

本书系浙江省哲学社会科学规划课题研究成果（课题编号16NDJC090YB）

网络信息 保存保护体系建设研究

陈清文 著

WANGLUO XINXI
BAOCUN BAOHU TIXI JIANSHE YANJIU



浙江工商大学出版社
ZHEJIANG GONGSHANG UNIVERSITY PRESS

本书系浙江省哲学社会科学规划课题研究成果（课题编号16NDJC090YB）

网络信息 保存保护体系建设研究

陈清文 著

WANGLUO XINXI
BAOCUN BAOHU TIXI JIANSHE YANJIU



浙江工商大学出版社
ZHEJIANG GONGSHANG UNIVERSITY PRESS

图书在版编目(CIP)数据

网络信息保存保护体系建设研究 / 陈文清著. —杭州 : 浙江工商大学出版社, 2017. 12

ISBN 978-7-5178-2419-0

I. ①网… II. ①陈… III. ①网络信息资源—资源保护—研究 IV. ①G255.76

中国版本图书馆 CIP 数据核字(2017)第 264778 号

网络信息保存保护体系建设研究

陈清文 著

-
- | | |
|-------|---|
| 责任编辑 | 白小平 |
| 封面设计 | 林朦朦 |
| 责任印制 | 包建辉 |
| 出版发行 | 浙江工商大学出版社
(杭州市教工路 198 号 邮政编码 310012)
(E-mail:zjgsupress@163.com)
(网址:http://www.zjgsupress.com)
电话:0571-88904980,88831806(传真) |
| 排 版 | 杭州朝曦图文设计有限公司 |
| 印 刷 | 虎彩印艺股份有限公司 |
| 开 本 | 710mm×1000mm 1/16 |
| 印 张 | 12.5 |
| 字 数 | 200 千 |
| 版 印 次 | 2017 年 12 月第 1 版 2017 年 12 月第 1 次印刷 |
| 书 号 | ISBN 978-7-5178-2419-0 |
| 定 价 | 38.00 元 |
-

版权所有 翻印必究 印装差错 负责调换

浙江工商大学出版社营销部邮购电话 0571-88904970

前 言

近年来,随着互联网在全球的迅猛发展,网络信息资源大量涌现,给人们的工作、生活带来了极大的便利,但是,网络信息资源生命周期极其短暂,消失速度惊人,给长期保存带来了困难。如何高质量地保存网络信息资源,使网络信息这种数字化的人类文化遗产得到长期保存与利用,是当前世界网络信息保存机构研究的重要课题。有鉴于此,本书把网络信息长期保存的诸多因素综合起来加以考察,从网络信息保存是一项复杂的系统工程入手,从整体上深入探讨了网络信息保存的采集策略、技术手段、法律框架、经济问题、管理策略等问题,并对改进我国网络信息保存能力提出了具体的建议。因此,本书的研究内容不仅对我国日益增长的网络信息资源长期保存的现实问题具有重要的实践价值,而且对于图书馆学、情报学、档案学的研究和实践也具有重要的意义,值得深入研究与推广。另外,本书还具有以下特点:一、选题新颖。本书运用系统论观点从网络信息保存体系建设的各个因素着手,深入探讨了网络信息长期保存的各种影响因素,并提出相关的解决策略与方法。从整体上研究网络信息保存的相关影响因素,有助于系统解决网络信息保存工作中存在的问题,克服分散研究、孤立研究的弊端,有利于网络信息保存各个主体、因素的协调发展,共同推动网络信息长期保存问题的解决。二、提出了改进我国网络信息长期保存的策略与措施。作者在充分借鉴国内外相关研究成果的基础上,深入分析了我国网络信息保存的现状,提出了改进我国网络信息长期保存的策略与方法,具有较强的针对性、实用性。三、创新点。作者在分析国内外大量研究资料与成果的基础

上,提出了图书馆、档案馆是重要的网络信息长期保存的责任者,多主体参与、合作保存是保持网络信息长期利用的最佳选择,呈缴保存是网络信息长期保存的重要方法,保存元数据是网络信息资源长期保存的重要技术因素,国家补偿、用户补偿和协议补偿相结合的网络信息保存补偿机制是重要的经济手段等重要的创新观点。

