

从世界数据库发展趋势看我国数据库的建设

Developmental Problems in China's Data Bank and Their Solutions Seen from the Developmental Tendency of the World's Data Bank

郑耀东 牧 心

(南开大学)

众所周知,信息是国家的重要战略资源,是社会、经济和科学技术发展的基础。数据库是电子化信息系统建设的核心。其发展水平越来越被看成是各国经济实力、文明程度和科技水平的重要标志之一。进入 70 年代中期以后,数据库发展势如雨后春笋。据统计 1975 年,世界上仅有数据库数十个,到 1990 年就已发展到数千个。我国数据库建设是自 1987 年国家信息中心成立以后正式开始的,到 1991 年初,已建成中文数据库 400 多种,其中文献数据库总容量超过 100 万篇。未来 10 年,将是世界数据库发展的盛潮时期,也是我国经济腾飞的关键时期。因此,密切注视世界数据库发展的新动向,积极吸收国内外先进经验和抓住时机加快我国数据库建设步伐,是具有重要战略意义的实践问题和理论问题。

一、世界数据库发展趋势和特点

纵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. 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。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%。目前,这种势头仍有增无减,估计至少要保持到下个世纪初期。

2. 库容量激增。在 80 年代以前,超过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。日本产业银行金融数据库,其数据存量在 1984 年就高达 1 280 万篇;美国国家图书馆数据库,每年收存资料 1 801 万篇,其数据量以万亿计。

3. 专业化。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单列出来的。

4. 社会化。数据库产品的内容日益社会化,是信息与经济相结合的产物,是与数据库的产业化特点紧密相关的。数据库从最初的科技文献型产品扩展到多品种体系,其服务范围从科技为主转移到科技、经济、管理、市场、娱乐等领域。据《Online Review》的统计,至 1982 年,科技数据库占全部数据库的一半以上,但是近年来,经济商业和其它社会科学数据库的发展突飞猛进,特别是市场、金融、保险、产业、出版等数据库大量涌现。在日本各类数据库中,商业类数据库所占比例最高,为 41.4%,科学技术类数据库占 29.7%。1989 年日本商业数据库又比上年增加了 9.4%,使用率增加了 40.8%。

5. 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。

6. 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。现在与初创时期不同,目前各国新建的大型数据库,一般都采用国际通用标准。标准化是数据库发展的必然趋势,锐不可挡。

7. 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》；英国于1984年颁布了《数据保护法》；日本也在1986年修改的《版权法》中将数据库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。

8. 产业化。由政府部门或学会、协会等非盈利机构生产、所有或资助建造的数据库利用形态的结束，标志着数据库商品化，进而实现数据库产业化的时机已经成熟。一些发达国家积极按照商品经济原则实行数据有偿服务、自主经营，并仿照其他产业模式对数据库行业进行生产、加工、流通、批发、零售的产业化改组，使得数据库生产和服务的产值也越来越大。1986年全世界数据库产值为65亿美元，比1984年增长25%，1987年仅美国的数据库产值就达到27亿美元。更值得一提的是，1989年对326家主要公司统计，日本数据库产业年销售额为1576.2亿日元，比上年增长50%，并计划到2000年达到2.5万亿日元。

9. 高度精细化。数据库的发展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深化上。具体而言：(1)在利用Landsat卫星收集资料数据的基础上，美国航空和航天局又在准备发射一批新一代搜集资料卫星。(2)对数据资料的加工精度从以整册图书、整篇书刊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。

此外，随着计算机科学技术的进步，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。

二、我国数据库建设的现状和问题

进入80年代以来，我国数据库建设已迈出了可喜的一步。目前，文献型、事实型、数值型等各种类型的数据库都已开始服务，其中在北京建成了世界最大的“自然资料数据库”，另外一些数据库还达到了国际水平，并在经济建设中提供了有效的信息服务，取得了较好的经济效益和社会效益。但是也应该看到，我国数据库建设中存在的问题不少，主要是：

1. 大型数据库数量少，不能满足社会发展的需要

我国国家级数据库只有“中国产品信息年鉴数据库”、“中国科学数据库”、“国家法规数据库”、“旅游资源数据库”、“台湾数据库”、“东欧十国数据

库”、“中国企业、公司及产品数据库”、“国家宏观经济数据库”、“国际收支数据库”、“固定宏观经济数据库”等10余个已经建成的大型数据库。这种状况与世界数据库发展速度和水平很不协调，也不能满足我国四化建设的需要。

2. 小型数据库标准化管理差，发展后劲不足

由地方或部门建立的一些小型数据库，普遍存在储存数据和资料方式不规范、不统一的问题。而这类数据库由于缺乏标准化管理，难于运用计算机等现代化手段管理和服务，并且也难于形成各数据库之间、地方数据库与中央数据库之间的联机并网使用，影响了小型数据库的发展壮大。

