

基于服务网格技术的中国科学院网站群设计方法

丛培民 (中国科学院计算机网络信息中心 100864)

摘要:在信息技术进步日新月异的今天,创新是互联网网站生存与发展的灵魂。本文以中国科学院网站建设为背景,提出基于服务网格形态建立中国科学院门户网站群的构想,旨在利用互联网中的虚拟与现实环境构造中国科学院网上宣传矩阵和面向互联网公众的服务网格形成一个内容丰富、风格独特、服务齐全,展示中国科学院风采的门户网站群。

关键词:网站 设计方法

1 引言

中国科学院“十五”信息化发展规划中鲜明地提出中国科学院网站(<http://www.cas.cn>),主旨是通过互联网络这一新生媒体介绍中国科学院,向社会公众提供科技信息服务,形成以信息快捷、内容丰富、风格鲜明为特征的中国科学院网络宣传窗口。在中国科学院“十五”信息化发展规划的带动下,中国科学院网站建设在最近一年来取得了很大成绩,其影响力、受众率、时效性等都有了显著提高。

作为中国科学院网站的运行者,深刻地意识到在信息技术进步日新月异的今天,创新是院网站生存与发展的灵魂。有了体制创新,院网站找到了立足之本,有了技术创新,院网站取得了今天的进步,只有坚持与时俱进、不断创新,院网站才能可持续发展。为此,以构筑连接科技界、沟通海内外、分享科研成果、共创学术繁荣、展示中国科学院风采的网络媒体阵营作为院网站创新战略目标,提出基于服务网格形态建立中国科学院门户网站群的构想。

2 实施知识创新工程给院网站的发展注入了活力

中国科学院是国内最早接入互联网的部门之一,其网站建设始于 1995 年。随着“百所联网”工程的实施,院属各研究单位也相继建立了各自的互联网站点,基本上构成了“网上科学院”的雏形。进入 21 世纪,在院知识创新工程全面推进、信息化建设快速发展的带动下,中国科学院网站建设列入了院“十五”信息化发展规划,在对其运作环境和管理机制进行了重要改革和调整,于 2002 年 11 月推出了全新的中国科学院网站。

新的院网站在整体结构、技术路线、版面风格、信息容量、服务功能等方面都进行了重新设计。功能设计上强化了信息利用效益的体现,内容设计上突出院的重点工作。信息资源表现由 19 大类 100 余个栏目构成,信息资源建设加大了政务史料的发掘。目前,院网站资源数据库累计的各类信

息内容与过去相比成倍增加。多媒体信息首次进入数据库,其信息量正在迅速增长。院网站上提供给访问者的信息利用途径越来越灵活,服务性栏目功能越来越完善。信息管理与播发除了采用当前网站开发的新技术外,在信息组织和编辑检索手段上也有了创新。新的院网站发布之后,经过对信息内容和表现形式的不断调整,在信息的时效性、资源的丰富性、技术的先进性方面得到了较大改善,跨入了国内政府部门网站建设的先进行列。

在院网站建设的带动下,院属各单位也相继更新改造了各自的外部网站。通过院网站的导航,公众可以在网上全面了解中国科学院院情及院属各单位风貌、分享丰富的科技信息资源、得到充分的科学数据服务。

目前的中国科学院网站逻辑结构如图 1 所示。

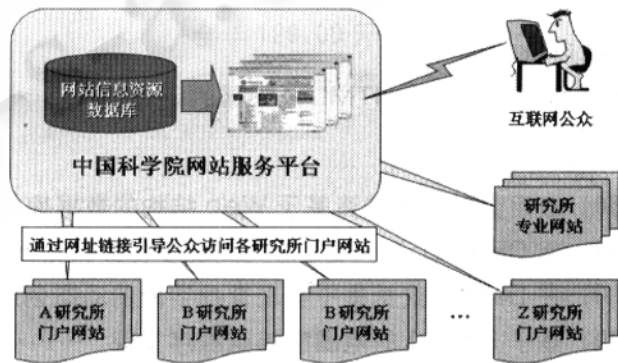


图 1 中国科学院网站逻辑结构

在中国科学院网站的物理环境中构造了信息资源管理的数据仓库和信息资源展现的系统服务平台。通过数据仓库,可以指示公众即时获得与院情所况相关的各类信息资源,通过系统服务平台,可以引导公众进入院属单位各自的网站内得到更加具体的信息服务。同时,还可以访问到与科

