

科研信息化发展初探

张亚明 唐朝生

(燕山大学经济管理学院, 河北秦皇岛 066004)

摘要: 科研信息化对科技和经济社会发展的巨大影响引起了全球的广泛重视, 发达国家纷纷把推进科研信息化作为提高国家科技竞争力的关键和迎接新一轮科技革命的战略举措。文章综述科研信息化的研究成果, 总结发达国家科研信息化的发展模式, 分析我国科研信息化的现状及存在的问题, 最后设计出我国科研信息化的发展路径, 为提升我国科研信息化水平提供有益的理论参考。

关键词: 信息技术; 科研信息化; 网格; 发展模式; 路径设计

中图分类号: G250.74, G31

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2012.02.014

Preliminary Study on e-Science

Zhang Yaming, Tang Chaosheng

(School of Economy and Management, YanShan University, Qinhuangdao 066004)

Abstract: The significant influence of e-Science on technology and economic and social development has been attracting worldwide attention. Many developed countries regard it as a key factor to improve national science and technology competitiveness, and also revolutionary measure to rise to the challenges of -new wavescientific technology revolution. This paper firstly reviews e-Science literature both home and abroad, summarizes its development models in developed countries, then analyzes the status quo and existing problems of e-Science in China, and finally puts forward developing path of e-Science in our country in order to offer reference to raise its informatization level.

Keywords: information technology, e-Science, Grid, development model, route design

在信息技术推动下, 科研活动中对信息和数据的处理速度和质量要求越来越高, 区域乃至全球性的科研交流日益频繁。科研信息化作为科研活动与信息时代的互动产物, 顺应了科研活动发展的这一态势, 有利于提高我国的科研发展水平, 增强我国科技实力和国际竞争力。加强对科研信息化的研究应成为我国科技发展引领性指标。值得注意的是, 科研信息化作为一个新的概念, 在国内外还都处于起步探索阶段, 在理论研究和实践中仍然有许多急需完善的地方。本文拟对科研信息化的国内外研究现状进行综述, 指出研究的经验和不足。同

时, 总结发达国家发展科研信息化的模式, 分析我国科研信息化的现状与问题, 最终提出我国科研信息化的发展路径。

1 科研信息化的内涵

科研信息化译自英语 e-Science。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, 甚至直接影响一些学科的发展。牛津大学认为, e-Science 的主要涵义包括 4 方面的内容, 即基于 Internet 的全球范围研究、存在大量可利用的数据

第一作者简介: 张亚明 (1962-), 男, 博士后, 留美学者, 燕山大学经济管理学院电子商务系主任, 教授, 博士生导师, 研究方向: 管理科学与信息系统。

基金项目: 河北省科学技术研究与发展项目“基于并行计算的神经网络物流智能预测系统的研究”(07213529)。

收稿日期: 2010年8月24日。

资源、存在T级规模的计算资源、能够实现高性能的可视化^[1]。欧洲生物信息研究所认为，e-Science就是依赖通信技术与计算发展，开发出一种对活动的各个方面提供支持的数字技术，从而给科学家的工作提供支撑。英国e-Science研究院主任Malcolm Atkinson教授则将其定义为：“利用先进计算思想的研究方式的系统发展。”这种研究方式，可以使研究者通过自己的桌面电脑访问和使用分布在各处的计算机上的资源，这些资源包括数据库、极大规模的计算能力、科学仪器和高性能可视化。也有研究认为，e-Science不仅仅是高带宽通信和高性能计算，更强调科研团体之间的协作研究。通常地，相关研究团体组成虚拟组织^[2]。

英国工贸部e-Infrastructure工作组的研究报告指出：科研信息化基础设施包括实施e-Science需要的网络、服务和软件，关键问题包括数据和信息的产生、保存、管理、服务、虚拟研究组织、网络、计算和数据存储设施、中间件、政策和法律环境等^[3]。Steven Wallace的研究认为，e-Science具有明显的技术性特征，即高度依赖网络技术以及对海量数据的处理，同时还涉及可视化基础设施的建设和世界范围内信息数据的无差别共享^[4]。Rouredd等研究了e-Science在样品分析中的工作过程，同时提出了基于市场的服务所有者-服务消费者模型^[5]。该模型中基于市场的服务所有者在特定的协议下为服务消费者提供服务。它们基于市场规则发生相互作用，而市场规则是由市场所有者提供。市场所有者可以是服务所有者、服务消费者或者是第三方中立的团体。Ian Foster等研究了信息的虚拟共享问题。他们认为信息共享应该具有条件约束，应该进行识别和限制，并且执行某些约束需要特殊的策略机制对信息资源的提供者和使用者进行身份认证，强调授权在共享机制中的一致性判定^[6]。

