

数字档案

集中管理研究

赵豪迈 著



陕西师范大学出版总社

数字档案集中管理研究

赵豪迈◎著

陕西师范大学出版总社

图书代号 ZZ16N1150

图书在版编目(CIP)数据

数字档案集中管理研究 / 赵豪迈著. —西安: 陕西师范大学出版总社有限公司, 2016. 6

ISBN 978-7-5613-8489-3

I. ①数… II. ①赵… III. ①数字技术—应用—档案管理—研究 IV. ①G270.7

中国版本图书馆 CIP 数据核字(2016)第 128693 号

数字档案集中管理研究

SHUZI DANGAN JIZHONG GUANLI YANJIU

赵豪迈 著

责任编辑 / 冯新宏 杜世雄

责任校对 / 李彦荣

封面设计 / 鼎新设计

出版发行 / 陕西师范大学出版总社
(西安市长安南路 199 号 邮编 710062)

网 址 / <http://www.snupg.com>

经 销 / 新华书店

印 刷 / 北京京华虎彩印刷有限公司

开 本 / 787mm × 1092mm 1/32

印 张 / 14.5

字 数 / 300 千

版 次 / 2016 年 6 月第 1 版

印 次 / 2016 年 6 月第 1 次印刷

书 号 / ISBN 978-7-5613-8489-3

定 价 / 45.00 元



读者购书、书店添货如发现印刷装订问题,请与本社高教出版分社联系调换。
电话:(029)85303622(传真) 85307864

摘 要

伴随着现代信息技术在档案行业的广泛应用,电子文件归档、馆藏档案数字化、档案信息网络化,对数字档案管理提出了新的挑战。本研究在分析数字档案的内涵、特点和发展状况的基础上,从集中管理概念、历史、前沿和现状等方面,对数字档案集中管理进行了深入探讨,论述了数字档案集中管理理论、技术、方法和机制,以及数字档案集中管理的制度与策略,在此基础上,提炼出了适合我国国情的几种数字档案集中管理模式。

本研究认为,电子文件具有非人工识读性、系统依赖性、易于传递与复制性等特点。数字档案从本质上来讲也是电子文件,但相对于电子文件,数字档案有明显的“档案”属性。数字档案虽然功能与传统的纸质档案基本相同,但由于它是新载体和新形态,相应地呈现出档案信息的数字化、物理结构的复杂化、信息内容的传输网络化等特征。数字档案馆则是档案界近年来提出的新概念,与数字档案有密切的关系,数字档案馆建设对数字档案管理具有基础性的作用,是近年来档案界推进数字档案管理的重要抓手。

研究认为,目前,我国电子文件管理实践仍然面临着缺乏制度规范、电子文件管理系统功能不完善、电子文件质量难以保障等诸多问题。即便是归档后的电子文件,因为缺少科学的管理措施,也面临无法正确划分保管期限、无法保证长期可读性、无法维

护其真实性等多重风险。因为前后端管理脱节、管理功能不完善、管理制度不健全、管理技术落后等一系列问题,大量的电子文件基本处于无序、无规的分散、混乱状态,导致我国电子文件管理面临严重的失控风险。

研究认为,我国目前对数字档案的认识与实践水平都比较低,大量数字档案失存、失信、失用,导致机构业务活动失去法律凭证与信息支撑,信息资产流失。从宏观来看,我国尚未建立起科学、有效的数字档案管理体系,相当数量的电子文件处于失控状态,使国家历史、民族记忆面临断层与割裂的危险,造成国家信息资源控制力的缺失。尤其在当前数字档案管理项目建设热潮中,由于缺乏国家层面的宏观规划、有效协调、科学指导,许多数字档案管理项目定位不清、功能不明、低水平重复、缺乏共建共享机制,在项目质量、效率和效益上都存在不少问题。

研究认为,数字档案对信息时代的国家和社会管理具有战略意义,根据我国国情,只有集中统一管理电子文件和数字档案,数字档案才能发挥更大的作用。我国目前关于数字档案集中管理的整体设计、制度规范和标准体系尚不完善,相互之间的衔接尚不紧密,对数字档案的集中管理尚没有提升到国家战略层面,造成一定程度上的管理分散和资源失控,亟须在宏观战略上进一步明确数字档案集中管理的国策,并在微观上通过制度建设、条件建设、环境建设和基础设施建设等方面予以落实。从国际经验来看,目前信息化走在前列的发达国家,如美、英、澳、加等西方发达国家在国家层面的电子文件管理上均采取了集中式保管,其中,澳大利亚——分布式保管的倡导者——的“易帜”更是强有力地向人们证明了集中式保管模式是各国电子文件管理的现实选择和大势所趋。

