

21

21世纪电子政务专业核心课程系列教材
全国高等院校电子政务联编教材

Managment of Digital Document of Government

政府电子文档 管理

陈次白 编 著



北京大学出版社
PEKING UNIVERSITY PRESS

中国档案学会“十二五”规划课题
中国档案学会档案学专业委员会
中国档案学会档案学专业委员会

Management of Digital Documents in Government

政府电子文档 管理

中国档案学会“十二五”规划课题

中国档案学会
中国档案学会

21 世纪电子政务专业核心课程系列教材

全国高等院校电子政务联编教材



政府电子文档管理

Managment of Digital Document of Government

陈次白 编著

G276
C380



北京大学出版社
PEKING UNIVERSITY PRESS

内 容 简 介

本书是电子政务系列教材之一,本书在讨论政府电子文档概念的基础上,较为详细、全面地介绍了电子政府文档及其管理中的管理理念、管理思想、法律地位与管理技术的许多问题。全书共分为8章,论述了关于电子政务与电子文件元数据的应用问题,多种生成政府电子文档的办公套件的功能,电子文档管理系统构建中应考虑的各种问题,政府机构中涉密电子邮件的管理技术和管理策略等问题,政府电子文档标准化和安全中的管理和技术问题,基于XML的公文管理系统和基于知识管理(含政府知识地图)的政府电子文档管理的问题。这些既涉及许多理论问题,又涉及许多相关的技术性问题,通过阅读本书,可以对政府电子文档的管理有全面、具体、深入的了解和认识。

本书可作为各类院校电子政务专业、信息管理与信息系统专业、电子商务专业及计算机信息管理等相关专业的教科书或参考书,也可供政府机关、企业部门和事业单位的领导及办公人员等读者阅读和参考。

图书在版编目(CIP)数据

政府电子文档管理/陈次白编著. —北京:北京大学出版社, 2007.5

全国高等院校电子政务联编教材

ISBN 978-7-301-11398-1

I. 政… II. 陈… III. 国家行政机关—电子文件—档案管理—教材 IV. G276

中国版本图书馆CIP数据核字(2006)第154494号

书 名: 政府电子文档管理

著作责任者: 陈次白 编著

责任编辑: 桂 春

标准书号: ISBN 978-7-301-11398-1/TP·0891

出 版 者: 北京大学出版社

地 址: 北京市海淀区成府路205号 100871

电 话: 邮购部 62752015 发行部 62750672 编辑部 62765126 出版部 62754962

网 址: <http://www.pup.cn>

电子信箱: xxjs@pup.pku.edu.cn

印 刷 者: 河北滦县鑫华书刊印刷厂

发 行 者: 北京大学出版社

经 销 者: 新华书店

787毫米×1092毫米 16开本 13.25印张 322千字

2007年5月第1版 2007年5月第1次印刷

定 价: 22.00元

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究

举报电话: 010-62752024; 电子信箱: fd@pup.pku.edu.cn

前 言

当今，信息化浪潮汹涌澎湃，席卷全球；知识经济热潮不断升温，方兴未艾。在这一波又一波的浪潮中，政府机构不可避免地卷入其中。事实上，信息化及其信息高速公路（国家信息基础设施）就是由政府正式提出并引导发展的。国家信息化与信息化社会的一个主要方面，就是政府的信息化。为此，“电子政府”应运而生，电子政务也同时出现，办公自动化空前发展。

电子文件，特别是电子公文是办公自动化的核心，是“电子政府”施政的记录，是电子政务行为的主要体现。电子文件的称谓由来已久，今天所说的电子文件应当是“数字化”的文件。但是，“电子文件”这一称谓已经不够充分了。由于“数字化”文件本身的特点以及多媒体信息的广泛应用，使其内涵已经更为丰富并发生了很大的变化。因为，现在的数字化信息往往是多种计算机文件（数字化的电子文件）的复合体。一份完整的“文件”经常是由多份不同类型的电子文件通过链接、嵌入或其他方式组合在一起的，它已经不是原来意义上的“文件”，而是所谓“文档”了。可以认为，文档即是这种复合文件的总称。此外，文档也内含档案之意，因为政府的“文件”最后都是要归档保存的。这也就是本书取名为《政府电子文档管理》的原因。当然，书中仍有“电子文件”的称呼，这就要看是在什么情况下了。

鉴于“电子文档”是建立在计算机系统处理之上，又需要政府档案管理理论的支持，因此，本书既介绍了电子文档及其管理的基本概念，又介绍了电子文档从生成、传输、处理到存档保管和利用全过程中的管理理论和管理技术。这些问题都是政府办公自动化中极为重要的甚至是生死攸关的问题，其中有些是信息化过程中出现的新技术和新问题，有些是涉及电子政务和办公自动化发展方向的重大问题。其中有些问题是作者与研究生们共同研究的成果，还有很多问题需要继续深入研究。也许，与其他同类书相比，本书论述的这些内容正是其具有自身特色的地方。

