

国家环境数据中心建设探讨*

Discussion of State Environmental Data Center Construction

■ 李顺 徐富春 孔益民 张波 (国家环保总局信息中心, 北京 100029)

摘要 结合我国环境数据共享现状, 对国家环境数据中心的建设思路、总体目标、技术框架、主要建设内容等进行了研究和探讨。国家环境数据中心由国家级数据节点和省级数据节点组成, 建设内容主要包括共享技术保障体系建设、数据资源建设、共享与服务网络平台建设三部分。

关键词 数据中心 环境数据 数据共享 电子政务

如何真正实现科学数据的有效共享, 保证这些数据发挥最大效益, 是一个长期困扰科技界的问题^[1]。2001年以来, 科技部在原来的“中央级科研院所科技基础性工作专项”等科技计划项目的基础上进行整合, 启动了“科学数据共享工程”^[2], 并逐步开始了试点工作。其目的是充分运用现代信息技术和国际资源, 搭建具有公益性、基础性、战略性的科技基础条件平台, 有效改善科技创新环境, 增强科技发展能力, 为科技长远发展与重点突破提供强有力的支撑^[3]。2004年底, 在科技部的领导和支持下, 国家环保总局启动了“环境科学数据库建设与共享”的项目, 目的是以环境质量、环境科研和生态环境数据为核心, 研制一批高质量的具有环保系统数据特征的标准数据集, 初步建立国家级环境科学分布式共享服务网络体系。国家环保总局信息中心作为项目的主要参与单位, 结合具体工作和“科学数据共享工程”的要求, 就国家环境数据中心的建设思路、总体框架、主要建设内容等进行了研究和展望。

1 建设思路与总体目标

1.1 建设思路 (见图1)

以满足国家、社会公众和各级环境管理工作对环境数据的共享需求为目的, 依托各业务司、直属单位成熟的业务技术体系, 以现有环境数据资源为基础, 逐步吸纳国内相关领域和国际数据资源, 通过整合集成。标准化和归一化处理, 形成一批以环境质量、环境统计、污染源管理、生态环境管理为核心, 涵盖环境保护范畴的数据集产品; 采用大规模分布式数据库技术、数据仓库技术、WEB技术、XML数据交换等技术, 建立与部门基础信息库相协调的分布式共享数据库系统, 分别建设基于因特网和环保总局电子政务外网的环境数据共享服务网络体系; 研究并制定各种面向用户的标准数据格式和存储模型, 开展环境数据应用服务技术研究, 建立持续稳定的环境数据共建共享运行机制; 培养一批从事环境科学数据研究、管理和服务的人才队伍。