陈清文

2017年9月

目 录

第一章 绪 论	001
第一节 问题提出	001
第二节 网络信息保存面临的主要障碍	002
第三节 研究意义	007
第四节 国内外关于网络信息保存研究现状综述	008
第五节 国外网络信息长期保存项目概览	012
第二章 网络信息保存基本理论问题研究	031
第一节 网络信息保存的概念、类型与特点	031
第二节 网络信息保存保护体系建设研究	037
第三节 网络信息保存标准发展概况	041
第三章 网络信息采集策略研究	061
第一节 采集主题的确定及相关问题	061
第二节 采集方式	062
第三节 网络信息采集策略与方法	065
第四章 网络信息保存技术问题研究	077
第一节 网络信息内容发现工具	077
第二节 网络信息保存技术方法	084

第三节	网络信息存储介质	090
第四节	网络信息长期保存的有效风险管理和灾难恢复机制	095
第五节	元数据	098
第六节	知名网络信息采集工具 WCT 分析	111
第七节	网络信息采集与保存平台建设	115
第五章	网络信息保存责任体系研究	117
第一节	国外网络信息保存责任体系研究综述	118
第二节	我国网络信息保存责任体系研究综述	120
第三节	存在的分歧	120
第四节	国外的做法	121
第五节	责任主体的确定	123
第六节	合作保存模式	131
第六章	网络信息保存相关问题研究	143
第一节	法律关系的确定	143
第二节	建立网络信息呈缴本制度	152
第三节	网络信息保存的管理与经济策略研究	160
第七章	改进我国网络信息保存工作的几点建议	179
第一节	改进我国网络信息保存工作的策略与措施	179
第二节	对访问与使用的优化	183
参考文献		186
后 记		191

第一章 绪 论

第一节 问题提出

近年来,随着数字信息技术的迅猛发展,互联网络深入到社会生活的各个方面,改变了人们工作与生活的方式,给人们的生活和工作带来了极大方便。特别是随着数字化文献的日益增多,科研人员信息获取的方式也发生了较大的变化,通过互联网获取研究文献信息日益成为广大科研人员获取信息的重要渠道,给广大科研人员开展学术研究带来了极大的方便。网络信息具有信息量大,可检索到大量正式出版系统中无法得到的“灰色文献”,能快速及时地了解当前学术界最新研究动态和研究趋势,方便与同行进行学术交流、共享研究成果等特点,深受广大科研人员的欢迎,通过互联网获取研究信息正成为广大科研工作者必不可少的学术文献获取方式之一。但同时,网络信息变化频繁,极易遭受网络病毒或黑客的破坏和攻击,加之,网络设备硬件故障等因素,使得大量有价值的网络信息随时面临永久消失的危险,使网络信息长期保存面临严峻的挑战,并成为制约其进一步发展的重要障碍。当前,网络信息长期保存问题并未引起全社会的足够重视,一旦发生信息丢失,则面临灾难性后果,无法恢复。特别是一些具有重要历史、文化价值的文献或研究信息,本身更新速度较快,如不及时保存,随时面临永久消失的危险。因此,加快对网络信息保存策略与方法的研究显得尤为迫切。另外,网络信息资源长期保存也是以计算机为代表的现代信息技术发展的必然要求,网络信息资源与传统纸质文献一样是人类共同的文明成果,是人类纸质和数字化文化遗产的重要组成部分,我们应当肩负起保存人类数字化遗产的重任,将网络信息资源进行整合,并加以长期保存,因此,该课题研究具有重要的现实意义和理论意义。一方面,网络信息长期保存是图

