

SHUZI
DANGANGUAN
GAILUN

数字档案馆概论

李国庆 主编

中国档案出版社

SHUZI
DANGANGUAN
GAILUN

数字档案馆概论

116

数字档案馆概论

主 编 李国庆
副主编 方燕 左维琪

中国档案出版社

809

责任编辑/刘 琛

封面设计/沧浪工作室

图书在版 (CIP) 数据

数字档案馆概论/李国庆主编. —北京: 中国档案出版社, 2003. 6

ISBN 7-80166-279-2

I. 数... II. 李... III. 数字技术—应用—档案馆
IV. G270.7

中国版本图书馆 CIP 数据核字 (2002) 第 089929 号

SHUZI DANGANGUAN GAILUN

出版/中国档案出版社 (北京市西城区丰盛胡同 21 号)

发行/全国新华书店发行

印刷/北京振兴源印务有限公司

规格/850×1168 1/32 印张/8.625 字数: 230 千字

版次/2003 年 6 月第 1 版 2003 年 6 月第 1 次印刷

印数/2000 册

定价/22.00 元

卷首语

档案信息化建设是国家信息化建设的一个重要组成部分。加强档案信息化建设,加快档案工作融入信息社会的步伐,是档案事业适应时代和社会发展的必然选择,是加速档案管理现代化的客观要求,是提高档案信息服务水平的必由之路。数字档案馆建设是档案信息化建设的一项重要内容,也是一项全新的工作。深圳市档案局的同志们结合深圳数字档案馆建设实际,及时把实践的经验上升到理论的层面加以总结和概括,这种刻苦钻研、勇于探索的精神值得称道。希望今后有更多的档案工作者进一步解放思想,开拓创新,努力学习、研究和运用当代先进的科学知识与科技手段,加快推进档案信息化建设,使档案事业实现跨越式发展,为全面建设小康社会做出新的贡献。

毛福民

2003年4月

目 录

第一章 数字档案馆概述	(1)
第一节 数字档案馆建设的背景	(1)
一、全球信息化的大趋势	(1)
二、国内的环境基础	(3)
三、深圳率先建设中国第一个数字档案馆的基本依据	(5)
第二节 数字档案馆的概念、分类和特点	(8)
一、数字档案馆的概念及含义	(8)
二、数字档案馆的分类	(11)
三、数字档案馆的特点	(11)
四、数字档案馆建设工程的创新性	(12)
第三节 数字档案馆在信息社会中的地位及作用	(13)
一、数字档案馆在信息社会中的地位	(13)
二、数字档案馆在信息社会中的作用	(15)
第四节 数字档案馆的主要功能	(15)
一、通过多层面的信息采集,形成社会综合信息资源库	(15)
二、海量数据存储	(16)
三、档案信息的有效访问与查询	(16)
四、档案信息的发布与传递	(17)
五、数字档案信息的安全保护	(18)
六、数字档案信息资源的系统管理	(19)
第五节 数字档案馆研究现状	(19)

一、国际研究情况	(19)
二、国内研究情况	(23)
第二章 数字档案馆的组织和管理模式	(25)
第一节 狭义数字档案馆的组织模式及其特征	(25)
一、狭义数字档案馆的组织模式	(25)
二、狭义数字档案馆的特征	(26)
第二节 广义数字档案馆的组织模式及其特征	(28)
一、广义数字档案馆的组织模式	(28)
二、广义数字档案馆的特征	(30)
第三节 数字档案馆的信息管理模式及其特征	(30)
一、从传统的数据库模式转变为智能知识库模式	(30)
二、信息处理方法从传统的逻辑运算向模糊运算方式 扩展	(31)
三、检索语言从人工语言向自然语言转变	(32)
四、信息检索模式从精确定位向概念搜寻转变	(32)
第三章 数字档案馆的建设内容	(34)
第一节 数字档案馆建设总体规划	(34)
第二节 基础设施建设	(36)
一、数字档案馆的网络环境建设	(36)
二、硬件基础设施建设	(40)
三、系统软件建设	(43)
第三节 信息资源建设	(44)
一、信息资源建设的原则	(44)
二、信息资源的来源和采集范围	(46)
三、传统档案馆在数字档案馆资源建设中的地位	(48)
四、数字档案馆提供的信息资源与互联网上信息资源 的不同	(50)
第四节 应用系统建设	(51)

