

档案管理与计算机

孙淑扬编著



中央广播电视大学教材

档案出版社

档案管理与计算机

孙淑扬 编著

档案出版社

一九八七年

档案管理与计算机

孙淑扬 编著

*

档案出版社出版

(北京市西城区丰盛胡同21号)

新华书店北京发行所发行

各地新华书店经售

一二〇一工厂印刷

*

开本: 787×1092毫米1/32 印张: 9.75 字数: 216千字

1987年3月第一版 1987年6月第一次印刷

印数: 1—50,000册

统一书号: 7283·072 定价: 1.40元

ISBN 7-80019-004-8

G · 0005

前 言

电子计算机技术是当今新技术革命的先导技术。自从把它应用于信息处理以来,电子计算机已成为世界各国国民经济中不可缺少的强有力的工具,被广泛地应用于各个领域。在实现档案工作现代化的过程中,计算机的应用有着巨大而深远的意义,采用以电子计算机为中心的新技术进行档案管理,是档案工作现代化的重要内容和主要标志之一。

目前,在我国档案管理中,计算机的应用开发刚刚起步。一些单位正在准备建立或者已经建立起了电子计算机档案管理系统。因此,如何向全国档案工作者普及这方面的知识,使他们掌握应用计算机的技能,已经迫在眉睫。近年来,许多大专院校及职工业余教育中的档案专业开设了关于电子计算机档案管理的课程,形成了档案管理科学中的一门新学科。本书正是为了满足档案工作与教学的需要而编写的。它作为中央广播电视大学档案专业教材之一,其直接读者对象是广播电视大学学员。同时,在目前高等院校档案专业暂缺本科教材的情况下,亦可作为本科学生的教材或参考读物。

全书共分十一章:第一章概述电子计算机在档案管理工作中的主要应用,作为学习后续章节的引导。第二、三、四、五章介绍了电子计算机的一般理论,硬、软件基本知识和计算机编程方法,这些章节是学习以后章节的基础。第六、七、八、九章全面地叙述了电子计算机档案管理的主要环节,如档案信息的前处理工作,档案信息在计算机中的存贮,档案数据库及机读档案目录的检索等。这部分内容是全书的重点。

第十章讲述了实现计算机档案管理工作的一项关键技术：汉字信息处理的基本原理与实现方式。第十一章阐述计算机档案管理系统研制过程。

本书是按照中央广播电视大学档案专业“档案管理 与 计算机”课程的教学要求编写的。在文字叙述上，本书为了做到通俗易懂、深入浅出，省略了一些冗长程序的范例和数学推导。

中国人民大学档案学院沈永年副教授和中国科学院计算中心金翠柏副研究员审阅了本书，并提出了宝贵意见。本书在编写过程中，参考了一些有关著作，吸收了这些著作中的有益成果。在此，仅向上述两位同志和有关著作作者致以衷心谢忱。由于本人水平所限，本书中疏漏谬误之处在所难免，祈请读者不吝指正。

作者

一九八六年七月

目 录

第一章 档案管理与计算机	(1)
第二章 电子计算机概论	(6)
第一节 电子计算机的分类及发展	(6)
第二节 电子计算机的用途	(8)
第三节 数据处理概论	(10)
第三章 电子计算机的基本结构	(16)
第一节 电子计算机的基本组成	(16)
第二节 控制器	(18)
第三节 运算器	(20)
第四节 存贮器	(21)
第五节 输入输出设备	(29)
第四章 电子计算机软件基础	(34)
第一节 电子计算机常用的数制及其转换	(34)
第二节 数据在计算机中的表示	(41)
第三节 计算机软件体系	(44)
第四节 程序设计语言	(52)
第五章 电子计算机程序设计	(58)
第一节 计算机的解题过程	(58)
第二节 流程图	(60)
第六章 档案的前处理工作	(67)
第一节 档案著录标引的作用与意义	(67)
第二节 档案的著录	(69)
第三节 档案的标引	(74)

第七章 档案数据的存贮技术	(93)
第一节 电子计算机数据结构.....	(94)
第二节 文档的基本概念.....	(135)
第三节 流水文档.....	(138)
第四节 顺序文档.....	(139)
第五节 链式文档.....	(153)
第六节 索引文档.....	(155)
第七节 倒排文档.....	(162)
第八节 直接存取文档.....	(165)
第九节 B 树文档.....	(169)
第八章 档案数据库	(172)
第一节 数据库的概念.....	(172)
第二节 数据库系统的组成.....	(176)
第三节 数据库的数据模型.....	(184)
第四节 微型机数据库.....	(191)
第五节 数据库的发展趋势.....	(200)
第九章 机读档案的检索	(205)
第一节 机读档案目录和机读档案检索.....	(205)
第二节 机读档案检索步骤.....	(211)
第三节 查找技术.....	(235)
第四节 联机检索系统.....	(241)
第五节 机读档案检索系统的评价.....	(247)
第十章 汉字信息处理	(251)
第一节 汉字的特点及汉字处理的发展.....	(251)
第二节 汉字的输入.....	(253)
第三节 汉字代码的转换及汉字库.....	(262)
第四节 汉字处理设备.....	(269)

