

领域科研信息环境构建研究*

□ 赵瑞雪 寇远涛 鲜国建 朱亮 / 中国农业科学院农业信息研究所 北京 100081

摘要: 新信息环境下科研人员对于集成服务、知识管理、知识共享以及交流协作等需求日益迫切, 面向科研机构、学科团队或者学科领域构建集知识、技术、工具、服务等一体化、集成化的数字知识环境是现代科学研究的需要, 也是图书情报机构深化服务、支持科研创新的有效途径。文章在调研国内外科研信息环境研究现状的基础上, 探讨了领域科研信息环境的建设思路、技术方案和应用实践等情况。

关键词: 领域知识服务, 科研信息环境, 知识组织, 本体

DOI: 10.3772/j.issn.1673—2286.2012.12.004

1 引言

信息技术的进步和知识经济的发展为加快科研环境的信息化、支持科研创新提供了新的工具与方法, 同时当代科学研究也呈现出许多新特征, 表现为: 科研资源的数字化、科研交流的网络化、科研工作的协同化以及科研知识的共享化等等。这些变化不仅影响到科学研究过程的方方面面, 而且改变着科研工作者的信息行为与信息需求, 如: 日益明显的集成服务需求、逐渐增强的科研产出保存与共享意识、面向解决问题的知识服务需求、科研交流与协作需求^[1-3]等。这些科研方式和用户需求的变化给信息服务带来了压力和挑战: ①传统普适的服务已经难以支撑科研精细化程度的不断增强; ②领域知识的分散、无序使得日益增长的专业化、知识化服务需求难以满足; ③面向特定专业领域的知识交流环境缺乏, 极大地限制了科研协作与共享。因此, 如何面向不同的用户视图, 构建紧密耦合其科研过程、集成各类资源和服务的个性化知识环境成为当前信息服务领域面临的关键问题。本文以面向学科领域、研究机构、学科团队提供知识服务为目

标, 探讨了领域科研信息环境的建设目标和建设方案, 对中国农科院院所科研信息环境实践进行了讨论。

2 国内外相关研究进展

目前, 科研信息环境的概念还不统一, 但是国内外都已启动了相关的研究项目。如, 美国在20世纪90年代就开始研究建设面向学科的协作研究体(Collaboratory), 重点研究分布式计算与数据资源的获取工具、学科化的分析工具、共享的工作空间和合作交流空间等, 明确提出要建设为科研与教育提供新的知识环境的整合基础设施。英国启动了虚拟科研环境(VRE, Virtual Research Environment)建设项目。VRE旨在集成科研团队涉及的各方面科研信息, 并发现领域内外支持科研活动的各种需求。澳大利亚在2005年成立了e-Research协调委员会, 开展相关的研究实践^[4,5]。

在国内, 中国科学院基于康奈尔大学的Vitro系统构建了专业领域知识环境SKE, 向领域内外的科研人员提供知识导航与研究合作支持^[6]; 中国农科院国家农业图书馆面向专业研究所进行资源组织和服务探

* 本文系国家“十二五”科技支撑计划项目课题“基于STKOS的知识服务应用示范”(编号: 2011BAH10B06)的研究成果之一。

索,构建了研究所科研信息环境^[7];中国科学院“地学e-Science应用示范研究——东北亚联合科学考察与合作研究平台构建”项目分析了地学研究对信息化科学环境的需求,提出了地学信息化科研环境的概念和技术架构,并构建了东北亚联合科学考察与合作研究示范系统^[8]。另外,华中师范大学、华东师范大学、北京邮电大学开始与开源虚拟学习软件Sakai合作^[9],不过国内的Sakai研究还主要集中在课程管理和兴趣小组间的知识共享。

总之,各国对于科研信息环境的研究和实践已经取得了一定进展,出现了一些可供借鉴的技术或工具,如基于本体的VIVO系统^[10,11]、基于SOA的体系架构^[12]、基于Sakai的虚拟科研环境^[13]以及哈佛大学Harvard Catalyst、哥伦比亚大学的Sciologer^[14]等,这些技术、工具、系统等为本文的研究提供了良好的基础。

