

构建数字城市的元数据服务体系^{*}

王 洪¹ 李 琦¹ 董宝青² 李浩川³

(北京大学遥感与 GIS 研究所数字地球工作室 北京100871)¹

(北京市信息资源管理中心 北京100088)² (国家信息中心 北京100053)³

Building the Metadata Services Architecture of Digital City

WANG Hu¹ LI Qi¹ DONG Bao-Qing² LI Hao-Chuan³

(Digital Earth Room of Institute of Remote Sensing & GIS, Peking University, Beijing 100871)¹

(Information Resource Managing Center, Beijing 100088)² (National Information Center, Beijing 100053)³

Abstract Metadata is one of the six key technologies of Digital Earth, which is also an important aspect in building and implementing the Digital City. Based on the technologies and standards of Web Services, the author came up with a metadata services architecture in the distributed heterogeneous network environment of Digital City. In this paper, the metadata services architecture is described and explained in detail, including the related standards, technologies, and the practical experiences in the 'Metadata Sharing Network Project of Digital Beijing' as well. It also points out that, based on the Web service, to build the metadata service architecture will be practically meaningful for organizing, managing and sharing the information resources in Digital City.

Keywords Metadata, Metadata services architecture, Digital city, Digital Beijing, Web services

1 引言

数字城市的建设将满足我国城市可持续发展的信息资源共享、综合决策和技术集成等需求,并最终构筑城市的数字化生存环境^[1]。元数据是数字地球的六大关键技术之一,Web 服务是当今涌现的新技术。基于 Web 服务技术所构建的元数据服务体系将支持数字城市的信息资源的管理、综合开发和利用,对实现城市多源、分布和异构的信息资源的共享、整合和交换起着关键作用。

2 元数据服务体系

2.1 元数据

元数据是关于数据的数据或关于信息的信息,它是实现信息共享的核心内容之一^[2]。尽管有很多不同的元数据标准,如 Dublin 核心元数据标准,地理领域的美国的 FGDC 地理元数据标准,ISOTC211 的空间元数据标准,我国的 NSII 空间元数据标准等,它们都具有以下几个作用:

1. 对信息资源进行高度结构化描述。
2. 可以用来管理和组织信息,支持对信息资源的维护和挖掘。
3. 帮助用户准确地查询检索所需要的信息,进而获取信息。

总之,元数据的建设为我们提供了信息共享的标准化的方法,可以避免部门间数据的重复建设,从而使得有效地进行信息共享和互操作成为可能。这些对于数字城市的信息资源管理均具有重要的意义。

2.2 元数据服务体系

元数据服务体系包括元数据标准体系、元数据服务模型、元数据节点服务体系三个方面。其中,元数据标准体系是由一系列元数据标准组成,元数据标准是由专门机构制定,并根据需求由各个元数据生产和所有者采纳并融入到各个元数据服务节点的具体应用之中,这在元数据的存储和转换、元数据的采集工具和支撑软件中得以体现。目前我国的元数据标准制订工作也在积极进行,其中包括地理领域的空间元数据标准、数字图书馆的中文元数据标准等的制订工作。

在数字城市中,元数据生产和管理单位是元数据的服务的提供者,它们是构成数字城市元数据服务的节点,是构成元数据服务的基本单元,并综合形成元数据节点服务体系。元数据节点服务的构成和具体形式将结合数字北京的元数据项目的实例进行描述。

元数据的服务模型是基于 Web 服务技术和标准的一套技术体系。下面主要讲述这个方面。

元数据服务模型是基于当前涌现的 Web 服务技术而构建的。Web 服务技术是为了解决当前在 Internet 环境下松散耦合的 Web 服务之间进行互相调用而设计的技术框架,其本质是一种支持互操作的协议。以 XML/SOAP/WSDL/UDDI 为主干的 Web 服务技术赋予了 Web 服务一个与传统对象调用技术相似但又不太相同的体系结构^[3~5]。

在 Web 服务的体系构架里有三个角色:服务提供者、服务注册中心和服务请求者。服务提供者是提供最终 Web 服务的提供商,它实现了一个 Web 服务,并放置在线服务器上供别人使用。服务注册中心是一个 Web 服务的注册地,汇集了很多在线的 Web 服务,一般来说,服务提供者将 Web 服务安装到在线服务器之后,会将 Web 服务发布到服务注册中心