一、信息采集模块	(51)
二、信息管理模块	(52)
三、信息利用模块	(52)
四、系统维护模块	(53)
第五节 标准规范建设	(53)
一、使用统一标准规范的目的和意义	(54)
二、数字档案馆标准规范建设的原则	(56)
三、数字档案馆标准规范体系建设的主要内容	(58)
第六节 人才队伍建设	(61)
一、数字档案馆所需人才类型及建设原则	(62)
二、数字档案馆人才的业务素质要求	(63)
第四章 数字档案馆应用系统	(67)
第一节 概述	(67)
一、系统设计原则及目标	(67)
二、系统性能	(70)
第二节 应用系统的结构与功能	(72)
一、数字档案馆应用系统的逻辑结构	(72)
二、数字档案馆应用系统的功能	(73)
第三节 系统的应用	(105)
一、信息资源加工	(105)
二、信息资源开发	(107)
三、档案资源及馆务管理	(107)
四、信息网络化传递	(108)
第五章 数字档案信息资源的建设与利用	(110)
第一节 数字档案信息数据的构成	(110)
一、文本文件	(110)
二、图形文件	(111)
三、图像文件	(111)

四、影像文件·····	(111)
五、声音文件·····	(111)
六、其它文件·····	(112)
第二节 数字档案信息资源的获取·····	(112)
一、馆藏档案的数字化·····	(112)
二、数字档案的接收·····	(113)
三、社会散存档案数字信息的收集·····	(118)
四、网上数字信息的收集·····	(118)
第三节 数字信息资源的整序·····	(119)
一、整序的概念·····	(119)
二、整序的内容·····	(120)
三、整序的标准和方法·····	(120)
第四节 数字信息的保管与迁移·····	(120)
一、数字信息保管与迁移的必要性和可能性·····	(120)
二、数字信息的存储·····	(122)
三、备份和异地存放·····	(123)
四、数字信息的鉴定与销毁·····	(123)
五、数字信息的迁移·····	(124)
第五节 数字档案信息资源的利用和服务·····	(124)
一、数字档案的利用和服务环境·····	(125)
二、数字档案利用和服务的对象·····	(126)
三、数字档案的利用和服务方式·····	(127)
第六章 数字档案馆的关键技术·····	(131)
第一节 概述·····	(131)
第二节 数字化信息生成技术·····	(133)
一、扫描技术·····	(133)
二、识别技术·····	(135)
三、图像处理技术·····	(139)

四、音频转换技术·····	(142)
第三节 数字档案的标识和存储技术·····	(143)
一、可扩展标记语言 XML·····	(143)
二、数字档案的压缩·····	(155)
三、海量数据存储系统·····	(162)
第四节 数据自动分类及智能化检索技术·····	(170)
一、自动分类技术·····	(170)
二、关键词检索技术·····	(175)
三、全文检索技术·····	(177)
四、多媒体检索技术·····	(184)
第五节 数字档案的安全和保护技术·····	(199)
一、数字档案的加密·····	(200)
二、数字水印·····	(208)
三、身份认证·····	(211)
第六节 系统体系结构所涉及的关键技术·····	(216)
一、中间件技术·····	(216)
二、集群技术·····	(219)
第七章 数字档案馆研究的相关问题·····	(223)
第一节 数字档案馆与传统档案馆的关系·····	(223)
一、数字档案馆不可取代传统档案馆·····	(223)
二、数字档案馆与传统档案馆互相依存·····	(224)
第二节 数字档案馆与数字图书馆的关系问题·····	(226)
一、数字档案馆与数字图书馆的异同·····	(226)
二、数字档案馆与数字图书馆的互补关系·····	(228)
第三节 工程建设的阶段性问题·····	(229)
一、初创阶段,构建基本框架·····	(229)
二、与时俱进,不断发展·····	(231)
第四节 法律问题·····	(231)