第五节	汉字处理软件.....	(272)
第十一章	电子计算机档案管理系统的研制.....	(277)
第一节	系统分析.....	(278)
第二节	系统设计.....	(286)
第三节	系统实现.....	(294)
第四节	系统的维护、运行与评价.....	(296)
参考文献		(299)

第一章 档案管理与计算机

档案是国家机关、社会组织和个人从事政治、经济、科学、文化等社会实践活动直接形成的文字、图表、声象等形态的历史记录。

档案工作是党和国家的一项重要事业。近年来,随着我国四化建设的发展,档案工作发生了很大变化,封闭式或半封闭式的管理方式正在向开放式转变。为了适应新的形势,迫切 need 实现档案管理现代化。

所谓档案管理现代化,就是指利用现代科学技术来管理档案。以电子计算机为中心进行档案自动化系统的开发是档案管理现代化的一个重要方面。

一、电子计算机在档案工作中的主要应用

根据近年来国内外档案工作部门的实践,目前,电子计算机在档案工作中的主要应用领域如下:

(一)电子计算机自动整理编目。档案整理工作是档案管理工作的重要环节,通过整理工作将零乱的档案系统化、条理化,而编目工作是整理工作的发展和提高,编目工作是档案管理部门开展各项业务工作的基础和必要条件。电子计算机通过相应的软件,可以对存贮在计算机中的档案信息进行整理,然后按管理和使用的不同要求编制各种目录。计算机整理编目与手工整理编目比较,既效率高又质量好,而且能自动进行。

(二)电子计算机自动统计。档案统计工作是通过统计数量的统计与分析,以数字的形式,揭示档案和档案工作状况的某些规律性的一种活动,计算机以它的极强的运算能力,能够胜任档案统计工作。统计工作主要包括两方面的内容,其一是对档案数量及有关内容的统计,计算机可以对机读档案目录按照不同的登录项,如某一作者、某一专题、某一类别等等进行登记和统计,并通过输出装置打印出来;其二是对档案利用情况的统计,计算机可以对某一天、某一月、某一年中的档案利用人次进行统计,并打印出日报、月报、年报表来,从而使档案工作者了解档案的利用频率和利用构成情况,进而总结出某些规律性,可以成为预测某些档案未来需要的依据,对于档案价值的鉴定也有一定的参考作用。

(三)电子计算机自动检索。档案的利用工作是档案工作的一个重要组成部分,随着档案工作重点和服务方向的转移,档案的利用日益广泛。档案工作的服务对象由过去主要为党政领导机关服务和为政治斗争服务发展成为社会的政治、经济、军事、外交、科学、文化、教育以及历史研究等各方面服务。所谓档案的电子检索就是使用电子计算机自动地从存贮在存贮器上的档案中找出用户所需的档案信息来。目前,国外大的档案馆已经有了一些实用的自动化检索系统,积累了一定的经验和有了较成熟的软件。目前我国档案自动化检索系统还处在实验阶段,现有的大多数系统能够从档号、标题、作者、时间、分类号、主题词等方面对机读档案目录进行检索。计算机自动检索与手工管理相比,最突出的特点是速度快、质量高,并且能以多种途径进行检索,以多种形式提供检索结果,从而满足各类用户对档案的利用要求。

(四)电子计算机用于档案馆的日常管理工作。计算机可

以代替档案工作者做一些繁琐的日常管理工作，如当档案借阅者借阅档案时，可以使用计算机识别用户的合法身份、办理借阅预约、借阅登记等手续，同时自动计算出归还时间，如到期未还则自动打印出催还通知。

(五)电子计算机库房自动管理。利用计算机对库存档案进行管理，掌握现存档案、调阅借出的档案、已归还的档案的情况，可以随时将情况打印出来；还可以利用计算机进行库房空间的安排；库房空气和温湿度的自动调节；库存消费品的管理；缩微胶片等专门档案的管理。自动控制照明、防火系统、报警装置等也是计算机库房自动管理的内容。

今后，随着电子计算机技术的不断发展，还会有更多的应用领域，如：档案的自动标引、全文存贮、语音识别、图形识别等等。总之电子计算机在档案管理中的应用有着广阔的前景。我国的档案自动化工作正在积极、慎重、有计划、有步骤地发展。

