

中国陆地生态系统碳收支集成研究的 e-Science 环境建设进展

于贵瑞¹ 何洪林¹ 黎建辉² 周园春²

¹中科院 地理科学与资源研究所 ²中科院 计算机网络信息中心

【摘要】以中国通量观测研究网络（ChinaFLUX）为基础平台，开展生态系统碳循环集成研究不仅是全球变化科学研究的重要内容，也是我国应对气候变化的重大科技任务。生态系统碳循环集成研究需要多台站联网观测、多源数据同化、多模型综合分析、以及跨区域、跨学科合作的协同工作环境，急迫需要开展 e- Carbon Science 应用和示范研究，建立 ChinaFLUX e-Carbon Science 平台。本文论述了 ChinaFLUX e-Carbon Science 目标、组成、关键技术及现阶段的研究进展，基本实现了通量数据采集—传输—存储—管理—处理—可视化—共享服务的一体化，并应用高性能技术，集成 GIS 和遥感数据以及生态系统模型，构建了区域/全国尺度的碳收支模型模拟平台，支撑并促进了我国陆地生态系统碳循环的研究，为引领我国野外台站网络的信息化提供了示范作用。

【Abstract】Based on the observation and research network of China flux (ChinaFLUX), integrated study of ecosystem carbon cycling is not only an important part of global change research, but also a major scientific and technological task to cope with climate change. Integrated research of ecosystem carbon cycling requires multi-station observation network, multi-source data assimilation, multi-model overall analysis, and trans-regional, transdisciplinary collaborative environment. There is an urgent need for the application and demonstration studies of e-Carbon Science and the establishment of ChinaFLUX e-Carbon Science platform. This thesis discusses the objective, composition, key technologies and recent research advance of ChinaFLUX e-Carbon Science, and basically realizes the integration of the flux data acquisition - Transportation - Storage - Management - processing - visualization - service sharing. With the application of high

performance technology and the integration of GIS, remote sensing data and ecosystem models, it constructs a platform for regional / national scale carbon balance model simulation to support and promote Chinese terrestrial ecosystem carbon cycle research. It plays an exemplary role in leading the informationization of our field station network.

【关键词】 生态系统碳收支、e-Science、陆地生态系统、中国通量观测研究网络 (ChinaFLUX)、中国生态系统研究网络 (CERN)

【 Keywords 】 ecosystem carbon balance、e-Science、terrestrial ecosystem、the observation and research network of China flux (ChinaFLUX)、Chinese Ecosystem Research Network (CERN)

1. e-Science 与 e-Carbon Science

e-Science 的实质是「科学研究信息化」，是信息时代中科学研究环境和科学研究活动的典型体现。它不仅包括采用最新的信息技术建设起来的新一代的信息化基础设施，更有在这种基础设施和相关支撑技术构成的平台上开发的科学研究的应用，以及科学家们在这样一个前所未有的环境中进行的科学研究活动。e-Science 的实现将为科学家们提供一个信息化的科学研究环境，改变他们从事科学研究活动的方法和手段，甚至直接影响到学科的发展。综观 e-Science 的发展，其表现的主要特征为：网络技术为核心支撑；与特定学科密切结合；开展合作为主线，共享为追求目标。科学研究的信息化水平正在成为国家核心竞争力和科学技术水平的重要标志之一。

全球气候变化是国际社会共同关注的重大环境问题，减少温室气体排放、抑制全球变暖进程是保护地球系统的重大国际行动，是人类社会共同承担的国际责任和义务。近 20 年来，陆地生态系统碳循环已成为地球系统科学和全球变化科学的前沿

问题，也一直是国际地圈生物圈计划（IGBP）、世界气候研究计划（WCRP）和全球环境变化国际人文因素计划（IHDP）等重大科学计划的研究主题。各国相继启动了许多大型碳循环研究计划，试图开展从区域乃至洲际尺度开展碳循环过程及其区域响应的集成研究。我国也启动了一系列的碳循环研究项目，如中国科学院重大知识创新工程项目「中国陆地/近海生态系统碳收支研究（2002—2006）」，国家 973 项目「中国陆地生态系统碳循环及其驱动机制（2002—2007）」；国家自然科学基金重大项目「我国主要生态系统对全球变化的响应与适应样带研究（2005—2010）」等。建立了以中国生态系统研究网络（CERN）为基础的包含 8 个通量观测站和 16 个箱法观测站的中国通量观测研究网络（ChinaFLUX），获取了大量的观测数据，在通量观测技术、典型陆地生态系统碳收支及其环境控制机理、生态系统碳水通量模型模拟等领域开展了大量研究工作，并取得了重要进展。

