



数字资源长期保存研究

SHUZI ZIYUAN CHANGQI
BAOCUN YANJIU

谢水宪 著



中国出版集团



世界图书出版公司

本书的出版得到以下项目的资助：

北京市教育委员会项目“数字资源长期保存策略

北京联合大学校级重点建设学科

北京联合大学校级骨干专业

北京联合大学校级优秀教学团队培育项目

数字资源长期保存研究

谢永宪 著

世界图书出版公司

上海·西安·北京·广州

图书在版编目(CIP)数据

数字资源长期保存研究/谢永宪著. —上海: 上海世界图书出版公司, 2011. 11

ISBN 978-7-5100-3843-3

I. ①数… II. ①谢… III. ①数字技术—信息管理—研究 IV. ①G203

中国版本图书馆 CIP 数据核字(2011)第 163800 号

数字资源长期保存研究

谢永宪 著

上海世界图书出版公司出版发行

上海市广中路 88 号

邮政编码 200083

南京展望文化发展有限公司排版

上海市印刷七厂有限公司印刷

如发现印刷质量问题, 请与印刷厂联系

(质检科电话: 021-59110729)

各地新华书店经销

开本: 890×1240 1/32 印张: 5.25 字数: 150 000

2011 年 11 月第 1 版 2011 年 11 月第 1 次印刷

ISBN 978-7-5100-3843-3/G·290

定价: 26.00 元

<http://www.wpcsh.com.cn>

<http://www.wpcsh.com>

前 言

如果不着手解决目前所面临的有关威胁,数字遗产将会迅速丢失,而且不可避免。^①

——联合国教育、科学及文化组织《数字遗产保护宪章》

数字资源是以数字形式存在的可供人类作为生存发展基础而利用的信息集合。数字资源作为国家的重要战略资源,产生于社会生产、生活的各个方面,是社会的“记忆”,是社会发展的基础。

在当今的信息化社会,数字资源每天都在海量生成。与由模拟技术生成的信息资源相比,数字资源在生产、复制、传输和利用等方面具有无可比拟的优势,同时数字资源在数量上已经处于优势地位。数字资源的大量生产与广泛利用已经改变了政治、经济、国防等人类社会基本组成部分的传统运转方式。数字资源正逐步成为人类信息资源的主体。

^① 《数字遗产保护章程》中认为数字遗产包括以数字方式生成的或从现有的模拟资源转换成数字形式的有关文化、教育、科学和行政管理的资源及有关技术、法律、医学及其他领域的信息。笔者认为,数字遗产和数字资源的主要区别在于认识角度不同,数字遗产强调对未来社会的影响,数字资源强调对现在社会的作用,而两者所包含的内容大体上是相同的。

实现数字资源的长期保存,有利于人类文明传承,有利于国家有效控制具有战略意义的信息资源,充分发挥信息资源对经济社会发展的战略价值。然而,现在数字资源流失现象非常严重。我国 1982 年第三次全国人口普查 99% 的原始数据已经丢失;我国中央单位 42.2% 的电子文件没有以任何方式有效留存;美国 1960 年人口调查数据、卫生与公共服务部的磁带、记录“越战”期间战俘及战斗失踪人员的战地伤亡文件等都已经无法读取……数字资源长期保存工作必要且紧迫。

数字资源的严重流失是多种复杂的社会因素综合作用的结果,实现数字资源的长期保存需要面对来自社会意识、法律、技术、资金等诸多方面的挑战。要解决这样一个社会问题,必须形成从根本上解决数字资源长期保存所需要的一切必备条件,任务复杂而艰巨。社会、机构、个人各个方面都应当为实现数字资源的长期保存而积极努力,其中政府的努力尤为重要和必要。联合国教科文组织 2003 年发布的《数字遗产保护宪章》认为:“会员国应鼓励采用法律、经济和技术手段来保护这种遗产,必将受益匪浅。当务之急是开展宣传工作,促使决策者和广大公众都意识到数字媒体的潜力和保存工作的实际问题。”