需要指出的是，由于政府电子文档，特别是电子公文往往有很多都涉及党和国家的机密，因此，我国的电子文档管理系统的开发和建设不能依靠国外的管理系统（包括 Office 套件、文档管理系统和操作系统），而应当建立在具有自主知识产权的标准和系统之上。基于这一考虑，本书着重介绍的是国内研究和开发的相关电子文档“标准”和管理系统。

本书从策划到成书，历经数年。其间，作者研读了大量的书籍资料，进行了多方面的咨询，并有多位研究生的研究成果融汇其中。因此，它是集体努力的成果和智慧的结晶。在本书的撰写过程中，得到了北京大学出版社黄庆生先生的大力支持和南京邮电大学姚国章老师的鼎力相助，研究生林磊、江静、王新昊、李毅博都做了许多工作，在此一并表示感谢。

社会信息化发展一日千里，信息技术不断更新。作者深感学识有限，书中论述恐多有疏漏与不当之处，望有识之士与读者不吝赐教。

作者于南京

2007年1月

目 录

第 1 章 文档概述.....	1
1.1 文档的概念.....	1
1.1.1 文件与文档.....	1
1.1.2 电子文档及其特点.....	3
1.2 政府文档与政府电子文档.....	7
1.2.1 政府文档的特征与作用.....	7
1.2.2 政府文档(公文)的格式.....	9
1.2.3 政府文档(公文)的分类.....	11
1.3 政府电子文档.....	12
1.3.1 政府电子文档的类型.....	12
1.3.2 政府电子文档应用中的法律问题.....	14
1.3.3 电子文件与文档及其格式发展.....	16
第 2 章 文档管理概述.....	18
2.1 文件与文档的生命周期.....	18
2.1.1 信息生命周期.....	18
2.1.2 文件生命周期.....	19
2.1.3 文档一体化与文件中心.....	23
2.2 文件与文档管理原则与基本内容.....	25
2.2.1 文件与文档的管理.....	25
2.2.2 文件管理的特征.....	26
2.2.3 文件管理的理论依据.....	27
2.2.4 文件管理的原则.....	28
2.2.5 文档管理的基本内容.....	29
2.3 电子文档管理概述.....	30
2.3.1 电子文件与文档管理的价值与原则.....	30
2.3.2 电子文档管理的功能需求.....	33
2.3.3 电子文档管理的法律法规与技术策略.....	43
第 3 章 元数据与政府电子文档.....	47
3.1 元数据.....	47
3.1.1 元数据的概念.....	47
3.1.2 元数据的作用与构建.....	50
3.1.3 元数据标准介绍.....	54
3.1.4 元数据在网络资源组织中的应用.....	59
3.1.5 各领域主要元数据发展趋势.....	59

3.2	元数据与政府电子文档管理	60
3.2.1	元数据与政府电子文档著录	60
3.2.2	美国政府信息指引服务 (GILS) 元数据	62
3.2.3	电子政务元数据标准 (E-GMS)	64
3.2.4	若干领域主要元数据	67
第4章	政府电子公文 (文档) 管理系统构建	72
4.1	政府公文 (文档) 管理概述	72
4.1.1	公文概述	72
4.1.2	传统方式的公文管理	73
4.1.3	办公自动化与电子政务	75
4.2	公文管理信息系统	78
4.2.1	公文管理系统概述	78
4.2.2	公文管理系统的功能	79
4.2.3	公文管理系统的模式结构	80
4.2.4	公文管理系统平台	82
4.2.5	公文管理系统工作流程与模块	82
4.2.6	公文管理系统存在的问题	84
4.3	基于 XML 的公文管理系统	85
4.3.1	基于 XML 格式的电子公文的优点	85
4.3.2	基于 XML 的公文管理系统的设计	86
4.3.3	基于 XML 的公文管理系统的工作模块	87
4.4	基于 XML 的公文管理系统核心模块设计	88
4.4.1	公文转换模块	88
4.4.2	公文流转自定义模块	89
4.4.3	安全与权限控制模块	90
4.4.4	公文查询与显示模块	91
4.5	政府电子文件系统的安全管理	93
4.5.1	电子政务系统中的文件管理风险与防范策略	93
4.5.2	典型的电子文件管理系统的功能需求	95
4.5.3	电子文件安全保护技术	96
第5章	政府电子邮件的管理	100
5.1	电子邮件概述	100
5.1.1	电子邮件的发展与地位	100
5.1.2	电子邮件系统简述	103
5.2	电子邮件文档的管理	107
5.2.1	电子邮件管理概述	107
5.2.2	电子邮件的管理内涵	108
5.3	政府电子邮件文件的安全管理	110
5.3.1	电子邮件传输中的安全隐患	110
5.3.2	电子邮件的安全需求与相关技术	110

