

中美商情数据库建设比较研究¹⁾

杨奕虹

(万方数据股份有限公司, 北京 100044)

武夷山

(中国科学技术信息研究所, 北京 100038)

摘要 美国是商情数据库建设起步最早、发展最快的国家, 因此研究美国商情数据库建设及其发展对加速我国商情数据库的建设有着实际的意义。本文通过对中美两国商情数据库发展概况的比较研究, 发现我国商情数据库建设中存在的问题, 并提出几点建议。

关键词 数据库 商情数据库 商情信息

Comparative Study on the Chinese and American Business Database Construction

Yang Yihong

(Wanfang Data Co., Ltd., Beijing 100044)

Wu Yishan

(Institute of Scientific and Technical Information of China, Beijing 100038)

Abstract America is the earliest country that started Business Database(BDS) construction and related data services. Today America is still leading the world with the fastest database development. Therefore, studying the BDS construction of America and its development trends would shed light on how to speed up BDS construction in China. Based on the comparison of the development course of business databases in USA and in China, the author gave recommendations on how to improve construction of BDS in China.

Keywords databases, business databases, commercial information.

有可能少走弯路, 取得较快速的发展。

1 引言

美国是世界上数据库业起步最早、发展最快的国家^[1], 其工业产品及商业经济的信息服务已有百年历史^[2]。尤其是近10年随着信息资源建设与网络发展的步骤加快, 他们在信息技术应用和信息管理方面均已获得极大的成功, 并积累了丰富的经验。如果能恰当吸取美国信息资源建设的经验, 我们就

2 商情数据库基本概念

商情数据库是指那些能提供与国际国内商务活动有密切联系的各类信息的数据库, 即有关公司、产品、市场行情、商业动态、金融活动、专利、技术标准及有直接关联的税法、国家政策等方面的信息数据库^[3]。

收稿日期: 2002年5月10日

作者简介: 杨奕虹, 女, 副研究馆员, 中国科技信息研究所情报学硕士。武夷山, 男, 研究员, 研究方向为情报研究与科技政策研究。

1) 国家自然科学基金资助课题: 70073007

3 美国商情数据库发展概况

3.1 美国商情数据库数量

美国不仅是世界上数据库建设及服务起步最早的国家,而且是发展速度最快、建库数量最多的国家,商情数据库目前已成为美国数据库生产商和信息服务系统的主流产品。根据 2001 年“Gale Directory of Databases” Online 版统计,2000 年全世界拥有数据库 11604 个,美国生产的数据库产品占 61%,其中商情类数据库占最大比例,达到 30%^[4]。根据世界著名联机数据库服务商 Dialog 网站所提供的数据库^[6]列表统计,该系统已有数据库 632 个,拥有文献记录 6 亿多条,114 万余用户遍及 120 多个国家和地区,其中商情数据库 351 个,占 56%。在 DIALOG 联机检索系统中一级分类有 8 个大类,且每个数据库根据内容可以分到多个不同的类目下,因此 DIALOG 中的 632 个数据库共涉及 1298 个类目。具体分布见表 1。

表 1 DIALOG 一级类目数据库数量统计

一级 类目	商情 信息	政府	知识 产权	医药	新闻	参考 工具	科学 技术	社科 人文	其他	类目 数
数据库数	351	100	82	125	158	96	270	54	62	1298
百分比 (%)	27	7.7	6.3	9.6	12.2	7.4	20.8	4.2	4.8	100

数据来源:根据 <http://www.Dialogweb.com/> 数据库列表自行统计

3.2 美国商情数据库建库单位类型与商业化程度

美国是商品经济高度发达的国家,大财团、大企业在经济、社会生活中起着支配作用。20 世纪 70 年代,随着美国政府大大削减对二次信息产品的无偿资助后,数据库生产者为了生存,及时采用了商业化的经营方式,数据库生产与经营逐渐向商业界发展。因而其产品与服务均围绕用户需求运作,以有偿服务为主,以盈利为目的。通过表 2 可以看到,商业性数据库生产商的比例直线上升,从 1977 年的 22% 增加到 2000 年的 82%,而且有逐年增长的趋势。由于在表 2 中的国际数据库生产商中美国占 63%^[4],故此表也能反映美国的情况。

3.3 美国商情数据库服务方式

美国商情数据库通常采用“同一种数据库,同时

提供多种载体服务”的方式,这一特点可从以下三个美国著名商情信息服务商所采用的商情信息服务方式看到。

表 2 国际数据库生产商类型统计

生产者类型	1977 %	1985 %	1990 %	1995 %	1997 %	1999 %	2000 %
政府	56	21	17	13	12	9	8
商业性	22	57	68	77	78	81	82
学术机构	22	11	12	8	8	8	8
机构联合	N/A	11	3	2	2	2	2
总计	100	100	100	100	100	100	100

