



# 网络信息资源管理初探

皮介郑

(中国科学院文献情报中心 北京 100080)

**【摘要】** 从宏观控制与微观组织两个方面讨论网络信息资源的管理问题。

**【关键词】** 网络信息资源 宏观控制 微观组织 **【分类号】** G250.73

## Discussion on Management of Network Information Resources

Pi Jiezheng

(The Documentation and Information Center of CAS, Beijing)

**【Abstract】** This paper discusses the management of Network Information Resources from the two angles of macorscopic control and microcosmic organization.

**【Keywords】** Network information resources Internet Information resources management

网络信息资源的特性与面临的诸多问题决定了网络信息资源管理这一重要命题的复杂性与艰巨性。这一命题已引起了许多信息专家的注意,但有关的研究与实践活动仍处于探索与实验阶段。本文从网络信息资源的宏观控制与微观组织两个方面作些讨论。

### 1 网络信息资源的宏观控制

网络信息资源赖以存在的因特网是一个松散、没有专门机构进行集中管理的网络,最大的特点就在于它的自由发展模式:自由连接,自由发布信息,自由扩展,自由增加服务方式和服务内容等。是自由发展造就了因特网,从而也最终形成了网络信息资源,但同时也带来了大量问题(有关内容已被广为讨论,在此不再赘述)。为解决这些问题,将网络信息资源的发展导向更合理的未来,采取一定措施对网络参与者的行为进行规范、约束与控制并辅于相关的技术手段势在必行。

#### 1.1 网络法规的建立

利用法律的力量解决网络信息活动中存在的各种问题,调整信息活动中的各种社会关系,规范和约束人们的信息行为已成为全球呼声。信息法律通过规范网络活动来协调各方的利益关系,解决矛盾,达到疏通信息渠道、保护信息产权、促进信息资源开发利用和信息交流的目的。以网络知识产权保护为例,由于网络的高度开放性和极大的自由度,给网上知识产权保护带来困难,在某种程度上妨碍了网络信息的正常交流。因此,完善网络知识产权保护受到各国关注。美国在《NII:行动计划》这一纲领性文件中十分强调网上知识产权和隐私

权的保护,英国信息基础设施推进委员会知识产权工作组的最终报告以及台湾资策会制定的《多媒体节目著作权法修法建议》都是有代表性的网络知识产权保护法规。这些法规明确了网络信息产权保护的范围及信息利用者的义务和规范,促进了网络信息交流和信息资源的有效利用。

在我国,公安部1997年12月30日发布了由国务院批准的《计算机信息网络国际联网安全保护管理办法》,就网络信息提供者的权限和义务(如政府信息的公开和保护、企业信息的公开和保护、个人信息的公开和保护等)和网络信息用户的自由权和获取信息权的界定,以及网络信息参与者应遵循的标准、原则、程序和条件等方面作了明确的规定。<sup>[1]</sup>1999年初,为了规范上海市统计信息的发布,保障统计信息的准确性、及时性,防止虚假、伪造统计信息的发布,上海市国民经济和社会信息化领导小组办公室和上海市统计局制定《利用计算机公众信息网络发布统计信息管理办法》。规定发布统计信息实行登记制度,由上海市信息办统一制发登记证,无登记证的单位不得发布统计信息;发布统计信息的单位应当持有有关资料,向市信息办申请发布统计信息登记证,经核准后,方可从事发布统计信息活动;发布统计信息,应遵守国家有关法律、法规或规章制度,严格执行国家安全保密制度;全市性重要经济、社会统计信息,由市统计局依照有关规定统一公布;凡从事统计信息开发或者传输活动的单位,应当具有《上海市计算机公众信息服务业准营证》,并检查发布统计信息单位的登记证。<sup>[2]</sup>类似的法规有助于保障网络信息的真实性和可靠性,从而提高网络信息资源的质量。

