

# 国家自然科技资源 e - 平台建设实践

王运红<sup>1</sup> 张 莞<sup>2</sup> 沈欣媛<sup>2</sup>

(1. 中国科学技术信息研究所, 北京 100038;

2. 北京万方数据股份有限公司, 北京 100044)

**摘 要:** 自然科技资源信息共享是促进实物共享的重要手段, 建设一个高质量的信息共享平台, 可以为自然科技资源的实物共享提供更加广阔的渠道和有效的保障。国家自然科技资源 e-平台建设是国家重要的科技计划项目之一, 2004年1月正式启动, 由中国科学技术信息研究所牵头承担建设。本文主要从设计和技术角度介绍国家自然科技资源 e-平台的建设情况。

**关键词:** 资源; 自然科技资源; 信息共享; 信息平台

**中图分类号:** G203 **文献标识码:** A **DOI:** 10.3772/j.issn.1674-1544.2008.04.004

## Research and Practice on National Infrastructure of Natural Resources for Science and Technology of China

Wang Yunhong<sup>1</sup>, Zhang Wan<sup>2</sup>, Shen Xinyuan<sup>2</sup>

(1. Institute of Scientific & Technical Information of China, Beijing 100038,

2. Wanfang Data Co. Ltd, Beijing 100044)

**Abstract:** Information sharing as a tool could be help to improve sharing of natural resources for science and technology. The high-quality information-sharing infrastructure can offer all kinds of ways for sharing of natural resources for science and technology and give a strong support to it. The construction of National Infrastructure of Natural Resources for Science and Technology is an important project of the national plan for Science and Technology. It started in January, 2004, and implemented by Institute of Scientific and Technical Information of China. In this article, we introduce how to design the infrastructure and the technologies.

**Keywords:** Natural resource, natural resources for science and technology, information sharing, information infrastructure

自然科技资源是在科技基础条件平台建设的大背景下提出来的, 是指自然生成、有人类科学研究活动加于其上, 并对人类科学技术研究和科学技术进步起基础作用和支撑作用的物质材料<sup>[1]</sup>。目前比较认可的自然科技资源的定义是, 自然科学资源“一般指经过长期演化自然形成(如化石、岩

矿)及人为改造(包括收集、整理等)的、对人类社会生存与发展不可或缺的、为人类社会科技与生产活动提供基础材料, 为科技创新与经济发展起支撑作用的重要物质资源”<sup>[2]</sup>。国家科技基础条件平台提出的自然科技资源主要包括植物种质、动物种质、生物标本、岩矿化石标本、人类遗传、微生

第一作者简介: 王运红(1971- ), 女, 工程师, 研究方向是科技资源管理、信息项目管理。

基金项目: 国家科技基础条件平台建设项目“国家自然科技资源 e-平台建设”(2005DKA21603)。

收稿日期: 2008年6月1日。

物菌种、实验材料、标准物质八大类资源<sup>[3,4]</sup>。

为扭转我国目前自然资源保存与利用中的无序、分散、流失、低效的局面,需要构建数字化和网络化的有序、系统、高效的保存利用体系<sup>[5]</sup>。国家自然资源 e-平台建设是国家重要的科技计划项目之一。e-平台是自然资源的信息共享平台。该平台的建立促进了自然资源共享。自然资源共享是一个复杂的系统工程,涉及资源单位、资源用户、政府部门、国家、社会等多个主体,发生整合、利用、监督、服务等多种行为,关系个人利益、国家利益和社会利益等多元利益。一个集成资源信息和共享服务的共享平台,提供了有利于共享和可持续发展的技术保障,从而促进共享参与主体的目标一致,减少和消除共享中矛盾和冲突,进而实现共享行为的巨大合力。

## 1 国家自然资源 e-平台综述

### 1.1 建设背景

国家自然资源 e-平台(以下简称 e-平台)建设工作,是国家科技基础条件平台建设的重要组成部分。自然资源综合数据库体系、国家自然资源子门户管理和应用系统(以下简称“子门户”)、国家自然资源虚拟博物馆三位一体构成了国家自然资源信息共享平台,即国家自然资源 e-平台。自 2004 年 1 月 1 日正式启动后,e-平台完成了系统规划、数据管理、环境建设、系统建设等工作。该平台最终将实现我国自然资源共享信息的高度集成和多方位服务,综合提高资源信息的利用效率和效益;把自然资源信息共享在数据、系统、管理等层面上做到高度融合。

平台建成后,可以从信息来源、技术支撑等方面为信息的共享提供强有力的支撑,弥补国内现有单一专业资源数据库系统和网站的许多不足,提升自然资源信息和实物的综合服务能力。