数据来源:《Gale Directory of Databases》2001 版

美国托马斯出版公司(Thomas Publishing Co)是一家专门为供应商和采购商提供工业产品采购指南的商情信息服务商。该公司已有 100 多年的历史。其最著名的产品《托马斯美国制造商企业名录》(Thomas Register of American Manufacturers)从 1905 年开始至今,每年更新出版一次书本式目录;到 1993 年开始每年同时提供 CD-ROM 版本;从 1995 年开始,在每年提供书本、CD-ROM 的同时免费提供 Internet 服务。在托马斯自身全方位提供信息服务的同时,还每季度给 DIALOG 系统提供一套《托马斯美国制造商企业名录》,由 DIALOG 在 535 文档提供联机检索服务,并出版光盘销往全世界^[2]。

美国盖尔公司(GALE GROUP)是一家具有世界领先地位的互联网信息出版商,是专门从事数据库生产的专业公司。目前盖尔公司自身可提供的数据库达 600 多个,以 7 种不同形式出版,形成 1282 种产品的体系。同时盖尔公司有 25 个数据库提供给 DIALOG 系统,其中进入一级类目商情信息(BUSINESS)类目的数据库有 19 个,占盖尔数据库提供给 DIALOG 系统数据库总数的 76%。从表 3 可以看出,盖尔公司不仅自有版权的数据库数量多,而且其数据库的载体出版形式也是种类繁多。还有一点令人注意的是盖尔公司书本印刷的产品在其产品体系中占近 50% 的比例,可见虽然电子信息迅猛发展,但印刷书本这一载体并没有退出历史舞台^[7]。

美国情报处理服务公司(Information Handling Service Inc)成立于 1959 年,最早以缩微胶卷的方式专门为航天工程研究人员提供产品目录数据库,目前可提供美国工业产品样本资料、技术规格与标准、

政府出版物等,还可以提供英国、加拿大、日本的产品目录服务^[8]。在 1988 年该公司缩微胶卷的生产达到 3 百万英尺,成为世界上最大的缩微制品生产商。它从 1985 年开始提供数据库的 CD-ROM 版本,1995 年提供 Internet 版本。现在 IHS 拥有 2.5 万个生产厂家提供的近 300 万页的各种产品样本,成为世界上最完整的工程技术信息提供商。

表 3 盖尔数据库产品体系统计表

产品类型	数据库数量(种)	百分比(%)	产品类型	数据库数量(种)	百分比(%)
在线查询	54	4.2	缩微胶片	57	4.5
打印、书籍出版	618	48.2	报纸信札	3	0.2
CD 光盘	91	7.1	记录/丛书	240	18.7
缩微胶卷(磁带)	219	17.1	总计	1282	100

数据来源: <http://www.gale.com/>

从以上介绍可以看到,美国商情数据库通常以印刷、光盘、网络三种载体提供给用户供其选择,有时还提供其他更多的载体。

4 中国商情数据库发展概况

4.1 中国商情数据库数量

中国商情数据库的发展始于 20 世纪 80 年代初。从 80 年代末 90 年代初到 90 年代中期,商情数据库的数量和类型随着数据库的发展大为增加。1990 年我国数据库总数为 655 个,经贸类占 5.49%;1991 年数据库总数为 805 个,经贸类占 16%;1995 年数据库总数为 1038 个,经贸类占

28.61%^[1]。目前国内企业仅仅数据库类的专项信息支出已从 1998 年的 9000 万元增长到 2000 年的 7.84 亿元。根据 IDC 的研究报告,目前该市场规模仍以 43% 的速度递增^[11]。从互联网上“中国数据库”网站^[12]提供链接服务的 254 个数据库内容分类可以看到(见表 4),商情信息数据库已名列第一,占 41%,比科技类数据库高出 10%,表明我国商情数据库已呈现良好的发展趋势。

4.2 中国商情数据库建库单位类型与商业化程度

为了弄清我国数据库建库单位类型与商业化程度,笔者选取了《中国数据库大全》^[14]的 574 种数据库作为研究样本。《中国数据库大全》是 1995 年国家计委、国家科委(现科学技术部)和国家信息中心等有关单位对全国数据库及电子信息网络资源的普查结果,是全面了解我国数据库发展的权威报告。虽然已过去 7 年的时间,但目前国家还没有做过更新更全面的统计。为了解我国数据库在这 7 年的发展变化,笔者从 1038 种数据库中挑选了有关公司名录、产品信息、市场行情、经济统计等属于商情类信息的数据库 574 种,通过电话、传真的联系方式对选定的数据库进行更新状况、市场服务以及资金来源等项目的调查了解,并对结果进行了统计分析。