* 国家科技基础条件平台工作项目: 环境科学数据库建设与共享 (编号 2004DKA20310)。

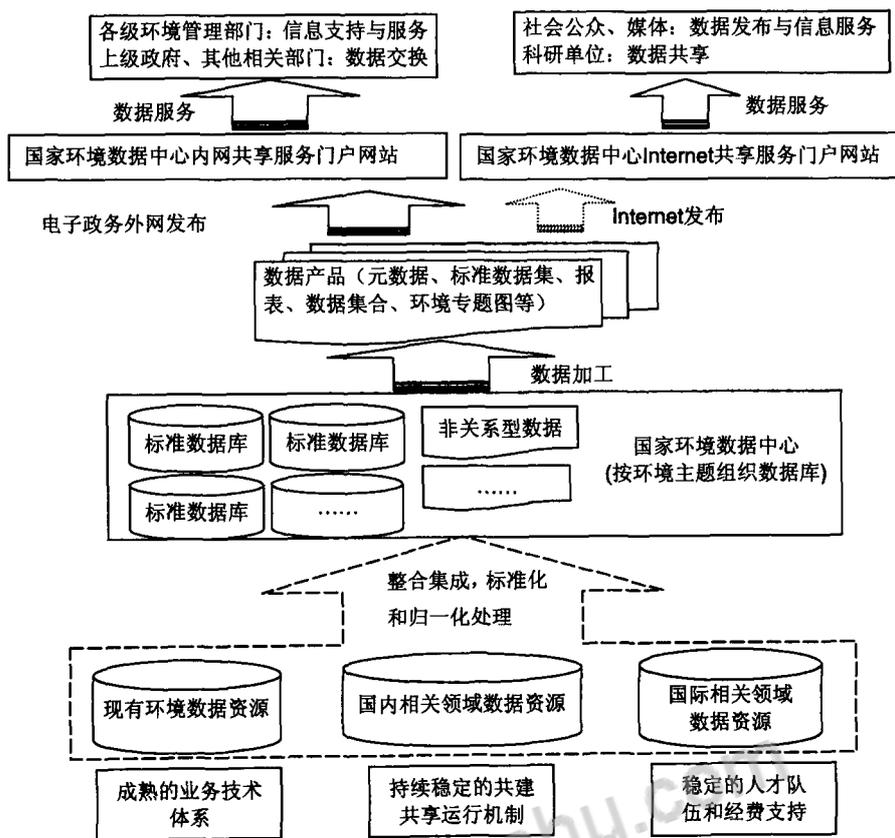


图1 国家环境数据中心总体建设思路

1.2 总体目标

通过5年（2005—2010年）的努力，初步建成国家环境数据中心，形成整个部门的共享网络体系，从而全面提高环境数据管理水平，极大增强环境数据共享服务能力，为环境管理、政府决策、环境信息公开提供全面、多层次的环境数据服务。具体表现在：

(1) 建立起稳定持续的环境数据共享机制和支持环境，包括全面的支持数据共享的政策法规、环境数据的规范标准和业务化的组织机构、高水平的人才队伍。

(2) 研制一批高质量紧密结合环境业务的标准数据集，使得国家环境数据中心成为环境数据的收集管理中心和共享发布中心。

(3) 分别建成依托环保总局电子政务外网和依托因特网的分布式共享服务网络体系，网络在线提供数据占共享总数量比例达到50%以上。

(4) 环境数据加工处理技术接近国内先进水平，

在多元数据综合集成、环境数据组织表现、数据分析挖掘等方面取得长足进步，使得国家环境数据中心成为环保总局内部和社会各级用户获取环境数据的“第一选择”。

2 总体结构与建设内容

2.1 总体结构(见图2)

国家环境数据中心是由国家级数据节点和省级数据节点有机组成的覆盖全国、分布式的网络化环境数据共享服务系统。国家级由国家环保总局环境数据中心、国家环境监测数据中心（中国环境监测总站）、环境卫星资料中心（建设中的国家环保总局环境卫星中心）、环境科研资料中心（中国环境科学

研究院、国家环保总局华南环境研究所等）、国际数据资源共享系统（主要通过URL定位和国际资料收集整理整合实现）组成，省级系统由各个省级节点组成。

2.2 建设原则与管理模式

2.2.1 建设原则

统筹规划、协调发展 国家环境数据中心是“国家科学数据共享工程”建设的重要组成部分，必须与国家科学数据共享工程总体推进，遵从统一规范和建设要求，采用一致的政策和标准。同时，国家环境数据中心也是环保部门整个业务技术体系的必要组成，也需要与整个环境保护事业现代化发展相衔接，紧密依托行业信息系统，与业务系统建设统筹考虑，协调发展。另一方面，国家环境数据中心的建设与一般科研项目的最大区别就是需建立持续稳定的环境数据共享服务机制，表现在具有稳定的数据追加更新能力，持续的在线共享能力和稳定的专业队伍支持，所以，国家环境数据中心的建设成果最终需转化为稳定的业务运行体系。

以需求为导向，边建设、边服务 国家环境数据中心的建设必须以国家、社会、环境管理部门对环境数据共享需求为导向，本着“边建设，边服务”的原则开展，在建设过程中注重服务效益的体现，不断完善环境数据管理与服务体系的布局，促进互联互通、资源共享。