正是在这样的背景下,通过对国内外相关研究和实践的深入分析,并引入政府治理理论,从有效留存国家具有战略意义的信息资源的角度出发,本书对政府如何有效应对数字资源长期保存这一重大社会问题进行了研究。

本书在以下几个方面有所突破:

第一,对数字资源长期保存所面临的各种问题以及解决这些问

题的所需要的必备条件进行了系统的分析。

第二,对政府治理在解决数字资源长期保存问题方面的有效性进行了充分的分析,明确了我国政府治理数字资源长期保存的目标定位和实现治理目标的具体方式。

第三,分析了政府治理工具在应对数字资源长期保存问题中的应用规律和特点,并对管制类、经济类、信息类治理工具在数字资源长期保存的政府治理过程中的基本安排进行了初步探讨。

本书试图通过对政府治理的研究,为解决数字资源长期保存问题提供理论与实践的指导,对数字资源长期保存活动起到推动作用。本研究属前沿课题,有一定的难度,同时作者学识有限,因而书中存在一些问题和缺点,恳请各位专家和同行批评指正,以便作者在今后的研究中有所长进。

谢永宪

2011年5月29日于北京联合大学应用文理学院

目 录

1	数字资源长期保存中政府治理的研究概况	(1)
1.1	时代背景	(1)
1.2	研究目的	(5)
1.3	国内外研究现状分析	(6)
2	数字资源长期保存的战略价值	(37)
2.1	信息资源在信息社会已经成为战略资源	(37)
2.2	数字资源逐步成为国家信息资源的主体	(45)
2.3	数字资源长期保存的战略意义	(53)
2.4	数字资源长期保存形势严峻	(55)
3	解决数字资源长期保存问题需要新形态的政府管理	(58)
3.1	数字资源长期保存问题的解决具有特殊的社会复杂性	(58)
3.2	数字资源长期保存问题解决过程中的政府责任	(68)
3.3	政府社会管理职能与传统管理职能的区别	(73)
3.4	有利于从根本上解决数字资源长期保存问题的政府 管理新形态是政府治理	(76)
4	数字资源长期保存政府治理的目标与方式	(78)
4.1	政府管理的新形态——政府治理	(78)
4.2	数字资源长期保存政府治理的有效性分析	(85)
4.3	数字资源长期保存政府治理的目标	(91)
4.4	数字资源长期保存政府治理的方式	(94)

5 数字资源长期保存政府治理的工具	(103)
5.1 政府治理工具及其功能	(103)
5.2 政府治理工具的分类	(105)
5.3 我国数字资源长期保存政府治理工具的选择	(110)
5.4 我国数字资源长期保存政府治理工具应用的初步设想	(119)
6 结论与展望	(139)
6.1 基本结论	(139)
6.2 未来研究展望	(141)
参考文献	(142)

1 数字资源长期保存中政府治理的研究概况

1.1 时代背景

1.1.1 问题的出现

信息技术的快速发展与广泛应用给数字信息资源(以下简称为“数字资源”,但为了方便行文,还需要在一些场合使用全称,两者含义等同)的长期保存带来了机遇与挑战。人类社会进入 21 世纪后,随着网络的普及,电子政务、电子商务的发展越来越快,数字信息被广泛的应用,其数量呈现出几何级数的增长态势。数字信息必须依附于某种载体存在,而存储介质普遍“短命”,因此,纸质信息载体比较容易实现的“百年存储”已经风光不再,信息转瞬即逝成为难以否定的事实。信息技术发展迅速,软硬件系统快速更新,使载体材料即使“耐久”也意义不大了,信息可能是“存在”的,但软硬件系统快速的升级换代,往往使相当一部分信息不再可以读取,信息事实上已经流失。这给人类文明的传承带来挑战,给人类对信息资源的有效控制带来挑战。

