

美国“合作实验室”的发展及其对我国科技政策的启示

Development of US Collaboratory and Its Inspiration to China's S&T Policies

秦 健

(美国伊利诺大学图书馆学情报学院)

一、合作实验室的概念

目前,风靡全球的 Internet 正在从根本上改变科学家工作的方式和方法,电子邮件及其他数据信息交流技术将科学合作大大推进了一步,而许多这些合作在过去都是非常困难甚至是不可能的。早在 1991 年底,美国科学基金会、美国国家研究顾委会下属的计算机与远程通讯部组成了一个“全国(科学)合作实验室委员会”,负责调查科学家对信息技术的需求和开发支持科研合作的信息技术。“合作实验室”(Collaboratory)的概念是由威廉·伍尔夫(William Wulf)在 1989 年首次提出的,他将英语中的“collaboration”和“laboratory”两字各取一半复合而成。伍尔夫定义“合作实验室”为:“没有围墙的[研究]中心,不同地区的研究者可以互相交流,操作仪器设备,共享数据及计算资源,并查询电子图书馆所存贮的信息”。从技术上说来,“合作实验室”是一个分散的计算机系统,它拥有计算机网络联结的实验室仪器与数据收集工作台,能提供科研合作所需的软、硬件和系统维护与发展用户技术所需的财物和人力,具有组织、管理数据和大规模数据共享的能力。总之,合作实验室提供的技术基础是特为支持科学家、实验仪器及网络数据之间的配合协同,克服地理区域的障碍而建立的。1991 年会议的结果之一是确定了 3 个学科作为试点:(1)海洋学,该学科在收集数据上的困难与费用、实验和模型构造上的相互独立性,都为合作提供了良好的动机。(2)空间物理学,该学科必须广泛应用计算技术来分析数据。(3)基因图谱与序列,该领域的研究已导致大规模数据库的建立及对其与日俱增的依赖。

合作实验室试点总结出这类科学辅助系统应具备的能力包括:(1)数据分享,即在不同地点研究同一项目的科学家能迅速、方便地查找本单位与外单位的数据库。(2)软件分享,即在不同地点的科学

家能方便地分享支持数据分析、图表可视、模型构造的软件。(3)远程仪器控制,即各地的科学家能操纵、控制那些位于人类难以到达的地方的仪器,如南极、宇宙空间。(4)远程学术交流,此地的科学家能与彼地的科学家进行跨越时空的实时交流。

二、合作实验室的技术

3 个试点的参与者总结出 3 项需要认真考虑的主要技术问题:(1)数据资源的相互联结。科学家认为,由于电子图书馆具有迅速查找文献的潜力,而且图书馆能提供检索信息数据的服务,因此建立电子图书馆是合作实验室的第一优先内容。(2)易于检索的数据档案库。随着试验及数据收集变得越来越复杂和昂贵,数据档案库已变得日益重要。(3)能支持“全天候”(即任一地点、任一或所有数据源)数据检索的综合系统。目前,虽然数据库的数据目录已有编出(如数据来源、数据类型、数据收集的时间等),但各个数据库的显示格式、检索方法,以及数据目录的可检索性等,都有很大的区别,特别是在不同的学科之间,区别更大。现在已有一些模型可以使用户进行跨数据库或多数据库检索,如布鲁斯·沙茨(Bruce Schatz)等研究的“蠕虫共用系统”(Worm Community System),该系统代表了解决一个较小研究领域对数据分享的需要的方法。如果一个学科领域里的所有或将近所有的信息可以用一致的方法检索,那么全部信息的总体可以为科研提供宝贵资源。但是,将各种不同的数据库联成一个统一的总体,牵涉到一系列标准化和数据转换问题。文字、图形、图像的转换在各学科可能大同小异,其他带有学科特殊性的数据,如基因组研究中的基因序列图,海洋物理学中的海流与温度,则会有一些难度。

合作实验室的另一重要方面是实现软件分享,科学家不仅能随时使用本单位已有的软件,而且还能随时使用外单位的软件。在软件日趋复杂的今

天,软件分享对科学家越来越重要。尽管软件分享在实行过程中会有许多困难,但这样的分享已经在一些小的地区网上实现,如多功能的PC软件已经把数据库、文字处理、图表、电子邮件集成一体,表格可以表达成图形并插入文字文件,也可以与数据库里的数据对应,文字文件可以和一个或多个数据库合并。这样的集成系统必须有一套通用格式来支持数据交换和转换。多功能PC软件(包括视像会议[Video Conferencing]和计算机支持的合作工具)的发展不仅使软件市场大为受益,也可能会使合作实验室系统的发展受益。