全球或区域生态系统碳循环研究具有多台站联网观测、数据多源异构、多模型综合分析、跨领域科学家协同工作等方面的特点，需要多个领域的科学家在不同尺度协同开展观测、数据分析、模型模拟等工作，需要高性能的计算环境支持，需要海量的、多种来源的、不同学科的数据支持，生态系统碳循环研究已成为数据密集型工作，比以往任何时候都需要跨学科、跨越时间、空间、物理障碍的资源共享与协同工作，急迫需要开展 e-Science 应用和示范研究。近年来，国际上也逐渐开展了以碳循环为主题的 e-Science 研究。其主要研究内容包括碳循环信息化环境的整体结构、数据的获取、管理和分析、碳循环模型研究等方面。如 2007 年美国 Los Alamos 国家重点实验室提出的 Carbon Cyberinfrastructure 概念，首次完整地给出了碳循环研究信息化环境的体系结构；美国国家大气研究中心（National Center for Atmospheric Research, NCAR）等完成的基于网格技术的生物地球化学碳循环模型开发（GRID-BGC）项目；由 Microsoft 资助、Microsoft 与 AmeriFlux 合作开展了以 CO₂ 通量数据为基础的 e-Science 项目等。极大地提高了碳循环科研信息化协同研究的水平，并进而推动碳循环综合研究的进展。目前，我国的生态系统碳循环集成研究还缺乏实现野外观测仪器—台站—综合中心连接的高速网络环境和数据的快速采

集、传输、分析、共享服务工具；缺乏驱动区域／全国尺度的碳循环模型运算的 GIS 和遥感数据自动化处理工具；缺乏服务于生态系统碳循环模拟分析的模型运行和可视化环境；缺乏跨领域科学家的协同工作环境。从而无法实现碳循环研究数据的有效整合、模型模拟的快速运算和可视化输出，更无法实现全国／区域碳收支状况的快速评估，影响了对国家外交谈判需求的快速响应，以及科研成果的快速产出。e-Science 的实质是为科学研究的各个环节提供信息化支撑，中国科学院为了推动我国 e-Science 的发展而设立信息化专项，开展了多领域 e-Science 应用示范研究，这为我国碳循环研究领域的信息化发展带来了新的机遇，也提出了更高的要求。因此，本文提出了「e-Carbon Science」概念，强调碳循环研究与信息化协同工作环境的有效集成。即以我国不同尺度生态系统的碳收支为科学研究重点，信息化技术贯穿于碳循环数据的采集、传输、存储和管理、分析以及模型模拟和可视化等多个环节，通过信息化基础设施的提升和信息化科研工具的开发及应用，构建生态系统碳循环联网观测和联网实验的数据采集与在线远程传输、数据分析与多源数据管理、模拟分析与可视化表达一体化、科学家异地协同工作的信息化环境，实现多站点联网自动化观测、分布式多源异构数据的有效共享、多模型快速综合分析、跨领域科学家协同工作，实现科学、快速、准确计算我国不同尺度生态系统（典型、区域、全国）的碳收支状况。

2. ChinaFLUX e-Carbon Science 组成

以 ChinaFLUX 为基础的 e-Carbon Science 主要由基础设施层（硬件平台）、e-Carbon Science 支撑环境层（e-Science 基础软件平台）和 e-Carbon Science 科学应用层（e-Science 专业集成软件平台）组成，如图 1 所表示。



图 1 ChinaFLUX e-Carbon Science 组成

2.1 基础设施层

主要是硬件平台，是 ChinaFLUX 「e-Carbon Science」 平台研发与实现的物理基础，包括各野外台站的基于涡度相关技术的碳水通量传感器、数据采集器、网络设施、视频设备以及计算机硬件设备以及综合中心的数据存储设施、高性能计算机和可视化设备等。

2.2 e-Carbon Science 支撑环境层

主要包括 ChinaFLUX e-Carbon Science 的数据集成与服务环境、模型模拟环境和可视化分析环境以及科研协同工作环境等四个信息化支撑环境。其中数据集成环境不仅包括 ChinaFLUX 通量观测数据、GIS 和生态遥感数据以及模型模拟结果数据等多源异构数据，同时也具备一系列的实现从数据采集—传输—存储—管理和分析的数据处理工具；模型模拟环境基于现有的高性能环境，利用集群技术，实现站