^{*}) 本文由北京市信息资源网中心工程总000800资助。王 洪 博士生,研究方向:数字城市、GIS、城市可持续发展。李 琦 教授,博导,主要从事空间信息科学、可持续发展,数字城市的研究工作。董宝青 北京市信息资源管理中心主任。李浩川 国家信息中心方位捷讯科技有限公司,总工。

去,从而使得服务注册中心包含越来越多的 Web 服务的技术信息。目前,服务注册中心即为 UDDI 商业注册中心。对于想要使用 Web 服务的服务请求者来说,一般情况下,它首先去查询注册中心,尝试在注册中心找到所需要的 Web 服务,当它发现了合适的服务之后,将从服务器注册中心获取这些 Web 服务的技术信息引用,通过这些引用找到真正的 Web 服务及其相关的技术信息,从而完成服务请求者和提供者之间的绑定。

对于这三个角色之间的任意两者之间的交互,使用的都是 Web 服务的交互方式,而这些被调用的 Web 服务,其调用界面都是用 WSDL 来描述的。

对于一个数字城市而言,其元数据服务提供者一般为数据生产者或者数据拥有者,如地矿局、测绘局、地税局等单位,他们将自己的元数据服务发布到数字城市的 UDDI 注册中心去。服务请求者为试图提供元数据查询等服务的单位或网站,它们通过查询 UDDI 获取相应的 Web 服务及其相关的技术信息从而完成服务请求者和数据生产者所提供的元数据服务的绑定。对于浏览器端的用户而言,这一切都是透明的,它获取的是通过这套机制而实现的元数据服务。

元数据服务内容由各个元数据服务提供者确定,一般分为:

1. 元数据目录服务:包括元数据的目录分类(Category)的管理,如分类的增加、删除、检索、修改。包括元数据目录的

导入、导出和备份。目录的检索服务支持目录的详细分类列表的获取。

2. 元数据注册服务:在指定的 Metadata Catalog 下元数据库或者元数据纪录的增加、删除、检索、修改。

3. 元数据的查询服务:接受元数据的查询请求,或者向其它元数据服务发送请求。其中可以通过元数据库的名称、关键字或者所在的 Category 进行查找。可获得所需要的元数据库的核心元数据信息或完整的内容元数据信息。

4. 元数据用户反馈服务:管理用户的应用元数据的反馈消息,如对元数据使用情况的评价报告等。

5. 元数据转换服务:有的元数据服务提供者会提供对不同元数据标准支持下的元数据库的转换服务,如空间信息交换中心需要将根据 FGDC 标准建立的元数据库转换成 ISO TC211 标准下的元数据库,这一服务对于空间信息交换中心对空间元数据的统一管理具有很重要的意义。

6. 用户认证服务:确定用户使用元数据的权限。这主要对于进行目录服务和注册服务的用户的权限进行管理。

总之,元数据服务提供者可以根据具体应用情况提供自己的服务,并将它用 WSDL 描述发送到 UDDI 中心里面进行服务注册。其中根据元数据查询检索的特点,协议为 SOAP/HTTP(S)/Z39.50 组成。

