

四川大学重点资助教材

DANG AN

XIN XI

JIAN SUO

档案信息检索

■ 陈丽 主编



四川出版集团 四川人民出版社
SICHUAN CHUBAN JITUAN SICHUAN RENMIN CHUBANSHE

四川大学重点资助教材

DANG AN XINXI JIANSUO

档案信息检索

四川出版集团 四川人民出版社
SICHUAN CHUBAN JITUAN SICHUAN RENMIN CHUBANSHE

图书在版编目 (CIP) 数据

档案信息检索/陈丽主编. —成都: 四川人民出版社, 2010. 4

ISBN 978-7-220-08071-5

I. ①档… II. ①陈… III. ①档案—情报检索 IV. ①G273

中国版本图书馆 CIP 数据核字 (2010) 第 039106 号

DANGAN XINXI JIANSUO

档案信息检索

陈丽 主编

责任编辑
封面设计
技术设计
责任校对
责任印制

杨 立
马欣晖
古 蓉
徐 英
丁 青 李 进

出版发行
网 址

四川出版集团 (成都槐树街 2 号)
四川人民出版社
<http://www.scpph.com>
<http://www.booksss.com.cn>
E-mail: scrmcbsf@mail.sc.cninfo.net

发行部业务电话
防盗版举报电话

(028) 86259459 86259455
(028) 86259524

照 排
印 刷
成品尺寸

四川胜翔数码印务设计有限公司
成都蜀通印务有限责任公司
170mm×240mm

印 张
字 数
版 次
印 次
书 号
定 价

20.25
300 千
2010 年 4 月第 1 版
2010 年 4 月第 1 次
ISBN 978-7-220-08071-5
25.00 元

■ 版权所有·侵权必究

本书若出现印装质量问题, 请与我社发行部联系调换
电话: (028) 86259624

前 言

信息检索是当前信息学界普遍关注的问题,档案信息检索也是关系着档案信息能否最终在更大范围内为社会所利用的现实问题,尤其是在计算机技术和网络技术为各类信息检索提供了广阔空间的今天,提高信息管理部门和普通民众的信息检索水平,更有利于各类信息在全社会得到共享。本书是在结合作者多年对档案信息管理研究成果以及档案信息检索教学经验的基础上,借鉴相关同类教材和专著的研究长处编写而成的。本书可作为高等院校以及高职高专院校信息管理类学生学习档案信息和其他信息检索使用的教材,同时也可以作为档案管理部门、政府机关内信息管理部门以及普通民众进行信息检索的参考书。

本书内容注重理论与实践相结合,全面详细地介绍了档案信息以及相关信息检索的基本知识,尤其关注计算机技术和网络环境下档案信息检索的现状与实施。本书共分十章,包括档案信息基础知识、档案信息检索基础、档案信息检索语言、档案信息著录与标引、档案信息检索类型、途径和程序、档案信息检索常用专业工具、档案信息检索常用中文工具书、计算机信息检索、政府信息检索、网络信息检索。本书在编写过程中力求突出几个特点,一是重视理论与实践相结合。本书作为高等院校学生用书,应该具有一定的理论高度,但本书也非常重视理论在实际中的运用,为相关管理部门和普通民众使用提供方便。二是考虑到了专业性与通俗性的结合。档案信息检索具有一定的专业性特点,本书在进行专业特点分析时尽可能进行相对通俗的解释和说明。三是兼顾了文理科的使用特色。考虑到信息管理类学科具有跨文理科的特点,本书在内容设计上也尽量两方面有所兼顾,为学生和教师使用提供方便。四是兼顾传统检索技术和现代检索技术的结合。在全面介绍传统的档案信息检索技术的同时重视对现代信息检索技术的介绍,这也是本书不同于以往档案信息检索教材的重要

特点。五是兼顾了社会对各类信息检索的需要。本书在重点介绍档案信息检索的同时对政府信息检索和网络信息检索也有所涉及。

陈丽负责本书的策划、书稿整体结构设计和内容组织,负责本书第1、2、3、4、5、6章和第7章第1节以及第8、9、10章部分内容的编写工作,并统稿全书。四川大学公共管理学院周亚参与第7章的编写,四川大学公共管理学院马英杰参与第8章的编写,四川大学公共管理学院唐婧参与第9章的编写,四川大学公共管理学院李昂参与第10章的编写。

在此特别要感谢的是四川大学公共管理学院博士生导师黄存勋教授对本书的热心关怀和积极帮助,黄老师对本书部分内容的结构设计提出了宝贵的意见和合理的建议,对此表示由衷的感谢!本书是四川大学重点资助教材,在课题申报中得到了四川大学公共管理学院的大力支持,对此表示衷心感谢!