5.3.3	涉密网邮件监控系统	112
5.4	电子邮件安全管理策略	117
5.4.1	电子邮件管理策略概述	117
5.4.2	垃圾邮件管理	118
第 6 章	政府电子文档管理与应用工具	122
6.1	电子文档的分类与存档	122
6.1.1	电子文件的类型	122
6.1.2	电子文件存档技术要求	127
6.1.3	档案数字化转换	130
6.2	Office 办公软件	132
6.2.1	微软公司的 Microsoft Office	133
6.2.2	中标普华 Office	133
6.2.3	Red Office	134
6.2.4	共创 Office	136
6.2.5	金山 WPS Office	137
6.2.6	永中 Office	138
6.3	电子公文文档格式转换	139
6.3.1	文档格式转换的提出	139
6.3.2	CEB 文档格式	139
6.3.3	电子文件收集中格式转换的要求	142
6.4	中文电子文档的 XML 格式	142
6.4.1	电子文档 XML 格式问题的提出	142
6.4.2	XML 文档格式标准	143
6.4.3	利用 XML 定义中文办公软件文档格式标准	144
6.4.4	coXML 设计原则	145
第 7 章	基于知识管理的政府电子文档	149
7.1	知识管理与电子政府	149
7.1.1	知识管理的概念与作用	149
7.1.2	知识管理的支撑技术	151
7.2	基于知识管理的政府电子文档管理	152
7.2.1	知识管理中的电子文档	152
7.2.2	基于知识的文档管理	153
7.3	政府知识地图	154
7.3.1	知识地图的作用	154
7.3.2	知识地图的构建	155
7.3.3	知识地图构建的基本原则与流程	157
7.4	政府知识库	160
7.4.1	知识库概述	160
7.4.2	知识库的构建	162
7.4.3	知识管理中知识库的构建	163

7.5	知识仓库	164
7.5.1	知识管理时代知识仓库的地位	164
7.5.2	知识仓库的架构	165
7.5.3	知识仓库建库与知识网站管理系统	168
7.6	知识管理系统的体系结构	170
7.6.1	知识管理系统在电子政务中所处的位置	170
7.6.2	知识管理系统构建的基本要求	171
7.6.3	知识管理系统构建的技术方法	172
7.7	知识文档管理与知识管理系统介绍	173
7.7.1	泛微知识文档管理	173
7.7.2	智识知识管理软件	174
第 8 章	政府电子文档管理系统解决方案	176
8.1	方正 Apabi 电子公文处理系统	176
8.1.1	方正 Apabi 电子公文处理平台概述	176
8.1.2	方正电子公文一体化解决方案	177
8.2	书生电子公文解决方案系列	184
8.2.1	书生公文传输系统	184
8.2.2	书生智能电子表单系统	185
8.3	索源电子文档管理系统与电子公文交换系统	191
8.3.1	索源电子文档管理系统	191
8.3.2	索源电子公文交换系统	194
8.4	安全智能电子文档解决方案	195
8.4.1	AIP 文档安全解决方案	195
8.4.2	AIP 的主要功能	195
8.4.3	AIP 文档的安全机制	196
8.4.4	AIP 的主要特点	197
8.5	政府智能传真系统解决方案	197
8.5.1	TraFax Server 智能传真系统概述	198
8.5.2	TraFax Server 的功能特点概述	198
8.5.3	TraFax Server 智能传真系统的应用方式	199
参考文献		201

第1章 文档概述

1.1 文档的概念

1.1.1 文件与文档

任何组织的管理活动,其实质就是对管理信息的处理过程。在诸多的管理信息中,文件与文档信息无疑占据主导地位,这是因为与其他的信形式相比,文件与文档具有权威性、时效性和规范性等特殊优势,并具有准确传递权威性信息和凭证性信息的独特功能。因此,文件与文档是国家机关和其他社会组织计划、决策、指挥、控制、协调工作的手段和方式,是联系上下、沟通左右的桥梁和纽带,在各项管理活动中发挥着特殊的效用。

1. 文件与文档的一般概念

(1) 对文件的理解。“文件”一词在英文中常用 document、file 或 record 来表示。其中,document 还可译为“公文”、“文献”,file 则还可译为“资料档”、“案卷”或“文档”,而 record 还可译为“文书”、“记录”、“档案”。

根据《中国大百科全书》的相关条目,对文件与文档的解释如下。

① 文件是国家机关、社会组织、企事业单位或个人在社会活动中为处理事务、交流信息而使用的各种载体的文字、图表、声像等记录材料。它是人们社会交往的工具,也是档案的前身。

② 文档是按一定的结构归档到计算机存储设备上的情报资料。这些资料以记录的形式存储,因此文档也可视为按一定结构组织的记录的集合。

③ 在计算机检索领域大多称 file 为文档,如顺排文档、倒排文档。文档管理系统是以文档记录为单位进行存取的。

而在《辞海》中则将“文件”定义为:计算机的一个专用词,指由若干相关记录构成的集合;指磁盘、磁带等存储的内容。它也可以是一般意义上表示目的、形式、内容上彼此相似的信息项的集合,还可以指内存中的一批数据。