点、区域和全国尺度的光合、呼吸模型以及遥感和过程模型的计算，并提供 web Services 服务，提供模型共享服务；可视化环境不仅包括野外观测设备和通讯网络的远程可视化监控，同时也对数据分析结果和模型模拟结果进行 2D/3D 展示；科研协同工作平台包括野外台站观测人员、数据管理人员和模型模拟人员以及项目负责人之间的沟通与协同，以及虚拟组织管理、协同科研活动组织（会议）、日志服务等。

2.3 e-Carbon Science 应用环境层

以不同尺度的生态系统碳收支研究为科学问题和应用目标，构建一系列的以上述四个工作环境为基础的、基于 WEB 的、典型生态系统碳源/汇季节变化及其机制研究、中国典型区域生态系统碳收支时空分布格局研究、中国陆地生态系统碳收支时空分布格局研究和东亚区域通量数据处理方法比对研究等四个生态系统碳收支评估应用系统，实现不同尺度生态系统碳收支快速评估。

3. e-Carbon Science 关键技术

3.1 基于 RBNB 的通量观测和视频数据的远程传输技术

为了保障通量和视频数据远程传输的实时性、稳定性、安全性。本文采用了基于内存和文件的环形缓冲池设计思想的 RBNB DataTurbine 数据流中间件技术，实现通量数据和视频数据的实时传输。RBNB DataTurbine 最早是由 Create 公司研究开发，提供了可靠的数据传输和对异构仪器访问的集成，特别是对分布的传感器流数据传输操作的支持。RBNB 服务器把数据保存在 Ring Buffer 中，嵌入式设备直接可以运行 DataTurbine 从而减少数据丢失的可能。RBNB DataTurbine 为通用的 Dataloggers 提供访问驱动，发布一些标准和 API 实现异构传感器集成。在碳循环研究中，温度、水汽、CO₂、风速等传感器以及摄像头时刻都在产生数据，如何对这些数据流建立起一个统一的管理机制非常重要，采用 RBNB DataTurbine 的数据组织方式，可以有效对多种异构的数据源进行有效管理，不同类型的传感器可以通过不

同的数据通道进行传输。

3.2 基于 VisualDB 的多源数据管理技术

碳循环研究中不仅涉及到小尺度的自动化观测获取的CO₂、H₂O通量观测数据和人工调查取样的生物量、土壤数据，同时，也包括大量的不同时空分辨率的GIS和遥感数据、文献资料数据，具有分布式、多源、异构、海量等特征。本文采用 VisualDB技术，建立元数据库和不同类型的数据库，实现台站和综合中心两个层次的碳循环研究数据管理与共享。VisualDB是一款由CNIC科学数据中心基于数据库连接池技术DBCP、表现层模板技术velocity、缓存技术、任务调度技术quartz、日志 logging技术、基于widget组件的门户技术等多项技术自主研发的WEB化数据库管理与发布工具。VisualDB为关系型和文件型数据库提供了安全可控的数据管理功能、强大的数据统计功能以及丰富的数据发布功能，整个使用过程完全可视化，高度可配置，用户基本可以脱离后台数据库系统体验数据库的建设，无需专业的RDBMS知识背景。基于VDK（VisualDB Development Kit）以及VTL（VisualDB Template Language）模版语言，可以很方便的实现二次开发，从而快速完成数据库应用的构建。此外，通过采用VisualDB作为数据访问中间件，可以方便的实现分布式数据的访问及互联。

3.3 基于WEB Services 和网格技术的数据分析和模型模拟技术

碳循环研究中，不仅涉及到大量的数据处理分析技术，也离不开复杂的、长时间的模型模拟运算。对运算性能要求不高的数据处理分析方法和模型-数据融合算法，本文采用WEB SERVICES技术将相关的数据处理方法和模型数据融合算法进行封装，提高数据分析方法的可重用性。Web Service是面向服务的体系结构，它使用XML、SOAP、WSDL、UDDI等一系列开放的标准协议进行开发，有利于将各个组织、应用程序、服务以及数据有机地连接在一起，为跨平台实现数据和服务共享提供了一套有效的方法。同时，大量的碳循环数据处理方法涉及到空间信息，本文采用OGC 标准构建空间数据处理服务。对运算性能要求高的复杂的模型模拟运算，本文采用网格集群的方式提高模型的运算速度。即应用网格集群技术在CERN综合

