息进行共享。

而且在开展元数据管理工作的过程中,能够对各类数据信息和技术进行集中的管理与控制,在管理的过程中如果发现存在一些问题会及时对这些问题进行分析和

作者简介: 王孝良(1993.07),男,汉,河南,本科,高级研发,企业团餐高并发系统建设。

改进^[1]。对企业的数据进行研究的过程中要保证企业数据的准确性和一致性,还要全面提高企业的整体发展效果,分析企业的数据质量。在对系统进行开发和研究的过程中,要考虑原数据关系的基本内容和相关规范,对其进行统一的描述,分析系统管理中存在的一些问题,要加大维护的研究力度和研究效率,真正实现系统的开发完善。系统开发人员要对计划内容进行合理的分析与规划,明确工作中的重点和难点,不断对改造和升级工作进行分析了解设计工作所产生的影响。构建完善的系统和环节,全面提高终身的效率,在开展元数据管理工作的过程中,还应该加大企业级的应用开发力度,制定现代化的数据仓库系统。开展元数据管理工作还能够全面提高业务人员的专业能力,让业务人员对于数据有更加深入的了解,也能够明确数据的相关标准和数据管理的基本内容。业务人员在工作的过程中经常会遇到一些问题,可能出现方案数据和实际数据不一致的情况。开展元数据管理工作之后,就能够对这些问题进行系统的调整和整顿。

2 元数据管理结构

从当前数据仓库系统的运行和发展情况来看,主要有三种典型的元数据管理结构,分别是集中结构,分散结构和邦联结构。

集中结构,对于一些中等规模的组织结构来说,采用单一的源数据库就可以满足团队的基本需求,以及对于元数据管理的基本需求。这些过程中需要用到特定的概念,主要是根据原数据的基本情况构建统一的模型,然后充分发挥模型的作用,对各种元数据的信息和定义进行管理。在这一过程中还应该考虑到原数据的存储问题,要将原数据信息和技术存储到中心元数据库中。后续系统在运行的过程中,所有的工具和数据的仓库会直接对中心源数据库进行访问,访问具有全面性和完整性的特点。

分散结构,在实际运行的过程中,不同的工具软件和组件等有构成的数据模型不相同,内部的格式和集中管理模式也不同^[2]。在对原数据库进行调整的时候,首先要保证原数据库能够还应该分析分散管理的基本内容,利用这种方式对局部的原数据库进行自我防治和治理,明确局部原数据库的基本内容,要根据当前企业发展的情况建立统一等数据库模型利用专业的元数据模型对全局的元数据进行管理。分析局部元数据库中元数据的子集,了解管理的极端情况,对现有的模型系统进行调整。

邦联结构,邦联方法主要是对前两种方法的一种整合,吸取了前两者的优点,进行系统的整合和研究。每一个工具在这一过程中都是有专有的原数据库的,能够充分发挥原数据库的作用,对其进行快速的访问和自我治理,也能够真正提供一定的交换接口,可以实现元数据的共享。

3 数据仓库系统中构建元数据管理

3.1 关系与作用

随着近些年社会的不断发展,企业要想全面提高自身的竞争实力,就应该切实开展企业管理工作。除了传统的人力资源管理工作之外,还应该迎合现代化的数据技术和信息化技术,建立完善的数据仓库,对数据仓库技术进行广泛的应用,将数据仓库技术应用于各项事务的处理中,既可以实现密集型的事务处理,也可以实现公共服务行业的应用。从当前的实际应用情况来看,所取得的效益是比较理想的,而且数据仓库数据源广泛性较强,所以在实际进行数据仓库系统构建的过程中,能够对企业内部现有的系统进行细致的了解和分析,可以将这些数据进行总结,这也是企业原设计管理工作中的重要一步。在这一过程中,用户还希望能够全面的了解本案数据的基本内容,了解数据的基本含义、结构来源和计算公式等,只有对这些问题进行系统的把握,才能够切实做好原数据管理工作。构建元数据管理的过程中,应该对仓库项目的构件进行分析,了解项目构建的主要内容。原数据技术能够有效的实现数据的集成工作数据仓库,在实际运行的过程中要切实做好数据集成,将不同的信息系统和数据进行采集,采集之后还需要借助特定的模式对其进行整理。完成这一系列的流程之后,要按照相关的模式对数据进行储存,将其储存到数据仓库中,在这一过程中所消耗的时间是比较多的,而且要有明确的地点信息,能够对原始数据和数据仓库中的对应关系进行校验,不断对各项规则内容进行转化和过滤。应该真正将这些数据存放在原数据中,这样能够给后续用户的访问带来一定的便利。原数据可以看作是应用集成的基础,所以在进行信息系统构建的过程中,要了解当前的一些问题,实现应用程序之间的数据转换。从当前的数据转化情况来看,整体的数据转换存在一定的困难性,很难从一个系统到另一个系统中进行数据的利用和转换,而且在这一过程中还可能会出现数据不一致的情况。此时就需要充分发挥原数据的作用,将数据仓库中的各个数据源系统进行分析,要真正实现元数据的共享和通信,实现企业应用数据的集成。