3 领域科研信息环境的建设方案

领域科研信息环境的建设目标是为一线科研人员提供一个高度集成的个性化知识环境,支持科研人员

的知识生产、知识管理、知识交流和知识利用等科研活动。因此,本文把领域科研信息环境的建设要素概括为:(1)需求和应用场景;(2)体系架构及用户接口;(3)资源组织与管理;(4)通用服务实现;(5)协议/标准等。下面从建设框架、服务功能、资源组织和平台建设四个方面论述领域科研信息环境的建设方案。

3.1 领域科研信息环境的建设框架

架构是服务实现的技术框架,支持内容和服务的综合集成,如基于SOA的体系架构。标准化和互操作性是架构设计的关键因素。本文以SOA框架为基础,依据分层服务思想,建立了领域科研信息环境的建设框架,如图1所示,包括支撑层、资源层、平台层、服务层和用户层4个层次。

(1)用户层:用户层位于整个架构的顶层,包括科研用户、管理用户和第三方用户。科研用户是主要的服务对象,可以是个人、团队或者机构。管理用户指领域科研信息环境的建设者和服务人员。第三方用户指通过数据接口使用领域科研信息环境的系统或服务。

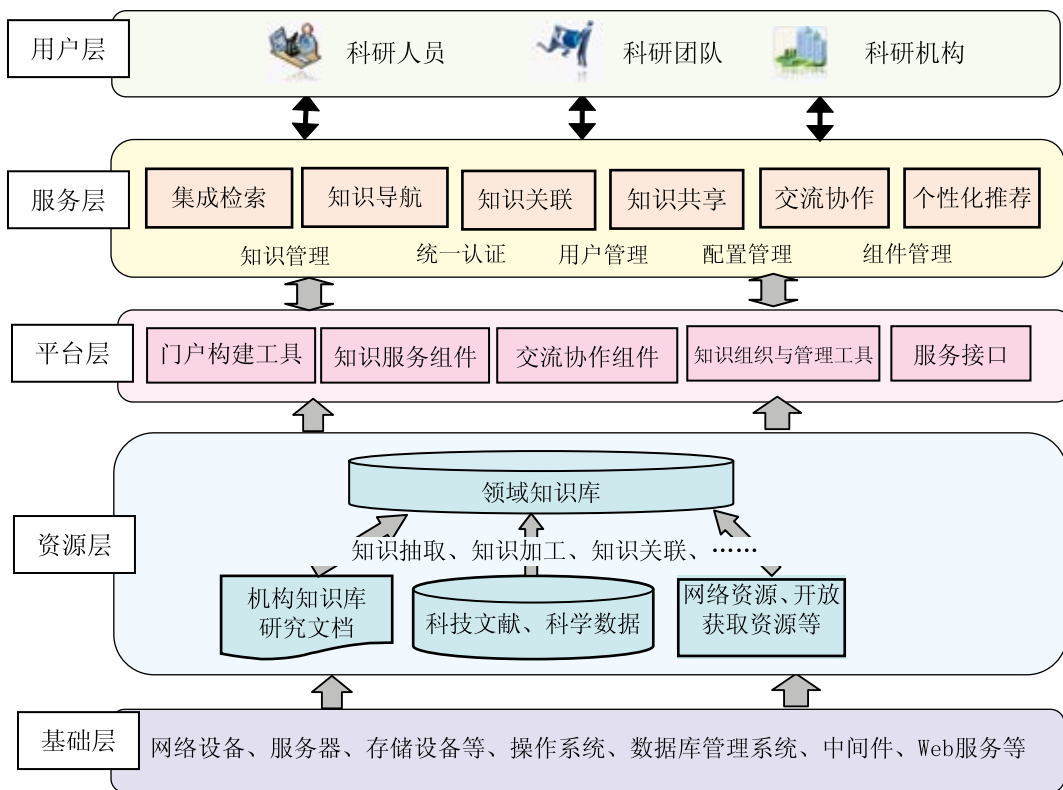


图1 院所科研信息环境建设框架

(2) 服务层: 服务层定义了领域科研信息环境提供的服务, 并通过一个门户框架向用户提供各种应用服务, 如资源检索、知识导航、在线交流等。这些服务的实现需要依托平台层提供的系统或工具以及资源层的资源支撑。

(3) 平台层: 平台层主要解决平台建设的问题, 以系统、组件、工具等方式实现领域科研信息环境的各种服务, 提供资源内容建设和知识组织管理工具, 并提供与现有服务系统和开源的第三方技术工具的集成机制。