书馆学、情报学与档案馆学研究的前沿课题。面对网络信息保存问题的巨大挑战,我们应积极行动起来,探讨适合网络信息长期保存的策略,建立合理的网络信息长期保存责任体系,深入研究网络信息保存的法律问题,加快相关网络信息长期保存法律体系的建设,尽早制定网络信息呈缴法律,使网络信息资源能够得到及时的保存。另一方面,面对纷繁复杂的网络信息资源保存的课题,我们还应积极探索适合网络信息长期保存的技术策略,与计算机专业人员一起合作攻关,尽早研制出适合中文网络信息资源搜集与保存的软件,并同国际网络信息长期保存标准融合,提高中文网络信息长期保存的质量与效果。另外,我们还应积极探索网络信息长期保存的经济效益与管理手段问题。经济是一切工作的基础,对于网络信息长期保存来说也不例外,网络信息长期保存是在高技术环境下利用计算机等现代信息技术设备完成的,其保存成本巨大,因此,注重网络信息长期保存的经济效益是使网络信息保存长期可持续发展的关键。我们应加强对网络信息长期保存成本的监控,通过严格的管理制度和手段,不断提高网络信息资源长期保存的经济效益。

第二节 网络信息保存面临的主要障碍

一、网络信息自身特点带来的挑战

网络信息数量巨大、格式多样、变化频繁,使得对其进行长期保存面临较大的困难。在进行网络信息搜集保存的过程中随时会遇到网络信息消失的问题,客观上影响了对网络信息的采集和系统跟踪,影响对有价值的网络信息的系统搜集和保存。如:当一个机构将新制作的网页放在网站上时,新的网页会代替旧的网页,致使原有的网络信息消失;或者,网站在周期性地重组网页内容结构时,许多“陈旧”的内容被删掉,导致其中的网络信息丢失,给网络信息保存工作带来挑战。

二、网络信息技术环境改变带来的挑战

网络信息技术的不断进步,一方面,带来了网络数字化信息利用方式的

不断变革,另一方面也给网络信息的保存带来了困难。网络信息的物理介质同一切数字信息一样,是以数字 0 或 1 的形式存储在存储介质中的,无论是磁盘、磁带、光盘、硬盘或是网络存储服务器,其致命的缺陷是容易受到各种内部乃至外界环境的干扰,其存储的寿命完全依赖于存储介质的寿命,与纸质文献相比较极其容易损坏并造成信息的长久丢失。这些缺陷包括:

第一,网络信息必须借助于特定的信息设备才能读取其中的内容,其保存的技术方法复杂多变,人们在保存网络信息内容的同时,只有保持原有设备使用所要求的技术环境特性,才能保证所保存的网络信息可以被长期使用。

第二,网络信息技术标准的多样性,使得网络信息检索与保存没有一个统一的标准可以执行。因此,人们在保存利用不同技术标准建立的网络信息的时候,就必须要进行必要的格式转换,才能保证所存储的网络信息可以得到长期的存取。

第三,网络信息的不安全性,也加重了对网络信息长期保存的负担。如网络计算机病毒、黑客入侵、信息的失真等不安全因素随时威胁网络信息的长期存取与利用。

第四,网络信息多数都是由文字、图形、声音、影像等构成的多媒体信息,因此,在进行网络信息保存的过程中,只有对网络信息的多媒体特性进行保存才能保持网络信息的原貌。然而,由于存储格式、技术标准、载体容量等限制,很难做到保存网络信息的多媒体特性,严重影响网络信息的保存质量。

三、网络信息保存工作没有得到应有的重视

近年来,随着信息技术的进步,图书馆信息处理手段得到了较快的发展,数字图书馆不断出现,各种数字资源日益丰富,推进了图书馆文献资源的广泛利用。但在图书馆数字化过程中,普遍没有把网络信息的保存工作提上议事日程,大多数图书馆对网络信息保存工作表现被动,更不用说制定一些保存网络信息的方案,或开展网络信息保存的试验了,严重影响网络信息长期保存工作的开展。在网络环境充分发展的现代信息社会,图书馆应当承担起保存网络信息资源的重任,并作为全社会网络信息长期保存的倡导者和领导者。只有这样,图书馆保存人类文化遗产的重任,才能发扬光

