

加强科研领域数据库建设 实现数据资源共享

邓永智

(厦门大学海洋与环境学院, 福建 厦门 361005)

摘要 简述了建设数据库和实现数据资源共享的必要性, 对照国外数据库的发展趋势和特点, 结合我国数据库建设和数据资源共享的现状, 提出了几点我国数据库建设和数据共享的建议。

关键词 数据库 资源 共享

中图分类号 TP392

科学数据是经济建设、社会发展和科学进步所需的重要基础信息, 是一个国家甚至全人类的宝贵资源。科学数据的收集、整理、存储、开发应用是科学技术进步和社会可持续发展的重要举措。由于当今各学科领域各分支学科相互交叉、渗透和综合的发展趋势, 以及人类所面临的灾害、环境、资源等重大问题的挑战, 我国科学界迫切要求实现科学数据和资源的共享, 这就必须要加强数据库的建设。

1 国外数据库建设和数据资源共享发展概况

据统计, 1975 年世界上仅有十几个可共享的数据库, 到 1990 年就已发展到数千个。综观世界数据库近十年来的发展状况、趋势和特点, 可概括如下:

1.1 数据库数量和容量高速发展

大多数国家数据库数量增加都呈越来越快的态势, 库容量也呈激增的趋势。如美国数据库数目, 在 1981~1985 年间共增长了 24 倍, 使用人数增加了 6.5 倍; 日本在 1982 年注册数据库 604 个, 1989 年就增加到了 3 096 个, 增长 4.1 倍, 年均增长率超过 50%。库容量方面, 在 80 年代以前, 超过数百万篇数据的数据库极少, 而今仅日本和美国的数据库, 超过百万篇数据的就有近千家。目前, 这种势头有增无减, 估计至少要保持到下个世纪初期。

1.2 专业化和高度精细化

美国最早建立的数据库为政府机关数据库, 而目前存在的数据库几乎涉及社会的所有方面, 并且种类越分越细。例如, 在人才数据库中又独立出拔尖人才数据库, 在资源数据库中又分化出矿物资源数据库, 全美闻名的马里兰州实验动物数据库就是从动物数据库中单列出来的。

数据库不仅表现在专业细分上, 也表现在对数据采集的高技术化和数据加工的精细化上。例如, 对数据资料的加工精度从以整册图书、整篇文章为单位发展到按章节段落乃至具体数据为单位; 从文字加工发展到图形、图像、声频信息等多媒体信息的加工; 事实型数据库、数值数据库、多媒体信息数据库等的比重越来越大, 书目型数据库不再是主要内容。

1.3 网络化

数据库自存数据量和内容总是有限的, 这就使其服务质

量受到影响, 而解决的最好办法是谋求数据库联机并网使用, 因此出现了国内联机并网化和国际联机并网化的趋势。例如, 日本在 1986 年虽然国内只有 296 个数据库, 但通过与美国联机并网, 实际可使用的数据库达 1 483 个, 而美国 1984 年就拥有各种联机数据库 2 453 个, 900 家公司从事联机数据库服务业, 终端用户达 40 万左右, 遍布全美和欧亚地区。

1.4 标准化

由于非标准化限制了国内和国际间的数据库并网使用, 不利于运用先进的手段检索、信息融通和管理, 影响服务的效率与收益。现在各国新建的大型数据库, 一般都采用国际通用的标准。标准化是数据库发展的必然趋势, 锐不可当。

1.5 法制化

1967 年, 美国率先颁布了《信息自由法》, 对使用数据库和保密等问题作了具体规定; 瑞典于 1973 年颁布了《数据资料管理法》; 法国于 1978 年颁布了《信息科学归档文件卡片与自由法》; 日本也在 1986 年修改的《版权法》中将书刊作为知识产权列入其保护范围。迄今为止, 西欧和北美大多数国家, 以及新西兰、澳大利亚等国都陆续颁布了数据保护法。

此外, 随着计算机科学技术的迅猛发展, 数据库向人工智能化、家庭化等方向发展的动向也应该引起我们高度重视。

2 国内数据库建设和数据资源共享现状与问题

我国在 80 年代初以前, 一些科学研究领域, 例如地质、水文、气象、海洋等数据的共享程度还是很不错的。但是, 这些办法并未随着我国改革开放的发展不断完善, 反而由于“自谋增加收入”政策的误导和数据管理体制的法制建设不及时, 导致数据共享日益困难, 造成数据管理、积累和应用仍然处于十分落后的状态, 特别是还有许多的数据资料分散在部门或个人手中, 不但难以充分发挥作用, 而且还有散失的危险; 许多珍贵的历史资料保存年久, 没有复制备份, 处于损毁边缘; 同时国家也没有较完善的数据资料共享与服务的政策法规体系, 使盲目的资料数据壁垒现象得以继续, 以至于低水平的重复屡见不鲜。国家对科学数据无法系统积累和共享, 直接影响科技进步和国民经济建设。