重点突出，逐步推进 国家环境数据中心建设

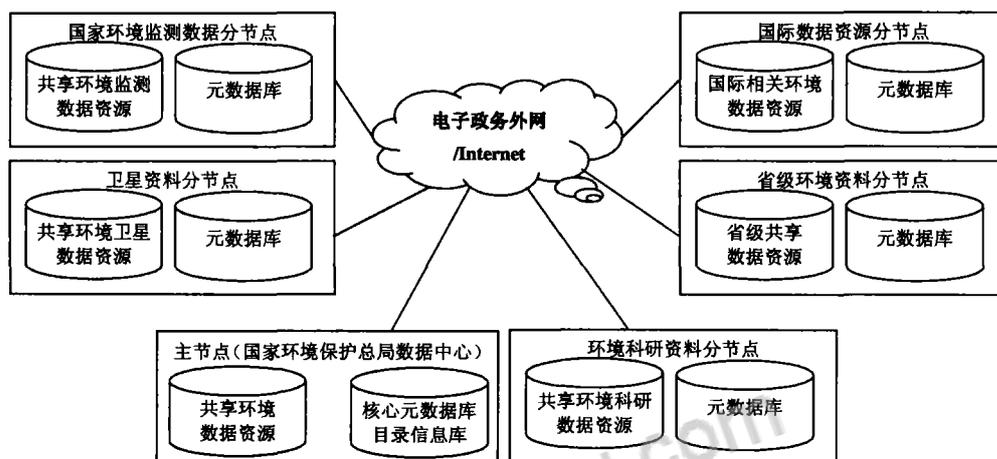


图2 国家环境数据中心总体结构

的重点是整合全部门的数据资源，不断吸纳行业和国际上的相关数据资源，利用国家和部门网络形成分布与集中相结合的能够快速提供共享服务的网络体系和共建共享、持续运行机制。根据部门内部实际情况和国家支持力度，遵循统一设计、分步实施原则，由易到难，由内到外、逐步推进原则。

2.2.2 管理模式

国家环境数据中心实行统筹规划、分层管理、数据共享的建设模式。其中由国家环保总局组织建设国家环境数据中心主节点（主中心），重要数据责任单位和省级环保厅局可以申请并经国家环保总局电子政务领导小组批准后建立分节点（分中心）。国家环保总局电子政务领导小组负责国家环境数据中心的规划建设与综合管理。国家环境数据中心主节点的软硬件平台由国家环保总局信息中心负责建设、运行与维护工作；列入环境保护重要业务数据目录的相关责任单位负责按照规定的时间、频次、格式和规范对数据进行采集、加工、传输、存储和发布；国家环境数据中心各分节点由各级责任单位按照主节

点统一的标准、规范组织建设实施。

国家环境数据中心数据共享服务采用分级用户认证授权方式。由主节点系统进行一级用户认证管理和系统访问授权管理，由分布式分节点系统进行各自平台和单位内部的二级用户认证管理和系统访问授权管理。用户可以通过主中心提供的数据共享服务平台系统，对整个共享网络系统内的数据资源进行透明的

访问并获得多种共享交换功能服务；用户也可以直接访问某个分平台系统，对分平台系统的数据资源进行透明的访问并获得多种共享交换功能服务。

2.3 建设内容

国家环境数据中心建设内容围绕三个方面开展，即：环境数据共享技术

保障体系建设，环境数据资源建设和环境数据共享与服务网络平台建设。

2.3.1 共享技术保障体系建设（见表1）

通过对数据分级分类、发布策略、数据格式、质量控制标准的研究和建设规范、运行管理制度的制定，构建环境数据中心的技术保障体系，形成环境数据共建共享和持续运行机制。

2.3.2 数据资源建设

数据资源建设最终将产生各类标准化数据集产品，为环境数据共享提供源源不断的资源，其主要任务是整理整合各类环境业务、环境政务、环境科研数据，形成标准格式的数据库或数据集，并对这些标准的数据库或数据集进行加工处理，形成可用于发布或共享的数据产品（元数据、标准数据集、报表、数据集、环境专题图等）。主要建设内容包括：环境元数据、环境法规与标准数据、环境文献与公报数据、环境质量数据、环境统计数据、环境背景数据、生态环境保护数据、生物多样性保护数据、辐射环境数据、其他环境管理相关数据等。

表 1 环境数据中心技术保障体系主要建设内容

类别	主要建设内容
总体架构设计	数据中心硬件平台架构、软件平台架构、基础通信平台、网络互联互通、数据节点间数据协同方式设计等。
通用技术标准研究	包括基本环境数据集元数据集标准、数据中心数据库开发设计规范、数据中心数据接口及基础数据交换标准、数据中心数据库数据字典规范、数据中心通用代码标准、数据集加工汇交流程、数据集说明文档格式标准等。
数据安全存储	数据库存储与数据备份方案、数据保护与灾难恢复制度等。
管理与评估规范	用户分级方案、环境数据及产品分级共享实施细则、环境数据共享管理办法、数据汇交管理办法、数据产品成本核算办法等。