研究认为,我国数字档案管理目前存在诸如管理基础薄弱、国家对数字档案掌控能力不足、数字档案开发利用严重不足、数

字档案安全问题突出等问题。同时,电子时代数字档案管理遇到诸多新情况,如数字档案集中管理的体制机制不健全,各种理念、规范、制度、人力,以及技术等多方面因素之间互相交错、互相重合等,导致当前在电子政务环境下我国数字档案集中管理存在着较多问题。有鉴于此,本研究结合电子文件生命周期理论、文件连续体理论、信息资源管理理论、面向服务的 IT 集约化管理理论等,为数字档案集中管理提供了理论依据。此外,本研究详细介绍了前端控制技术、分布式多元控制技术、档案元数据处理技术、封装包技术、数字档案目录索引技术、信息安全与等级保护技术、异构信息整合技术等数字档案集中管理技术。

数字档案管理与大数据关系密切,本研究提出了分类管理方法、分级管理方法、分布式网络管理方法、元数据管理方法、数字档案利用服务方法,以及数字档案长期保存方法等六种数字档案集中管理方法,以实现对数字档案的集中管理。研究认为,面对数字档案管理混乱失控的状态,建立具有中国特色的数字档案集中管理机制尤为重要。这一数字档案集中管理机制建设就是要使当前数字管理机构能在国家宏观管理规划的前提下,统一顶层设计,统一标准,统一管理,统一开放方式,实现我国数字档案集中管理的协调发展。研究认为,数字档案集中管理制度建设是个系统工程,并从传统档案集中管理制度体系概要、数字档案集中管理制度的新挑战、我国数字档案集中管理制度问题、数字档案集中管理制度创新分析、数字档案集中管理制度建设准则、数字档案集中管理制度保障因素、数字档案集中管理制度体系建设等七个方面,提出了数字档案集中管理制度建设的创新思路。

研究认为,数字档案有其自身的特性,也面临数字时代特有的风险,按照文件生命周期和文件连续体等理论,在借鉴其他国家先进经验,结合我国具体情况,参考数字档案集中管理的相关技术、方法、体制等因素的基础上,本研究构建了理顺管理体制机

制、完善管理标准体系、强化集中统一管理、实行正确管理策略等我国数字档案集中管理的策略,以提高我国数字档案集中管理的质量与水平。

研究根据数字档案分类管理和分级管理相结合的管理策略,参考信息技术领域的分布式网络管理理论、模块化系统设计与集成理论、工作流管理理论等,提出应在数字档案集中管理模式设计中采用工作流模型管理,并详细介绍了这些模型的功能。这些模型既相互独立,又可相互集成,方便实现数据共享和信息交换。这些模型的分化组合,既适应了数字档案集中管理中不同阶段和不同环境的需求,又形成了数字档案集中管理的不同模式。这些模型包含角色分配模型、移交归档模型、数据采集模型、信息交换模型、分类鉴定模型、数据管理模型、归档存储模型、检索服务模型和系统集成模型。

研究认为,数字档案集中管理服务模式,是指数字档案集中管理服务系统的各组成要素(数字档案管理人员、数字档案信息服务人员、数字档案信息资源、数字档案用户、数字档案服务平台及服务策略与方式等)之间相互关系的组合。在数字环境下,由于数字档案信息的特殊属性,档案实体可集中,亦可分散,档案信息可集中,亦可分散,从而形成了数字档案集中管理的不同模式。

按照数字档案实体集中与信息集中的不同强弱程度,本研究把数字档案集中管理区分为国家文件档案中心模式、文件档案条块结合模式、数字档案文档分离模式、数字档案备份呈缴模式等四种集中管理服务模式。同时,本研究认为,数字档案集中管理还可以探索出其他创新模式,并会兼顾上述模式各自的不同特点,或者有比较独特的管理服务内容,形成独特的数字档案集中管理模式,如数字档案登记备份模式,同时,详细分析了这五种类型的数字档案集中管理服务模式的不同特征。