### 1.2 建设目标

按照国家科技基础条件平台的整体部署,e-平台的建设目标是围绕“整合、共享、完善、提高”

的建设思路,按照统一的数据描述标准和数据质量规范,建立完善的国家自然资源信息共享体系,把分散在全国各地的八大类自然资源信息进行整合、高度集成为一个集中的、虚拟的自然资源“库”或“馆”,以便于国家更好地管理和社会更有效地利用自然资源。利用平台提供的图片、音频、视频、3D 等多媒体信息进行自然资源的可视化和多元化表达,构建自然资源信息共享网络服务系统,开发平台在科技创新、人才培养、科学普及等领域的服务功能;实现平台与国际接轨,形成一个集成、高效、开放、连通的,具有信息整合共享、信息导航、资源可视化虚拟展示的国家自然资源信息平台,实现我国自然资源信息的全面共享和应用服务,进而推动自然资源的实物共享<sup>[6]</sup>。e-平台的建设目标,如图 1 所示。



图 1 e-平台总体建设目标

## 2 平台设计思想

e-平台在总体设计上,采用成熟的主流数据整合和信息服务技术,在统一的资源描述规范和数据整合规范指导下,建立一个详情信息分布式存储,共性信息集中式管理,提供一站式信息服务的多层协同、高速并行的信息共享平台。e-平台的总体建设工作分为如下 4 个部分:①实物资源信息化:由各资源保藏单位根据共性描述规范的要求,完成本单位资源的共性描述信息的加工处理,然后上报给 e-平台;②数据整理和整合:对资源保藏单位上报的资源共性描述数据进行加工、整理、整合、质量控制和数据管理;③建设平台数

数据库体系：一套完善、科学的数据库系统，保障数据的存储和有效利用，e-平台的数据库体系分为八大类自然资源共性描述信息数据库、分类库、发布信息数据库、资源关联信息数据库、统计信息库、管理信息库六大类；④建设平台门户系统和后台管理运行系统：将整合后的自然资源资源信息进行多角度发布和展示，为用户提供多元化的信息和知识服务，后台管理系统为整个平台的管理提供支撑。实物资源信息化工作是整个信息共享平台建设的基石，是信息共享的前提；数据整合是数据集中发布的数据基础，而数据发布前的质量管理是使数据科学、正确的有力保障；平台数据库提供了数据的存储保障。这些基础工作完成后才能进行后台管理系统和平台门户的建设。这几部分工作环环相扣，缺一不可。

### 3 技术解决方案

目前，我国资源数据共享尚缺乏健全机制保障，资源单位考虑到数据保密和知识产权保护等诸多因素，无法将本单位全部详情信息发布到互联网上进行完全共享。e-平台通过共性描述规范的制定和应用，已经从一定程度上解决了元数据统一性问题，但平台数据整合范围非常广泛，整合数据量也将达到千万级别。考虑到系统效率等诸多方面因素，采用详情信息分布存储、共性描述信息集中发布的方案是合理有效的解决办法。技术方案见图2。

各资源单位的数据库存放资源的详情信息与基于共性描述规范加工整理的共性描述数据，e-

平台通过数据转换工具、数据文件汇交等方式获取共性描述数据，整合到中心数据库，然后通过平台门户集中发布。用户通过门户检索资源数据，保证了资源的检索效率。详情信息是资源的特性属性信息，存放在资源单位的数据库。当用户查询到一份资源的共性描述信息后，用户还可以发出进一步的查询请求，查询该资源的详情信息。

为保证国家自然资源平台服务的持续与稳定，平台采用双机热备技术，提供应用服务器及数据库服务器的冗余备份服务，保证从服务器角度提供不间断的服务响应。服务器构架如图3所示。

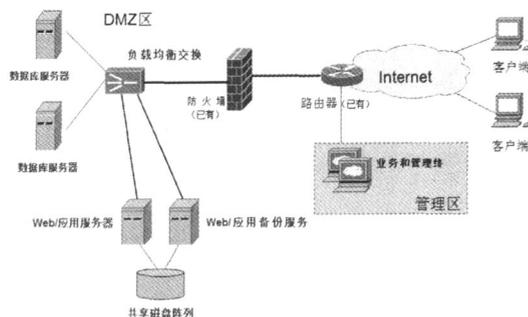


图3 服务器部署图

系统由2台Web应用服务器采用双机热备方式，并配备一台磁盘阵列，用以存储大量多媒体数据；2台数据库服务器放置于交换机DMZ区，以确保数据的安全性；所有服务器均安置于专业硬件防火墙后，确保系统整体安全性。同时，为了保证分散在不同网段的专家在进行数据评价时，能拥有更快的网络响应速度，还在科研网段进行了评价服务器镜像的部署。