一、电子文件的法律效力问题·····	(232)
二、数字档案馆系统的知识产权保护问题·····	(232)
三、数字档案馆在可能从事的商业行为中的法律约束 问题·····	(236)
第五节 安全问题·····	(237)
一、物理安全·····	(237)
二、网络安全·····	(238)
三、信息安全保密·····	(239)
四、安全保密管理·····	(243)
附录一：相关标准规范（题录）·····	(245)
附录二：重要名词解释·····	(247)
附录三：主要参考文献·····	(256)
后记·····	(262)

第一章 数字档案馆概述

第一节 数字档案馆建设的背景

人类社会迈入 21 世纪,信息技术革命继续高涨。经济全球化、社会信息化的浪潮,不仅加速了人类社会历史发展的进程,而且日益深刻地改变着人们的工作、生活及思维方式。信息经济已成为世界经济发展的新动力和新增长点,信息产业升级为发达国家的支柱产业,信息化成了当代国际竞争的战略制高点,是当今世界经济和社会发展的的大趋势,也是我国产业优化升级和实现工业化、现代化的关键环节。数字档案馆建设项目的提出具有厚实的国际、国内环境基础。

一、全球信息化的大趋势

1993 年,美国总统克林顿在上任不久就提出了建立“信息高速公路”的信息化发展计划,其中包括投资 4000 亿美元建立美国“国家信息基础结构”。紧随其后的是英国、法国、日本、加拿大等国家先后提出了自己的“信息高速公路”建设规划。1994 年,美国又提出兴建“全球信息基础结构”的建议,这是一个全球性的高科技发展目标,其目的是实现全球范围的信息联网和各国之间的信息共享。1998 年 1 月 31 日,在美国加利福尼亚科学中心召开了 Open GIS Consortium 年会,在会上,副总统戈尔提出了“数字地球”(Digital Earth)计划,旨在建立以地球坐标为依据的、具

有多分辨率的、有海量数据组成的、能立体表达的虚拟地球。该计划将由美国宇航局（NASA）协同其它部门组织实施，计划至2005年实现，在2020年正式建成。

无论是“信息高速公路”，还是“数字地球”，这种信息化都是以数字化、网络化、智能化、可视化为基础的，是建立在计算机体系和通信网络基础上的一种数字时空。其中，数字化是一切的基础，只有数字化以后，各种信息资源才能够在网络空间得以生存和利用。从上世纪90年代以来，各种数字概念和名称已经充满了我们的生活，数字中国、数字城市、数字政府、数字长江、数字社区、数字家庭等众多概念通过各种媒体和渠道进入了我们的大脑。与此同时，计算机技术和网络技术的发展已经进入了第三代。以高速、宽带主干网为基础的信息网络正迅速建立起来，信息技术的发展正向5W的方式转变：即任何人（whoever）可以在任何时间（whentime）、任何地点（wherever）都可以和任何一个其他人（whomever）通过网络进行通信，已达到传递任何信息（whatever）的目的。

所谓“数字城市”是综合运用GIS（地理信息系统）、遥感、遥测、多媒体及虚拟仿真等技术，对城市基础设施、功能机制及其运行的各种信息，进行自动采集、动态监测管理和辅助决策服务的技术系统。该系统在城市规划、建设与运营管理以及城市生产与生活中，运用数字化信息处理技术和网络通讯技术，将城市的各种数字信息及各种信息资源加以整合，并充分发挥其在城市各领域中的作用。