元数据服务模型可用图1表示。

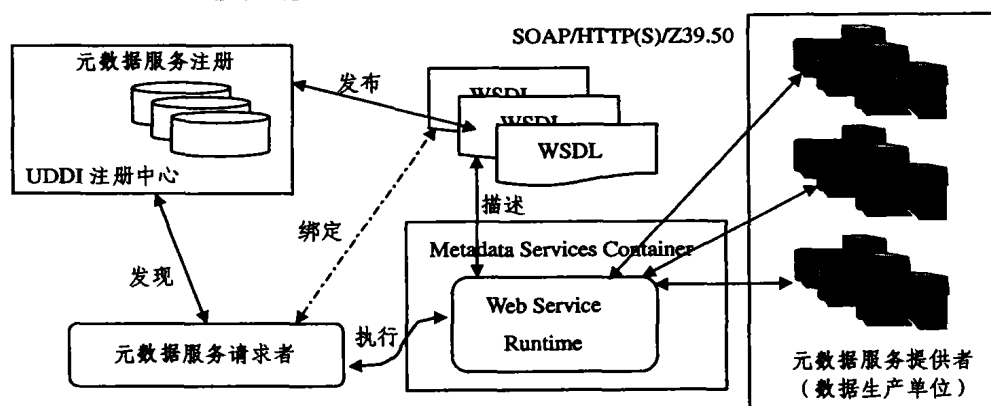


图1 元数据服务模型

3 基于 Web 服务的数字城市元数据共享服务建设

上节已经阐述了数字城市的元数据的服务体系是由元数据标准体系、元数据服务模型、元数据节点服务体系组成。下面结合北京市信息资源网元数据建设实践说明元数据服务节点的应用服务的构建,并对元数据标准做一个初步的探讨。

3.1 北京市信息资源网元数据库建设背景和意义^[6]

北京市信息资源网元数据库建设是北京市信息资源网中心工程的一部分。北京市信息资源网中心工程是首都信息化“十五”规划中确立的21项重大信息化工程之一,目标是建成北京市信息资源网,形成数字北京的基本框架。北京市信息资源网是跨部门、跨领域、跨系统的网络信息系统平台,包括一个中心节点、130个分中心节点、1000个主题数据库、20000个专题数据库,超过20万个共享数据项。数据库分布在基层单位、分中心节点和决策中心。在基层单位设有元数据库、专业库和基础库;在中心节点设有元数据库、市情库、基础库和主题库;在决策中心设有决策库和元数据库。其中专业库是指各部门的专业信息库;基础库是作为信息整合的基准数据库,包括遥感

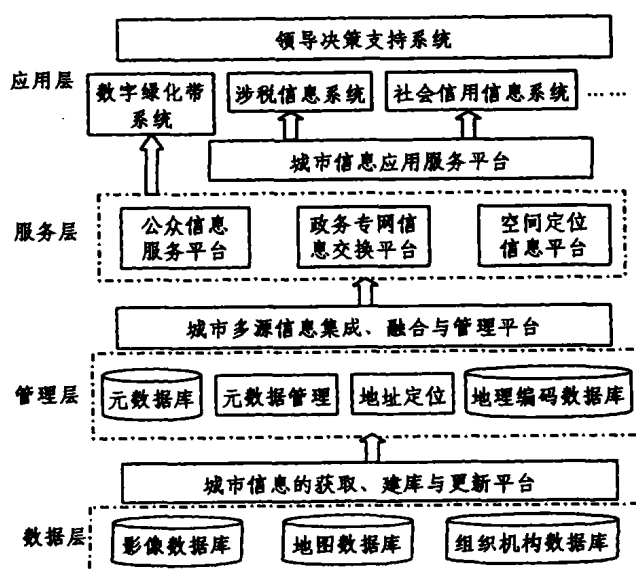


图2 北京市信息资源网中心工程体系结构

影像库、电子地图库、人口基本单元数据库等;主题库是按主题组织的数据库,如信用数据库、税务数据库等;决策库是为支持决策而经过分析综合的信息库;元数据库是根据相关的元数据标准对各数据库进行描述的信息库。

北京市元数据工程建设是北京市信息资源网的基础性工程之一(图2)。它是信息资源网络各种应用和增值服务的基础和前提条件。同时也是信息资源网实施服务的必要条件。其具体目标为:

1. 建设信息资源元数据库群。在北京市各委办局及区县级机构建立符合规范的信息资源和服务资源元数据库群。

2. 建立信息资源元数据共享网络。基于首都公用信息平台,建立连接所有机构的元数据共享网络。实现对北京市各类信息资源的规范化管理,推动信息资源共享。

3. 推动应用服务整合。在 Web 服务技术的支撑下,统一网络应用服务的描述规范,在元数据共享网络和元数据服务体系的支持下,提供网络应用的发现服务,进一步推动应用整合。

3.2 北京市信息资源网元数据节点服务体系的构建^[7]