研活动相关的各种专业网站提供的信息资源。

通过以上对院网站逻辑结构的描述可以看出,以中国科学院网站为首的院网站群框架已经形成,全院的互联网网络媒体站点建设已经有了相当的规模,为进一步提升院网站群的建设水平创造了良好条件。

3 公众的期望对院网站建设提出了更高要求

虽然院网站建设经过各方面的努力发生了显著变化,但从网站访问者对冠以“中国科学院网站”的信息需求角度看,在信息内容的组织模式、公众服务的完善程度、信息资源的专业化程度等方面均需要进一步的强化。

3.1 信息内容的组织模式

目前,中国科学院网站信息资源数据库中已经存储了大量信息。这些信息有的来自院属各单位的直接上报、有的来自网站编辑摘编的相关媒体报道、有的来自与同业网站的信息交换、还有的来自采用信息抓取技术手段搜集到的各类网上信息。这些信息通过各种信息组织模式从不同角度提供给访问者。尽管如此,仍然不能充分满足访问者对信息关注深度的要求。

为了拓展信息空间,网站加大信息展现力度的做法是采用网址链接技术,将关注更加深入信息的访问者带到相关网站上来满足资源获取需求。为了表现中国科学院的整体形象,充实院网站信息资源的方法也是将分散在院属各单位门户网站和专业网站通过网址链接起来,从信息管理角度构成了分布式信息资源管理环境,从组织模式角度构成了院网站群概念。

这种方法虽然从形式上扩展了访问者的视野,但是由于被链接网站中表现的内容时效、资源储备、技术含量等参差不齐,无法充分表现中国科学院网站群体的良好形象。从对院网站链接院属各单位相关网址进行的分析结果看,信息资源鲜活、网站形象完好、技术含量高的网站不足四分之一;一些网站信息内容更新缓慢;还有个别网址有名无实。这与中国科学院网站所代表的国家科技队伍的形象尚有差距。

3.2 信息服务的完善程度

网站技术的发展已经不再局限于单纯的信息发布,一个明显的趋势是网站与访问者的互动性逐渐增强,电子商务、电子政务、网上办公等互动服务形态应运而生。院网站在这方面也做了大量工作。包括通过网站进行各种业务信息的采集、学术会议的组织、大型仪器设备的共享共用、人才的求职招聘、热点话题的网上调查等等。这些服务有些是由院网站直接提供,有些则是由链接网站提供。由于服务形式相对分散,服务形态各异,网站访问者对其服务内容的完善程度

还有更多的期望。

3.3 信息资源的专业化程度

作为中国科学院网站,公众期望得到的是专业化的科技信息资源,包括中国科学院科研工作的各项进展以及发展态势,知识发现的最新信息乃至国内外的科技动态;期望得到的是专业化的、全面的科技咨询,以及知识转移和知识传播服务等。目前院网站在这方面还有很大的发展空间。

4 建设中国科学院门户网站群的构想

解决院网站发展过程中出现的问题,真正形成以院网站为主体的网络媒体宣传阵营,除了加强和改进管理和运行机制外,采用新的技术路线加大对院网站整体形态的改造是十分必要的。在基于网格形态的信息服务平台上建立中国科学院门户网站群,是实现院网站管理和技术创新的重要手段。

4.1 网站群框架

中国科学院依托院所两级法人治理结构,其网站建设也形成了院网站和院属单位网站两个层次。在院层面有为社会公众服务的院网站和为院属各单位服务的院办公网站,在院属单位层面有各自的公共网站和科研专业网站,从而构成了院门户网站群的基本框架。无论访问者关注科研动态还是了解院情所况,都可以得到与之想对应的信息资源。这些信息资源,有的是来自院网站,有的是来自院属单位网站群节点。总之,新构建的院门户网站群包括院、研究所及其他院属机构和专业网站,其节点数量是十分充分的,所涵盖的科技信息资源是非常广泛的。

4.2 门户网站群逻辑结构

提出基于网格形态建立院门户网站群,是量变到质变的过程。通过网格形态形成无缝的、集成的信息处理环境。

在信息服务平台上构造服务网格的过程,实际上是网站资源网格化的过程。中国科学院网站网格形态的逻辑结构如图2所示。实现网站资源的网格化,包括虚拟网格和现实网格的结合。虚拟网格实际上是与院门户网站联合构造的,一方面可以完全覆盖院属各单位,消除院属单位网站的盲点;另一方面便于集中统一管理。当虚拟网格单元提供的服务不充分时,通过院属单位现实的物理网格单元进行补充和扩展。此外,还有部分与院门户网站主题思想相关的站点,仍然通过网格单元网址链接的方式表达。