(2) 对文档的定义。对文档的理解也有多种,各有不同的侧重点。《Merriam-Webster》字典对文档(document)的定义如下。

① 作为活动、事件的根据、证据或支持的、原始的或正式的文件。这一定义中的“paper”指的是纸质的文件,强调了文档的证据作用。作为证据用的东西,如照片或录音带等。这一定义强调了文档的物质多样性。

② 含有信息的作品。该定义引入了信息的概念,说明文档的作用是保存信息。

③ 带有公认的标志与符号以描述思想的物质实体,如石头、硬币等。这与第二个定义有很多相似之处,但更多地强调了文档载体物质的多样性。

(3) 对文档的理解。根据以上定义,对文档的理解应该包括如下三个方面。

① 文档是信息的载体。文档保存信息和思想,文档有证据、凭证的作用,具有法律效力。

② 把文档视为管理对象的比较权威的理解是信息和图像管理协会(The Association for Information and Image Management, AIIM)的定义,即文档是信息的容器,也就是信息的载体。国际标准化组织(ISO)对其定义为:文档是指可为人类感知的信息,是用户与系统之间进行信息交流的基本单元。

③ 文档作为信息的载体,并不一定是纸张。从广义上讲文档可以是任何能提供有用信息的物体,学术界甚至有人把砖块也看作是文档,砖块所蕴涵的信息是通过它的形态以及标签上的信息来反映的。随着社会信息化的发展,文档已经不仅仅是以纸张的形式存在,人们更多的是用计算机来产生和修改文档,由于电子文档容易传输,在信息流通过程中的作用也越来越大。

实际上,人们对文件与文档的理解常常是难以严格区分的。在办公领域,文件即公文,或者称文书,文书是各种文件的统称。在计算机领域,文件是按名字存储在计算机外部存储设备上并按名读取的相关信息的集合(注意,计算机外部存储设备上还保存有不是按名字存取的信息或数据);特别地,为了处理方便,还常常将一些常用的外部设备(硬件)定义为文件,例如,打印机的设备名为 Prn。存储在计算机外部存储设备上的文件可以分为记录式文件和流式文件两种存储格式。

在计算机领域对文件与文档往往还是有所区分的。例如:计算机软件是由计算机程序、数据和相关的文档组成的,其中的文档是指程序开发、维护和使用中所涉及的资料(如设计报告、维护手册和使用指南等),软件必须有完整、规范的文档作为支持。在这里,文档主要是指说明文字、图表等可读性资料。

又如:在使用文字处理器(如 Word)时,创建一个新文档后,系统会自动打开一个名为“文档 1”的空文档(连续打开则命名为“文档 2”)。这时,常常又把正在编辑处理的信息称为文档,而把存储到外部设备(如磁盘)上的信息集合称为文件(如磁盘文件)。

2. 文件与文档及其基本特点

文件与文档的概念是随着科技和社会的发展而发展的,正规的文件与文档是社会组织或个人在各种管理活动中直接形成、接收和使用的具有规范体式 and 法定效用的信息记录。该信息记录由足以为其活动提供凭证的内容、背景和结构所构成,而不论记录的方式或载体形态如何。通常,“文件”概念包括以下几层含义。

(1) 文件形成、接收和使用的主体是各种社会组织或个人。只要是依法成立的合法组织,不论其性质如何,只要管理活动需要,均可以形成、接收并使用文件。公民个人管理自身事物过程中也可以依法制作使用文件。

(2) 文件形成、接收和使用的目的是为了满足各类社会组织或个人记载事物、表达意志、交流信息、联系工作、处理各种事务的需要,具有最直接的应用性,是处理事务的必不可少的一种工具。

(3) 文件具有一定的规范体式,包括法定和约定俗成的体式。文件体式具有长期的稳定性和固定表达方式,以维护文件的严肃性、郑重性和权威性。

(4) 文件伴随各项管理活动同步产生,是对该项活动的原始记录,对一定的组织或人员

具有法律所规定的约束力和影响力,在法律上具有不可否认的凭证属性。也就是说,文件(records)是由官方记录下来作为证据或证词的信息。因而文件不仅具有信息价值,而且还具有证据价值(evidential value)。证据价值是在证明接收或产生它们的机构的组织、职能与行为方面有用,具有证据价值的信息经鉴定后,可能会长期保留。