中心构建高性能计算环境，通过网格计算聚合计算资源来满足计算能力的需要。

3.4 基于 Duckling 的协同工作环境套件技术

支持 e-Science 的协同工作环境套件 (Duckling) 是专为科研团队提供的综合性的资源共享和协作平台，由中国科学院计算机网络信息中心开发。其核心工具集包括：文档协同工具 (DCT)、用户管理工具 (UMT)、文档库管理工具 (CLB)、活动组织工具 (AAT) 和日志服务工具 (DLOG)。该套件支持多语言，并具备「应用插件定制开发框架」这一框架所提供的应用插件开发机制可以便利地构建统一的科研协作平台。如台站、综合中心的人工数据汇交系统、会议注册系统等。

4. 主要研究与应用进展

4.1 通量和视频数据的实时传输

目前，在 ChinaFLUX 所属的 8 个台站中，已经在江西千烟州、山东禹城和广东鼎湖山实现了基于 RBNB 技术的通量和视频观测数据的远程传输。在台站的传感器和数据采集器端，通过串口服务器将 RS232 串口转换为 RJ4 网口，实现了数据采集器的 TCP/IP 协议支持功能，三个野外台站部署安装数据采集、数据传输程序。在北京综合中心部署安装 RBNB 服务器和 Web 服务器，通量数据、视频图片数据和设备状态数据实时传输到北京综合中心并以 Web 的方式对外提供访问服务，图 2、3 是系统的运行界面。通过对通量数据进行实时传输、连续监测和自动报警，有效减少了由于仪器故障导致的数据缺失，提高了观测数据质量。同时，也逐渐转变野外工作模式，特别在鼎湖山站观测人员人工定期获取数据的频率由以前的每 2 天 1 次转变为每 7 天 1 次，降低了野外工作强度。



图 2 通量数据实时显示



图 3 视频实时显示

4.2 碳循环多源异构数据的管理和共享系统

ChinaFLUX 碳循环多源异构数据管理与共享系统将易用性作为一个主要设计目标，由基于 ViisD 技术、B/S 结构的台站和总中心两个分系统构成。在数据资源上，台站分系统主要包括元数据库、通量观测数据库和视频数据库；综合中心分系统不仅包括元数据库、通量观测数据库和视频数据库，同时还包括站点、区域和全国尺度的 GIS 和生态遥感数据库以及模型模拟结果数据库。在系统功能上，台站分系统面向台站观测人员，主要作为数据管理的工具，具有实时监控、数据检索、数据管理和设备控制等功能；而总中心分系统面向广大科研和社会公众用户，主要作为我国开展碳循环研究的数据支撑平台，具有实时监控、数据查询与下载、可视化分析，系统管理（包括用户管理、数据分级）等功能。目前，该系统已经基本完成，正处于边完善边运行阶段。图 4 是综合中心和台站分系统的相关界面。



图 4 综合中心和台站数据管理与共享系统

4.3 通量数据分析处理工具

如何处理和利用基于涡度相关通量观测技术获取的不同生态系统类型的碳水通量数据，从而向研究全球变化相关问题的科学家和建模者们提供完整可靠的生态系统通量数据产品，已成为全球通量观测研究领域所面临的一个严峻挑战。ChinaFLUX 在参考FLUXNET、AmeriFLUX、EuroFLUX和AsiaFLUX等通量数据的处理流程的基础上，研究并建立了一套较为完整的CO₂通量数据处理流程。本研究应用科学计算软件MATLAB、科学工作流Kepler和WEB Services技术完成了C/S模式和B/S模式的ChinaFLUX通量数据分析处理工具的开发，实现了ChinaFLUX 8个野外台站通量观测数据产品的自动化生成。同时，也构建了关键算法的web services库，为通量数据处理算法的重用和改进提供了一个支撑平台。图5为基于WEB SERVICES的通量数据处理系统界面。