#### 1.2 建立健康的网络伦理

收稿日期:2000-10-31

网络伦理指在网络信息活动中被普遍认同的道德观念和应遵守的道德标准。由于网络信息资源污染问题日益严重,网络伦理问题已引起普遍关注。一些专门的研究组织纷纷成立,并开展各种学术研讨活动。有的机构或组织还制定了约束网络行为的“网络守则”。健康的网络伦理有助于人们正确回答诸如“应在网上发布何种信息”、“如何复制网上信息才是正当的”、“可否在网上窥探他人的隐私”之类的问题,从而推动网络信息资源的健康发展。

### 1.3 技术手段

技术手段的使用是控制网络信息资源不可或缺的重要方面。如开发特殊的网络软件用于过滤或删除含有黄色、暴力、种族歧视等不健康内容的网络信息,利用防火墙、电子密钥等技术保护信息安全等。因特网在技术方面存有的缺陷对网络信息资源的质量、安全、利用及其它方面都是一种巨大的威胁,但它本身是现代高新技术的产物,许多问题还得靠技术的手段加于解决。

## 2 网络信息资源的微观组织

### 2.1 基层组织

网络信息资源整体上的混乱与无序并不说明它缺乏任何形式的组织,而是说它已有的组织程度对满足各种信息需求还没有达到令人满意的程度。

Internet 在其发展过程中逐步建立了许多信息查询工具(或服务功能),它们对网络信息资源构成了一定程度的组织手段。如 Archie 可自动地查询 FTP 服务器,将其中的文件索引创建到单一的、可搜索的数据库中。该数据库就是 Archie 所能查到的所有服务器中文件目录数据的索引。用户只要给出希望查找的文件类型及文件名,Archie 服务器就会指出在那些 FTP 服务器上存放着这样的文件。而 Gopher 将网上信息组织成联机的菜单系统,以方便用户浏览感兴趣的题目。它将用户的请求自动转换成 FTP 或 Telnet 命令,在菜单的引导下,对 Internet 上的远程联机信息系统进行实时访问而不必知道所访问机器的地址。Wais 则是用户查询互联网上各类数据库的一个通用接口。用户只要在给出的数据库列表中用光标选取希望查询的数据库,并键入查询关键字,系统就能自动进行远程查询,读出相应的数据库中含有该查询的所有记录,再根据查询关键字在每条记录中出现的频度进行评分,用户可以进一步选择感兴趣的记录。然而最神奇的还是 WWW 系统,它是一种基于超文本方式的信息查询工具,能够将位于全世界 Internet 上不同地点的相关数据信息有机地编织在一起,使用户能够在 Internet 上查找已经建立的 WWW 服务器站点(Site)所提供的信息资源。WWW 把各种类型的信息(静止图像、文本、声音和影像)天衣无缝地集成起来,并提供图形界面下的快速查找,使用同样的图形用户界面可与 Internet 上的其它服务器对接。

但是,上述工具所提供的主要是方便地浏览网上信息的手段,并不能提供从茫茫的网络信息资源汪洋中迅速找到特定信息的方法。搜索引擎正是为解决这一问题而产生的信息

查询工具,它对网络信息资源具有二次组织的作用。

### 2.2 搜索引擎的二次组织作用

搜索引擎(Search Engine)是一种在互联网产生后伴随着网上用户快速查询信息的需求而产生的能提供网上信息查询服务的系统,即提供信息检索服务的计算机系统。它本身也是以网站的形式存在,有自己的网址,与一般网站的区别在于其目的是为了帮助人们查找网络信息,是一种信息查询工具,如同图书馆里的目录,或者是电信局的 114 查号台。其工作原理是以自动或人工的方式收集 Internet 中的信息资源并提供一条检索途径。概言之就是将计算机网络信息资源做一完整的集合、整理、分类,从而便于用户能轻易地查找到所需要的信息。