在对系统进行应用的过程中，要从根本上提高系统的灵活性，了解当前企业对数据仓库系统的实际需求^[3]。元数据具有一定的智能性，能够对整个系统中的数据的基本情况详细的记录，利用元数据之后，能够对配置进行系统的管理也可以分析相关的通知信息和机制内容。对这些内容进行系统的分析之后，可以给技术人员的数据仓库系统开发工作提供一定的帮助和支持，可以有效的对整个业务的信息流程进行管理，提高系统的灵活性和拓展性。

3.2 构建相关管理模式

原数据管理中所涉及到的内容比较多，既包含数据仓库的构建运行也包含后续的维护工作，所以是企业级数据仓库构建的重要环节。构建数据仓库时，相关人员需要对数据源进行系统的分析，了解ETL过程的基本内容，明确数据库的基本结构，还应该对数据的模型进行分析。了解数据实际应用过程中出现的一些问题，针对数据仓库的基本情况，构建完善的元数据管理架构。从当前数据仓库员记录的实际内容来看，既包含原系统的信息也包含ETL的整个过程。能够对系统外部的应用模式进行操作，了解源数据管理的基本模式，分析系统的运行环境和运行内容。还应该切实做好技术数据的结构分析工作，了解业务数据的结构模式，将这些结构模式真正纳入到元数据的管理范围内。研究ETL语言数据管理主要内容是应该要明确主要的任务，对ETL规则进行收集和管理，切实做好系统数据管理工作，了解数据仓库的映射关系，切实做好数据的转换和清理工作。还要构建数据仓库的逻辑数据模型，对企业原数据的基本内容进行分析。逻辑数据模型的建立也是企业级元数据管理的终极目标，所以相关人员应该要明确数据仓库的基本结构，对仓库内部的元数据管理内容进行系统的研究了解ODS和DW数据结构之间的内容明确相关的业务规则。分析程序数据定义和变量模块，切实做好视图的定义管理工作。

4 元数据管理及其在数据仓库中应用应注意的问题

4.1 与应用系统开发相结合

企业在进行信息化建设的过程中，需要了解行业发展的趋势，根据仓库项目的基本情况制定完整且现代化的源数据解决方案。企业在发展的过程中要想构建元数据管理模式，就需要对当前存在的一些问题进行系统的调整。元数据管理在实际实施的过程中复杂性较强，工作强度较大，所以在实际开展元数据管理工作的过程中，相关管理人员首先应该将元数据管理工作与企业的应用

系统进行开发，而且应该真正加元数据管理和应用系统的开发结合在一起，寻找两者之间的切入点。元数据管理对于企业的软件开发能够起到一定的推动作用，而且在这一过程中可以充分的调动领导的管理意识和管理责任思想^[4]。企业在发展的过程中，应该结合当前原数据系统管理的实际情况，研究当下市场和行业的发展趋势，结合其他企业的应用模式和应用发展经验，对现有的数据仓库进行系统的共建和优化。在对原数据管理模式进行拓展的过程中，还应该构建完善的企业级的元数据管理系统。找出管理系统中存在的一些问题，加大不同系统和软件之间的融合力度。

4.2 建设相关数据模型

企业在开展原数据管理工作的过程中，应该考虑到企业其他系统中的it系统和各个信息技术模式的运输情况，IT系统不同于其他的平台，在实际进行建设运行和使用的过程中，可以对不同的数据结构和数据类型进行定义。即使将其应用到数据仓库之后，元数据管理也可以在短时间之内呈现不同层次的内容。而且这些不同层次的元数据管理，常常是由不同公司和不同产品来进行提供的，在后续使用的过程中，某些工具中也具有一定的原数据管理的功能。在对产品的元数据内容和相关的数据信息进行研究时，要保证元数据的同步性，是比较有挑战性的。相关人员应该基于这种情况进行系统的分析，要想从根本上解决这一问题，首先企业应该了解当前企业发展的基本特点和优势，根据企业的实际情况，构建完善统一的企业级的数据模型，还应该对不同系统的数据进行系统的整合和请求访问。其次，在实际选择源数据管理工作的过程中，相关人员还应该对工具类型进行严格的把控，严格按照原数据的通用标准进行选择构建系统及现代化的源数据模型。