(1)建库单位类型

从表 5 可以看到,我国商情数据库的建设主体,主要集中在信息所、专业信息中心和图书馆,约占 60%—70%;第二大部分集中在研究院所、政府机构、协会。

(2)调查结果分析

根据《中国数据库大全》上提供的联系方式与

表 4 中国数据库网站数据库分类统计

分类	商情信息	科学技术	医药	新闻	参考工具	知识产权	社科人文	总计
数量	104	77	26	22	17	6	2	254
百分比	40.94%	30.31%	10.24%	8.66%	6.69%	2.36%	0.79%	100.00%

数据来源:根据中国数据库网站提供的数据库名称自行分类统计

表 5 《中国数据库大全》商情类数据库建库单位类型统计

单位类型	专业信息中心	信息所	政府	专业公司	专业研究院所	信息公司	媒体	银行	图书馆	其他	总计
数量	267	105	82	38	31	16	10	6	1	18	574
百分比(%)	46.5	18.3	14.3	6.6	5.4	2.8	1.7	1	0.2	3.1	100

数据来源:根据《中国数据库大全》提供的资料自行统计

574 个数据库的建库单位进行了试联系,能联系上(打通电话或传真有答复)的数据库有 318 个,占总数的 55.4%。具体调查统计结果见表 6。

表 6 中国商情数据库建库单位类型
与商业化程度调查结果统计

调查内容	数量	占调查对象数 574 的百分比
建库单位仍能联系上	318	55.40%
数据库在更新	147	25.61%
数据库对外服务	98	17.07%
政府提供资金	48	8.36%
政府加自筹资金	31	5.40%
自筹资金	19	3.31%

数据来源:根据《中国数据库大全》提供的联系方式进行调查后自行统计

以上调查表明,574 个数据库中仅有 3.3% 的数据库已走入“以库养库”、以市场为导向的数据库建设道路。这说明,没有政府资金注入的信息公司其生存状况要比有政府支持的建库单位更加艰难,但从美国数据库业发展的历史进程来看,数据库由信息公司承建并走向市场这是发展趋势。不管怎样,我国商情数据库中现在已经有 3.3% 走向了竞争激烈的市场,这是我国商情数据库商业化程度的标志,是我国商情数据库发展的希望,也是我国数据库产业市场运作的起点。

4.3 中国商情数据库服务方式

为了了解我国数据库服务的载体形式,选取了

四种来源的数据作为研究样本。(1)《中国数据库大全》574 个商情类数据库中能继续更新且可对外服务的 98 种数据库。(2)近 2 年内国内以印刷或光盘出版、其准确率在 80% 以上的 93 种商情信息资料包括企业名录、产品大全、行业年鉴等。(3)“万方数据资源系统”^[13]提供的 120 个数据库中能够提供更新版本的 88 种数据库。(4)“中国数据库”网站提供链接服务 254 个数据库中可以继续提供服务的 173 个数据库。

通过对以上 4 种来源的数据库服务载体形式的统计分析(详见表 7),可以看到,我国数据库服务的载体形式比较单一。

(1)可同时提供印刷版、光盘版和网络版三种载体的数据库数量屈指可数,最多不超过 5.8%。

(2)以出版书本为主的商情信息同时上网服务的比例较大(占同类信息数量的 50.5%),但同时出版光盘的比例则小得多(仅占 18%)。

(3)以上网服务为主的数据库同时提供光盘服务的比例远远高于同时提供书本印刷的比例,高出 10%—20%。

总之,普遍存在的一种现象是,提供光盘和网络服务的数据库不提供印刷版本;能提供印刷版本同时提供上网服务的名录数量大于可同时出版光盘的名录数量。应该承认,在我国目前信息资源仍然十分缺乏。计算机及网络的普及率尚没有达到十分优越的条件下,多种载体形式的出版更有利于信息内容的服务与传播。而且,多种载体的存在有助于用户的选择,可更好地满足用户的需求。

表 7 数据库服务载体方式统计

来源	数据库 总数	仅有书		仅有光盘		仅有网络		书、光盘		书、网络		光盘、网络		书、光盘、网络	
		数量	比例 %	数量	比例 %	数量	比例 %	数量	比例 %	数量	比例 %	数量	比例 %	数量	比例 %
中国数据库 大全	98	13	13.3	30	30.6	37	37.8	1	1	3	3.1	14	14.3	0	0
可购买的 商情名录	93	38	40.9	1	1.1	0	0	1	1	47	50.5	2	2.1	4	4.3
万方资源系 统网站提供	88	0	0	0	0	40	45.4	0	0	10	11.4	34	38.6	4	4.5
中国数据库 网站的链接	173	0	0	0	0	121	69.9	0	0	3	1.7	39	22.5	10	5.8