2.3.3 共享与服务网络平台建设

国家环境数据共享与服务技术平台具有环境元数据导航、信息管理、分析、查询和发布功能，开展多层次、多目标的环境数据分析、处理、共享与应用服务，并能通过广域互联网方式提供具有综合性、全面性、权威性和实时性的环境数据服务，全方位的支持国家环境数据的发布，为各级政府部门、科研机构和

社会公众迅速方便地获取环境信息及其技术服务，促进环境信息大范围、高效率的共享和应用提供急需的数据基础和技术手段。

环境数据共享服务网络平台开发建设主要包括 4 部分内容 (图 3): 共享数据集与元数据汇交系统开发建设、数据中心数据库管理系统开发建设、环境数据中心建设 (软硬件环境搭建及部分数据资源整合)、

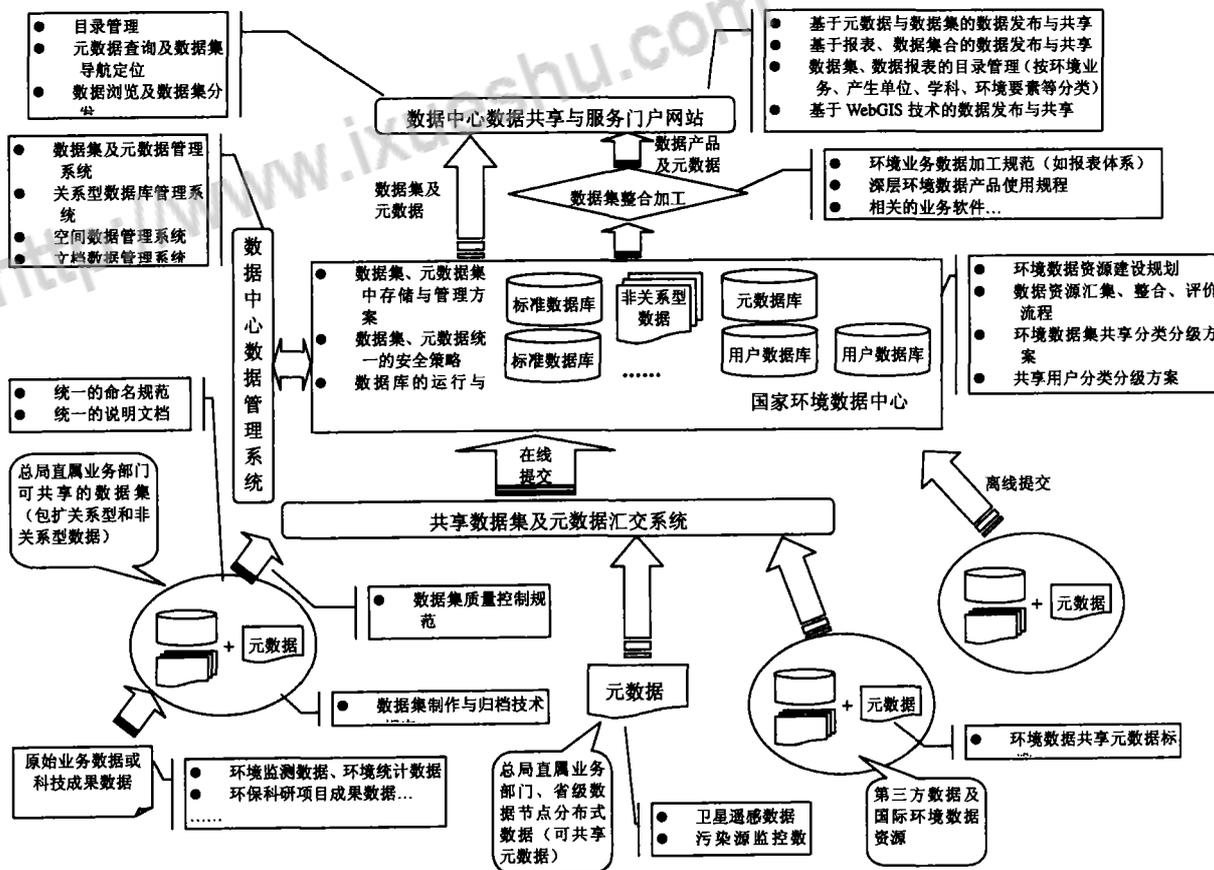


图 3 环境数据共享服务网络平台的技术框架

表2 环境数据共享服务网络平台主要建设内容

建设内容		主要功能	面向用户	涉及标准规范
共享数据集及元数据汇交系统		元数据的上载,用户通过上载接口,提交元数据到共享网络中心; 真实数据的上载,用户除了上载相应元数据之外,还可以通过上载接口上载元数据描述的真实数据; 数据的检查审核,以及数据的实时更新和数据的后期处理。	数据中心维护人员; 各级数据提供单位或个人。	数据集制作与归档技术规定; 数据集命名规范与说明文档; 数据集汇交办法; 数据集质量控制规范;
数据中心数据管理系统	元数据管理系统	元数据模板定义; 元数据描述; 元数据维护、查询、发布等。	数据中心维护人员	环境数据元数据标准; 数据库设计规范; 环境数据集共享分类分级方案; 共享用户分类分级方案;
	基础数据管理系统	主要包括数据的逻辑组织与管理; 基础数据的维护(数据报表、视图的定制与管理,批量更新、数据查询等); 数据统计;数据装载与导出; GIS接口管理等。		
	数据集管理与发布系统	数据报表、数据集、其他非关系型数据单元的逻辑组织与管理、审核与发布。		
	数据安全管理系统	主要包括用户管理、权限设置与认证、分节点元数据的同步、数据备份、日志管理等。		
分布式环境数据中心(主节点、分节点)		各级汇交数据资源(包括元数据资源)的存贮与管理(信息中心目前主要负责软硬件环境的搭建及部分已有数据资源的整理发布)	数据中心维护人员(包括各数据单位数据维护人员)	数据集、元数据集中存储与管理方案; 数据集、元数据统一的安全策略;