(4) 资源层: 资源层是通过资源整合和知识化组织形成的领域数字知识体系, 包括科技文献、科学数据、机构知识仓储、互联网开放获取资源, 以及通过知识构建形成的领域知识库等。

(5) 基础层: 基础层负责软硬件环境搭建, 包括网络设备、服务器、存储设备等物理硬件以及操作系统、数据库管理系统、中间件等系统软件, 是整个体系架构的基础。

3.2 领域科研信息环境的核心服务

据调查, 目前专业信息服务机构的生存危机主要表现为其提供的服务难以满足用户的需求, 无法维系用户对其的依存或者其资源和服务逐渐被其他信息服务机构所淹没。因此, 一项新的服务, 必须充分考虑其对应用环境的适应性和用户需求的应答度, 并且需要提供一种动态更新和扩展机制以适应不断发展的科研需求和社会信息环境的变化。图2给出了领域科研信息环境的主要服务, 包括数字资源服务、领域知识服务、科研交流与协作服务、机构仓储和开放获取服务、系统服务以及数据交互服务6大类。

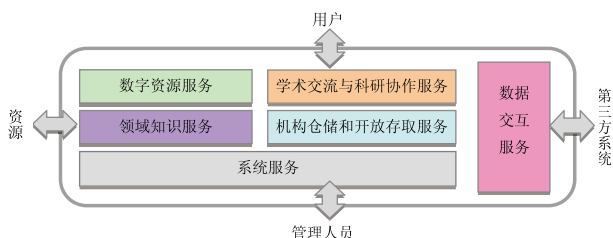


图2 领域科研信息环境的服务构成

统服务以及数据交互服务6大类。

(1) 数字资源服务。数字资源服务是以内容管理、资源检索为核心的服务, 提供用户个人文献管理、分布

式领域资源、服务和系统的整合以及资源检索浏览服务, 支持用户个性化信息的订阅与推送。

(2) 知识化服务。知识化服务是指以知识管理、知识共享、知识呈现和知识利用为核心的服务, 主要包括知识导航、知识关联、知识地图等服务形式。

(3) 学术交流与科研协作服务。为科研人员提供在线交流、学习以及协同科研环境, 支持科研人员隐形知识共享, 支撑和管理多人协作完成科研项目, 实现项目内的各项任务、文档、课题的协同编辑、共享和发布, 支持在线讨论、交流和学习。

(4) 机构仓储和开放存取服务。通过开放存取机制, 存取领域开放的学术资源, 支持研究机构、科研团队、个人等学术成果的收集、保存、发布和共享。

(5) 系统服务。指系统管理、用户管理、用户认证等服务, 满足系统的运行需求。对系统用户、角色、权限等进行分配与管理; 对系统访问日志等维护管理。支持数据统计分析。

(6) 数据交互服务。数据交互服务是以服务扩展、数据融合为目标的服务方式, 目的是提升领域知识的可见性、可用性, 支持外部数据调用, 通过数据接口实现为外部检索、内容展示以及数据关联。

3.3 领域科研信息环境的资源建设

领域科研信息环境的资源建设包括两个层面: 一个是领域综合科技信息资源的搜集、加工和保存, 二是对领域知识的关系计算、揭示, 实现知识关联。资源建设是科研信息环境建设与服务的基础, 必须明确科研信息环境各要素资源的分布特点、获取途径以及整合策略。资源的质量以及有序化、知识化组织水平则是保证服务深度和水平的关键。

(1) 知识内容与来源

领域科研信息环境的知识内容由数字资源和领域知识组织体系两部分组成: ①数字资源是通过购买、自建和整合等方式建立的领域综合科技信息资源。其中, 购买资源主要是商业数据库资源, 自建资源主要包括机构知识仓储和各种专题数据库, 整合资源则包括互联网开放获取资源以及其他可共享资源。②知识组织体系是解决知识资源的组织、揭示和利用的重要手段, 是实现知识服务的支撑。本文在国家“十二五”科技支撑计划项目组的支持下, 构建了面向专业领域的知识组织体系框架, 以期实现对领域知识及其关系的揭示。