国外学者的研究指明了e-Science所必须具备的先进的计算机和光纤传输等硬件设施、可视化的和交互性强的人机界面、极大量的信息资源存储、信息的共享理念、协作的信息平台等。同时，又从信息化环境的建设、科研信息资源的开发与应用等方面进行了广泛探索，从实际应用角度提出了一系列理论和实践模型，特别是对网格技术的研究，如数据/计算网格、信息网格、知识网格、语义网格等，从信息技术方面解答了e-Science深入发展

和实际应用问题，覆盖了从基础知识创新到生产生活的多个层面，形成了初步的研究体系。但是，这些研究成果的技术倾向严重，造成非专业人员理解上的障碍，客观上抬高了公众了解e-Science的门槛，不利于e-Science的普适性推广。

2 国外科研信息化的发展模式

美、英、欧盟以及日、韩等国发展e-Science呈现出一定的规律性特征，可归纳为以下4个方面。

(1) 政府主导推动e-Science的实践应用。主要表现在：政府成立专门的管理机构对e-Science相关项目进行统一指导和规范，为促进e-Science发展制定多种科技规划，同时也为e-Science提供大量的资助。例如，英国政府专门成立了一个e-Science项目指导委员会和由学术界和工业界人士组成的e-Science技术咨询组，负责监督管理e-Science核心计划^[7]。该计划得到215亿英镑的研究经费，成为近年来英国政府在信息技术领域支持的最大研究计划；美国e-Science的发展主要得到了国家科学基金会(NSF)和能源部(DOE)这两大政府科研机构的资金支持。例如NSF支持下的NMI计划(国家磁浮倡议计划)涉及金额4145万美元，对ITR计划(信息技术研究计划)的资助额度达到13244万美元。DOE得到了美国国会近6亿美元的拨款，主要用于支持DOESciDAC计划(能源部先进计算计划)和DOE National Co-laboratories(能源部联邦实验室)。欧盟委员会在第五研究框架计划(FP5)中第一次开始资助网格研究项目，总资助额度达到5800万欧元，在2002-2006年的欧盟第六研究框架计划(FP6)中，资助额度超过FP5的两倍，仅FP6第一批网格计划资助金额就已达5200万欧元。

(2) 专业性科学协会组织对发展e-Science进行具体指导。例如，英国生物科学研究理事会负责的e-Science计划主要涉及生物信息学及网格技术在改进生活质量，尤其是基因群组、生物结构、细胞动力学和生物多样性等方面的应用^[7]。英国研究实验室理事会负责管理大型科研设备，并力图使实验设施网格化；德国的e-Science计划项目则由著名的研究机构FIZ Karlsruhe和马普学会共同承担；荷兰阿姆斯特丹科技中心和阿姆斯特丹大学等4个科研机构共同负责了面向e-Science的虚拟实验室

项目。

(3) 合作促进 e-Science 的应用。这种合作既存在于国家之间，又包括不同的科研主体。例如美国发起的 Access Grid 就是一个国际性的 e-Science 计划，参与的国家有美、英、法、日、韩、俄罗斯等国；中国和英国也在进行着 e-Science 计划的合作，合作的内容包括建立网格中间件研究所 OMII-China、网格数据存储维护技术、e-Science 和网格应用的合作以及促进科学家的交流^[8]。

(4) 应用网格技术推动 e-Science 的进步。e-Science 网格大致可以分为 3 个层面，即数据 / 计算网格、信息网格和知识网格，关注的重点分别是数据和计算机共享与协作，信息的存储、访问与保存以及知识的使用、抽取、发布和维护^[9]。对网格技术的研究成为 e-Science 发展最具活力同时也最具发展潜力的领域，衍生出了许多新的研究领域，如语义 Web 和网格相结合产生的新研究领域。