元数据服务节点是由提供元数据服务的数据生产单位或者管理单位构成的。这些服务提供者将他们的服务发布到 UDDI 注册中心供服务请求者查询,在 SOAP/HTTP(S)/Z39.50 协议的支持下对他们的元数据网关所注册的节点分发查询请求,或者直接对元数据发布服务器分发查询请求,并最终返回查询结果给用户。下面结合信息资源网元数据共享体系结构图来说明元数据节点服务体系的组织和运行概念。

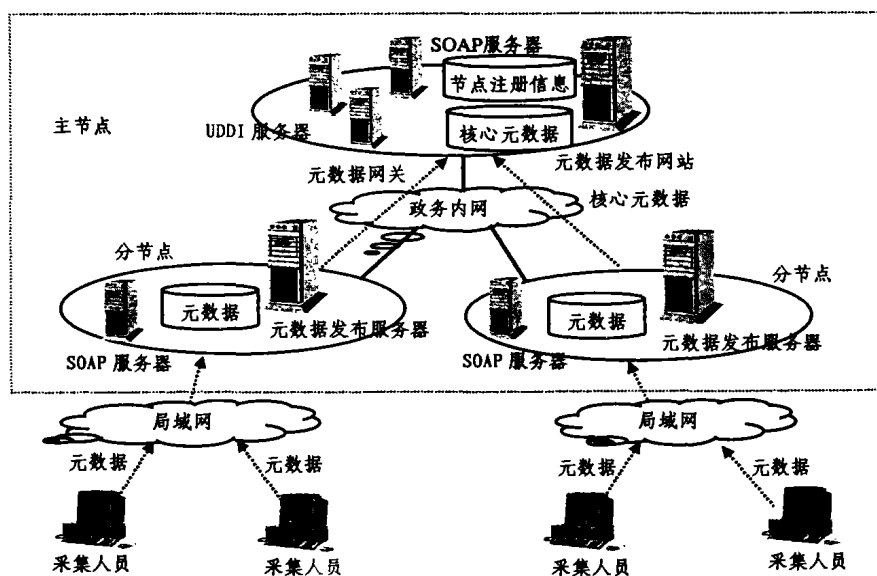


图3 北京市信息资源网元数据共享体系结构

在图3中体现了主节点和分节点的两级结构,虚线表达了数据采集和推送的过程和内容,实线指网络连接。其中主节点是由北京市信息资源管理中心进行管理,分节点是由各个委办局的信息中心进行管理。

在这个元数据共享网络中,元数据划分为两级:核心元数据和内容元数据。其中,信息资源网的核心元数据是描述分节点的基本注册信息和元数据类别的描述信息的集合,这由信息资源管理中心集中管理,由各个分节点产生并提交。内容元数据是完整的信息资源元数据,由各分节点进行采集和管理。

元数据网关是信息资源网主节点与各分节点上元数据发布服务器的连接枢纽,其主要作用是管理各元数据发布服务器,向发布服务器分发门户网站提交查询请求和返回查询结果。元数据网关主要功能有:Z39.50协议查询代理、元数据发布服务器注册管理、元数据发布服务器连接控制、访问日志等。

元数据发布服务器是支持 Z39.50 协议的服务器软件,用于提供网络信息搜索和提取服务。该软件利用抽象数据库机制保证了查询过程与后台数据库实现的无关性,可以有效地支持对多标准元数据联合查询。元数据发布服务器在主节点中用于提供核心元数据和应用服务描述信息的查询提取服务。

在主节点和分节点均有 SOAP 服务器,来支持各节点对 UDDI 注册中心的服务注册以及基于 SOAP 的消息传输机

制。在元数据服务体系中,对主节点和分节点的元数据查询请求均遵循 SOAP/HTTP(S)/Z39.50 协议。

UDDI 服务器支持整个信息资源网的应用服务注册。元数据服务是其中的一种服务。根据信息资源网工程的建设内容,它还应包括地理编码服务、地图服务等。UDDI 注册中心所注册的元数据服务主要是元数据发布服务器和元数据网关所提供的服务内容。

北京市信息资源网元数据共享网络系统的主要组件可划分为元数据采集和管理、元数据网络发布和用户交互三个层次的9个组件。

第一层为用户操作类组件,包括元数据发布网站、中心管理系统、节点管理系统和元数据编辑器,这些组件分别被网络用户、中心节点和分节点管理员、数据单位的数据采编人员使用;