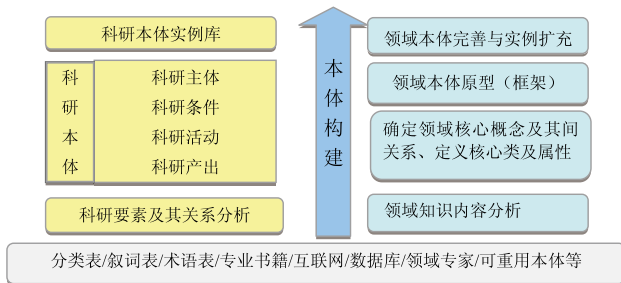


图3 基于本体的领域知识组织体系框架

(2) 领域知识组织体系建设框架

知识组织的概念是在基于文献单元的分类法基础上发展起来的，目前正朝着基于概念单元或知识单元的本体 (Ontology) 方向发展，以适应计算机海量信息的处理和网络信息资源的获取、组织和利用要求。本体描述了某一知识领域的概念（术语）及概念间的关系，在知识工程、信息检索和知识管理等领域具有广泛的应用。图3给出了基于本体的领域科研信息环境知识组织体系框架，该框架是在项目总体组提出的知识组织体系总框架的基础上，结合领域知识服务应用需求而建立的，由知识源、科研本体和领域本体构成。

知识源是本体库构建的基础，可以是已有的叙词表、术语表、专业书籍等知识素材，也可以从互联网、数据库中抽取，或者查找可重用本体、寻求领域专家支持等。科研本体揭示了科研活动主体及各科研对象之间的关系，可以实现基于科研本体的资源导航和关联检索，是不同科研人员协作研究、相互交流的基础。本文在调研了VIVO科研本体的基础上，采用原型法初步建立了领域科研信息环境科研本体的核心概念、概念属性以及概念间的关系，基本涵盖了科研过程中主要的资源类型。其中的核心概念包括科研主体、科研条件、科研活动、科研产出以及一些辅助类。

领域本体是对领域概念及概念之间的关系的结构化描述，用于领域知识资源的组织和共享。由于应用领域的不同，领域本体建设更具复杂性和多变性，而领域中包含的大量概念和对象也使得领域本体的建设不可能迅速而全面完成。因此，建议采用软件工程中原型法思想先建立领域本体原型（或框架），然后再进行完善和本体实例扩充。笔者单位基于已有的农业科学叙词表，采用OWL语言开展了叙词表向本体的转换和映射工作^[15]，并在此基础上根据具体的专业领域进行了核心概念析取和部分扩充，建立了部分专业领域初级本体，但是仍然需要进一步的补充和语义关系调整。

3.4 领域科研信息环境技术平台的构建及关键技术

技术平台是支撑领域科研信息环境内容建设、知识组织、并开展知识服务应用的重要依托。本文对国内外同类系统、工具以及部分相关的开源软件进行了调研测试，并重点分析了康奈尔大学的VIVO系统。VIVO集成了功能强大的本体管理工具，通过构建科研专家 (People)、学术活动 (Education and Training)、研究工具 (Research Tools)、仪器设备 (Facilities)、学术机构等本体，采用Jena推理系统实现科研对象的关联导航与检索^[10,11]。但是，VIVO目前还是一个以“人”为核心的知识关联系统，在支持汉化、领域知识资源的组织以及服务扩展等方面还有待进一步提升。本文吸收了VIVO本体驱动机制，同时考虑到系统的灵活性、扩展性和开放性需求，构建了本体驱动、组件化、开放式的技术平台架构，如图4所示，并采用以自主开发为主、开源软件改造相结合的技术路线开发了一套基于开放的标准协议、可扩展和复用的领域科研信息环境构建平台 (SIE)，支持科研机构、学科组、学科领域等快速部署个性化科研信息环境应用系统，同时支持与现有的平台或工具以及第三方工具及服务的嵌入和集成。SIE共由7部分组成，各部分可独立应用，也可集成应用：

(1) 集成服务门户构建工具。目前，一些开源的机构门户软件，如uPortal等已经广泛应用。考虑到已有的开发基础，本文采用Java开发语言和开源全文检索引擎Lucene实现了领域科研信息环境集成服务门户构建，提供资源检索和系统管理服务，支持面向不同的用户群进行个性化门户的定制和用户特色资源的建设与集成应用。

(2) 分布异构资源整合发现工具。目前较成熟的资源整合发现工具有ProQuest公司的基于海量元数据仓储的新一代学术资源发现与获取系统Summon、以色列艾利贝斯公司的统一资源发现与获取系统Primo、EBSCO公司的EDS发现检索工具。但是，价格都很昂贵。国内也出现了一些自主研发的系统，大多不很稳定，维护工作繁重。笔者单位也经历了自主研发—合作开发—产品购买的尝试过程，现在主要基于购买的Primo产品提供服务。