本书在编写过程中参考了大量档案信息管理、信息检索相关论著的科学观点以及互联网上的一些优秀成果,有明确出处的都已在参考文献中列出,没有明确出处或因作者疏忽未能列出的,特向有关作者致歉并表示衷心感谢!同时还要向为本书的出版提供了大力帮助的四川人民出版社的周颖女士和杨立女士表示感谢,因为她们的热心支持本书才得以顺利付梓!感谢我先生对本书的出版给予的大力支持和帮助。由于作者水平有限,疏漏和欠妥之处在所难免,诚望读者指正。

作者

2010年1月于四川大学

目 录

前 言	(1)
第一章 档案信息基础知识	(1)
第一节 信息的概念	(1)
第二节 档案信息的特点	(2)
一、原始记录性	(2)
二、广泛性	(3)
三、可识别性	(3)
四、可存储性和可传递性	(3)
五、共享性	(4)
第三节 档案信息的类型	(4)
一、按照档案信息形成时间划分	(4)
二、按照档案信息形成属性划分	(4)
三、按照档案信息发挥作用划分	(5)
四、按照新中国成立前后时间划分	(5)
五、按照档案不同的载体划分	(6)
第四节 档案信息与其他信息的关系	(6)
一、档案信息与文书信息	(6)
二、档案信息与图书	(6)
三、档案信息与资料	(7)
四、档案信息与文物	(7)
本章思考题	(8)
第二章 档案信息检索基础	(9)

第一节 档案信息检索的基本工作原理	(9)
一、档案信息整序	(9)
二、档案信息检索	(10)
第二节 档案信息检索的现实价值	(10)
一、直接向社会提供有价值的各种档案信息	(10)
二、提高档案信息检索质量	(11)
三、提高档案信息管理标准化、规范化、科学化水平	(11)
四、促进档案管理的基础工作	(11)
第三节 档案信息检索工作现状	(12)
一、出版了一批档案检索专业著作	(12)
二、重视档案分类检索理论的研究	(12)
三、重视档案信息检索语言的进一步科学化和规范化	(13)
四、重视档案著录与标引技术的研究	(15)
五、对档案检索工具、检索方法和技术的研究更加深入	(16)
本章思考题	(17)
第三章 档案信息检索语言	(18)
第一节 检索语言的种类	(18)
一、分类语言	(19)
二、主题语言	(20)
第二节 档案信息检索语言	(21)
一、档案信息检索语言的基本含义	(21)
二、档案检索与图书检索的区别	(21)
三、档案信息检索语言的基本构成原理	(22)
第三节 《中国档案分类法》的结构和使用	(22)
一、《中档法》的编制原则	(23)
二、《中档法》的体系结构	(24)
三、《中档法》的编号方法和标记符号	(24)
四、《中华人民共和国档案分类表》的使用	(25)
五、《中档法》中其他几种分类表	(28)
六、《中档法》中关于几种特殊问题的处理	(28)
七、《中档法》中常用的辅助符号	(29)
第四节 《中国档案主题词表》的结构和使用	(30)
一、主题法简介	(30)

二、《中国档案主题词表》简介	(31)
三、《中国档案主题词表》的选词原则和选词范围	(31)
四、《中国档案主题词表》的结构体例及其使用	(32)
第五节 《中国图书分类法》简介	(39)
第六节 《汉语主题词表》的结构和使用	(40)
一、《汉语主题词表》的结构	(40)
二、《汉语主题词表》与《档案主题词表》的关系	(41)
本章思考题	(42)
第四章 档案信息的著录与标引	(43)
第一节 档案信息著录	(43)
一、档案信息著录的基本含义	(43)
二、档案著录的基本作用	(44)
三、档案著录规则	(44)
第二节 档案分类标引	(52)
一、档案分类标引的基本概念	(52)
二、档案分类标引的基本规则	(53)
三、档案分类标引的工作程序	(56)
四、档案分类标引的质量管理	(58)
第三节 档案主题标引	(61)
一、什么是档案主题	(61)
二、确立档案主题的依据	(61)
三、档案主题标引的步骤	(61)
第四节 档案标引中应注意的问题	(63)
一、分类标引和主题标引的关系	(63)
二、关于内容交叉档案的标引	(63)
三、关于多主题档案的标引	(64)
第五节 档案信息计算机著录标引	(65)
一、档案信息计算机著录的特点	(65)
二、档案信息计算机著录的基本要求	(69)
第六节 电子文件著录	(69)
一、电子文件著录与电子档案著录的关系	(69)
二、电子文件著录的基本原则	(70)
三、电子文件著录的特点	(74)