(5) 只要符合文件的定义,不管其产生的技术环境,如记录的形式或载体如何,都应当称为文件。一般说来,文件可以以模拟(analog)记录方式形成,也可以以数字(digital)记录方式形成。模拟是传统的记录方式,它使用连续的物理量来表示信息,例如:书写、雕刻、绘画、录音、传统的摄影等,并采用石头、金属、纸张、绢帛、胶片等制作材料,传统文件一般都是以模拟方式形成的。这些记录信号在某一信号间连续变化,因而能够精确地表示原始信息,以及原始信息的连续关系。由于这类连续变化的物理量很容易被检测与鉴别,因而由模拟记录方式形成的文件可以通过技术方法来鉴定其真实性与完整性。数字记录方式是将信息用“0”或“1”的二进制状态数据表示,通常采用磁带、磁盘、光盘等保存,用磁的不同极性或某些状态的变化来表示,这类信息是需要通过计算机及其相关设备进行处理后才能识别的,这类文件与文档由于其记录信号的不连续性,是难以通过技术手段识别其真实与完整的。

(6) 文件与文档必须包括三种基本要素,以保证其凭证价值。这三种要素是:内容(content),背景(context,亦称“上下文”)和结构(structure)。内容是指文件与文档中所包含的表达作者意图的信息;背景是指能够证明文件与文档形成环境、形成过程、存在状态以及和文件与文档之间相互关系的信息,包括表示文件与文档来源与传送目的地的信息、文件与文档形成和存在相关的其他信息,如作者、签发人、成文日期、收文者、形成文件与文档的活动说明、文件与文档生成的软硬件环境、文件与文档状态改变的说明(如修改、迁移等)、相关文件与文档名称及其关系说明等;结构是指文件与文档内容信息的组织表达方式,如文字的段落安排,电子文件与文档所使用的代码、格式,以及载体、附件等方面的信息。

1.1.2 电子文档及其特点

1. 电子文档的概念

(1) 对电子文档的理解。人们对电子文件与文档的认识也是处在一个不断深化的过程中,从 Machine-readable Document(机读文档)、Machine-readable Archives(机读档案)到 Electronic Records(电子文件)、Archival Electronic Records(归档的电子文件),再到 Electronic Archives(电子档案)。国内外学者对“电子文件”或“电子文档”有着多种不同的定义,但是总的来说有四种主要的模式,即“文件说”、“代码序列说”、“信息组合或信息集合体说”、“文字材料说”。解释如下。

① 国际档案理事会电子文件委员会对电子文件与文档的定义是:“电子文件是一种通过数字电脑进行操作、传输与处理的文件,并具有文件的一般定义。”

② 国家档案局《电子文件归档与电子档案管理规范》中的定义是:“能被计算机识别、处理和存储在磁带、磁盘或光盘等介质上,并可以在网络上传递的代码序列。”

③ 从信息的角度对电子文件与文档进行定义,认为电子文件与文档的本质就是信息。电子文件是人们为了某一目的,使用计算机,采用一定的格式和处理方式生成的信息的组合。

④ 从文字材料的角度出发来定义电子文件,即所谓电子文件是指人们在社会活动中形

成的，以计算机磁盘和光盘等为载体的文字材料。它主要包括了电子文书、电子信件、电子报表、电子图纸等。

由上可见，尽管对电子文件与文档这一术语存在不同的表达方式，着眼点与侧重点各有不同，但基本内涵是一致的，即电子文件与文档必须满足以下两个基本条件。

第一，电子文件与文档首先是文件，必须满足文件的所有条件。国内外在定义电子文件与文档时，都将“文件”作为电子文件与文档的上位概念，认为电子文件与文档是文件的一种类型，应当具有文件的各种属性。这是电子文件与文档与其他形式文件与文档的共同点，也是电子文件与文档与其他类型数字信息的基本区别所在。

第二，电子文件与文档是由计算机生成与处理，并可经由计算机网络进行传播的，计算机及其网络是电子文件与文档赖以“生存”、“活动”的基础。

严格地说，“电子”信息也可以用连续的模拟量来表示（也就是说存在模拟计算机），但这些模拟信息往往也不是人们可以直接识读、检测与鉴别的。因此，书中所谓电子文件或电子文档应当是指用电子数字计算机生成和处理的文档，亦即比特文档。

(2) 对电子文档的再认识。随着 IT 技术的发展，特别是数据库技术的广泛应用，组织中的信息呈现出结构化和非结构信息两种形式。结构化的信息通常可以作为数据记录（Data Record）存储在如层次型数据库、关系型数据库等结构化的数据库中。结构化的信息所描述的是事实对象，即与组织业务过程紧密相关的实体，而实体是通过其属性来描述的，如单位员工的工号、姓名、年龄等数据。

非结构化的信息所描述的主要是组织中的概念、思想、方法或知识，如会议决议、施工图纸、施工规范等。非结构化的信息记录在文档中，或以组织成员的经验的形式的存在于员工的大脑中。数据记录与文档的区别如表 1-1 所示。

表 1-1 数据记录与文档的区别

项 目	数据记录（结构化的信息）	文档（非结构化信息）
描述的对象	实体	思想或概念
对象的属性	字段	字符集
对象的属性集	记录	逻辑段落
相关的对象集	数据库文件	文档
描述方法	树形结构或关系型结构	关键字或超文本
管理系统	数据库系统	文档管理系统