(3) 知识服务功能组件。知识服务的实现程度取决于领域知识库的丰富程度和知识服务工具的强弱。考虑到与领域知识组织工具的衔接，本文采用Java为

主要开发语言实现领域知识导航、知识关联、知识地图等知识服务功能组件。

(4) 科研协作支持组件。支持协同工作环境的创建和知识资源的充分共享，全部基于开源软件开发。包括：基于微软的开源dsoframer控件和FengOffice系统实现科研协同、基于HDWiki构建领域知识百科、基于ThinkSNS构建学者网络等。各功能相对独立，采用组件式嵌入或链接接入，基于Ucenter技术实现不同开源产品之间的信息同步和账户集成。

(5) 领域知识组织与管理工具。采用Java开发语言自主研发，用以领域科研信息环境的内容建设，支持科研本体和领域本体中的类、数据属性、对象属性等的管理维护和本体实例的构建，以及领域综合科技信息资源的采集、加工和保存。提供信息自动采集工具。

(6) 机构仓储和开放获取软件。基于开源Dspace软件进行改造和扩展，支持各类组织进行数字资源的收集、保存、发布和开放获取。

(7) 数据服务接口。提供了TCP接口、Http接口、WebService接口以及基于Java和.Net两种语言的二次开发组件，实现领域科研信息环境的可扩展、可延伸性，支持对领域科研信息环境资源和服务的第三方系统调用，扩展领域科研信息环境的应用服务模式。

4 应用实践

为了进一步完善领域科研信息环境的建设方案，本

文将已经形成的研究成果应用到了中国农科院院所科研信息环境建设项目，该项目的目标是面向中国农科院的专业研究所、重点创新团队和承担国家重大科技专项的课题组建立个性化知识服务平台。目前，利用领域科研信息环境构建平台SIE，已经部署了中国农科院作物科学研究所、北京畜牧兽研究所、生物所、蔬菜所、资源区划所、水稻所等10个研究所的个性化科研信息环境的示范系统，实现了“一站式”资源检索、知识导航与知识关联、知识共享、科研交流与协作、个人学术空间以及个性化推荐等服务功能。图5是作物科学研究所科研信息环境首页，除提供包括科技文献、科学数据以及开放获取资源的分类导航检索和集成检索服务外，还支持作物学科领域的专业期刊、会议、网站、机构、人员、项目、成果以及科研设施等领域知识的维护、扩充和发布，初步实现了基于本体的领域知识呈现，支持领域科研动态信息的自动采集和发布，支持研究所建立本所机构知识仓储，支持研究团队协同工作环境的搭建。图6是作物种质资源保护与创新学科组的个性化信息环境首页，展示了包括学科组介绍、科研团队、学科动态、项目成果、研究进展、共享资源以及专业信息等。对于学科组团队成员，登录之后会进入团队协同工作平台，支持团队成员之间资源共享和协同科研。

目前，国家农业图书馆学科化服务团队正依托该平台进行学科资源整理和学科化服务的推进，相信随着试点范围的扩大和服务工作的开展，平台会不断完善，内容会不断丰富，科研人员也会渐渐喜欢并习惯这样的科研助手。