造成上述问题的原因主要有三点: 其一, 长期以来人们对

科学数据库问题的重要性认识不足,数据库建设和数据资源共享的问题没有得到社会重视,数据库建设没有纳入国民经济建设和社会发展的轨道,此类公益事业无经费渠道,无法有组织有计划地进行数据库建设和资源共享政策法规的研究;其二,管理手段十分落后,如果说我国的数据源、数据产品以及数据整理方法方面都具有较高的质量的话,那么数据的管理仍然处在分散的落后管理状态,无“法”可依,很难对数据进行高层次的综合管理和充分开发应用;其三,不少国家重大科研项目立项时,没有明确科学数据管理考核目标,项目结束后无法形成统一管理和共享的数据系统。因此,只有提高对数据库建设的认识,通过国家支持建设数据库,才能从根本上解决这一方面的严重问题。与此同时,应该全面利用现代科学技术的最新成果,提高科学数据的获取、整理、存储、开发应用的现代化水平,特别是计算机网络技术和通信手段的现代化,从方法上解决数据共享的途径,使我国科学数据库真正实现科学管理和全社会共享,从而促进经济建设和科学发展。

应该说,通过多年来各部门的努力,国内已经打下了实现现代化网络条件下数据库建设和数据共享的一定基础和条件,再通过国家制订科学数据的规范标准和共享政策与法规,提高全社会对科学数据库建设及其共享服务的认识,完善科学数据库系统的建设,研究数据库管理的基本理论与方法,就能使科学数据(包括国际交换所获得的数据)充分发挥其应有的作用,使我国在科学数据库管理与共享领域的落后状况,得到迅速的改观。

3 当前应抓的主要工作

为逐步完善我国科学数据库建设,实现数据资源共享,当前首先要做好以下几方面工作。

(1)总结我国自己的经验,吸收国际的经验,研究制订我国科学数据库建设和数据资源共享的政策与法规。

(2)以数据标准化改造为龙头,使一系列准数据库机构发

展成标准数据库。我国目前标准化数据库虽然很少,但准数据库,既有储存数据、资料的任务和职能,而又不符合数据库规范的机构很多。对于这类机构,只要进行标准化改造,实现规范化的收集、加工、储存、检索过程,就能较省力地把成千上万的准数据库改造成为标准化数据库。

(3)以效益为核心,推动数据库的自身改造和发展。我国数据库建设必须避免单纯追求数量的外延式发展,从一开始就紧紧围绕“结构优化、布局合理”这一轴心,一方面对现有数据库的数量、布局、数据性质和容量进行摸底、登记,进行有计划、有步骤地挖潜、调整、改造和提高,使原有数据库上水平、上等级。根据数据库数据资料的客观真实性和不断吐故纳新的要求,及时调整和补充数据,不断扩大存储量,保持数据的真实性、及时性和可靠性;另一方面,尽快制订发展数据库的产业政策,指导和调整数据库建设,在规模、布局、质量及数据结构上把好审批、验收和监督关。此外,一支训练有素,具有相当业务水平且决心投入数据库建设的专职队伍,是我国数据库建设效益型道路的根本保证,数据库生产和销售部门必须全力抓好后继人才培养和在职人员培训,努力提高管理水平,确保数据库生产和服务走上高效、有序良性循环轨道。

(4)以横向联合为纽带,推进数据库生产服务向集团化发展。通过横向联合,带动政府、企业及各学术团体对数据库建设的资金、技术、智力投入,拓展数据库信息资源,沟通数据库产销渠道,改善数据库服务条件,提高数据库服务效率和质量,有利于充分发挥数据库资源共享的特点和提高数据库系统的整体能力。当然,这种横向联合也应不断向国际间联合发展,争取我国有更多的数据库进入国际信息循环系统,为数据库应用开辟更加广阔的前景。

实现我国科学数据共享的目标不是一蹴而就的,要逐步发展完善,但也不宜“久战不下”,应以一个五年计划为限。

(责任编辑 张春强)

收稿日期:1999-01-26

· 科技进步文摘 ·

北京调查领导干部科技素养

为贯彻落实中央有关提高干部科技素质的指示精神,把领导干部的培训工作推向新高度,中共北京市委组织部、中共北京市委党校和北京市科协,联合对北京市领导干部的科技素养进行调查,并于今年2月公布了调查报告。

调查表明,北京市领导干部中高达86.7%的人充分肯定了科技对我国生产力的发展、综合国力的提高和人民生活水平的改善所起的巨大推动作用。他们最希望科技解决的问题依次是环境保护,疑难病症的医治和预报自然灾害;认为北京应优先发展的科学技术排在前三位的次是人口健康与环境保护、电子信息技术和农业与食品技术。此次调查表明,北京市领导干部科技常识的了解程度较好。从综合得分情况看,以39分为满分(每一个正确答案计一分),这部分的平均分为23.23分。

调查结果显示,领导干部对科技政策的了解和掌握程度

普遍偏低。

本次调查显示,北京市领导干部获取科技知识的渠道和方式比较单一。他们中的绝大多数是依靠书报和广播电视学习科技知识的,通过公共设施和场馆以及参加科普活动等有效而生动的教育形式获得科技知识的比例较低。调查结果还显示,领导干部在获取科技知识时呈现较大的随意性。电脑和网络,是当今世界获取各种知识和信息最快最有效的途径,但是,他们中能熟练使用电脑的人数并不多。

调查报告在“相关建议”部分中指出,在今后的干部科技知识普及教育中,加强科技政策教育具有极其重要的意义。要帮助领导干部从战略高度认识科学技术是第一生产力和科教兴国战略,增强领导干部和科学意义和对科学技术的理解能力。建议组织部门和科普机构要针对这一科普工作中特殊而重要的群体,专门制定相应的科普计划。

(摘自《科技日报》1999-03-25)

● 焦点话题——国家创新体系的构建 ● 13