四、电子文件著录的组织	(75)
五、电子文件的著录项目	(76)
本章思考题	(77)
第五章 档案信息检索类型、途径和程序	(78)
第一节 档案信息检索的类型	(78)
一、按照检索的内容划分	(78)
二、按照检索结果的表现形式划分	(80)
三、按照检索结果的内容范围划分	(82)
第二节 档案信息检索的一般途径	(83)
一、以全宗为检索线索	(83)
二、以时间为检索线索	(83)
三、以事件为检索线索	(84)
四、以人物为检索线索	(84)
五、以地区为检索线索	(85)
六、以档案上各种标记为检索线索	(85)
七、“求书八法”的基本内容	(86)
第三节 档案信息检索步骤	(86)
一、接受检索要求	(86)
二、分析待检任务	(87)
三、制定具体检索步骤	(87)
四、实施检索	(88)
五、对检出结果的处理	(89)
六、档案信息检索工作的效率统计和测定	(89)
本章思考题	(90)
第六章 档案信息检索常用专业工具	(91)
第一节 检索工具的类型	(91)
一、条目式检索工具	(91)
二、卡片式检索工具	(96)
三、叙述式检索工具	(98)
四、其他类型的检索工具	(100)
第二节 档案检索工具常用排检法	(101)
一、字顺法	(101)
二、类序法	(108)

三、自然顺序法	(109)
本章思考题	(109)
第七章 档案信息检索常用中文工具书	(111)
第一节 中文工具书在档案信息检索中的作用	(111)
一、对档案中出现的特殊字、词、句进行解释	(112)
二、对档案中出现的某些人物进行解释	(112)
三、对时间、地点等进行解释	(113)
第二节 综合性字典、词典	(114)
一、字典、词典概述	(114)
二、字典、词典的基本结构	(115)
三、古今字典、词典举要	(115)
第三节 综合性百科全书	(123)
一、百科全书概述	(123)
二、百科全书的分类	(124)
三、综合性百科全书举要	(125)
第四节 人名及人物资料查询工具书	(128)
一、与人物有关的工具书特点	(128)
二、一般性人名辞典举要	(128)
三、专门性的人名辞典举要	(131)
四、人物传记	(135)
第五节 地名及地理资料查询工具书	(145)
一、地名和地理工具书的作用	(145)
二、地名和地理工具书举要	(145)
第六节 政书	(155)
一、政书概述	(155)
二、政书的类型及作用	(156)
三、重要政书简介	(157)
四、政书的查检	(160)
第七节 类书	(164)
一、类书概述	(164)
二、类书的特点及作用	(166)
三、重要类书简介	(167)
四、类书的查检	(171)

本章思考题	(172)
第八章 计算机信息检索	(173)
第一节 计算机信息检索概述	(173)
一、计算机信息检索的概念	(173)
二、计算机信息检索的特点	(174)
三、计算机信息检索发展	(175)
四、计算机信息检索发展趋势	(177)
五、我国计算机检索技术的发展	(178)
第二节 计算机信息检索原理及系统构成	(179)
一、计算机信息检索原理	(179)
二、计算机信息检索系统构成	(181)
第三节 计算机信息检索系统的类型	(185)
一、联机检索系统	(185)
二、光盘检索系统	(190)
三、国际互联网信息检索	(195)
第四节 计算机信息检索常用方法和步骤	(195)
一、计算机信息检索的方法	(195)
二、计算机信息检索的步骤	(199)
第五节 档案信息计算机检索	(200)
一、档案信息计算机检索的基本过程	(200)
二、档案信息计算机检索的必备条件	(200)
三、档案信息计算机检索的前期准备工作	(201)
四、建立档案信息数据库	(202)
五、完成档案信息计算机检索	(202)
第六节 档案馆计算机检索工作建设	(202)
一、档案馆计算机检索建设现状	(202)
二、我国档案馆计算机检索发展状况	(205)
三、档案馆计算机检索项目建设的目标与任务	(211)
本章思考题	(222)
第九章 政府信息检索	(224)
第一节 政府信息手工检索	(225)
一、政府信息手工检索的工具	(225)
二、政府信息手工检索工具的排检技术	(226)