在计算机世界，多媒体信息的产生，并通过对象的链接或嵌入技术以及超文本技术，人们可以将若干分散的电子文件整合成为一个电子文档，从而结合在一起形成一个整体。从这一角度看，文档与文件是有所不同的。但在一般情况下，可以不对文档与文件做严格的区分。

结构化的信息通常占组织信息的 10%~20%，而非结构化的信息则占 80%~90%，并且大部分的非结构化信息都存在于文档中，因此文档在组织中具有重要的作用。文档的主要作用如下：

- ① 作为组织的最终产品或生产产品的支持资料；
- ② 作为组织中人、群体或组织之间进行信息交流的一种介质；
- ③ 作为组织活动过程记录的载体；
- ④ 组织记忆的一个重要组织部分。

2. 电子文件与文档的特点

从以上的定义,可以归纳出电子文件与文档的基本特点如下。

(1) 信息的非人工可读性。传统文件与文档的内容以字符、数字或图表等形式记录于纸张等载体之上,属于“原子”层次,因而往往人工可直接识读。电子文件与文档的信息以比特(bit)的形式存在,而比特没有体积、没有重量、没有颜色、没有形状,能以光速传播。它是信息的最小单位,是一种存在的状态,是数字化计算中的基本粒子(或称为信息的DNA)。因此,电子文件与文档是一种人类无法凭借自身的器官直接来识读的信息。

(2) 电子文件与文档的系统依赖性。电子文件与文档信息存储的时候需要进行编码,信息检索和阅读的时候需要进行解码,编码和解码过程中则要借助于一定的硬件设备、系统和应用软件,否则电子文件与文档既无法产生也无法阅读,电子文件与文档真正的形成者和读者是程序。电子文件与文档对系统的依赖性特别是它对特定系统(电子文件与文档产生时运行的软、硬件系统)的依赖性电子文件与文档的一大特点,也是管理中的难题。

(3) 电子文件文本内容是离散性的合成。传统书写方式形成的文本内容与形式一旦写入纸张等载体,就结合为稳定的复合体,其内容与形式是一个彼此不可分离的统一体,并一起构成了文件与文档原件。而在电子语境中,文件文本内容各个组成部分处于离散状态。内容只是计算机能识读的数字串记录,其全文形式不再以固定的形态整体存在。信息内容不会永远和特定的载体相连,信息内容易于复制、迁移,载体仅仅成为每一个特定时刻文件与文档内容的承载物,它已无法固定某些用以鉴辨文件与文档原始性的外部信息。

(5) 电子文件文本失去了书写的物质痕迹。传统的书写方式语境中,在纸张上书写,用打印机打印出的文件文本等都具有一种不变的痕迹,如油墨、字迹、图形、记号等,个性特征清楚,改变或磨擦掉这些痕迹总会留下印记,如墨水的氧化会留下时间的印记等。然而在电子语境中,当文件内容录入计算机,显示在屏幕上,这种代码是以计算机能够识读的、人无法直接看见和触摸的“非物质”方式存在,并可以根据需要任意改动,而表面上不留痕迹。这样,传统文件写作中的“草稿”、“定稿”、“正本”等概念的界限在电子环境中变得模糊起来,在同一内容的多版电子本文稿中,从外在的形式上无法直接发现具有“原始性”或“本真性”的文本。目前,通过一些专门技术如数字签名、加密等方式可以确认信息内容的真实性,但是,其验证的过程仍是由一系列的计算机程序实现。

(6) 信息的可操作性和不稳定性。电子文件信息是灵活的,可以根据不同的需要进行各种操作,如修改、粘贴、删除,而且还可以不留痕迹。信息可以直接进入计算机辅助制造系统(CAM)进行生产。信息的质量,如清晰度,也可以进行调整。电子文件的可操作性有时会带来难以分辨或预测的易变性,这主要是由信息增、删、改的便捷性,信息系统自动更新和补充功能,计算机病毒的侵袭等原因造成的。电子文件载体性能的不稳定性有可能造成文件中信息的改变,在文件“定型”保存之后,这种易变性容易导致人们对它的真实性和可靠性产生怀疑。此外、电子信息技术的发展,新的信息编码方案、存储格式、存储载体和系统软件的更新换代更加对电子信息稳定性造成巨大的冲击,这要求将原文件迁移到新的技术环境之中,而迁移过程中信息的损失、变异也是难以避免的。

(7) 电子文件形成系统的异构性。异构性是指系统的数据处理与描述方式多样化和不一致性。形成电子文件的系统性质、功能及功能实现方式是多种多样的,建设单位的业务特色、工作环境差异、开发者的技术差异等,都会导致具体系统的概念定义、功能定义、结构方式、