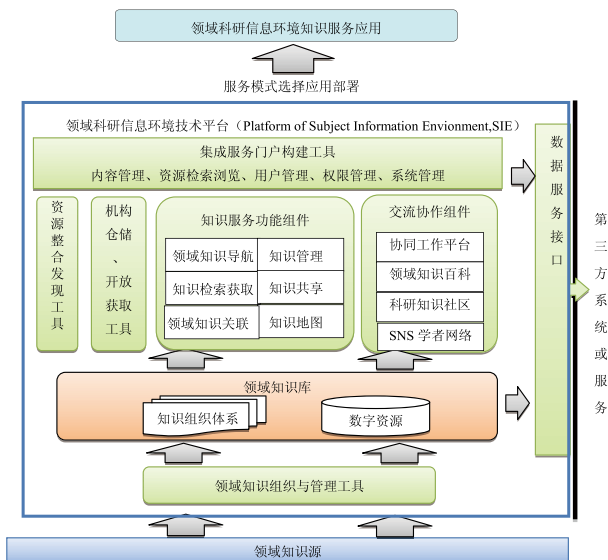


图4 领域科研信息环境平台逻辑架构



图5 作物科学研究所科研信息环境



图6 作物种质资源保护与创新学科组信息环境

5 结语

领域科研信息环境作为一种个性化、学科化服务模式，其建设尚处于探索之中。有很多问题需要在今后的工作中不断完善：（1）如何准确把握用户需求并使之真正成为用户科研创新的常态支撑平台；（2）如何保障领域知识内容的持续更新和知识组织体系的进化；（3）如何不断跟踪互联网服务技术的发展并使相关服务能够集成融合；（4）如何进一步协调好技术、应用和服务的高效耦合机制；（5）如何实现跨领域的知识交换、知识应用与知识服务等。总之，我们将继续关注该领域的发展动态，不断完善其功能和服务。

参考文献

- [1] 陈成鑫.E-science环境下用户信息需求与信息服务研究[J].情报科学,2009,27(1):108-112.
- [2] 杨颖,宁方.基于数字环境的专业图书馆建设及信息服务[J].情报杂志,2011(6):184-156.
- [3] 肖仙桃,王丹丹.用户信息环境、信息行为及信息需求发展趋势[J].图书馆理论与实践,2010(1):40-42.
- [4] LAWSON I, BUTSON R. eResearch at Otago- A report of the University of Otago eResearch Steering Group[R], 2007.
- [5] BORDAA, CARELESS J, et al. Report of the Working Group on Virtual Research Communities for the OST e-Infrastructure Steering Group[M]. London, UK: Office of Science and Technology, 2006:33-36.
- [6] 宋文,张士男.专业领域知识环境建设的理念与实践[C]//数字图书馆高层论坛2010年年会论文集.北京:国家科技图书文献中心,2010:19-23.
- [7] 赵瑞雪.数字环境下面向专业研究所的知识服务探索[J].情报杂志,2011,30(6):189-193.
- [8] 诸云强,孙九林,宋佳,等.地学信息化科研环境研究与应用示范[J].科研信息化技术与应用,2009,2(4):42-51.
- [9] 胡敏,刘清堂,李浩,等.开源软件Sakai的开发及应用研究[J].中国教育信息化,2010(21):70-73.
- [10] KRAFFT D B, CAPPADONA N A, CARUSO B, et al. VIVO: Enabling National Networking of Scientists[C]// Web Science Conf. 2010, April 26-27, 2010, Raleigh, NC, USA.
- [11] DEVARE M, CORSON-RIKERT J, CARUSO B, et al. VIVO: Connecting People, Creating a Virtual Life Sciences Community[J]. D-Lib Magazine, 2007(13):7-8.
- [12] ALLAN R, ALLDEN A, BOYD D, et al. Roadmap for a UK Virtual Research Environment[R]. Report of the JCSR VRE Working Group, 2004.
- [13] YANG X, ALLAN R. Web-Based Virtual Research Environments (VRE): Support Collaboration in e-Science[C]// Proceedings of the 2006 IEEE/WIC/ACM international conference on Web Intelligence and Intelligent Agent Technology. Washington, DC, USA: IEEE Computer Society, 2006:184-187.
- [14] [EB/OL].[2012-01-23].http://swl.slis.indiana.edu/repository/systemdoc_files/comparison_with_VIVO%2012-01.pdf.
- [15] 鲜国建.农业科学叙词表向农业本体转化系统的研究与实现[D].中国农业科学院,2008.

作者简介

赵瑞雪 (1968-), 博士, 研究员, 博士生导师, 研究方向: 信息管理与信息系统、信息资源管理、知识组织及数字图书馆。E-mail: zhaorx@mail.caas.net.cn

Construction of Subject Research Information Environment

Zhao Ruixue, Kou Yuantao, Xian Guojian, Zhu liang / Agricultural Information Institute of CAAS, Beijing China, 100081

Abstract: Scientific researchers' increasingly urgent demand for service integration, knowledge management, knowledge sharing, coordination and communication under the new situation, makes it an effective service way to build an incorporate knowledge environment that integrates knowledge, skills, tools, and service of certain subjects. This paper investigated the status of research information environment at home and abroad and put forward some building solutions and practice situation.

Keywords: Subject knowledge service, Research information environment, Knowledge organization, Ontology

(收稿日期: 2012-08-01)