陕西师范大学优秀著作出版基金资助出版

国家社科基金项目“数字档案集中管理策略与模式研究”（项目编号：09BTQ030）结项成果

目 录

第一章 数字档案集中管理概述

- 1.1 数字档案的基本情况 (1)
 - 1.1.1 电子文件的概况 (1)
 - 1.1.2 数字档案的定义 (7)
 - 1.1.3 数字档案的特点 (9)
 - 1.1.4 数字档案馆 (12)
- 1.2 数字档案的发展状况 (15)
 - 1.2.1 电子文件状况 (15)
 - 1.2.2 数字档案状况 (29)
- 1.3 数字档案集中管理概况 (38)
 - 1.3.1 数字档案集中管理概念 (38)
 - 1.3.2 数字档案集中管理分析 (43)

第二章 数字档案集中管理历史、前沿与现状

- 2.1 数字档案集中管理的现实意义 (51)
 - 2.1.1 数字档案集中管理的战略意义 (51)
 - 2.1.2 数字档案集中管理的现实需要 (53)
 - 2.1.3 数字档案管理模式与路径选择 (56)
- 2.2 国外数字档案集中管理历史与前沿 (67)
 - 2.2.1 国外数字档案管理的历史发展 (67)
 - 2.2.2 国外数字档案集中管理的现状 (69)

2.2.3	国外数字档案集中管理的经验	(79)
2.3	我国数字档案集中管理的现状	(83)
2.3.1	我国数字档案管理基础薄弱	(83)
2.3.2	国家数字档案掌控能力不足	(87)
2.3.3	数字档案开发利用严重不足	(92)
2.3.4	数字档案安全问题令人担忧	(94)
2.4	电子政务下的数字档案集中管理	(95)
2.4.1	我国电子政务发展现状	(95)
2.4.2	电子政务下的数字档案集中管理	(97)
2.4.3	电子政务中制约数字档案集中管理的因素	(99)

第三章 数字档案集中管理理论、技术、方法与机制

3.1	数字档案集中管理理论	(103)
3.1.1	传统文件生命周期理论	(103)
3.1.2	电子文件生命周期理论	(110)
3.1.3	文件连续体理论	(114)
3.1.4	信息资源管理理论	(118)
3.1.5	面向服务的 IT 集约化管理理论	(124)
3.2	数字档案集中管理技术	(130)
3.2.1	前端控制技术	(130)
3.2.2	分布式多元控制技术	(135)
3.2.3	档案元数据处理技术	(137)
3.2.4	封装包技术	(148)
3.2.5	数字档案目录索引技术	(154)
3.2.6	信息安全与等级保护技术	(162)
3.2.7	异构信息整合技术	(167)

3.3	数字档案集中管理方法	(173)
3.3.1	数字档案分类管理方法	(173)
3.3.2	数字档案知识组织方法	(179)
3.3.3	数字档案分级管理方法	(187)
3.3.4	数字档案分布式管理方法	(190)
3.3.5	数字档案元数据管理方法	(196)
3.3.6	数字档案利用服务方法	(204)
3.3.7	数字档案长期保存方法	(216)
3.4	数字档案集中管理机制	(224)
3.4.1	数字档案集中管理机制分析	(224)
3.4.2	统一顶层设计	(234)
3.4.3	统一标准	(239)
3.4.4	统一管理	(264)
3.4.5	统一开放方式	(272)

第四章 数字档案集中管理制度与策略

4.1	数字档案集中管理制度	(280)
4.1.1	传统档案集中管理制度体系概要	(280)
4.1.2	数字档案集中管理制度的新挑战	(285)
4.1.3	我国数字档案集中管理制度问题	(291)
4.1.4	数字档案集中管理制度创新分析	(298)
4.1.5	数字档案集中管理制度建设准则	(305)
4.1.6	数字档案集中管理制度保障因素	(308)
4.1.7	数字档案集中管理制度体系建设	(312)
4.2	数字档案集中管理策略	(322)
4.2.1	理顺管理体制机制	(323)
4.2.2	完善管理标准体系	(328)