技术路线、信息构建等多方面的异构性。解决异构性问题的主要方法是数据处理方式和描述方式的标准化。

3. 电子文档的优势

用电子文档进行信息管理与沟通相对于使用纸质文档有很多优越性,主要表现在以下几个方面。

(1) 节约成本。利用计算机生成的电子文档进行信息沟通,最直接的好处是节约纸张和打印的费用,对于费用昂贵的工程图纸更是如此。并且电子文档的复制不需要使用复印机,尤其是当文档需要进行很多份的分发时,纸张和复印费用的节约是非常可观的。

(2) 节省存储空间和信息存储的可压缩性。电子文档以数字化的形式存储在存储设备中,如 U 盘、磁盘、光盘等,可以节约存储纸质文档所需的空间。将纸质文档扫描而成的图形化电子文档与原纸质文档所需存储空间的分析对比如下:

1 张标准的 A4 纸张=50KB (黑白两色扫描,使用 CCITTG4 压缩技术) 文字存储量

1 个文件柜 (四个抽屉)=10 000 张纸=500MB=1CD (CD-ROM 或 CD-WROM)

10 个文件柜=5GB=10CD=1DVD 光盘 (DVD-WROM)

1 张 A0 图纸=16 张 A4 纸=800KB 文字存储量

计算机直接产生的电子文档比图形化的电子文档占用的空间更小。通常与一张写满中文的 A4 大小的纸张文件 (1 000 字左右,字号是小四) 相当的 Word 电子文件需占用 25KB 的磁盘空间,包括这一页所包含的文字、文字格式和应用程序对文件整体格式的描述语言所占的空间。一个 100 页的纯文字的电子文件包含大约 10 万小四号大小的中文字,占用的空间为 300KB。到目前为止,不但一般计算机的硬盘已经达到了几十 GB 以上,而且光全息信息存储容量将更加惊人。

电子文件与文档的数字化特征使其通过特殊的编码可以进行压缩处理,使巨量信息可以压缩存储,这不但可以节省大量存储空间,而且可以减轻网络传输的带宽占用 (这在政府机关召开视频会议时是很有意义的)。

(3) 传输方便迅速。因特网的出现使信息能以更加低廉的成本迅速地进行传输。文件管理系统通常通过上传的方式把电子文档存储到连接在因特网上的服务器中,服务器收到文件后在很短的时间内用可以在 WWW 上发布或通过 E-mail 通知所有与此文档相关的人员,在收到 E-mail 通告信后,收信人按照 E-mail 中所提供的路径访问文件服务器以获取此文档。从文档作者上传文档到接收人看到文档的整个过程可以在很短的时间内完成,从而提高了信息流动的速度。当然,Internet 的数据传输速度受网络连接设备及采用的传输技术的不同而有很大的差别。

(4) 电子文件与文档的可重用性。存储起来的电子文件可以被反复调取、使用;可以对所存信息整体进行拆分,或提取信息的属性、建立关联、组配使用;通过数据挖掘等方式可以对结构化文件中蕴含的信息进行分析和调用,以发现数据资源库数据中的信息规律。

(5) 电子文本撰写的协同性。在传统文件写作活动中,撰写者始终是写作活动的主体,文本结构的安排、语言的运用、材料的选择等工作都由其独立完成,而且同一文本在同一时间只能由单一撰写者进行写作。而电子文件的写作动摇了文件撰写者的稳定地位。首先,计算机不仅是文件撰写者的工具,而且成为其写作活动的合作者。使作为客体的计算机与主体的书写行为合而为一,计算机按照自身的信息方式参与到文件的书写活动中,它与撰文者共

同成为电子文件的创作者。其次,为同一文本的多作者写作活动提供了方便,除了可以通过网络集体讨论和写作,甚至还可以利用协同软件“同时性写作”,即多个撰文者可以在不同的地点同时对同一份文本进行写作和编辑,共同完成文件的写作。

4. 电子文件与文档功能特征

在办公自动化环境下,电子文件从功能上看,有其显著的特征。

(1) 便于简化文档的处理和管理,实现无纸操作。所有的电子文档都被归档到一个集中的库中,授权的用户可以在任何时候、任何地方进行访问,这将减少有纸办公和储存的空间。另外,可以将不同版本的数据存储在不同的地方,避免了数据之间的矛盾。

(2) 可以进行版本控制,保持文档的完整性。可以提供版本号控制(经验证明,实际5个版本就已经足够优良),因此,用户能够设置并保存在数据库中的首选版本号,有利于保持文档的完整性。

(3) 可实现集中的文档访问。通过同步多站点服务器,支持多站点访问,这样,在任何一间办公室工作,都可以实现对文档的访问。

(4) 文件查找将更加便捷。检索系统的文件查找功能能帮助工作人员以最快的、最友好的方式准确地访问及找出文件。系统可以设计很多搜索策略实现快速、准确的查询,例如:按标题、按日期、按主题词等。