在院门户网站网格逻辑结构图示中,院门户网站的物理环境是网格逻辑的集中体现。将集中管理的网站资源以内部逻辑关系进行虚拟划分,构造以院属各单位为单元的虚拟门户网站节点。实现信息资源融合、处理能力集中、风格表现统一。院属单位都将在院网站的虚拟网格环境中形成各

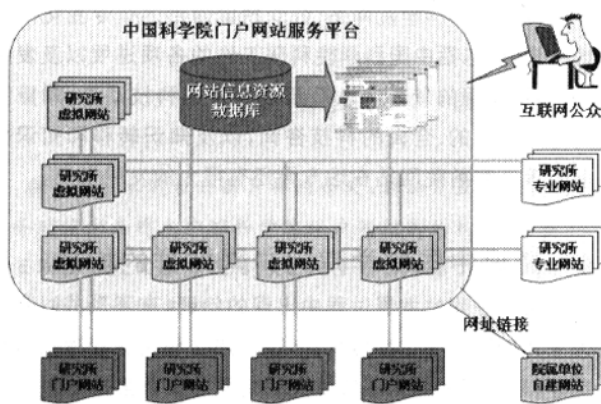


图 2 中国科学院门户网站网络逻辑结构

自的节点。其信息内容的分类按照共性展现需求进行多维度设置,保证信息网格的一致性。对信息资源管理进行统一协调,保障各类资源的归属与分类相对应,保证网格存储的整体性。提供公众服务的功能在节点中均衡分布,保证网格计算的均衡性。

院门户网站群的物理环境是网格逻辑的现实体现。个性化的表达要求、专业化的服务功能均可在各自节点的物理环境中展现,充分发挥网格逻辑的优势。

4.3 网格形态功能组成

通过策划院门户网站群的网格框架,其目的是要在院门户网站群的基础上,整合信息处理能力形成院网站的计算网格——构造一个覆盖全院的分布式网站服务器群;协调信息存储资源形成院网站的存储网格——在分布式网站服务器群中统一调度信息存储资源;分散信息采集渠道形成院网站的管理网格——在各网格节点网站上建立专业化的编辑队伍;规范信息指标体系形成院网站的信息共享网格——对非结构化的信息资源按照统一的结构化标准进行规范使其可共享;构造资源利用环境形成院网站的服务网格——将专业化的服务建立在充分的信息挖掘基础上。

5 实施策略

5.1 建立统一的域名体系

中国科学院网站已经注册了 cas.cn 域名,在这一域名体系下首先设置了 www.cas.cn 作为院门户网站群服务平台的统一入口。建设院门户网站群的服务网格框架,为院属各单位用其英文缩写在 cas.cn 域中建立域名。例如:中国科学院自动化研究所的域名为:ia.cas.cn,中国科学院植物研究所的域名为:ib.cas.cn 等。这些域名统一由 cas.cn 解析服务器管理,为院门户网站提供域名解析服务。其

他节点将按照其原有域名设置。

5.2 建立标准的节点域名

cas.cn 下的节点构成院网站虚拟网格,而作为院门户网站群的节点统一命名为 WWW。各节点域名按照应用性质统一分配和命名,采用这一方法对今后院门户网站群成为中国科学院资源规划(ARP)项目的公众服务窗口是十分有益的。院门户网站群节点域名如下表所示:

单位名称	域名	门户网站节点域名
中国科学院自动化研究所	ia.cas.cn	www.ia.cas.cn
中国科学院植物研究所	ib.cas.cn	www.ib.cas.cn
.....		

5.3 建立规范的网站节点信息资源分类

院门户网站群中虚拟网格节点的信息资源以表现本单位、本领域相关信息为主,信息内容统一管理;而现实网格节点的信息资源应充实,特色突出,成为院门户网站在此技术领域的重要补充。由院属各单位网站构成的现实网格节点应表现下列内容:

机构简介:包括本单位的历史沿革、现状、发展远景、治理结构及综合实力的概要介绍;

新闻报道:包括本单位科研活动和政务活动的新闻信息;

人才培养:包括科研队伍建设、人才培养、研究生教育等;

合作交流:包括国际合作、院地合作、科技副职等;

招聘:包括硕博招生、人才引进、人才计划等;