大,否则,图书馆将失去自身存在的价值。

四、网络信息保存的法律法规严重滞后

网络信息的长期保存工作除了应用各种先进的现代信息技术实施保护以外,网络信息的保存责任、保存主体的确定、保存法律、法规的制定等非技术手段同样具有重要作用。当前,网络信息长期保存的法律、法规还不健全,已经成为制约网络信息长期保存的最大障碍。在网络信息保存工作中,法律法规的严重滞后,致使许多网络信息长期保存工作没能很好地开展起来。各保存机构开展的网络信息长期保存工作,由于没有相关的法律法规约束,不能较好地履行自己应尽的网络信息保存责任,网络信息保存工作处于一种分散、孤立的情况,这一局面急待改变。因此,建议在今后制定图书馆法时,应当把有关网络信息保存工作的内容列入其中,以便使网络信息保存工作走上法制化、制度化、正常化的轨道。网络信息长期保存的整个保存过程都与知识产权密切相关,从网络信息资源的采集、下载到存储和长期保存及提供服务,都会遇到知识产权相互冲突的地方,因此,在网络信息长期保存过程中,就要首先解决保存信息的知识产权保护问题,以便使网络信息保存工作有法可依,保证网络信息保存工作的长期可持续发展。然而,严酷的现实是目前还没有关于网络信息长期保存的专门法律,其他的相关法规也不健全,无法保证网络信息长期保存的实践能顺利地开展,因此,法律障碍也成为开展网络信息长期保存过程中亟待解决的问题。

五、经济因素是影响网络信息保存的重要因素之一

网络信息的长期保存是在高技术条件下进行的,其所需的软、硬件和人力、物力等条件要求较高,需要大量、长期的经费支持才能顺利进行,因此,经济因素是我们在进行网络信息保存工作时必须要加以认真考虑的问题。如何高效、低耗、经济地保存网络信息是我们在进行网络信息保存工作时必须始终关注的重点。以最低成本获得最大的保存效益始终是我们在网络信息保存工作中追求的目标。单从网络信息的特点来看,其维护费用受到技术、人力等因素的影响。在具体网络信息保存的工作中,每一个保存的步骤和环节,以及每一种保存策略和方法的选择都与经济因素息息相关。效益较好的保存策略和方法可以节省网络信息保存所用的成本,以较高成本消

耗来换取低效的网络信息资源保存成果是不可取的,因此,在网络信息保存工作中,提倡注重经济效益的观念是十分重要的。

六、技术因素是影响网络信息长期保存的重要因素

网络信息长期保存是在高技术环境下利用计算机和网络技术手段实施保存的复杂技术工作,每一步都与保存技术手段息息相关。随着网络信息技术的不断发展,软硬件技术过时,保存介质老化,造成新旧技术不能兼容,使网络信息保存技术操作面临许多困难。如:读写数字化信息的计算机技术过时问题,网络信息存储载体的老化问题,以及由于缺乏相关的标准及协议形成的网络信息存储格式多样化等问题。针对这些问题,保存机构必须经常性地对已经保存的网络信息开展周期性的更新、复制、迁移等保存操作,所有这些都使网络信息保存面临严峻的挑战。具体包括以下几个方面:

1. 存储载体问题

当前,多种网络信息保存内容都是以 0 或 1 的二进制方式存储在一定的物理载体上,如各种不同类型的磁盘、光盘等,这些网络数字信息的寿命就取决于这些物理载体的寿命。一般认为磁盘的使用寿命平均为 10—15 年,光盘 30 年左右,耐用光盘可达 100 年,但价格昂贵,无法普及使用。因此,即使最耐用的光盘与保存几千年的印刷文献相比也相差太大。另外,由于网络数字化保存信息载体容易发生物理或化学变化,对存储环境要求严格,长期保存十分困难,容易发生信息丢失或永久消失的问题,网络数字信息在保存载体中一旦发生信息消失现象,结果将不堪设想,因此,解决网络信息存储载体长期保存问题是网络信息保存工作的重要挑战,需要认真地加以研究。