(5) 多媒体电子文件信息的集成。以往的文件都是平面,文字和图形在平面的纸张或其他载体上呈现。而电子文件是多媒体的、立体的,文字、声音、图像和视频可以融合到一份文件中,或通过超级链接组织在一起。运用多媒体技术,可以使电子文件声形并茂,真实地再现当时的活动情况,从而强化了文件对社会生活记忆和再现功能。可以说,电子文件是一种全方位的记忆和再现,实现了文件功能的革命性变革。

(6) 电子文件管理流程变化。电子文件归档、鉴定、著录环节提前;著录将贯穿于电子文件的编制、处理、归档、迁移、利用等整个生命周期。电子文件管理内容变化:如实体整理仅为对脱机保存的磁盘、磁带、光盘的归类;电子文件管理中,可实现归档与检索的集成、逻辑归档工作与检索工作的集成、归档与鉴定的集成、保护与其他流程之间的集成。电子文件管理顺序发生变化,由串行向并行、线性向非线性变化。

1.2 政府文档与政府电子文档

1.2.1 政府文档的特征与作用

机关团体、企事业单位等依法成立的社会组织用来办理公务、有一定格式的文件即为公文,全称公务文书,对于政府而言就是政府文档。公文办理公务,就是以文字的形式实施管理。公务,应该包括内务(内部管理)和外务(处理与其他社会组织的关系)。公务还分为政务和事务:政务是社会组织领导所主管的人员安排和组织发展的大事;事务是社会组织全体成员从事的业务性、日常性的工作。

1. 公文特征

公文的特征具体描述如下。

(1) 公文内容的公务性。公文的“公”，指的是社会组织。领导者个人只有在其职务的前提下，并且只能是社会组织的最高领导才能代表社会组织发出公文。公文的内容必须是反映和传达社会组织的公务信息。党纪国法都规定了公文内容的范围和性质。

(2) 公文格式的规范性。公文写作都是有格式的，公文的格式，有惯用的格式和法定的格式两种。惯用的格式，是约定俗成的，没有严格的限制，如普通公文中计划和总结的格式。法定的格式则是权威机关规定的，因而必须严格按照格式写作。法定公文中的行政公文和党务公文，更是由最高权力机关及有关部门通过法规性公文规定了严格的格式。需要进一步说明的是，公文格式同时又是程式，呈现出公文写作和办理的程序性。

公文格式的规范性，是公文本质特性的发展，是公文写作和办理的需要。公务具有公众性和同一性，对社会组织成员产生一致的认可、制约和指挥，否则社会组织就不可能运作。相应地，反映和办理公务的公文，也就形成了格式和程式，显著提高了公文写作和办理的效率。随着时代的发展、技术的进步和社会组织的完善，公文的规范及格式会更加科学、严谨，以至于公文写作和办理会由高度先进的电脑软件来实现电子化、数字化、自动化、网络化、多媒体化和智能化。

(3) 作者和读者的指定性。公文的作者只能是法定的社会组织及其法人代表（或者称为第一领导人）限于有权利进行公文写作的社会组织。这个法定的社会组织及其第一领导人，就成为公文的法定作者。至于动笔起草公文初稿的人，如秘书，应称为起草人，不是法律意义上的作者。公文的读者是特定的，在公文格式上有专门规定，如“主送机关”、“抄送机关”和“传达（阅读）范围”等。当然，有的告知性公文如通告，指定的读者应包括发出公文的社会组织之外的社会大众。

(4) 法定权力的制约性。公文只能由法定的作者发出。法定的作者即社会组织的机关及其部门都规定了隶属关系和职权范围，公文是这种隶属关系和职权范围的反映。写作公文和办理公文都有一定的规定性。也就是说，对于作者和读者，公文具有法规给予社会组织职权相应产生的制约性。制约性在不同的公文中有不同的情况。行政公文的命令，对于公文的接受者具有强制性，如果接受者不按命令办理，就会受到法律的制裁。发出命令的政府机关有权依照法律规定，动用军队或警察等国家机器进行处罚。行政公文的决定，具有国家指挥性和约束力。行政公文的通知，具有规定性、指挥性和指导性。经济公文合同，依照《中华人民共和国合同法》对于缔约各方具有确定的制约性，如奖惩、期限等。普通公文的讲话稿，对听众具有领导者所拥有的指导性和指挥性。法规性公文如法律、规章，制约性更是不言而喻的。公文正因为有制约性，才能产生现实的管理作用。

2. 公文的作用

(1) 具有法规和准绳作用。政府公文，特别是法规性公文有宪法、法律和规章（如条例、办法等），在发出公文的社会组织的管理范围中有法规和准绳作用。全体成员都要遵守，否则就要受到法规所规定的处罚。其他行政公文也有一定的法规和准绳作用，特别是直接发布规定的命令、训令、决定或者通知。

(2) 具有领导和指导作用。上级机关发出的下行文，对下级机关单位具有领导和指导作用。