数据中心数据共享与服务门户网站		目录管理、元数据查询及数据集导航定位、信息浏览与数据产品分发等。 其中数据发布方式主要包括:基于元数据与数据集的数据发布与共享;基于报表、数据集的数据发布与共享;数据集目录管理(按环境业务、产生单位、学科、环境要素等分类);基于WebGIS技术的数据发布与共享等。	各级用户	环境业务数据加工规范(如报表体系); 环境数据产品分发与使用规程;

环境数据中心门户网站建设。其内容见表2。

结语

环境科学数据是国家基础信息资源的重要组成部分,随着全社会对环境问题的日益关注,社会各部门和公众对环境数据共享与服务的需求也越来越迫切。建设国家环境数据中心,整理整合并共享各类环境数据,不仅是社会经济发展和科学创新工作的需要,也是我国环境保护工作的必然要求。与国家气象局、国土资源部相比,环保总局在国家环境数据中心建设和环境科学数据共享工作开展方面存在环境管理业务体系不够成熟、数据资源相对薄弱;环境数据资源存储管理分散,加工处理能力较弱;还未形成数据资源共

享的体制和方式等困难。对国家环境数据中心的总体结构、建设内容、管理模式、技术框架等进行研究,对于全面推动我国环境数据共享有着深远意义。^①

参考文献

- [1]孙九林,黄鼎成,李晓波.我国科技数据管理和共享服务的新进展.世界科技研究与发展,24卷第5期,15-19.
- [2]林业科学数据中心编.林业科学数据库和数据共享技术标准与规范(第一辑).中国林业出版社.
- [3]中华人民共和国科学技术部 国家科技基础性工作专项 科学数据共享——当代科学界的战略选择 <http://jcxgz.most.gov.cn/shownews.asp?newsid=1155>.



论文写作，论文降重，
论文格式排版，论文发表，
专业硕博团队，十年论文服务经验



SCI期刊发表，论文润色，
英文翻译，提供全流程发表支持
全程美籍资深编辑顾问贴心服务

免费论文查重：<http://free.paperyy.com>

3亿免费文献下载：<http://www.ixueshu.com>

超值论文自动降重：http://www.paperyy.com/reduce_repetition

PPT免费模版下载：<http://ppt.ixueshu.com>