应对挑战的方式方法多种多样,政府机构、社会组织和个人各个方面都应当为问题的解决积极努力,其中政府的努力尤为重要和必要。政府应当注重充分发挥治理的作用,以有效的政府治理手段,从

社会、机构等方位,从管理、技术等角度,从法律、政策、标准、道德、素养等方面,以“硬治理”、“软治理”等方式,积极应对挑战,完成历史赋予我们这一代人的使命:有效传承人类文明;强化在信息社会具有高度战略意义的信息资源的控制力,充分发挥信息资源对经济社会发展的战略价值。

1.1.2 数字资源长期保存问题上的政府缺位和失效现象

随着数字资源数量的不断增长和社会经济发展对信息资源依赖性的不断增强,世界各主要发达国家的政府对数字资源的长期保存问题已经有所觉悟。

1991年5个北欧国家的国家档案馆(瑞士、挪威、丹麦、芬兰和冰岛)就电子文件的保护与存取问题进行了调研,并最终颁布了《电子文件的存取与保护》研究报告。1994年12月美国保护与存取委员会(Commission on Preservation and Access, CPA)与美国研究图书馆(Research Library Group, RLG)共同组建了数字存档特别工作组。1996年,该工作组完成了《保护数字信息:数字信息归档特别工作组报告》(*Preserving Digital Information: Report of the Task Force on Archiving of Digital Information*)^①。以上两个研究报告,对数字资源长期保存中存在的问题做了比较全面的分析,并提出了建设性意见。这两个报告引起了世界有关国家政府的重视,世界各主要发达国家相继成立了专门研究机构,采取了一系列应对措施。

1993年,澳大利亚成立了“朝向2001年,保护澳大利亚数字信息调研组”(Towards Federation 2001, Preserving Access to Digital Information, TF2001PADI)^②,对数字信息长期保存的基本情况进行了调研,并在1995发布了调研报告《电子档案、文件、出版物管理

① CPA, RLG, *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information*. [2008-06-20] <http://www.ifla.org/documents/libraries/net/tfadi-fr.pdf>.

② PADI. [2008-06-20] <http://www.nla.gov.au/padi/about.htm>.

指南》(*Guidelines for the Management of Electronic Records, Documents and Publications*)^①。1995年,澳大利亚国家图书馆启动了保存澳大利亚文化遗产项目(Preserving and Accessing Networked Documentary Resources of Australia, PANDORA)^②。1996年,澳大利亚国家图书馆制定了“澳大利亚电子出版物的国家策略”,提出了澳大利亚长期保存电子出版物国家策略的基本框架。到2004年澳大利亚已经形成了一个包括国家图书馆和全部州立图书馆在内的数字资源国家合作保护体系。

2001年11月到2004年10月,由欧洲委员会和瑞士政府资助120万欧元,荷兰、英国、意大利和瑞士的几个学术机构合作完成了电子资源保护和存取网络项目(Electronic Resource Preservation and Access Network, ERPANET)。该项目对OAIS、数字保存政策、WEB存档、元数据、文件数字保存格式等做了大量探索性研究,并将数字存储机构(如图书馆、档案馆)、研究机构、政府部门等联合起来,建立了一个可以扩展的数字信息保存联盟。^③

在世界主要发达国家对数字资源长期保存进行政府投入与管理的同时,我国政府也采取了应对措施。从2000年开始,在我国各级政府的支持下,深圳市档案馆、珠海市档案馆、青岛市电子文件中心相继进行了基于OAIS(Reference Model for an Open Archival Information System)参考模型的数字信息长期保存探索实践。他们在保证OAIS的六大功能的基础上,对OAIS参考模型进行了扩展,实现了结合本馆特点、本地区情况的对数字资源的有效存取和利用。他们的成功为其他档案馆开展数字信息长期保存活动提供了学习的经验。

经过理论的探索和实践的总结,我国分别在2002年和2006年

① National Library of Australia, *Guidelines for the Management of Electronic Records, Documents and Publications*. [2008-06-20]<http://www.nla.gov.au/niac/meetings/npo95kp.html>.