4.2.3	强化集中统一管理	(332)
4.2.4	实施正确管理策略	(334)
第五章 数字档案集中管理模式		
5.1	数字档案集中管理工作流模型	(342)
5.1.1	角色分配模型	(344)
5.1.2	移交归档模型	(347)
5.1.3	数据采集模型	(351)
5.1.4	信息交换模型	(354)
5.1.5	分类鉴定模型	(358)
5.1.6	数据管理模型	(366)
5.1.7	归档存储模型	(369)
5.1.8	检索服务模型	(373)
5.1.9	系统集成模型	(378)
5.2	数字档案集中管理服务模式	(392)
5.2.1	国家文件档案中心模式	(395)
5.2.2	文件档案条块结合模式	(402)
5.2.3	数字档案文档分离模式	(407)
5.2.4	数字档案备份呈缴模式	(411)
5.2.5	文件档案登记备份模式	(415)
5.2.6	集中管理服务模式比较	(419)
参考文献		(423)
后记		(456)

第一章 数字档案集中管理概述

1.1 数字档案的基本情况

1.1.1 电子文件的概况

电子文件是数字档案的前身,数字档案本身从其基本特性来看也是电子文件。因此,要了解数字档案,先得了解电子文件。

一般来讲,电子文件是指在数字设备及环境中形成,以数字形式存储于磁带、磁盘、光盘等载体,依赖计算机等数字设备阅读处理,并可在通信网络上传送的文件。^①简单地来说,电子文件就是“电子”加上“文件”的和,或者说是“数字化”与“文件”的集合,其中“文件”概括了电子文件的社会功能属性,“数字化”表明了电子文件的技术属性。我们可以这样理解:(1)电子文件是数字化的文件。电子文件是文件的一种,它具备文件的各种属性,特别是具有特定的用途和效能。这是电子文件与一般数字化信息之间的区别,也是电子文件与传统文件的共同点。(2)电子文件

^① GB/T18894-2002,电子文件归档与管理规范(S).

是文件的数字化。电子文件是由计算机生成和处理的,其信息由0或1两种二进制数字代码记录、表示并存储于计算机存储器中。这是电子文件与传统文件的区别,也是电子文件与一般数字信息之间的共同点。

与以往的任何一种记录方式和载体形式的档案相比,电子文件都表现出很多不同的性质和特点。例如,采用二进制数字代码,非人工识读性,生成和传递过程对计算机系统依赖性强,易于传递和复制,等等。深入考察这些性质和特点,可以发现这些都与电子文件信息与载体的关系这一问题有着紧密的关系。^①具体来讲,电子文件因其自身的特殊性而具有如下特点。

(1)非人工识读性

电子文件是由计算机产生的信息,所有的这些信息都要被转化成为仅可由计算机识别的二进制代码,即由0或1组成的一连串数据,电子文件在计算机处理器中的运算,以及在存储器中的存储,都是以这种二进制代码的形式来进行。这些由0或1组成的代码数据,即便是再高级的计算机操作人员也很难直接从中读出信息。人们如果想要识别、读懂这些代码信息,就必须借助计算机系统来实现。只有通过计算机系统的各种输出设备,存储器中的那些由0或1组成的代码数据才能被转化成人们可以接受的文字、声音、图像、动画等信息形态。而所有这些信息形态并非是电子文件本身,其只不过是电子文件的数据信息经由输出设备转化以后呈现出的输出形态。这与纸张文字、图片等可由人工直接

^① 谢光耀. 动态的、相对的分隔和绝对的依附——电子文件信息与载体关系及原始性研究[J]. 档案学通讯, 2001(3): 62-65.

识读的传统文件有着明显的不同。

(2) 对计算机系统的依赖性

电子文件对计算机系统的依赖性是指相对于传统文件对人的依赖性来说的。在对传统文件进行管理时,对于传统文件的制作、整理、归档、存放等活动,管理人员可以通过多种途径和手段进行灵活的控制和管理,对电子文件的管理则有所不同,管理人员如果不借助于计算机系统就无法对电子文件的任何活动进行干预。这里所说的依赖性包含了两重含义。一方面是指电子文件的生成、传输、处理、存储、输出都需要通过计算机系统来实现,没有计算机系统,人们便无法对电子文件施加任何的影响,不仅电子文件无法被识读,甚至于根本无法产生,也就更谈不上保存和利用。另一方面则是指电子文件的数据必须要与使用它的计算机系统相匹配、相兼容。举例来说,如果要管理一种特定阅读格式的电子文件,必须要配备可以将其打开的计算机软件才行,这就要求在电子文件的管理过程中,不但要确保电子文件本身的有效性,还要保证计算机系统的有效性,保证计算机系统的不断升级和具备有效的兼容能力。