联系方式:地址、电话、电子信箱。

研究机构网站还应包括以下内容:

研究领域:本单位的科研方向、科研项目、重大课题等可公开的信息;

科技成果:高科技产品展示、专利技术、成功案例、高技术产业化的开发应用等;

科研动态:以研究室为单元的科研方向介绍及以首席科学家、课题组长、研究员为线索的领域课题介绍等。

学术活动:各类学术会议,讲座,学术来访、出访,学术报告等;

学术团体:相关学会、国家级或国际团体组织的介绍等;

科技新闻:本领域的相关科技新闻等;

主要专家:本单位的专家学者、杰出人才、学术带头人、两院院士等。

5.4 建立信息类别分析系统

当一条信息被采集进入到信息资源数据库后,除编辑人

员为其指定栏目隶属关系外,系统需要提供信息类别的自动分析功能。以一条报道某研究所在生物学领域作出一重大成果的新闻信息为例,这条信息的分类由编辑人员指定属于“科研”栏目“生物领域”分类,同时这条信息还被指定放入“今日要闻”栏目。然后,信息类别分析系统对此信息进行分析判断,将此条信息归类到“一线报道”栏目,同时划分到对应研究所虚拟网站中的新闻类栏目和科研进展类栏目中。这样,如图 3 所示,一条信息可以通过多重渠道在院门户网站群中展现。

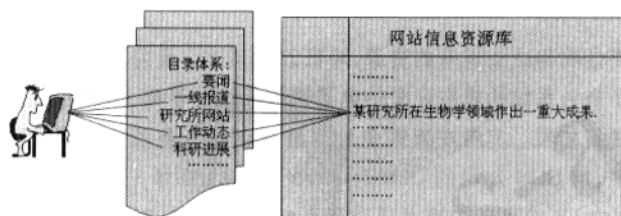


图 3

5.5 建立组件库

在院门户网站群建设中构造各种网站应用功能的组件,形成一个网站功能组件库,可以方便地进行各种组件的选择、组装和集成。有了组件库之后,应用开发人员就可以利用现成的软件组件装配成适用于不同领域、功能各异的应用系统,提高网站生产效率和质量。

5.6 建立模板库

对院门户网站群进行统一的 VI 设计。按照不同的研究领域,如基础科学、生物科学、资源环境科学、高技术领域等建立相应的网站版面模板;按照不同的单位性质,如管理部门、技术开发部门、控股公司等建立相应的网站版面模板,形成模板库供用户选择。以此统一院门户网站群的风格形象。

5.7 开放虚拟网站的维护权限

院门户网站群中虚拟网格节点上表现的信息内容,除经过资源调度和整合进行统一资源导航表现外,将信息采集与维护的权限授予网格单元的管理者,充分发挥各方面的优势力量,形成院门户网站群的运作体系,提高信息发布的时效性和灵活性。

实现在基于网格形态的信息服务平台上建立中国科学院院门户网站群的构想,将使中国科学院网站建设提高到一个新的水平,对于加强和改进院网站管理运作机制、整合和利用信息资源、强化服务功能具有重要作用。目前采用该设计方法的院门户网站群正在建设之中。

参考文献

- 1 段沫毅,“知识服务网格:新世纪的信息服务基础设施”专题报告。
- 2 中国科学院办公厅:中国科学院主页建设项目任务书。
- 3 杜义华,中国科学院属单位网站调查报告。



《电脑开发与应用》月刊

明年扩版

欢迎订阅

本刊集信息、知识、趣味、可读性于一体,以计算机实用技术见长,博采、精选国内外电脑研究、开发与应用的精华。具有军事、兵器控制色彩,军用计算机的开发,将用相当的篇幅刊登企业信息化、信息化建设、信息传输处理与管理方面的信息及文章,关注 IT 产业。尤其是刊登 Internet、WWW、网友、软硬件二次开发、电脑测控、CORBA、开放式与微内核技术、柔性与敏捷制造等方面的内容,跟踪报道世界最新技术。

本刊为大 16 开(A4)、64 页、定价 6 元/册、全年 72 元之电脑月刊。

全国各地邮局均可订阅,邮发代号:22-96 国外代号:M4257

05 年新扩版,页码增加,信息量增大

欢迎订阅,欢迎赐稿,诚征广告!

联系地址:太原市 193 信箱《电脑开发与应用》编辑部

邮 编:030006 电 话:(0351)7023553-2182

传 真:(0351)7022975

E-mail:DNKF@chinajournal.net.cn