第二层包括元数据网关、元数据发布服务器、SOAP 服务器和 UDDI 服务器,构成元数据共享网络的发布服务层,完成元数据的网络发布服务以及 UDDI 元数据服务注册;

第三层由数据库访问组件、元数据库和环境数据库构成,是元数据共享网络的数据存储层组件,用于元数据、核心元数据和各类系统数据的存储和管理。

3.3 对北京市信息资源网元数据标准体系的思考

北京市信息资源网元数据标准体系是一种由多领域、多(下转第107页)

作是一个给定空间内的搜索问题。但是,与普通的搜索问题不同,参与协商的 Agent 既是相互合作,也是相互竞争,它们之间的交互既要交换信息,同时交换的信息中也含有欺骗,这使得协商过程复杂化。

近年来,拍卖方法作为一种规则明确、便于操作的协商协议逐渐受到研究者的关注。Jenings 在 2002 年将协商分为三类:基于拍卖的协商、一对一协商和劝说式协商,并认为拍卖是 MAS 中最主要的协商手段。拍卖方法可以快速有效地完成一对多和多对多的协商。

Rina 在 2000 年为解决多服务器(每个服务器为一个 Agent)之间的数据分配问题,引入组合拍卖方法,提高了系统的性能。将组合拍卖应用于多 Agent 旅行商问题,也取得了一定的效果。

结语 拍卖是一种快速有效的资源分配方法,具有较强的可操作性,可使参与拍卖的卖方和买方均获得理想的效用。本文系统地介绍了已有的拍卖方法,并提出了一个可以涵盖大多数拍卖方法的形式化拍卖模型。拍卖方法的研究包括拍卖规则、买方和卖方采取的策略和对拍卖方法性能的分析。作为一种主要的协商手段,拍卖方法可以应用于 MAS 完成一对多和多对多的任务分配、资源分配。不同的拍卖方法具有不同的特点,应根据不同多 Agent 系统的具体环境选择合适的拍卖方法。选择时主要考虑效用、实时性、Agent 的计算能力、Agent 的通信能力、资源与任务的互补性与可替换性等因素。例如单边拍卖时,若资源与任务互补性与可替换性强、卖方计算能力较强时则选取组合拍卖,否则选择单一物品拍卖方法;若多个 Agent 具有共同目标,互相信任并合作求解问题,则可以放宽对抗假名叫价鲁棒性的要求。

(上接第 87 页)

行业的元数据标准的集合。其中在地理空间领域,北京大学数字地球工作室、国家信息中心和国家基础地理信息中心在空间元数据标准研究和实践方面做了大量的工作。在非空间领域,北京数字图书馆和北京大学图书馆等单位对中文元数据标准进行了研究和探索。但至今尚未有一个国家级的元数据标准颁布,因此在元数据标准建设方面相对滞后,基本还停留在探索 and 科研阶段。针对北京市信息资源网的建设情况,我们认为,元数据标准体系的建设内容应该包括:

1. 元数据目录体系的标准化建设。元数据的目录是元数据的元数据,是对元数据提供的服务进行的分类,是支持元数据服务的重要方面。目前,我们将元数据从总体上划分为空间元数据和非空间元数据,并对这两类元数据进一步分类以期形成数字城市的元数据目录分类体系。

2. 核心元数据标准的建设。核心元数据是描述节点数据库的基本注册信息和元数据类别以及节点元数据服务内容的描述信息的集合。可以参照 Dublin 核心元数据标准并根据元数据服务内容进行扩展以制定符合数字城市信息资源特征的核心元数据标准。

3. 内容元数据的标准化建设。内容元数据标准确定了对某种信息进行规范化描述的基本内容和模式。它需要提供关于所描述的信息的可用性、适用性、获取或使用方式的详细信息,为用户查找、理解、获取和使用数据提供方便。内容元数据标准是根据元数据服务对象和领域,针对特定的元数据服务内容进行的元数据标准化建设,如空间元数据内容标准、税务元数据内容标准等。