2. 网络信息存储格式多样化问题

网络信息资源数量巨大,信息格式种类繁多,缺乏统一的技术标准,严重影响网络信息长期保存工作的开展。随着网络信息资源存储技术的不断发展,各电子出版商为了保证自己出版的网络信息的版权,采用专门的网络信息存储格式生产网络电子出版物,如清华同方光盘股份公司采用 CAJ 格式存储、阅读,该公司出版的网络电子期刊必须使用该种格式的浏览器;而超星电子书公司也采用自己开发的电子图书存储格式保存利用自己的网络电子图书等。因此,众多的网络信息存储格式使得各种网络信息之间的数

据交换带来困难,给广大用户利用网络信息资源带来麻烦,影响网络信息资源的共享与利用。

3. 技术过时存在的问题

网络信息技术及其产品的不断更新换代,一方面,使得网络信息处理能力增强、成本降低,另一方面,也为网络信息存储利用带来了新的困难。由于网络信息是一种数字化的电子信息资源,必须通过专门的计算机设备借助一定的软件才能使用,因此,随着这些软、硬件技术的不断升级更新,新旧版本之间的软件无法兼容,利用旧有版本技术存储的网络信息无法阅读,造成网络信息的丢失,失去使用价值,影响网络信息的长期保存。

4. 网络信息安全问题

近年来,随着互联网的迅速普及与发展,各种网络信息大量涌现,人们可以方便、自由地在网上阅读、浏览、检索、下载各种网络信息,给人们获取信息带来了极大的便利。但是互联网上有大量的计算机病毒滋生,严重影响网络信息的传播与保存。据报道,目前每天全世界都有计算机病毒产生,病毒的种类有数千种之多,严重干扰了网络信息保存工作的顺利开展。同时,网络黑客也乘虚而入,他们利用计算机系统本身的缺陷和大量存在的薄弱环节进行攻击,轻者使计算机不能使用,重者造成网络的瘫痪,使网络信息保存面临日益严重的安全问题,造成网络信息的永久丢失,成为网络信息保存工作面临的主要威胁之一。

七、网络信息采集面临的问题

由于网络信息更新频繁,数量巨大,因此,对所有网络信息实施保存显然不可能做到,而对网络信息进行选择性保存就显得十分重要。然而,选择什么内容进行采集保存就十分困难,不同的采集标准都会造成标准以外的一些网络信息被排斥在保存之列外,因此,选择适合的网络信息采集的标准就是一件非常头痛的事情,我们选择的标准无论多么完善,总会造成没有被选进的网络信息的永久丢失。从采集网络信息的频率来看,由于网络信息的相互链接及动态更新的特点,使得我们采集网络信息的频率难以确定。目前大多数网络信息保存项目都是按照事先统一设定的采集频率进行网络信息的采集,这也会人为造成网络信息采集的不全面,影响保存质量。

八、经费保障问题

网络信息资源长期保存与传统的印刷文献保存有很大不同,无论在网络信息的采集、长期保存、更新还是日常管理等方面都需要有充足的经费做保证,加之前期网络信息资源长期保存基础设施的建设,所需经费巨大。并且网络信息资源长期保存的投资很难在短期获得较快的经济回报,也很难得到各类商业投资的支持,因此,如何保证对网络信息资源长期保存的经费支持,以便保证网络信息资源长期保存项目的可持续发展,是网络信息长期保存工作面临的主要难题之一。

九、保存主体障碍

网络信息资源数量巨大,主体分散,加之保存费用巨大,到底由哪些机构来承担保存的责任还没有明确的答案,纵观国内外网络信息保存项目的实践,也没有较为一致的看法。有学者认为网络信息长期保存应该由信息的生产者或出版者承担;也有学者认为网络信息长期保存应该由图书馆、档案馆等传统文献保存机构承担。但无论哪一个机构承担网络信息长期保存责任,都无法承担巨大的经费需求,因此,以法律的形式构建起科学合理的网络信息资源保存的责任体系,形成像纸质文献保存体系一样的网络信息长期保存的全国保存体系显得十分重要。