4.3 实现制度内容变革

不断对现有的数据管理制度和管理内容进行变革，了解当前市场环境和行业环境的发展情况，对发展的各种数据进行系统的分析，要将不同的数据分散的分布在不同的部门中。在后续开展元数据管理工作的时候，还应该对这些数据进行仔细的研究，对数据进行系统的整合，保证数据的完整性和规范性^[5]。在这一过程中应该要考虑到具体的整合模式和整合方法，考虑到各部门和系统之间的差异性。要从企业发展的实际情况出发，构建完善的数据模型，控制数据访问的频率和次数。安排专业的部门开展这些工作，不断对企业数据的基本内容进行完善和优化。了解源数据管理的基本目标和基本内

容,完成数据管理的目标,切实推动企业的发展与进步,提高元数据管理的效果。

4.4 重视方法和过程变革

企业在对原数据管理模式进行开发和研究的过程中,除了对现有的内容和系统进行使用之外,还应该在现有的基础上不断进行创新和优化。从当前我国发展的情况来看,原数据管理是一种新型的事物,相关人员在研究的过程中,应该充分发挥和把握原数据管理的具体价值和具体使用价值,真正将元数据管理内容纳入到企业的信息系统建设。将其纳入到企业的信息建设之后,能够对开发人员传统固守的思维进行一定的转变。在这一过程中应该开展系统的培训工作和思想引导工作,要对开发人员的思想进行切实的影响。通过耐心的讲解之后,让开发人员能够意识到方法创新和变革的重要性。在后续进行开发研究和建设的过程中,也能够及时了解行业发展的实际趋势,加大信息化创新和建设力度。

4.5 实现功能多元化发展

元数据管理可以看做是为企业的数据管理工作指明一定的方向,了解企业发展实际情况的基础上构建完善的发展计划和发展体系,能够真正为企业内部的数据管理工作规划理想的蓝图。在发展的过程中为了能够实现美好蓝图的发展,需要在项目中赋予更多的功能和内容,在进行数据仓库项目元数据管理工作的过程中,应该考虑到不同用户的实际需求,要对现有的管理模式和发展模式进行创新,通过创新和优化之后,可以真正给不同的用户提供不同的信息。充分体现和发挥原数据管理的价值,达到更加理想的使用效果。例如,研究数据仓库项目的时候,相关人员应该对原数据管理的模式进行调整,加大开发人员的数据研究力度,提供数据源系统结构。还需要仔细观察ETL的映射过程,对这一过程中的主题内容和各个环节进行系统的分析^[6]。了解主题中所涉及到的一些业务和逻辑,对这些内容进行系统的整合之后,还要考虑到项目的开发效率和开发问题。进行元数据管理工作的过程中,相关人员应该开展系统的维护

工作要求,对系统故障中的一些定位信息进行迅速的把握,提供定位信息之后,可以在短时间之内找出系统运行的故障。开展元数据管理工作的过程中,还应该切实做好系统的升级工作,了解数据的变化情况,分析数据变化之后给系统所造成的影响,对这些影响进行系统的分析之后,还应该对项目的进度进行合理的安排。袁书记的管理者应该要了解数据管理的主要模式以及数据应用者的具体能力和管理方式。要对原有的系统进行调整,真正提供一定的方便信息,可以让用户在短时间之内对某些信息进行访问,不断对信息访问的入口进行管理,充分发挥整个系统的作用和价值。

5 结束语

总之,对于企业的发展而言,必须采用元数据管理工作,将元数据管理工作和数据的仓库系统融合在一起,全面推进企业的城市化进程和发展。实现两者之间的互相促进与融合,在进行数据仓库系统建设和完善的过程中,应该要构建完善的元数据管理模式,并且积极引进现代化的管理体系,全面提高数据仓库的质量和水平。

参考文献:

- [1] 平静, 平林瑞. 元数据管理及其在数据仓库中的应用研究[J]. 平原大学学报, 2006, 23(4): 3.
- [2] 吴文忠, 汤小苹. 元数据管理系统在数据仓库建设中的应用[J]. 新金融世界, 2011(12): 4.
- [3] 徐博艺, 谢诚, 蔡鸿明. 领域本体方法在数据仓库元数据管理中的应用研究[J]. 计算机应用研究, 2010(11): 3.
- [4] 徐博艺, 谢诚, 蔡鸿明. 领域本体方法在数据仓库元数据管理中的应用研究[J]. 2022(11).
- [5] 彭飞. 关于数据仓库中元数据的研究与应用[J]. 硅谷时代, 2009, 000(011): P.75.
- [6] 仇道霞, 石涛, 宋楠. 浅析数据仓库中的元数据管理技术[J]. 山东轻工业学院学报: 自然科学版, 2009(2): 58-62.