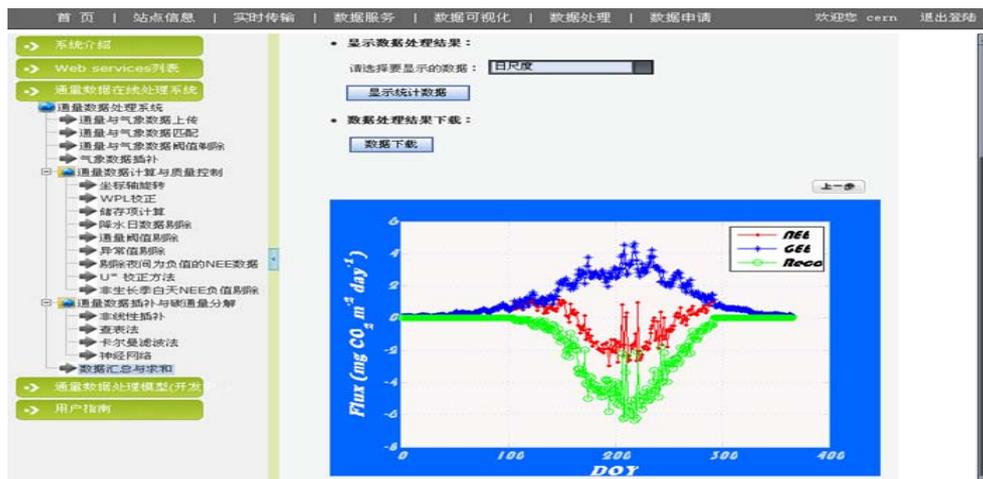


图5 基于WEB 的ChinaFLUX通量数据分析处理系统

4.4 区域/全国尺度的碳循环模型模拟平台

本项目应用 mpich2 作为 MPI 中间件，GlobusToolkits5.0 作为网络中间件进行模型模拟系统环境的搭建，以 Python2.6 为主要语言进行开发。其中，科学计算调用 Numpy、GDAL、mpi4py 等函数包进行并行计算，使用 PyWPS 和 Apache2 封装模型运算为网络处理服务（OGC WPS）并发布。目前，已经整合了站点尺度的光合和呼吸模型、区域/全国尺度的碳收支遥感模型以及数据模型—融合方法（MCMC 方法和 Klamann Filter 算法）。用户可以在线远程应用并计算典型、区域尺度的碳收支模型，同时获取模型模拟结果。

4.5 协同工作平台和虚拟科研组织

基于 Duckling 协同工作环境套件，我们构建了 ChinaFLUX 数据汇交虚拟组织，实现了台站观测人员、综合中心数据管理人员、项目负责人三者之间的有机协同工作；建立了 e-Science 项目开发组协同工作环境，实现了科学院地理所何计算机网络信息中心两所组成的开发团队关于工作计划、内部交流、资料共享与讨论；搭建了 AsiaFLUX2010 年研讨会会议管理平台（包括 AsiaFLUX 年会会务组织、信息获取、交流与沟通等），实现了用户注册、摘要提交与摘要集的自动编排出版、酒店预订、管理和维护用户信息、会议信息发布等功能，提高了会议组织效率。

5. 结语

推进 e-Science 发展，引领我国 e-Science 应用，是科学院「十一·五」信息化工作的重要目标。本文提出的「e-Carbon Science」研究工作还处于起步阶段，需要准确和把握碳循环研究对信息化环境建设的需求。需要生态学、地学和计算机等多个领域科学家的共同、长期努力，以碳循环研究领域信息的快速获取、资源整合共享、高性能模拟以及跨领域协同研究为核心，深入研究信息化技术发展对传统碳循环研究模式的影响，构建 e-Carbon Science 环境。从而建立我国第一个陆地生态系统碳循环观测网络（ChinaFLUX）信息化基础设施，实现野外台站 Flux 观测设备的网络互联，实现野外观测数据的实时传输、存储、处理与共享服务。该项研究工作将为我国生态系统碳循环观测网络的信息化提供示范作用，逐步改变我国生态系统通量观测的研究方法和手段，推动和促进我国生态信息学和碳循环研究学科的发展。

参考文献

- [1] Liese Coulter, Pep Canadell, Shobhakar Dhakal, A GCP report for the ESSP, CARBON REDUCTIONS AND OFFSETS, 2007, <http://www.globalcarbonproject.org/index.htm>
- [2] 2003 The U.S. Climate Change Science Program. Vision for the Program and Highlights of the Scientific Strategic Plan. A Report (41 pages) by the Climate Change Science Program and the Subcommittee on Global Change Research.
- [3] Yu GR, Fu YL, Sun XM et al., 2006a. Recent progress and future directions of ChinaFLUX. Science in China Series D, 49 (Supp II): 1–23.
- [4] 于贵瑞，孙晓敏等. 2006. 陆地生态系统通量观测的原理与方法. 北京：高等教育出版社.
- [5] 于贵瑞，孙晓敏. 2008. 中国陆地生态系统的碳通量观测技术及其时空变化特征,

北京：科学出版社.