二、档案自动化给档案工作带来的影响

电子计算机在档案管理中的应用，国外称为Archives Automation即档案自动化。七十年代中期，美国、英国、法国、苏联等国在国家级的档案馆相继建立档案自动化系统，进行档案的自动编目、自动检索等管理工作，至今已经有了一些比较实用的档案自动化系统。

在我国，七十年代末、八十年代初也开始了档案自动化的研究工作。档案事业的发展促进了档案自动化，而档案自动化又反过来推动了档案工作的发展。与此同时，也给档案工作和档案工作者提出一些新的问题：

(一)档案工作者要研究机读档案的管理问题

对档案原件进行加工，然后将有关信息存入计算机的存储设备中，这就构成了二次档案信息——机读档案和机读目录。传统类型档案的载体以纸张为主，而机读档案则存储在磁性介质上。载体的改变带来了处理手段、保管方法的一系列改变。例如，档案的前身是现行文件，今后随着办公室自动化的发展，文件将在计算机上生成，通过计算机通讯网络传递，档案的收集、归档方式也将发生变化，由于直接收集到的是电子型的档案。那么档案的整理、编目、保管等一系列管理环节都将相应地发生变革。档案工作者必须研究如何管理存储在磁性载体上的机读档案的问题。

（二）档案学理论需要进一步发展

电子计算机在档案工作中的应用构成了档案自动化系统（以下简称系统）。对系统的开发、研究、设计、实施、评价等工作都需要有新的理论做指导，这就要求档案工作者在原有档案学理论的基础上，应用经济学、管理学、控制论、信息论、系统工程学、运筹学，计算科学中的许多概念、原理和方法，发展档案学理论，以适应新技术在档案工作中的应用。

（三）档案工作标准化是实现档案管理自动化的前提

为了适应电子计算机管理，档案馆的基础工作要加强，只有规范化的档案信息才能用电子计算机进行处理。也只有在档案馆内部和馆际之间实行了标准化，才能获得可供交流的机读档案。目前我国的档案标准化工作刚刚开始。新技术的应用必须以档案工作的标准化为基础。

（四）档案工作者要学习新技术，以适应电子计算机管理的新的工作方式

随着科学技术的发展，电子计算机的应用日益广泛，可

以肯定，电子计算机将成为档案工作不可缺少的重要工具，到那时，微型计算机或计算机终端将成为档案工作者办公桌上的一种标准的档案管理设备，就象实验室里的科学仪器一样。这就要求档案工作者的知识结构必须发生变化，要学习计算机科学的有关理论与技能，学会操作计算机的方法，进而胜任档案自动化系统的开发和管理工作的。

总之，电子计算机等新技术在档案工作中的应用，将促使档案工作从理论到技术，从干部队伍建设到专业教育发生一个大的变革，它将成为推动档案事业现代化的巨大动力。

第二章 电子计算机概论

第一节 电子计算机的分类及发展

电子计算机有两种类型，即电子数字计算机和电子模拟计算机。

电子数字计算机也叫数字式电子计算机。它是以数字形式的量值在机器内部进行存贮和运算的电子计算机。所谓数字，是指表示数量的一串符号，如人们笔算时用0，1，2，…9等十个符号来表示数量。而电子数字计算机中用电信号作为表示数量的符号，通过电信号在机器内的传递变化来实现数字的加、减、乘、除等各种运算。电子数字计算机就是人们通常所讲的电子计算机或电脑。

电子数字计算机按其规模可分为巨型电子计算机、大型电子计算机、中型电子计算机、小型电子计算机和微型电子计算机(微型机)五类。表2-1列出了区分这几类计算机的一般标准。需要指出的是，这些标准并不是绝对的，随着各类计算机的发展，各项指标将不断有所改变，如超小型机VAX-11字长为32位。

电子计算机除按规模分类外，还有很多其他分类方法，如按用途分为通用计算机(能适合于科学计算、数据处理、过程控制等多方面应用的电子计算机)和专用计算机(为解决某一特定问题而专门设计的电子计算机)。

表2-1 电子计算机按规模分类

规 模 常用标准	巨型计算机	大型计算机	中型计算机	小型计算机	微型计算机
运算速度(次/秒)	五千万以上	十万— 五十万	百万—千万	百万	十万
字长(位)	64	32—64	32	16	4—16
主存容量(字节)	>128MB	32M— 128MB			

(注: 1MB = 10^6 B, 即百万字节)