但拍卖方法也有局限性。考虑 Agent 策略时,可将拍卖看作一种博弈。较复杂的拍卖方法中虽然博弈存在 Nash 平衡点,但由于 Agent 的计算能力有限与信息的不完全,求解最优策略往往是不现实的。一般只能通过各种学习方法辅助 Agent 决策,希望在多次行为后接近或收敛到最优解。拍卖方法不适合多问题的协商,也不能完成一对一的协商。此外,单边拍卖尤其是组合拍卖中过分依赖卖方(Auctioneer),导致这样的多 Agent 系统不能看作是完全分布式的系统。

参考文献

- 1 Andersson A, Tenhunen M. Integer Programming for Combinatorial Auction Winner Determination, ICMAS, 2000. 39~46
- 2 Boutilier C, Goldszmidt M, Sabata B. Sequential Auctions for the Allocation of Resources with Complementaries, IJCAI, 2001
- 3 Yokoo M, Sakurai Y, Matsubara S. Robust Multi-unit Auction Protocol against False-name Bids, IJCAI, 2001. 1089~1094
- 4 Yokoo M, Sakurai Y, Matsubara S. Robust Combinatorial Auction Protocol against False-name Bids, AAAI, 2000. 110~115
- 5 Yokoo M, Sakurai Y, Matsubara S. Bundle Design in Robust Combinatorial Auction Protocol against False-name Bids, IJCAI, 2001. 1095~1101
- 6 Tennenholtz M. Some Tractable Combinatorial Auctions, AAAI, 2000. 90~97
- 7 Sandholm T. An Algorithm for Optimal Winner Determination in Combinatorial Auctions, IJCAI, 1999. 542~547
- 8 Sandholm T, Suri S. Improved Algorithms for Optimal Winner Determination in Combinatorial Auctions and Generalizations, AAAI, 2000. 90~97
- 9 Sandholm T, Suri S. CABOB: A fast optimal algorithm for combinatorial auctions, IJCAI, 2001
- 10 Vickrey. Counterspeculation, auctions and competitive sealed tenders. Journal of Finance, 16, 8~37
- 11 Fujishima Y, Leyton-Brown K, Shoham Y. Taming the Computational Complexity of Combinatorial Auctions: Optimal and Approximate Approaches, IJCAI, 1999. 548~553

结论 基于 Web 服务的元数据服务体系的构建将支持数字城市的信息资源的管理、综合开发和利用,对实现城市多源、分布和异构的信息资源的共享、整合和交换有着重要的现实意义。针对我国数字城市元数据建设的状况,需要进一步加快元数据标准的研究和制定的步伐,在管理上、法律上和技术上共同推进元数据服务体系的建立和实施。

参考文献

- 1 王洪,李琦,承继成. 数字城市与城市可持续发展[J]. 中国人口、资源与环境, 2001(2): 114~119
- 2 中国可持续发展共享示范项目组. 地理信息 Metadata 研究、标准和实施(论文集). 1998
- 3 <http://www.webservice.org>
- 4 <http://www-106.ibm.com/developerworks/webservices>
- 5 <http://www.uddi-china.org>
- 6 北京市信息资源管理中心. 北京市信息资源网中心工程研究报告. 2002.4
- 7 北京市信息资源管理中心元数据项目组. 北京市信息资源网元数据共享网络体系总体设计报告. 2002
- 8 国家信息中心. 国家空间信息基础设施(NSII)空间元数据标准(草案). 2000
- 9 承继成,王宏伟,等. 走向城市信息化时代—数字城市的理论、方法 and 应用[M]. 中国城市出版社, 2001
- 10 承继成,林晖,等. 数字地球导论[M]. 北京:科学出版社, 2000
- 11 李琦,吴少岩. 数字地球—人类认识地球的第三次飞跃[M]. 北京:北京大学出版社, 1999
- 12 陈爱军. 空间信息共享与标识语言研究.[北京大学博士论文]. 2000
- 13 赵永平. 基于国家空间信息基础设施的 Metadata 研究及其共享示范体系的建立.[北京大学博士论文]. 1998
- 14 <http://www.digitalbeijing.gov.cn>
- 15 <http://www.cybergis.org.cn>
- 16 <http://www.w3.org>
- 17 <http://www.opengis.org>
- 18 <http://msdn.microsoft.com/webservices/>