数据来源:根据调查结果自行统计

5 中美商情数据库总体建设比较结果

通过上述分析比较可以看出,中国商情数据库建设与美国相比存在以下几方面差距:

1)与美国商情数据库数量相比,我国商情数据库的建设起步晚,发展速度比较慢。

美国 1989 年到 1998 年 10 年的时间里,数据库从 330 个增加到 4000 个,增加了 12 倍;中国 1990 年到 1995 年 6 年的时间里,数据库从 655 个增加到 1038 个^[5],增加 1.6 倍。一般来说,商情数据库数量和类型是伴随着数据库总量的发展而不断增长的。从商情数据库建设来看,1995 年中国商情数据库数量有 297 种,而 1995 年美国商情数据库就已达 1934 种^[10],这说明中国商情数据库的数量与美国相比差距非常大。

(2)与美国商情数据库建库单位的性质及商业化程度相比,我国商情数据库生产商所生产的数据库数量偏低,商业化程度不高。

美国在数据库发展初期,政府支持力度较大,但随着数据库市场逐渐成熟,政府支持的数据库数量逐年减少,从 1997 年占 56% 平稳减少到 2000 年占 8% 的水平,目前建库主体已是商业性数据库的生产商,所生产的数据库比例已由 1977 年的 22% 增加到 2000 年的 82%。中国目前数据库建库主体主要集中在信息所和专业信息中心(占 70%),而且建库所需资金主要依靠政府(占 89%)。

3)与美国商情数据库服务方式相比,我国数据库载体种类相对单一。

尽管美国信息产业如此发达,计算机拥有率及网络速度和普及率比较高,但同一种商情信息内容一般可同时提供印刷、光盘、网络三种载体,甚至还提供缩微胶卷或缩微平片。但我国数据库载体种类则相对单一,能同时提供印刷、光盘和网络三种载体服务的数据库屈指可数,几乎都不提供缩微制品。

6 中国商情数据库发展建议

商情数据库的建设是一个投资大、工期长、任务重的长久工程。中国商情数据库的发展,需要信息企业自身的奋斗,更需要政府强有力的政策干预和扶植,需要一个有利于信息服务业发展的宏观规划。为此,参照上述比较结果,对我国商情数据库建设提出以下 3 点建议。

1)制定宏观发展规划,建立全国数据库发展中心

国家应制定出一个推进商情数据库建设发展的宏观发展规划和整体发展战略,建立一个国家数据库发展中心,制定数据库行业标准、质量认证并推动数据库标准化工作,规范行业行为^[15]。从建库论证、建库步骤设计、建库标准制定到数据收集、文献标引、著录、数据录入等都由中心负责质量监控。

2)加大政府投资力度,制定优惠政策

在现阶段中国商情数据库建设仍然需要政府加大投资力度,制订优惠政策。目前,我国商情数据库的建设主要是由传统的信息服务机构来承担。由于经费来源和信息来源的局限性,新技术应用相对落后,很难制作出数据量大、涉及范围广、内容详实、更新及时、查询便捷的大型综合性商情数据库。要达到目前国际先进商情数据库的规模和水准,政府的政策和资金支持是必不可少的。

3)中国商情数据库产品载体多样化建设

现阶段我国计算机普及率和网络环境有了突飞猛进的发展,但由于发展的不均衡及网络速度的限制和不稳定,同一种数据库提供多种载体形式也是现阶段提高我国数据库使用效率的有效方式之一。

参 考 文 献

- 1 冷伏海等.中国商情数据库发展现状及对策研究.图书情报工作,2000(1):70~73
- 2 Thomas Publishing Company Celebrates 100 years of Bringing Buyers and Sellers Together, <http://www.thomaspublishing.com/press/1998/100release.html>
- 3 项艳.我国商情数据库建设的几点设想.国外情报科学,1997,(1),43~47
- 4 Williams M E. The State of Databases Today: 2001. In: Gale Directory of Databases. Detroit: Gale Co., 2001
- 5 朱爱群.中美商情数据库比较.情报杂志,1998,(1),23~25
- 6 <http://www.dialogweb.com>
- 7 <http://www.gale.com>
- 8 <http://www.ihs.com>
- 9 蔡庆芳.国外资源共享系统对我国信息资源建设的启示.情报杂志,2000,(6),7~8
- 10 中国科技信息研究所等.我国科技电子信息资源的开发和利用研究.国家软科学研究项目,1998
- 11 植万禄.新华社欲做信息服务业盟主.北京青年报,2001.7.23 第 19 版
- 12 <http://www.chinadata.com.cn>
- 13 <http://www.wanfangdata.com.cn>

-
- 14 国家计划委员会.中国数据库大全.北京:中国计划出版社,1996 2000,(6),408~411
- 15 孔桃.国内外商情数据库发展之比较.情报理论与实践,

(责任编辑 马兰)