电子模拟计算机也叫模拟式电子计算机, 简称模拟计算机。这种电子计算机不直接对数字进行运算, 而是运用一些物理量, 例如电压、长度等来表示被运算的数字, 根据相似原理来解答各种问题。模拟机不适合于数据处理领域中的应用。而数字计算机解题精度高、通用性强、便于数据的存贮, 因此在数据处理领域中使用的都是数字计算机。本书所提到的计算机, 非特别指明都是指数字计算机。

自从一九四六年第一台电子计算机“埃尼阿克”问世以来, 到现在只有四十年的时间, 计算机得到了迅猛的发展, 其发展速度之快, 可以说是日新月异。电子计算机的发展经历了: 电子管、晶体管、集成电路和大规模集成电路四代, 现在具有超大规模集成电路的第五代电子计算机正在研制之中。与物理器件发展的同时, 计算机在体系结构, 使用方式, 软件等方面也都经历了深刻的变化(见表2-2)。

我国计算机的发展始于五十年代。一九五八年研制成我国第一台电子管计算机, 一九六四年研制成晶体管计算机, 一九七一年国产集成电路计算机问世。八十年代初千万次向量计算机“757”和亿次计算机“银河”首次制成, 这是我国计算

表2 2 电子计算机各代划分和特征简表

计算机代	起止年份	代表机器	硬 件			软 件	使用方式特点	应用范围
			逻辑元件	主存储器	外存储器			
第一代	1947—1957	IBM-704 UNIVAC-1	真空管	延迟线	基本无	符号语言 汇编语言	个人独立机器	科学计算
第二代	1958—1964	IBM-7090 ATLAS	晶体管	磁芯	磁鼓 磁带	程序设计语言 多道程序设计 管理程序	多道程序	科学计算, 数据处理, 事务处理
第三代	1965—1970	IBM-360 CDC-6000 PDP-8 NOVA	中小规模 集成电路	磁芯	磁盘 磁带	操 作 系 统	分时 运行	实现系列化、标准化, 广泛应用于各领域
第四代	1970年后	IBM-4300 FACOM	大规模 集成电路	半导体存储器	磁盘 磁带	分布式操作系统 数据库	计算机 网络	微处理机和计算机网络应用, 更普及深入到社会生活各方面

机发展史上的重大事件。

第二节 电子计算机的用途

由于电子计算机具有运算速度快、精度高、存贮信息和逻辑判断能力强以及自动进行运算等特点, 使它得到广泛的应用。归纳起来, 电子计算机的用途可分为五类。

(一)用于科学技术计算。例如人造卫星轨迹的计算, 水坝应力计算, 房屋抗震强度的计算等。早期的电子计算机主要用于数值计算, 因此而得名“计算机”。现在, 虽然计算机的用途有了很大的发展, 但数值计算仍是计算机的一个重要用途。计算机广泛用于科学和工程中的许多复杂问题的方程求解,

多项式、微分和积分的计算以及概率统计计算等。人口预测、密码破译、地震统计预报、天气数值预报等工作都是借助于计算机通过对大量数据进行复杂的计算和分析而完成的。

(二)用于数据处理和信息加工。由于计算机具有逻辑判断能力,它可以对非数值的数据如字母、符号、表格、单据、资料、图形、图象乃至文字、语言、声音进行处理。所以计算机的应用早已突破了单纯的数值计算的范围,而发展到非数值应用领域。例如国民经济计划管理、仓库管理、财务管理、统计工作、情报检索、档案管理、办公自动化等。

(三)用于自动控制。计算机除了用于纯数值计算和管理工作以外,还可以应用到自动控制的系统中。特别是工业、交通的自动控制。如化工生产过程自动化,机械制造工艺加工自动化等。一个由计算机控制的钢厂,年产量一千万吨,只需一万名工人。一台带钢热轧机,使用计算机控制后,产量是人工控制的一百倍,而且质量明显地提高。在铁路运输中,列车的自动驾驶不仅减轻了司机的劳动强度,使列车安全准时地运行,还能合理地使用机车牵引动力。电子计算机可以在恶劣环境中代替人工作。

(四)计算机辅助设计CAD(Computer Aided Design)。这是近年来新兴的计算机应用的重要分支。目前计算机辅助设计已成为应用计算机改造传统工业、促进企业技术进步的具有战略意义的重要方向,不少国家将计算机辅助设计技术的广泛应用作为推动国民经济、特别是工业发展的重要因素之一。随着CAD的应用领域不断扩大,人们能够利用计算机部分地代替人工进行飞机、汽车、船舶、机械、电子、房屋、水坝、电路以及服装等的设计,它不仅把设计人员从繁琐的制图中解放出来,提高了设计人员的工作效率,而且在设计阶段同