- [6] Yu G.R., Zhang L.M., Sun X.M., 2008. Environmental controls over carbon exchange of three forest ecosystems in eastern China. *Global Change Biology*. 14, 1–17, doi: 10.1111/j.1365-2486.2008.01663.x
- [7] Zhao L, Li YN, Xu SX et al., 2006. Diurnal, seasonal and annual variation in net ecosystem CO₂ exchange of an alpine shrubland on Qinghai-Tibetan plateau. *Global Change Biology*, 12: 1940–1953.
- [8] Zhou GY, Liu SG, Li Z et al., 2006. Old-growth forests can accumulate carbon in soils. *Science*, 314: 1417.
- [9] 李克让, 王绍强, 曹明奎, 2003. 中国植被和土壤碳储量. *中国科学 (D 辑)*, 33: 72–80.
- [10] National Science Foundation, 2006. Cyberinfrastructure Vision for 21st Century Discovery. www.nsf.gov/od/oci/ci_v5.pdf.
- [11] Randy Butler, Mark Servilla et al, 2007. Cyberinfrastructure for the analysis of ecological acoustic sensor data: a use case study in grid deployment . *Cluster Computing*, 10:301~310
- [12] Thornton, P. E. 2005. Biome-BGC: Modeling Effects of Disturbance and ClimateModel product. Available on-line [<http://www.daac.ornl.gov>] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A.
- [13] Environmental Cyberinfrastructure Needs For Distributed Sensor Networks; workshop held August 12-14, 2003; information available at http://www.lternet.edu/sensor_report
- [14] Gretchen Miller, The AmeriFlux Cyberinfrastructure Prototype – Introduction and Applications . bwc.berkeley.edu/home/presentations/Ameriflux_wg3

[15] Paul M. Rich, Gordon N. Keating et al .2007. A Vision for Carbon

Cyberinfrastructure.www.bigskyco2.org/presentations/chapman-cyberinfrastructure-18jan05.ppt

【作者简介】

于贵瑞 男

职 称： 中科院 地理科学与资源研究所 研究员

职 务： 中科院 地理科学与资源研究所 副所长

研究领域： 主要从事陆地生态系统通量观测理论和方法、生态系统碳水循环及其管理、植物生理生态、SPAC 系统水分运动及生态数据管理等领域研究工作。

个人简介： 先后获沈阳农业大学农学博士和日本千叶大学环境学博士学位。中科院「百人计划」入选者和国家基金委「杰出青年基金」入选者。主持中科院重大项目、国家基金委重大项目和重大国际合作项目、国家 973 计划项目课题等 10 余项科研项目。发表论文 180 余篇，其中 SCI 论文 63 篇，主编专著 5 部。

联络电话： 86-010-64889268

联络邮箱： yugr@igsnr.ac.cn

何洪林 男

职 称： 中科院 地理科学与资源研究所 副研究员

职 务： 中科院 地理科学与资源研究所「生态系统研究网络综合中心」
副主任

研究领域： 生态信息学与地理信息系统、遥感应用

个人简介：获华东师范大学地理信息系统博士学位，中科院地理科学与资源研究所博士后。先后主持和参与国家自然科学基金、科学院方向性课题、国家 973 课题等 10 余项。发表论文 20 余篇，其中 SCI 论文 4 篇，专著 2 部。

联络电话：86-010-64889599

联络邮箱：hehl@igsnrr.ac.cn

黎建辉 男

职 称：中科院 计算器网络信息中心 研究员

职 务：中科院 计算器网络信息中心「科学数据中心」主任

研究领域：海量数据存储技术、大规模数据处理、语义集成等

个人简介：任中科院计算器网络信息中心科学数据中心主任、研究员级高级工程师及硕士生导师，发表论文 30 余篇，主持完成了一项国家标准的研制，参与了 2 项国际标准的研制。近年来主要从事海量存储技术、海量数据处理技术与科研应用、语义集成等方向的研究工作。

联络电话：86-010-58812518

联络邮箱：lijh@cnic.cn

周园春 男

职 称：中科院 计算器网络信息中心 助理研究员

研究领域：数据挖掘、数据库技术、进化计算及 workflow 技术等

联络电话：86-10-58812561

联络邮箱：yczhou@sdb.cnica.cn