远程仪器控制与数据收集的价值取决于一定环境中接触到仪器和数据收集的难度。远程仪器控制需要可靠的网络、足够的数据传送速度来支持实时科学交流。通过网络来控制仪器,要求能捕捉到仪器的输出,并将其传送给操纵仪器的人,还要求网络能执行用户指令并将其作为输入传送给仪器。这一整套过程对标准化有较高的要求。目前空间科学领域正着手制定远程仪器控制的标准,这项工作有来自26个国家的空间科学机构的代表参加。

作为合作实验室的第一优先发展项目,美国科学基金会、国防部、航空航天局(NASA)在过去两年联合拨出巨款,分别在全美6所一流大学,试点筹建数字图书馆。这已成为科学界瞩目的大事。建立数字图书馆的目的,是开发新的计算机技术,将浩瀚的网上信息组成一个统一的信息资源。这个资源——数字图书馆——以Internet作为数据交换的工具,可以大大有助于每个领域里的科学家和学者的研究工作。在这6项研究试点中,各个学校的学科侧重点、试验用的信息载体(如数字录像、地图、期刊、音像资料等)各有不同,参加试验的合作单位也来自各行各业,形成一个前所未有的跨学科、跨行业、跨单位的大协作。

三、政策建议

笔者在去年作的一项调查中发现,中国科学家

(上接第28页)

有关农业持续发展问题,国外学术界近10年来就其定义、内涵和实现途径展开了广泛的讨论和探索。一些科学家很早就清醒地意识到,如果光是“坐而论道”,空谈“持续农业”的重要性而提不出有效对策和技术方案,来实际解决可持续性农业面临的经济、生态环境和社会效益三者之间的矛盾,则持续农业最终仍不过是一种“美好诗意”。因此已有一批科技人员和农民实践者全力以赴地研究相关

的研究,26%是单干,高出国际水平(13%)的一倍;45%是大学系内合作,也比国际水平(35%)高;其他形式的合作(校内系之间、国内不同单位之间及国际)均只及国际水平的一半。在科学高度发展的今天,越来越多的科学问题需要广泛的合作,尤其是学科之间的合作。中国目前致力于发展与Internet联接的设施,这对发展中国国内和国际的科学合作,促进科学发展,有着重要的意义。但是,仅有联网的能力并不是联网的最终目的。联网的最终目的是通过联网把用于科学交流的配套软件、网上信息联成一体,使中国科学家既能方便地接收世界最新信息,也能将他们的最新研究成果通过网络迅速地传播;既能与同一地区的科学家合作,也能与不同地区的科学家进行实时交流并合作,以便从根本上改变中国科学家从事科研的条件。在发展中国Internet联网能力的同时,应该着重考虑:

1. 调查研究科学家交流的方式、方法、工具及其存在的问题,弄清他们的需要,如中国科学家在正式交流过程中,引用的国外信息占有所有引用信息的80%左右,而中国科学家的研究成果只有10%左右进入世界科学信息交流的主要渠道。除了语言障碍之外,我们应该了解还有哪些因素造成了这种双向交流的不平衡。这对建立些什么样的合作实验室和数字图书馆是至关重要的基础工作。

2. 大力发展大学、研究机构图书馆的自动化,使现有科技信息资源能早日进入Internet,为建立中国的数字化图书馆和促进科学交流提供充足的物质基础。

3. 在开发信息技术、建立为科研服务的计算机网络的过程中,政府应鼓励和支持工商界、大专院校、研究所及政府部门之间的协作,使开发过程中所需的设备、人才、资金能在协作者之间得到满足,从而减少不必要的行政与管理上的耗费。

(责任编辑 王宏章)

的适用技术,并在病虫害综合防治(IPM)、持续农业信息系统和决策支持系统等方面取得了突破。探索和实现具有中国特色的可持续集约农业,难度更大。需要一大批有志的科技人员,针对我国农业面临的大量实际难题,努力创造适合中国国情的适用技术。相信在争取“农业科技率先跃居世界先进水平”的奋进进程中,中国式的可持续集约农业将扮演重要的角色。

(责任编辑 蔡德诚)