3 国内科研信息化现状

江绵恒指出，e-Science 不仅包括采用最新的信息技术如网格等建设起来的新一代信息基础设施，而且包括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下进行的科学研究活动。黄国彬等指出，e-Science 是一种全新的科研协作模式，是一种科研信息平台，其着眼点是“关键科学领域的全球性合作以及实现这种全球性合作的下一代基础设施”。通过开发和应用各种信息技术，让科学家能够利用更加先进的技术方法更高效地从事科学研究。他总结出 e-Science 的研究特点：利用信息技术进行研究数据的采集、获取、保存和分析问题；数字环境下被研究对象的建模和仿真问题；动态的、分布式的、虚拟的研究团队的组建和工作模式问题；网格的研究与应用以及由此引发的计算力等资源的共享等问题^[10]。宋琳琳认为，e-Science 指的是通过互联网实现分布式的、全球性合作的大规模日益增长科学，实际上是一种基础设施，提供了一种信息化的科学研究环境和平台，使不同学科领域的研究和科研活动能够有针对性地开发特定的科学研究与应用^[8]。渠岸杉等对支撑 e-Science 应用的基础体系结构进行了讨论，对 e-Science 中需要解决的关键技术问题及研究方向进行了分析，最后对我国 e-

Science 研究的前景进行了探讨，指出应从我国科学研究的实际出发，以中国大学 e-Science 网格为突破点，通过实际应用促进我国 e-Science 的研究进程，在此基础上，经过试点和推广，最终全面实现我国 e-Science 计划^[9]。

国内学者的研究具有明显的针对性。首先，和国外学者一样，认识到发展 e-Science 必须重视基础设施和信息环境的建设，也注意到网格技术对科研信息化的推动作用并进行一定程度的研究。其次，对国外 e-Science 发展的现状和经验进行了深刻的总结，对科研信息化进行理论溯源，梳理了发达国家为推动 e-Science 所采取的多种政策措施，并提出相应的对策建议，但是其可操作性方面仍然欠佳。再次，有选择性地将 e-Science 的应用作为研究重点，特别是对高校教育、图书馆建设、农业信息化平台、卫生事业等科研信息化平台建设进行了探讨。这些研究尚显薄弱，不够全面，需要在实践可行性方面进行深化。

此外，从基础设施及技术储备来看，我国已具备了发展 e-Science 的良好硬件。例如，经过技术升级之后，我国首台千万亿次超级计算机“天河一号”二期系统性能已实现飞跃，以每秒 4700 万亿次的峰值性能和每秒 2507 万亿次的实测性能位居榜首，是目前美国“美洲虎”超级计算机性能的 1.425 倍。同时，我国拥有覆盖面最广的互联网干线，拥有 CNNIC 这样的“科学数据库及其信息系统”。中国科技网 (CSTNET) 已覆盖全国 20 多个省、市、自治区。这些在信息化建设过程中所积累的丰富经验为 e-Science 的发展提供了有益的借鉴。另外，不少地方和部门已经在生产实践过程中开始了 e-Science 的应用，例如浙江省政府与科学技术部、国家知识产权局合办的浙江网上技术市场，使科研产品更能切中市场的需求点，有力地促进了当地经济的发展，形成了“全国资源浙江用”的发展新模式。教育部实施的中国教育科研网格计划，充分利用中国国家教育科研网 CERNET 以及高校的大量计算资源和信息资源，将分布在教育和科研网上自治的分布异构的海量资源集成起来，实现 CERNET 环境下资源的有效共享，从而形成国家级的、高水平低成本的科研教学服务平台。

我国的 e-Science 存在的问题也较为突出。首先，对 e-Science 的理论研究多是概念引入性质的介绍，自主创新性质的研究成果相当匮乏，基本是处

于跟随学习阶段,亦步亦趋。其次,信息技术相关的工艺技术与国外差距还非常大,如集成电路技术相对落后,集成电路芯片80%都依靠进口。再次,e-Science的应用只在传统的高校和科研院所形成一定规模,没有在全社会形成普遍性的应用,信息交流作用有限,难以实现科研成果的共享与信息服务。此外,我国地域广袤,各地方对e-Science的理解和接触层次也不同,形成了严重的“数字鸿沟”,不利于e-Science的发展。