由于搜索引擎具有重要作用,短短几年内已有大量的搜索引擎被开发和建立起来。不同的搜索引擎可能会有不同的检索界面,不同的侧重内容,但有一点是共同的,就是它们都有一个庞大的索引数据库。这个索引库是向用户提供检索结果的依据,其中收集了因特网上数百万甚至数千万主页信息(网站和网页信息),包括该主页的主题、地址、包含于其中的被链文档的主题,以及每个文档中出现的单词的频率、位置等等。为了获得这些索引信息,每个检索工具都配备有自动检索程序(Robots 或 Spiders),随时都会顺着超文本之间的链接跟踪网上新加入的主页、节点,为其建立索引,并送回集中管理的索引数据库。以 AltaVista 为例,它对因特网上上亿个主页的内容进行了索引,建立了约为 200GM 的庞大索引数据库,其 Web 部分的索引为每日更新,Usenet 部分的索引为持续更新,整个索引数据库定期重建。为了使用户在其庞大的数据库中迅速找到信息,AltaVista 设置了 14 个检索入口,有网址、域名、主页标题、图象文件、文本内容等,便于用户根据需要查找信息。可见,搜索引擎对网络信息资源的组织作用十分明显,它们在一定程度上实现了对网上信息的控制,在逻辑上序化和优化了网络信息资源,为充分开发利用这种资源提供了前提条件和可能。

### 2.3 更高级的组织形式

目前的搜索引擎还有很大的局限性,首先是在建立资源索引时针对性不强,面向大众的资源覆盖面广,而面向科学技术的则相对非常少;另外,检索结果中包含大量相关性很小的内容,用户必须花大量时间进行剔除,检准率很低。搜索引擎所能达到的组织程度还远远不能满足针对学术性资源的专业性较强的深度信息需求。而且,网络中的各种检索系统虽然功能十分强大而易于使用,但它们对资源不具有选择和价值判断的能力,对网络信息资源更高级别的组织仍然需要图书情报机构在信息分类编目等方面的专业知识和经验。美国图书馆界的许多单位已在这方面进行了积极的尝试和努力,典型的做法是用机读编目格式(MARC Format)来组织和整理网络信息资源。

美国图书馆协会于 1992 年年底提出议案,建议修改书目信息的内容以容纳因特网中电子期刊和电子文献等各种电子

信息。此议案主要是修改机读编目格式,新增 865 字段(Field 865)即“电子位置和存取”(Electronic Location and Access)。866 字段的内容有:存取方式(如电子函件、FTP、远程登录、Gopher 或 HTTP)、主机名称、路径、文档名称或其它可以协助用户通过网络获取电子信息的方法等。1993 年 1 月此议案经修正获得通过,目前已成为美国机读编目格式(USMARC)中的正式字段。

运用机读编目格式组织和整理网络信息资源的研究项目中,最有代表性的是 OCLC 的因特网编目计划(简称 Internet Cat)。它旨在发动全国图书馆界密切合作,共同为网络信息资源建立一个可供检索而且形式与 AACR II 和 USMARC 保持一致的目录,并提供资源的地址和利用。1996 年 2 月,OCLC 又推出一个收录网络资源的数据库 Net First,采用自动收集、验证技术和传统的标引与制作文摘技术,从各种网络资源中选择并记录一批被判定有长远保留价值的信息,排除了质量不高的和有关个人的信息。收录范围包括万维网主页、Gopher 服务器、FTP 站点、电子论坛、Listservs、电子期刊和图书馆目录等。

除了 OCLC 的代表性项目外,还有一些单位在网络资源收集整理方面做了尝试。美国依阿华州立大学图书馆发起了网络资源服务计划“CYBER Stacks”,目的是组织和整理网络中科技方面的参考型工具信息;丹佛大学图书馆学和信息服务学系利用标引方法对网络中有关工程方面的信息资源进行组织、分类标引和提供检索等。

在我国,一些图书馆和情报单位利用搜索引擎筛选网上信息资源,在交通、军事、医学等领域建立了专业性的网络信息资源主题指南。颇具代表性的是交通部科技信息研究所 1995 年 8 月开始的“Internet 国际互联网交通信息资源开发利用研究”这一项目。该所利用 Yahoo 在万维网上找到一批“交通信息源”,如联机出版物、邮政名单(Mailing List)、新闻组、虚拟图书馆、光盘服务、Dialog 联机服务和与交通有关的学术机构、政府机构、非盈利机构等,然后按照中英文切换的