② PANDORA(Preserving and Accessing Networked Documentary Resources of Australia). [2008-06-20]<http://pandora.nla.gov.au/>.

③ ERPANET. [2008-06-20]<http://www.erpanet.org/>.

颁布了国家标准《电子文件归档与管理规范》(GB/T 18894—2002)和《中国档案机读目录格式》(GB/T 20163—2006),在 2003 年颁布了《电子文件归档管理暂行办法》。这些相关标准和指导性文件的发布与贯彻,促进了电子文件管理意识的加强和管理方法的普及,对我国数字资源的长期保存起到了推动作用。

中国和欧盟在数字资源保存方面的合作,是中国政府和欧洲委员会之间科学与技术合作的重要部分。2000 年中欧双方签订了科学与技术合作协议,同意在双方共同感兴趣的领域合作开展研究活动,参与双方基金项目。2004 年 7 月 14 日至 16 日,在中国国家科学图书文献中心、中国国家图书馆、中国高等教育文献保障系统、中国科学院国家数字图书馆、德国哥廷根大学图书馆、奥地利国家图书馆、国际电子图书馆联盟大力支持和共同组织下,“中欧数字资源长期保存国际研讨会”在中国科学院文献情报中心召开。作为图书馆情报界前沿领域专业性的国际性对话会议,此后每年召开一次。

就目前而言,总的情况是,各国政府对数字资源长期保存并非不重视、但收效不明显、不理想。以我国为例,尽管我国相关政府部门在数字信息在政府信息管理领域还没有占据主体地位的 1999 年,就开始以行政命令的方式,要求处理好这些信息资源的归档长期留存问题,而且几乎每年都通过行政命令手段不断予以强调,但据中国档案学会组织的《电子文件管理机制研究》研究课题组在 2007 年所做的抽样调查数据显示:我国中央国家机关及直属企事业单位,42.2%的电子公文没有以任何方式留存;74.4%的机构没有采用任何措施存留数据库、电子邮件、多媒体文件、网页文件等类型的电子文件。^①这说明政府的努力与客观成效之间并没有建立正相关关系。产生这种情况的原因可能是多方面的,笔者认为,其中的一个重要原因,就是政府虽然努力了,但没有真正站对位置,缺位现象严重。

缺位的表现,一方面是该政府出面的还没有出面,该管、该问、该

^① 见张宁《我国电子文件管理现状调查与思考》,《档案学通讯》2008 年第 6 期,第 15—16 页。

解决的问题,相当一部分还没有考虑,更遑论采取有效措施。比如,我国至今还没有建立有效的数字资源长期保存的管理体制,在全国各行各业实际运行中的信息系统大都没有留存功能,数字资源长期保存所需要的关键性技术方法的开发与推广一直无人问津。数字资源必须全程管理,但是就目前制度法律看,连数字资源的有效性都基本上无法保证,更不用说全程管理了。

缺位的另外一个方面表现,就是政府并没有把数字资源长期保存问题作为一个普遍的社会问题来认识,在解决这样一个复杂社会问题的过程中,依旧是在试图通过传统政府直接管制的简单手段解决问题,而没有发挥政府治理的作用,幻想依赖简单的直接针对数字信息资源保存的“命令”、“规定”,就从根本上解决问题。

1.2 研究目的

《数字遗产保护宪章》第三条认为:“世界上的数字遗产面临着消失和失传的危险,造成这种状况的因素包括:能创制这种遗产的软件和硬件迅速淘汰,维护和保存这类遗产的资金、责任和方法都不确定,以及缺少相关的立法。观念的转变落后于技术的更替。数字技术的发展日新月异,耗资巨大,各国政府和有关机构无法及时制定出相应的保存策略。数字遗产具有经济、社会、知识和文化价值,是建设未来的重要元素,但正面临着威胁,而且对这种威胁的认识尚不充分。”