在“数字城市”中，电子政府的建设正在凸现。北美、欧洲、日本等IT业发达国家将通过网络进行的政务活动称为电子政府，其含义涵盖了政府这个角色该做的所有事情。我国“十五”科技攻关计划从两个方面揭示电子政务概念的含义：一是内部办公，实现内部信息共享、内部资源整合、内部决策支持；二是对外提

供公众服务，凡是当前政府应该做的事情，如网上工商、网上税务、网上社保等，均涵盖于电子政务之中。

二、国内的环境基础

全球信息化的大趋势促使以计算机技术、通信技术、网络技术以及多媒体技术为基础的网络环境的形成，也推动了我国国家信息化的进程。近年来，我国主要城市的信息化建设取得了长足的发展。上海正在发展“信息港”；香港提出建设“数码港”；“数字北京”轮廓初现，北京的城市宽带骨干网络和接入网络基本覆盖全市，首都公用信息平台基本建成，中央在京及市属部门的各类数据库相继启用。“首都之窗”是涵盖人大、政协、市委、市政府的统一的国家机关网站，其面向百姓的信息服务受到群众欢迎。全市公用的电子商务服务平台——首都电子商城初步解决了电子商务涉及的CA认证、安全配置、支付平台等关键环节。科教网的开通为全市科技信息交流、科技普及、成果发布创造了良好的环境。此外，数字保障信息系统中的医疗保险系统已经试运行，北京市民卡已试点发放，银行工程做到全市联网，市政交通一卡通已启动建设，城区服务网络及呼叫中心已经建成，一批信息化小区已发挥了示范作用。在改革开放最早，经济较发达的广东省，信息化已经有向基层普及的趋势。广州市已进入了数字城市的实施阶段，如广州市的一些老牌企业——珠江啤酒、广药集团、广州市邮政等，通过充分运用信息化高科技手段，发展迅速，业绩斐然，有力地展示了“数字广州”的风采。广东省的县级市——顺德、开平、南海市，都在加紧数字政府的建设，进展很快。尤其是南海市推进信息化建设的经验，引起社会各界的普遍关注。南海市政府全面应用、推进网络化，成为这个地区推行信息化的龙头。该市目前已有70%的政府职能部门通过统一的信息平台进行联网，实行无纸办公、网上传递和处理政务，电子政务在这里已

成为现实。

在此基础上，中国电子政务应用示范工程也逐步进入人们的视线。“电子政务示范应用工程”站在国家的高度，从全局考虑，计划构建一个统一标准、统一业务规范，能够实现信息共享，全部采用自主信息安全技术、彼此之间可以相互认证和鉴别的电子政务平台。为确保国家安全、信息共享、避免投资浪费，国家正在通盘制定电子政务的总体发展思路 and 标准，并确定深圳市作为建设电子政务的试点城市之一。

国家信息化的推进，电子政务的应用，正在对档案和档案工作产生着巨大的影响，促使档案管理的信息化。档案信息化是档案管理模式由传统的以档案实体保管为重点，向以档案实体的数字化信息为主要形式面向社会提供服务的重心转变的过程。随着数字政府的建立，政府各部门将大量产生和应用电子文件，使档案信息的形态从原来以纸质文件、照片、录音录像等为主，转变为以数字形式为主。人们获取档案信息的需求也将从到馆调阅案卷转变为从网上索取信息；从查阅历史文献型档案信息，转变为广泛索取各种时效性较强的有序的现行文件信息。面对这样的形势，档案部门能否顺应发展，积极应对，使档案和档案工作融入信息社会大环境，充分运用数字文档、数字视频能够通过网络迅速实现信息传递的特点，研究出一套相对完善的数字档案管理方案，建立数字档案管理的基础设施，不仅会影响和制约各领域信息化的进程，同时对档案机构的生存、作用的发挥及其社会地位也将产生决定性的影响。

正是在这样的背景下，由国家档案局科研所、深圳市档案局、深圳世纪科怡科技发展有限公司共同研究、开发的全国第一个数字档案馆科研建设项目才应运而生。通过数字档案馆的研究和建设，将能圆满地实现档案信息管理和提供利用方式的转变，突破传统档案管理和提供利用的局限，真正使档案馆从封闭走向开放，