3. 社会缺乏发展和使用数据库意识

由于对数据库作用估计不足，关于数据库的有偿服务、收费标准、发展措施、鼓励性政策、发展战略规划等很少有人专门研究分析和统筹安排。大众对使用数据库更缺乏认识，仍习惯于传统的自我服务方式和碰运气求得所需数据、资料；有些领导者守着一大堆现成材料，决策却依然依靠老经验和想当然，有些发展规划确定或项目可行性研究，不是求助于有关部门或利用数据、资料条件，而是找几位各方面的名家走马观花地考察一番便作结论，致使现有的数据库运行不充分，建成后使用者寥寥。

4. 缺乏法律保护

我国目前还没有建立专门的数据保护法，各种垄断、侵权、泄密事件等时有发生，既增加了法律部门受理和断处的困难，又制约了数据库的建设和发展。

5. 数据结构与需求失衡

数据库建设和发展中，数据结构失衡问题日益严重。具体而言，“七·五”期间生产的数据库产品与社会需求不协调，结构不对称；数据库之间存在较严重的重复现象；大型、中型和微型数据库之间的布局 and 比例，国内建库数与引进数据库及联机终端之间的布局 and 比例，科技、法律、经济等各专业类数据库之间的布局 and 比例，以及各类数据库内部文献数据库、数值数据库和事实型数据库的结构比例都存在不同程度的失衡问题。这种重建轻管理、重数量轻结构的做法，已经使一些数据库陷入“死库”境地，根本发挥不了应有的效益。

6. 体制改革滞后

我国数据库建设受以行政为主的条块式管理方式的影响，出现了机构重迭，职能混同，数据库系统与管理系统互不协调，工作体制、运行机制、服务方式等“软建设”与信息技术采用的“硬建设”不相匹配等方面，给数据库建设事业带来了一定的困难和阻力。

7. 投入严重不足

数据库建设的资金来源主要依赖于政府的拨

款,在国内资金普遍不足的情况下,对数据库建设投资不景气是显然的。因为,一方面国家用于整个信息产业的投資尚不足 GNP 的 1%,用于数据库建设的投資就更少;另一方面,我国采取产销合一的建库方式使许多数据库的经营者都逃避不了亏本的厄运,增加了数据库建设的投资风险。

8. 专业化、社会化和产业化进程缓慢

表现在:数据库主要集中在中科院和图书情报部门,公司企业尚属空白;长期强调计划经济,无偿服务,而忽视了市场对数据库发展的调节作用,缺乏适宜的竞争环境和强烈的竞争意识;数据库生产单位都是自产自销,自己开展服务,摆脱不了“小而全”模式的规模不经济,如此等等,都使数据库建设和发展举步维艰。

此外,我国数据库行业人员年轻力壮的向外企转移,剩余者老龄化,面临退休,处于后继无人的局面。宏观调控不力也是数据库建设中不可忽视的重要问题,需要进一步在实践中寻找解决这些问题的对策。

三、我国数据库建设的对策选择

总的来说,我国数据库发展正处在起步阶段,要使其进入蓬勃发展的广阔天地,必须紧扣世界数据库产业发展的“脉搏”,根据数据库发展的一般特点要求和我国数据库建设的条件及存在的问题,选择发展对策。

1. 以知识和信息商品化发展为契机,带动数据库的迅速发展

我国最近几年知识和信息服务商品化发展的速度很快,其中特别是技术商品化更快。自 1985 年以来,技术成果交易额平均每年以超过 30% 的速度递增。数据库发展走商品化道路,既能为自身发展准备强大的物质和技术基础,又能把自身纳入到整个知识和信息服务商品化行列。从国外情况看,数据库服务基本是有偿化,全世界数据库日新月异发展的同时营业额也在与日俱增。所以,我们应该抓住目前知识和信息服务商品化迅速发展的有利机会,由国家包下,无偿服务方式,通过改造,逐步过渡到有偿服务和自我发展的道路。

2. 以政策鼓励为动力,推动数据库迅速发展

发展数据库也必须实行国家、集体和个体一齐上的原则,建立各种形式和不同专业内容的数据库。要做到这一点,国家除了兴建一批重点数据库外,还应当对非重点和地方、部门兴建数据库实行鼓励性政策,包括在投资、税收、信贷等方面提供方便或优惠条件,对有关人员的工资、职称、荣誉称号等方面给以具体指导和管理。通过鼓励性政策,激发各方面建设数据库的积极性,起到加快建设步伐