数字资源长期保存活动中存在的问题错综复杂,它涉及数字资源的生产者、利用者、保存者、研究者、系统设计者、软件开发者的合作与制约,涉及各相关政府部门、社会组织、私人部门、企事业单位的利益与责任,需要社会各个方面的努力与合作。因此,有效解决这样复杂的问题,需要政府在法律、经济和技术等多方面的积极努力与投入,需要政府将数字资源长期保存作为社会问题予以治理。

本书的研究目的就是明确政府在应对这样一个具有时代特征的重大社会问题中的定位,探索以政府治理解决这一重大社会问题的

基本途径。

1.3 国内外研究现状分析

在广泛阅读相关主题文献基础上,笔者了解了数字资源长期保存中的政府治理方面科学研究的基本状况与最新发展态势。

1.3.1 文献调研

1.3.1.1 国外文献调研

笔者查询了 ProQuest Digital Dissertation(简称 PQDD,博硕士论文数据库)、ProQuest 检索平台数据库中的 Academic Research Library(学术研究期刊数据库)和 EBSCO 系列数据库中的 Academic Source Premier(学术期刊数据库)。检索结果中与本文主题相关的文献如表 1-1 所示。

表 1-1 外文文献调研情况^①

数据库 \ 检索词 篇数	“D”+“P”	“E”+“P”	“I”+“P”	“G”	“digital resources” +“policy”
博硕士论文数据库 (ProQuest Digital Dissertation)	32	15	141	344	0
学术研究期刊数据库 (Academic Research Library)	892	567	365	16 581	23
学术期刊数据库 (Academic Source Premier)	378	174	1 242	142	1

其中 D(digital)、P(preservation)、E (electronic)、I(information)、G(governance)

^① 检索时间为 2009 年 10 月 8 日,检索项为“主题”。

以上检索结果表明,国外学者对于“数字资源长期保存”方面的研究较多,对于“政府治理”的关注更多,讨论数字资源保存的政策和策略问题的文章很少,而专门从政府治理角度研究数字资源长期保存的文献尚未发现。

1.3.1.2 国内文献调研

在对国内文献的调研中,笔者查询了中国期刊全文数据库、中国优秀硕士学位论文全文数据库、中国博士学位论文全文数据库、中国重要会议全文数据库和中国重要报纸全文数据库,检索结果如表 1-2 所示。

表 1-2 中文文献调研情况^①

数据库 \ 检索词 篇数	“数字资源”和“长期保存”	“数字信息”和“长期保存”	“电子文件”和“长期保存”	“政府治理”	“数字资源保存”和“政策”
中国期刊全文数据库	245	320	111	1 812	34
中国优秀硕士学位论文全文数据库	36	64	26	162	14
中国博士学位论文全文数据库	10	13	5	14	3
中国重要会议全文数据库	3	6	2	51	0
中国重要报纸全文数据库	2	2	0	39	0

以上检索结果表明,国内学者与国外学者的研究重点有相似之处,“数字资源长期保存”的研究成果丰富,而“政府治理”的研究成果更多,但只有较少的文章讨论了数字资源保存的政策问题。在对数字资源长期保存的研究过程中,一些文章涉及了政府政策问题,但是还没有专门明确从政府治理角度研究数字资源长期保存的文献。

^① 检索时间为 2008 年 10 月 8 日,检索项为“题名”,匹配条件为“模糊”。

1.3.2 数字资源长期保存研究

1.3.2.1 数字资源长期保存的研究焦点

从目前笔者所掌握的文献看,国内外学者们研究的数字资源长期保存的焦点主要集中在数字资源长期保存的相关主体、保存原则、经济、技术、法律与合作等几个方面。

1. 数字资源长期保存的相关主体方面

宛玲提出数字资源长期保存运作的主要实施者包括国家图书馆、高校图书馆、出版商和非营利信息服务机构。^①

赖朝新认为数字信息长期保存的主体有出版者、国家图书馆和其他国家机构、研究图书馆和基金会、非营利性机构以及个人等。她提出了“理想的保存主体联盟”应该包括: a) 数字图书馆与其他图书馆的联合,包括国家图书馆、大学图书馆、研究图书馆等; b) 数字图书馆与图书馆之外的机构的联合,其理想的合作关系应该是数字图书馆、非营利性基金会和 1 个或者多个有一定技术优势的机构,或者是数字图书馆、国家机构和有一定技术优势的非营利性机构的联合; c) 数字图书馆内部工作人员的合作。^②