4 促进我国科研信息化的思考与建议

(1)发挥国家政策工具的强大推动力,以“点”“面”结合的方式推动e-Science的梯度发展。e-Science的实质是实现分布在不同地方的仪器、科学数据、信息资源、计算资源、科研工具等各种科学研究资源的共享,实现虚拟科研团队科学研究活动的协同。e-Science虽然有强大的潜力,但是发展还远未完善,短期内仍需要国家政策的扶助。就目前情况来看,我国科研经费少,e-Science所需的科研人力资源、设备以及知识储备都严重不足,全面进行e-Science建设尚不现实,因此,必须以“点”“面”结合的方式有步骤、有计划地进行e-Science的梯度建设。所谓“点”,就是利用国家专项资源,在关系国计民生的重点领域和与国外先进水平相比差距最小的领域开展e-Science研究与建设,在保持与国际水平同步的情况下,尽快使e-Science在诸如信息产业、化工、机电等国家重点行业实现普遍应用,从而在根本上增强国家竞争力。“面”是指国家相关部门要出台促进e-Science发展的战略规划和相关的法律法规,营造发展e-Science的良好制度环境,在宏观上促成发展e-Science的社会氛围,培养全社会建设e-Science的共识。同时,在推进e-Science发展过程中,要注意区域研究向纵深推进,形成梯度建设模式,在发展基础条件好的地方先试行并总结相关经验,因地制宜地进行区域e-Science建设,做到既重视基础理论又抓住前沿问题,使e-Science在我国的发展既稳步进行又立足实际,可以有针对性地解决现实问题,使有限的科研资源取得最优的效果。

(2)以产业化和盈利性为切入点,调动产业单元的积极性,加速e-Science的商用化。通过发展e-Science,可以将海量的信息分布式存储,借助大容量通信传输系统和高性能的计算处理设备,

把之前相互封闭的信息互联并激活,使之具有应用价值,形成一个具有无限存储量的“数据库”,通过研发成熟的中间接口或“人-机”交互式界面软件,使信息需求方可以便利地进行诸如文献数据的调用、合作伙伴选择、科研或商业项目开发等生产经营活动,使e-Science从单纯的科研活动向实用性的生产活动转化。因此,加强实用性的应用研究,组织与e-Science关系密切的高新技术企业,利用现有的技术条件进行e-Science系统的开发,使之不仅能够用于科研活动,而且可以进一步服务于产品研发、方案规划、投资决策等生产实践,发掘e-Science系统及其衍生知识的盈利潜力,进而加速“产学研”一体化,从根本上促进e-Science的健康发展。

(3)以既有的公共信息终端为跳板,夯实e-Science资源基础,尽快形成e-Science发展的实体雏形。e-Science的发展需要大量基础资源,特别是科研数据、科技文献、自然资源、科研交流平台、网络科研支持系统以及科研设备分布等资源,而这些资源大多存在于高等院校、科研院所、公共图书馆、政府网站以及各种形式的专用信息数据库。因此,在促进e-Science发展的过程中,要重点考虑这些部门和机构的e-Science建设,并在此范围内优先进行适度规模的e-Science环境集成,短时间内促进一部分e-Science实体的较快发展。具体来看,一方面要充分发挥这些信息终端在科技人力、物力等方面的传统优势,继续加大投入,加强其在信息收集、处理和存储方面的作用,为e-Science的快速发展提供丰富的信息支撑;另一方面要统一标准,在不同的科研主体之间进行信息共享和资源的交流以及相互间的协作,实现优势互补,防止出现科研活动方面的重复建设。

(4)促进e-Science技术研究的多头并进与重点突破相结合。多头并进,就是不能轻易放弃对任何一个学科的研究,要保持与国际主流技术的衔接,防止出现“短板效应”;重点突破,就是专门集中科研人力、财力和物力,对e-Science发展的关键技术进行集中攻关,在一个或者几个关键技术领域取得领先地位,赢得e-Science发展的话语权。就目前发展趋势来看,网络技术仍然是支撑e-Science发展的关键技术,需要着力研究。我国通过中国国家网络、中科院计算技术研究所“织女

(下转第110页)