作用。

3. 以标准化改造为龙头,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。

4. 努力扩大数据库服务的业务

目前应力求数据库服务业务类型有重大突破,以引起社会对发展、使用数据库给以广泛和高度重视。充分利用数据库特有的优势,在现有内容的基础上,扩大服务的业务种类。例如,开办咨询服务业务,向社会提供力所能及的咨询服务,以满足社会对咨询服务发展的迫切要求;增添为科学研究需要深加工服务业务,象为科研单位或个人提供专项数据、资料,或者把有关数据、资料整理加工到要求程度等;开展项目可行性论证业务,由数据库专业技术人员根据具体部门提供的项目和提出的要求给以技术、经济等方面的论证;附设为个人或单位代存数据和资料业务。有些数据和资料,由于价值性、保密性、经济性方面的原因,个人或单位可能更愿由数据库代为存管,同时,对关于数据库也十分有利。通过扩大业务种类,既能增加收益,又能扩大影响,是一举两得的好办法。另外,对关于数据库建设和发展及其信息服务的舆论宣传,必须给予高度重视,才能唤起全社会的关注,才能为信息服务大开方便之门。

5. 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要求客观真实性和不断吐故纳新的要求,及时调整和补充数据,不断扩大存储容量,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,是我国数据库建设走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、有序的良性循环轨道。

(下转 20 页)

贸易洽谈,随后旅游一程,在我们离罗马返国前签署了双边贸易意向协议书。这种集科技交流、经贸活动、游山玩水为一体的安排,就是来自对方的一种多重目标行为。

四、旅游业可为科技经济发展带来更多的机会

综上所述,旅游业确能为国家科技、经济发展带来不同层面的好处。但因我国旅游业发展时间不长,社会服务功能还未充分发挥,故继续开发旅游软、硬件,加强管理和完善法规、制度势在必行。为了使旅游业在其运行中更好地体现为经济建设中心服务的方针,我们在这里仅就它与科技、经济工作有关的部分,提出对其发展的几点建议:

1. 为使旅游业能更好地发挥大众传播媒介的作用,今后的组织、宣传和导游工作应更加充实,不仅是景点、文化遗产、民情习俗、饮食购物等方面的介绍,还应加强对我国现代化建设成就、改革开放成果和各地区、部门科技经济发展情况的介绍。这样,可以发挥“大窗口”的作用,使外国人不仅了解中国的昨天,而且了解中国的今天,还能展望中国的明天,从而增强与中国发展科技经济往来的吸引力。

2. 在旅游服务工作中要增强导向意识,要运用自身信息灵的优势,抓住时机,及时帮助中外旅游者之间、旅游者与旅游地区之间疏通科技、经济联系的渠道。在宾馆中、在游览中、在旅游购物中都存在着这样的大量信息和机会。有心境、有心计的热

心服务者过去已经有过这种疏通联系渠道的行为,起过导向作用,今后还应积极提倡。自然,不是说要让旅游业去做份外的科技、经济方面的工作,那样就舍本逐末了。导游者的主体行为应是导游,这是毫无疑问的。但是,如果在完成主体行为的同时,辅之以牵线搭桥的导向工作,似乎也在情理之中。要做到这一步,一是旅游部门应加强与旅游地区科技、经济部门的联系;二是要提高旅游服务人员的素质,使他们有一定的科技、经济知识和导向意识,从而有利于多方位服务;三是从事科技经济开发的科技工贸人员,乃至参谋决策人员也要大大增强利用旅游业,渗入旅游业的自我意识,这也是当今世界发展大交叉大交融趋势中的又一新兴大趋势。

3. 旅游业应进一步发挥其乘数效应的优势。据旅游部门介绍,旅游业对世界经济的贡献,不仅是它本身提供的产值和就业岗位,它还通过从其它产业采购原料,通过利润消费、股票分红和员工收入等方面的贡献,刺激其它各门各类产业的发展。这种现象被称为乘数效应,即它以投资的乘数为基数,相应地引起收入、消费、税金、就业等各类乘数效应的发生。一般说来,任何一种投资都会产生乘数效应,但旅游业的乘数效应远高于平均乘数效应。1990年世界旅游业对所需原料和服务的采购金额为15 000亿美元,这使得全世界经济总产值增长73 000亿美元,并总的提供了2.95亿个就业岗位,这是一个了不起的数字。我国旅游业应借鉴国外的经验,不断在提高乘数效应上下功夫,以便为国家科技、经济发展作出更大的贡献。

(责任编辑 冷远猷)

(上接第38页)

6. 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

整体功能。在这方面,武汉市和国内近30家图书情报单位联合发起组建的“武汉科技情报数据库联合推广应用中心”不失为一个成功的尝试。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。

(责任编辑 宋义荣)