主题指南,通过超文本置标语言(HTML)加于组织,在补充国内外有关的交通信息源基础上,形成完整的“世界交通站点”的框架,并且进行及时的搜寻和更新。<sup>[3]</sup>在联合国教科文组织的支持下,中国科学院化工冶金研究所建立了 ChIN(The International Chemical Information Network),其 Web 主页除了对 ChIN 的系列学术会议等进行介绍外,重点对 Internet 上重要的化学资源进行了系统的收集和整理,使 ChIN 主页成为 Internet 化学资源的窗口。<sup>[4]</sup>另外,我国著名的情报检索专家陈光祚教授目前正在进行一项图书情报学虚拟图书馆建设的科研项目,拟利用 Alta Vista 等搜索工具对有关该学科的网站网页进行搜索和收集,加于鉴定核实,并对核实后的网址进行合理组织,使之成为能够提供检索、浏览和链接的信息集合。在近一年的时间中,已从网上搜集了近 1000 个有关网站网页的网址,为其中某些条目进行了标引及撰写摘要,同时正编制虚拟图书馆的软件,进行初步的调试。计划将在武汉大学大众传播与知识信息管理学院的网站上用这个虚拟图书馆向公众提供服务。<sup>[5]</sup>这必将为利用图书情报专业的网络信息资源提供极大的便利。

因此,根据条件和需要,对因特网上某领域的信息资源进行收集、组织、整理和有序化的资源重组,建立为特定用户群所需要的信息资源指南系统,是提高管理网络信息资源能力的有效方式。

#### 参考文献

- 1 <http://www.cnnic.netcn/Internet>
- 2 消息. 中国信息导报,1999,(1)
- 3 隋利玲 郭瑜. Internet 上专业性信息资源索引库的建设. 现代图书情报技术,1997,(2)
- 4 李晓霞等. Internet 上的化学数据库. 科学数据库与信息技术论文集. 科学出版社,1998 年 6 月
- 5 陈光祚. 论“图书情报学虚拟图书馆”的建设. 中国图书馆学报,2000,(1)

#### 《现代图书情报技术》(1980—2000)光盘珍藏版征订通知

ISBN 7-980036-63-8/G. 11

由中国科学院主管、中国科学院文献情报中心主办的信息管理专业技术刊物《现代图书情报技术》(原名《计算机与图书馆》1980 年创刊)杂志。“全文检索数据库系统”光盘版已由中国新闻出版署批准将近期出版。该光盘汇集了《现代图书情报技术》创刊 20 年来(1980—2000)编辑出版的全部内容,集中展现和凝练了我国图书情报领域学习采用现代化技术的启蒙—学习—试验—起步—普及—发展—应用—提高—进步的各个历史发展阶段的进程概况。《现代图书情报技术》光盘版集创刊号至今,共出版 102 期。数据盘以全文录入方式压缩制作,真实记载了期刊原貌,并可任意查找、检索、浏览。《现代图书情报技术》光盘版是跨世纪的数字化进程中具有独特历史意义的电子出版物,无论是实用性还是珍藏性,《现代图书情报技术》光盘版都是不可多得的收藏珍品。

目前,《现代图书情报技术》光盘版,由《现代图书情报技术》编辑部独家销售,每盘售价:人民币 180 元,另加邮寄费 20 元,总计每份人民币 200 元。欲购者可通过邮局或银行汇款至北京中关村科学院南路 8 号《现代图书情报技术》编辑部,款到发货。

联系人:李春源 联系电话/传真:(010)6255167 E-mail:lichy@mail.las.ac.cn

收款单位:中国科学院文献情报中心《现代图书情报技术》编辑部

开户银行:北京市工商银行海淀分理处 帐号:0200004509088129221 户名:中国科学院文献情报中心

邮政编码:100080 地址:北京中关村科学院南路 8 号

(汇款单务必写明收件人详细地址、邮编、姓名、份数、字迹清楚,以免发生差错)