第三节 研究意义

随着互联网的快速发展,网络信息增长速度极快,我们可以称之为信息爆炸,但同时网络信息消失的速度也令人吃惊,同印刷文献相比较,网络信息固有的更新频繁、动态性等特点使其丢失或永久消失更加不易察觉,长期保存更加困难,相当一部分有价值的网络信息随时面临永久消失的危险,因此,对网络信息进行长期保存显得非常重要。图书馆、档案馆在纸本文献长期保存方面积累了丰富的保存经验,对于网络数字化文献这种新的文献形式的长期保存,也将发挥积极的作用。网络信息长期保存是全人类面临的重要课题,图书馆、档案馆等传统文献保存机构应该积极行动起来,探索网

网络信息资源长期保存的策略与方法;探索网络信息保存的技术策略;建立网络信息长期保存的责任机制;根据中文网络信息的特点,研究制定适合我国国情的网络信息长期保存的策略与措施,使网络数字化文献这种新文献形式可以得到长期保存与利用。同时,随着数字化、网络化步伐的加快,网络数字化文献数量日益增多,成为人们获取文献信息的主要途径。然而网络数字化文献由于数据格式多样、技术标准的区别,以及网络数字化文献变化频繁,更新速度快,对保存环境和保存技术手段较之传统纸本文献要求高等特点,导致网络信息资源面临随时消失的危险。因此,我们要积极行动起来,采取有效措施保存网络信息资源,以避免网络数字文化遗产永久消失的危险。我们应树立网络信息资源也是人类数字文化遗产重要组成部分,也是重要的“原生性”信息,和印刷文献一样是我们人类宝贵的文化遗产,应当得到长期保存的重要观念,迅速行动起来,积极探讨保存网络信息的最佳方法与途径,使网络信息资源这种人类重要的文化遗产的新形式能得到长期保存,能为子孙后代长期利用。特别是当前,网络信息资源长期保存与利用问题正成为世界各国数字图书馆发展过程中急需解决的重大课题,引起了各国图书情报工作者的高度重视,这一重要课题的解决,不但对数字图书馆的发展极为重要,同时也是保存人类网络数字化文化遗产的需要。因此,我们要在深入了解国内外网络信息资源保存研究的现状和进展情况的基础上,集中研究力量、节省研究时间和经费、避免重复劳动、重点解决网络信息资源保存过程中的关键问题,促进网络信息资源保存问题科学合理地解决。因此该课题具有重要的学术价值和实际意义,值得深入探讨。

第四节 国内外关于网络信息保存研究现状综述

一、我国网络信息长期保存研究进展

随着以互联网为代表的现代信息技术的快速发展,网络信息长期保存与利用问题正成为世界各国信息保存机构发展过程中急需解决的重大课题,引起了各国信息保存工作者的高度重视。这一重要课题的解决不但对网络信息检索与利用的发展极为重要,同时也是保存人类网络数字化文化

遗产的需要。为了能充分了解我国关于网络信息保存研究的进展情况,笔者利用清华同方光盘公司出版的《中国学术期刊网全文数据库》和《超星发现数据库》等中文网络数据库进行相关检索。为了能够充分了解中文网络信息资源保存研究情况,笔者选取了“网络信息保存”作为检索词进行检索。首先采用主题词检索,在《超星发现数据库》中限定图书情报范围检出文献后,继之以相关著者为检索词进一步检索,查看是否有遗漏;然后再辅以《中国学术期刊网数据库》的电子信息专辑进行补充,检索时间 2007—2016 年。结果如下:共检索到发表论文 2323 篇。总被引频次 4181 次。见图 1。