William LeFurgy 认为数字信息长期保存相关主体应该包括图书馆、档案馆、政府机构等,并结合 NDIIPP 项目进行了具体说明。^③

韩丽、梁爱东、恭维英认为数字信息资源的生产者、出版者和图书馆、档案馆等传统保存机构,都有对数字信息资源的长期保存责任。^④

翟慧杰认为数字信息长期保存的责任主体包括档案馆、图书馆、

① 见宛玲《国外数字资源长期保存的最新发展及对我国的启示》,《中国图书馆学报》2004 年第 2 期,第 22—26 页。

② 见赖朝新《数字信息长期保存的主体研究》,《图书馆理论与实践》2005 年第 2 期,第 47—48 页。

③ William LeFurgy, Building Preservation Partnerships: the Library of Congress National Digital Information Infrastructure and Preservation Program, Library Trends. Urbana: Summer 2005. Vol. 54, Iss. 1, pp. 10, 163.

④ 见韩丽、梁爱东、恭维英《数字信息资源长期保存的责任体系探讨》,《曲靖师范学院学报》2007 年第 11 期,第 122—124 页。

出版商、大学、学术团体、基金会以及个人。^①

2. 数字资源长期保存的原则方面

Jeff Rothenberg 认为“真实性”是指保证数字信息的完整性、完全性、正确性、有效性、忠实于原始含义,并且与原有目的相适宜。^②

刘家真提出了保存数字文献的要求:将数字文献确定为一个离散的主体加以保存,保存数字文献的媒体,保存数字文献的内容,保留数字文献外观,保存数字文献的功能性,真实可靠的保存数字文献,保存数字文献的出处和保存数字文献的背景信息。^③

张晓林认为数字信息长期保存就是“a) 保存数字比特流,即通过对数字存储媒介的保护或转移,确保存储其中的物理数字数据能被准确完好地读出。b) 保存数字格式与处理信息,即通过保存有关数字信息编码、格式、标记、结构、压缩、加密等方面的技术方法信息,确保能够识别和解析数字信息内容。c) 保存数字信息处理环境,例如相关的软件甚至硬件系统,确保能拥有相应的技术工具来识别、利用数字信息。d) 保存数字信息的内容校验、身份认证、版本演变、知识产权管理机制,确保能可信赖、可靠和合法地鉴别使用被保存的数字信息。e) 保存数字信息知识组织体系信息,保存数字信息的组织利用环境”。^④

张绍武、吴建西认为数字信息的长期保存必须要保存“数字信息比特流、数字信息的功能性、数字信息的出处、数字格式与处理信息和数字信息的管理手段”。^⑤

韩毅、邓小昭、杨晓琼认为“长期保存系统的基本原则是:确保任意时间和空间上产生的信息资源都应得到保存;所保存的信息资

① 见翟慧杰《数字信息长期保存的责任主体及其合作模式》,《档案》2008年第1期,第57—60页。

② Jeff, Preserving Authentic Digital Information. [2008-11-20] <http://www.clir.org/pubs/reports/pub92/rothenberg.html>.

③ 见刘家真《保存数字文献的问题与思考》,《图书馆杂志》2000年第8期,第28—31页。

④ 见张晓林《数字信息的长期保护问题》,《图书馆》2001年第5期,第7—12页。

⑤ 见张绍武、吴建西《数字信息长期保存的要求、难题及策略》,《情报理论与实践》2002年第5期,第457—460页。