技文献平台中心网站,形成山西省兵器科技文献资源保障体系与服务体系,在资源整合和开放服务上取得了一定成效。但目前平台服务功能还存在很大局限性,如何进一步提升平台服务能力,满足科技创新需求还需要做大量的工作,需要结合学科领域特点,从技术手段和服务模式上进行创新,在更大范围、更深层次上推进平台信息服务工作,以带动平台平滑过度到长效运行阶段,逐步走向资源数字化、服务网络化、运行规范化的发展轨道。

参考文献

- [1] 张群,彭奇志.应用JSP技术开发学科导航数据库[J].情报科学,2004(5):587-589.
- [2] 付年钧.湖北省科技信息共享服务平台的建设[J].中国科技资源导刊,2009(9):56-60.
- [3] 曹如国,徐兴余,谭永钦.利用TPI自建特色数据库

的探索[J].大学图书馆学报,2008(6):89-92.

- [4] 贺培凤,何小峰.地方医学科技文献平台建设机制分析与研究[J].医学信息学杂志,2008(5):34-37.
- [5] 刘冬梅.图书馆信息资源整合问题研究[J].图书馆工作与研究,2009(6):22-25.
- [6] 贺培凤,王秀平,冯毅,等.山西省医学科技文献信息资源与服务平台建设与实践研究[J].医学信息学杂志,2009(10):65-69.
- [7] 李演军.甘肃省科技文献信息共享平台联盟服务模式探讨[J].甘肃科技,2009(21):81-83.
- [8] 韩瑞平.基于网络的科技信息服务平台建设研究——国内省级科技信息服务网站建设的调查与评价[J].科技管理研究,2009(11):407-409.
- [9] 李晓.云南省科技文献公用平台管理机制创新实践[J].现代情报,2010(1):147-150.

(上接第88页)

星网格计划”等,已在网格技术的研究方面积累了相当的经验,为e-Science技术取得突破性进展奠定了良好的基础。在今后一段时期内,强化对网格技术的深入研究应该成为我国发展e-Science技术的重点。

(5)重视e-Science科研群体的培育。适应e-Science发展的科研群体应该具有开放性和延展性,可以是一个或几个科研团队,也可以是多个科研团队之间的相互融合。这样的群体应该是动态的,其个人成员或团队可以自由加入或离开,以利于在充分发挥专业人才的知识能力和积极性的同时,使新观点和新思维得到自由发展;这样的群体应该具有良好的成员构成,既有高水平的专家学者,又有精通多学科知识的通才,使整个科研群体相互协作,呈现复合性特点。高水准的科研群体的形成,有利于在整体上促进e-Science的健康、均衡和全面发展,提高e-Science的整体发展水平。

参考文献

- [1] 陈沈斌.地球科学的e-Science探索[C].第七届科学数据库与信息技术学术讨论会,2004:19-24.
- [2] Foster I, Kesselman C, Tuecke S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations[J].Inter-

national Journal of High Performance Computing Applications, 2001, 15(3):200-222.

- [3] OSI E-infrastructure Working Group. Developing the UK's-Infrastructure for Science and Innovation[R]. Edinburgh: National e-Science Centre,2007,3:22.
- [4] Steven Wallace. Global Terabit Research Network: Building Global Cyber Infrastructure Browse[EB/OL]. [2004-10-09].http://archive.dante.net/conference/globalsummit2002/html/3-3wallace.sld003.htm.
- [5] Roure David De, Nicholas Jennings, Nigel Shadbolt. Research Agenda for the Semantic Grid: A Future e-Science Infrastructure[R]. Edinburgh: National e-Science Centre, 2001, 12:3.
- [6] Ian Foster, Carl Kesselman, Steven Tuecke. The Anatomy of the Grid: Enabling Scalable Virtual Organizations [EB/OL]. [2008-08-20].http://www.globus.org/alliance/publications/papers/anatomy.pdf.
- [7] 王小飞.英国全面开展e-Science研究[J].全球科技经济瞭望,2002(6):34-36.
- [8] 宋琳琳.e-Science发展情况简介[J].图书馆学研究,2005(7):21-25.
- [9] 渠岸杉,黄理灿,陈卫东,等.e-Science技术综述[J].计算机测量与控制,2003(9):641-646.
- [10] 黄国彬,孙坦.e-Science的特点及文献情报机构的应对措施[J].图书馆杂志,2005(9):22-24.