走向社会，最大限度地满足政府决策、运作和社会大众的需求，使自身地位得到巩固，技术得到进步，水平得到提高，与社会协调发展。凭借数字档案馆还可以用我国丰富的档案信息去改变目前国际互联网上中文信息偏少，源自中国的中文信息更少的现象，参与解决国际信息环境中的中、西文化比例失调问题，营造有利于我国现代化建设和发展的国际环境。由此可见，数字档案馆的开发、研究和建设，意义是十分重大的。

三、深圳率先建设中国第一个数字档案馆的基本依据

由深圳市率先建设全国第一家数字档案馆，是国家档案局在充分比较了各地信息化建设和各方面情况后的正确选择。其根本原因在于深圳市档案信息化建设具有四大优势：

1. 环境优势。作为我国改革开放的窗口和试验田，深圳已经国务院批准成为我国信息化建设的试点城市。深圳市政府对信息化建设极为重视，近年来出台了一系列信息化建设管理规范；并且逐年增加信息化建设经费，在实施信息化建设中取得了引人注目的成就，从而为档案信息化建设营造了良好的整体环境氛围。

“九五”期间，深圳是我国信息化的带头和示范城市。深圳市在建立符合中国国情的信息化理论和实施信息化的发展战略上，敢于理论创新和实践突破；在探索信息化的相关基本概念、理论原则、发展战略、操作方法、重点项目选择和实施等方面，初步形成了一整套有中国特色的信息化发展模式，并在1997年首次全国信息化工作会议上得到了充分肯定和高度评价。深圳率先建立了信息组织领导体系和信息法规框架，初步建成了宽带IP城域网和交互有线电视网，信息资源的开发和利用取得了显著的进展。政府各部门普遍开展信息化建设，提高了政府机构的办事效率；多数大中型企业建立了信息系统，信息应用遍及金融、税收、公安、

人事、劳动、社保、交通、口岸、海关、城管、科技、教育、卫生等 20 多个重要社会公共领域，建成各类数据库 200 多个。

“十五”期间，深圳市信息化建设的总目标是：把深圳信息化推进到世界发达国家水平，把深圳建设成用信息化加速现代化的示范城市，成为国家信息化创新试验城市，建成信息化深圳。具体目标是：建成具有国际先进水平的信息基础设施；建立电子商务综合服务平台；基本实现政务信息化；以信息技术应用带动经济结构的调整升级；基本形成城市信息化综合服务体系；形成以移动通信、光电子、计算机及网络、软件、集成电路、数字视听和信息服务业等七项重点产业为主体的信息产业特色；信息终端应用普及率达到新水平；建立信息技术及产品创新的研究开发基地。

2. 政策优势。深圳市被广东省档案局确定为率先基本实现档案管理现代化的示范市。新近制定的《深圳市国民经济和社会信息化“十五”规划》（2001—2005），确定深圳市政务信息化建设以“三网一库”（各部门的办公业务网、政府办公业务资源网、政府公众信息网、政府办公业务资源数据库）的建设为目标，开展广泛的电子政务应用。将建成电子政府信息网、城市运行监控及应急指挥中心、口岸管理综合信息网络、对外经济贸易管理信息网络、人口管理信息系统、社会治安信息系统、财政综合业务办公管理系统等。深圳市数字档案馆被纳入到政务信息化建设中，成为电子政府建设的重要组成部分之一，并作为“十五”期间市信息化重点建设项目之一。数字档案馆将作为政府信息资源库，成为政府信息资源中心和数据转换站，成为市民查询有关政府文件、法律法规和政府办公信息的窗口。在建的“市民中心”是全市行政、文化中心和商务中心，是一项综合性、高科技、智能化的系统工程建设项目。市政府将市数字档案馆建设纳入了“市民中心”建设总体规划，并要求与“市民中心”建设同步进行。由市