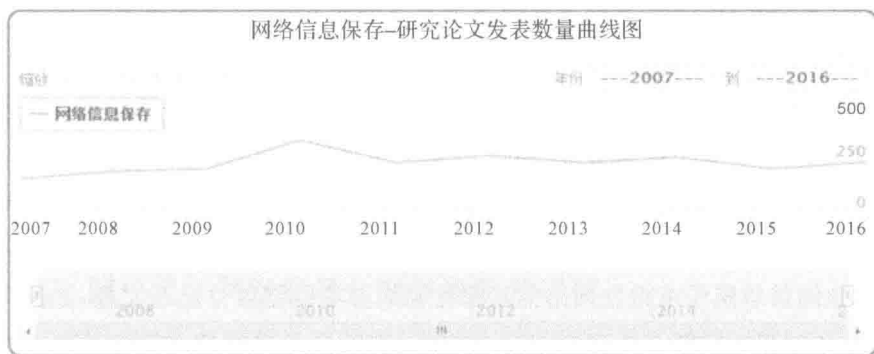


图 1 近十年我国网络信息保存研究论文发表数量曲线图

从以上检索结果来看,2007 年至 2016 年 12 月,总共检索到 2323 篇该主题的研究论文,总被引频次 4181 次。从研究论文的发表轨迹来看,近十年,我国关于网络信息保存问题研究的论文呈现平稳发展的态势,其中 2010 年出现一次研究的高峰,发表论文 344 篇,之后论文发表数量保持平稳,维持在每年 200 篇以上的状态。计量结果表明,网络信息资源保存研究已经引起我国图书情报界的关注,研究成果保持平稳发展趋势。同时,论文统计也表明:我国对于网络信息保存问题的研究虽然已经起步,但研究还不够深入,像赵俊玲博士的《国外关于网络信息保存研究的综述》这样专门论述网络信息保存问题的论文还较少。但从最近十年来我国关于网络信息保存的研究来看,我国学者已经开始认识到网络信息长期保存问题的重要性,已有像刘家真教授、张晓林教授、吴振新教授、宛玲博士、赵俊玲博士、杨道玲博士等专家学者对网络信息保存问题的探讨和论述。特别值得一提的是为了改变我国网络信息保存研究的落后状况,我国中科院文献情报中心与

中国国家图书馆、中国高等教育文献保障系统等单位于2004年7月和2007年10月在北京先后举行了两次中欧数字资源长期保存国际研讨会。这两次会议的召开对于了解国外关于网络信息长期保存问题研究的进展情况,借鉴国外网络信息保存研究方面的成功经验,促进我国网络信息保存问题研究的不断深入,缩小我国同发达国家在这一领域的差距具有重大的意义。此后,每年的年度会议都有中国学者的身影。^①

另外,从网络信息长期保存研究项目来看,我国先后开展了一系列网络信息长期保存的研究项目,如2002年的中国web信息博物馆项目;2003年国家图书馆开展了网络信息资源采集试验项目,对中文网络信息资源进行采集与保存的研究;2007年我国加入国际互联网保存联盟;2009年我国成立了国家图书馆互联网信息保存中心,对中文网络信息资源实施保存。在我国网络信息保存研究领域,2003年国家图书馆项目研究小组的建立以及开始实施“网络信息资源采集与保存试验(WICP)”和ODBN(on line database navigation 网络数据库导航)项目具有重大的开创性意义。此外,近年来,中国科学院图书馆在网络信息保存领域多项研究项目已在实施,如网络信息长期保存国际合作研究项目(CSDL)和数字化科技信息资源长期保存体系与政策机制项目(NSDL项目),该项目的研究运行为进一步了解、掌握网络信息保存工作中应解决的问题提供了实证性的研究数据,使网络信息保存研究理论同网络信息保存实践结合得更加紧密,使研究结果更有说服力。另一个正在实施的网络信息保存项目——“中国Web信息博物馆”也积累了许多网络信息保存的实践经验。中国Web信息博物馆(Web Info-Mall)是在国家973和985项目支持下,北京大学网络实验室开发建设的中国网页历史信息存储与展示系统。系统提供如下几种功能:(1)浏览过去存在的网页;(2)典型历史网页展示;(3)历史事件专题回放。Web InfoMall搜集了从2001年以来约5亿个中文网页,正以每天数十万网页的速度增长。特别是对网络信息采集、组织、保存与服务方面需要研究探讨的问题,该项目的研究更加具体,针对性更强,也使得网络信息保存问题的研究更具有深远的历史意义。

近年来,我国广大图书情报工作者还对网络信息长期保存课题进行了

① 陈清文. 网络信息资源保存研究综述[J]. 山东图书馆学刊, 2006(1):18-21.