(3) 信息与载体之间动态的、相对的分离和绝对的依附性

对传统文件来说,其信息与载体是不可分离的,传统文件中的信息一旦脱离载体,信息本身也将不复存在。这一现象使得电子文件发生了变化,关于电子文件信息与载体的关系,主要存在以下几种观点。一种是以冯惠玲教授为代表的“可分离”观点,其主要内容是电子文件信息与载体之间可以分离,可分离性是电子文件诸特点中最为核心的特点,派生并影响其他特点。可分离观

点又可细分为绝对分离和相对分离两种。前者认为,“电子文件中的信息不再具有固定的物理位置,也不再对原记录载体‘从一而终’,载体的转换不仅是可能的,也是必须的”^①。后者认为,“信息一刻也离不开载体,所谓信息与载体的可分离性是相对和有条件的……就辩证唯物主义方法论而言,电子文件信息与载体是不可分离的……信息可以通过光纤网络等介质传输流动,并与原载体暂时分离,但它最终还是要存储到某一载体上”^②。这种观点,认为电子文件不再拥有固定的物理存放地点,它可以在不同的载体之间复制、编辑、修改,也可以随时改变其存放的位置,扩展或缩小其存放的空间;可以简单地从一个载体转移到另一个载体,也可以通过不同的网络传递到一个或多个已知或未知的终端。信息与特定载体的可分离性,使得电子文件本身不再具有物理意义上的实体形态,而成为了非实体形态的信息,以非实体形态存在的电子文件,其本身也不再是不可分割的整体,电子文件以全文形态存在的文件形态可以被随时、任意地分割为若干个符合逻辑关系的信息要素,这些信息要素不仅可以按照原来的逻辑关系组成原有的电子文件,还可以通过新的逻辑关系被组合成其他数字化信息。另一种是以成永付、唐小宁等为代表的“不可分离”观点,其主要内容是:“电子文件依旧是信息内容和信息载体的统一……载体是电子文件形成的物理基础,电子文件不可能脱

① 冯惠玲. 拥有新记忆——电子文件管理研究[J]. 档案学通讯, 2003(1): 57-59.

② 黄世喆, 刘勇. 论电子文件信息与载体的可分离性[J]. 广西民族学院学报(哲学社会科学版), 2001(5): 138-140.

离载体而凭空产生,载体也不能移出信息内容而独立存在。”^①

笔者认同谢光耀的观点^②,认为电子文件信息与载体存在可分离性,但这种分离是相对的、动态的,不能离开信息对载体的依附而单独地、片面地谈论分离,也不能只谈依附而不谈分离;电子文件信息与载体的关系既不是单一的可分离性,也不是单一的依附性,而是一种动态的、相对的分离和绝对的依附的辩证统一的关系;这种关系决定和影响电子文件的原始性,必须在档案的范畴内思考电子文件信息与载体的关系。

(4) 信息内容的易变性

计算机系统的存储器诸如硬盘、磁盘、优盘、磁带、可擦写光盘等都具备可读写的功能,因此存储在这些载体中的电子文件很容易被修改、增删并整理,而这些操作行为有时很难留下明显痕迹。另外,电子文件所依托信息系统的更新和补充也使得电子文件的信息内容会发生改变,这为以后鉴定电子文件的真实性、确保电子文件的完整性带来了难题。

(5) 载体的不稳定性

虽然随着信息技术的不断发展,信息存储设备的制造技术不断更新,但与传统的纸质、胶片载体相比,电子文件的载体无论是硬盘、磁盘、优盘、磁带、光盘等,其在稳定性方面的表现依然较差。计算机载体长期使用易自然磨损的问题,在常温、常湿环境

① 黄世喆,刘勇.论电子文件信息与载体的可分离性[J].广西民族学院学报(哲学社会科学版),2001(5):138-140.

② 谢光耀.动态的、相对的分离和绝对的依附——电子文件信息与载体关系及原始性研究[J].档案学通讯,2001(3):62-65.