### 4 取得成果

#### 4.1 引导信息共享工作走入标准化、规范化轨道

技术标准已经成为国际科技规范化的大趋势，e-平台项目组制定数据制作和上报的技术标准和操作规范，从根本上解决了数据汇交和数据质量的问题，数据整合工作效率明显提高。图像规范的研制工作有效带动资源图像整合与利用工作

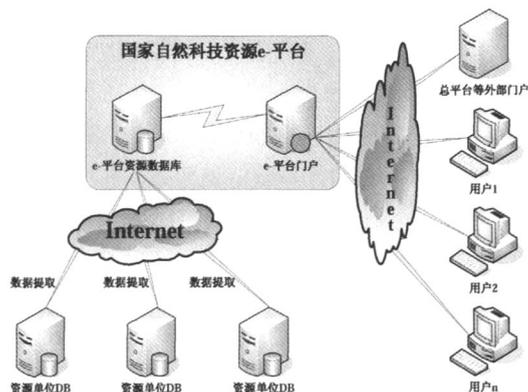


图2 分布存储、集中发布的技术方案

的开展,通过《图像规范》,资源建设单位对图像数据工作的认知程度和重视程度有了明显提高,资源单位多媒体信息采集加工水平显著提高,图像数据质量实现了飞跃。

#### 4.2 通过数据整合工作彻底改变我国自然科技资源数据分散的局面

经过近4年的建设,e-平台整合的数据从无到有,逐年加速递增,直至形成目前百万级的规模。截至2008年3月底,累计共性描述数据量达到4103357份,累计多媒体数据281212份。数据来源于全国20多个省市的100多家平台参建单位。

#### 4.3 双重审查机制切实保障信息共享平台数据内容的质量

为保障全面、准确、完整的高质量资源数据对外提供信息共享服务,e-平台开发了格式审查软件和数据上报软件,利用合理的工作流程和机制来保障数据的质量。组建了一套专家评价系统为主的进行资源数据的质量控制系统。专家评价系统针对通过格式审查的上报数据,聘请各领域知名专家对上报数据内容的科学性进行评价,确保数据的权威性。

#### 4.4 实现八大类自然科技资源信息共享

建成完善、统一的资源信息共享服务门户。通过项目的实施,建成自然科技资源e-平台信息共享管理与应用系统,同时作为自然科技资源领域的门户网站,对外提供相关服务。e-平台先后完成植物种质资源等八大模块资源信息检索系统及其他总计28个可执行子系统研制开发。建成资源保藏地信息展示系统,实现资源多元化揭示。

#### 4.5 与各级门户互联互通,访问量大幅攀升

在共享与互通方面,e-平台处于一个承上启下的重要位置,它既要实现与国家科技基础条件平台总门户的无缝链接,也要通过一套接口,实现与各资源单位(指已经建有独立子平台的单位)子平台每份资源详情的互访。目前,e-平台网站(www.ninr.cn)已经成功开通,并拥有了一定数量的用户群。随着网站影响的不断扩大,网站用户数量及响应次数也在不断扩大,产生了广泛的社会影响。

#### 4.6 虚拟博物馆推动自然科技资源领域科普工作蓬勃发展

在自然科技资源领域,e-平台是最早提出虚拟博物馆的概念并付诸实践的平台之一。4年时间里,e-平台坚持不断推广虚拟博物馆概念,目前有多家单位已建、在建及计划建设各自的“网络化博物馆”,通过热点专题等信息组织形式,将多种网络多媒体技术应用于其中。e-平台在领域内起到了良好的带头示范作用,成功促使行业内网络化科普工作蓬勃开展。

#### 参考文献

- [1] 王东阳. 自然科技资源共享政策法规研究[M]. 北京: 科学出版社, 2005.
- [2] 徐尚义. 自然科技资源共享机制研究[D]. 博士学位论文, 2005, 4.
- [3] 科技部, 财政部, 国家发展和改革委员会, 教育部. “十一五”国家科技基础条件平台建设实施意见[R], 2005
- [4] 曹一化, 刘旭, 等. 自然科技资源共性描述规范[M]. 北京: 中国科学技术出版社, 2006, 1.
- [5] 李新男. 关于国家科技基础条件建设的若干思考[J]. 中国科技资源导刊, 2008(1), 10-11
- [6] 杜占元, 刘旭, 等. 自然科技资源共享平台建设的理论与实践[M]. 北京: 科学出版社, 2007, 12.