三、政府信息手工检索的方法	(226)
第二节 政府信息计算机检索	(226)
一、政府信息计算机检索的主要方式	(227)
二、政府信息上网工程简介	(227)
三、政府信息上网工程的实施步骤	(228)
四、政府信息上网工程的意义	(228)
五、政府信息上网工程的实施	(229)
六、政府信息上网工程取得的成果	(232)
第三节 政府网站建设	(233)
一、政府网站建设的基本要求	(233)
二、政府网站建设评估	(234)
第四节 政府网站信息检索	(241)
一、政府网站信息的类型	(241)
二、政府网站功能模块构成	(242)
三、政府网站信息的检索方式	(243)
四、政府网站信息检索的现状	(243)
本章思考题	(245)
第十章 网络信息检索	(246)
第一节 国际互联网 Internet 简介	(246)
一、互联网的起源	(246)
二、互联网的发展	(247)
三、互联网在中国	(248)
第二节 Internet 基础知识	(249)
一、万维网 (WWW)	(250)
二、TCP/IP 协议	(251)
三、Internet 网地址和域名	(252)
四、超文本 (Hypertext)	(254)
第三节 Internet 的服务功能	(255)
一、远程登录 (Telnet)	(255)
二、文件传输协议 (FTP)	(256)
三、电子邮件 (Electronic Mail, 简称 E-mail) 和邮递表 (Mailing List)	(256)
四、Gopher	(257)

五、博客 (Blog)	(257)
六、RSS	(258)
第四节 网络信息资源	(258)
一、网络信息资源的含义	(258)
二、网络信息资源的特点	(259)
三、网络信息资源的类型	(261)
第五节 网络信息资源检索	(269)
一、网络搜索引擎简介	(269)
二、网络搜索引擎的特点	(270)
三、网络搜索引擎的分类	(272)
四、网络搜索引擎的工作原理	(278)
五、常见网络搜索引擎的使用方法及特点	(279)
第六节 档案信息资源网络建设	(297)
一、我国档案网站建设概况	(298)
二、我国档案网站的特色服务	(302)
三、国外档案网站建设概况	(303)
四、档案网站建设中需要注意的问题	(304)
本章思考题	(307)
主要参考文献和网络信息源	(308)

第一章

档案信息基础知识

第一节 信息的概念

信息普遍存在于自然界、人类社会以及人们的思维活动之中，它已成为人们须臾不可离的重要资源。关于信息的概念到目前为止有多种说法，具有代表性的观点有以下几种：

- (1) 信息是人们在适应外部世界并且将这种适应反作用于外部世界过程中，同外部世界进行交换的内容和名称；^①
- (2) 信息是能够用来消除不确定的东西。信息就是负熵^②，是组织程度、有序程度，它能使系统的有序性增强，减少破坏、混乱和噪声；^③
- (3) 信息是客观世界中各种事物的变化和特征的反映，是客观事物之间相互作用和联系的表征，是客观事物经过传递后的再现；^④
- (4) 信息是经过某种方式加工或更具意义的形式提供的数据；^⑤
- (5) 信息是信息用户与文献或数据^⑥的一种交互作用，是用户根据特定的

① N. Wiener. Cybernetics and Society. Houghton Mifflin Company, 1950.

② 负熵是物质系统有序化、组织化、复杂化状态的一种量度。熵是用以表示某些物质系统状态的一种量度或说明其可能出现的程度，或者说是描述一个孤立系统中物质的无序程度。（作者注）

③ C. E. Shannon. Mathematical Theory of Communication. Bell System Technical Journal. 1948.

④ 刘昭东：《信息与信息化社会》，科学技术文献出版社 1994 年版。

⑤ （美）斯蒂芬·哈格、梅芙·卡明斯、詹姆斯·道金斯 著，严建援等译：《信息时代的管理信息系统》，机械工业出版社 2000 年版。

⑥ 数据是指那些未经加工的事实或是着重对一种特定现象的描述。（作者注）

时间、特定的目的、具体问题环境和自身知识对文献或数据的主观认识和描述,并将随着目的、问题环境和自身知识的变化而变化。^①

根据以上各种观点可以看出,对信息的定义的理解主要基于三种不同的角度,即通信技术的角度、客观的角度和主观的角度。

档案是社会信息中一种特殊的信息记录方式,其内容即档案信息,是对客观事物的反映,其记录手段不同于一般的通信手段,具有更多的主观性,因此,档案是技术、客观、主观等多方面因素的集合体,在众多社会信息中具有独特的含义:档案指过去和现在的国家机构、社会组织以及个人从事政治、军事、经济、科学、技术、文化、宗教等活动直接形成的对国家和社会有保存价值的各种文字、图表、声像等不同形式的历史记录。^② 根据档案的定义,我们认为档案信息的概念应该这样理解:档案信息即档案中记载的内容,它是社会组织或个人在以往社会实践活动中直接形成的,具有清晰、确定含义的、被固定在各种载体上的原始记录。准确理解档案信息有利于对档案信息的管理,对档案信息检索而言具有更为重要的实际意义。它一方面可以明确档案检索工作目标,增强档案检索的目的性;另一方面,可以明确档案检索的研究范围,提高档案检索的有效性,提升研究水平。

第二节 档案信息的特点

对档案信息特点进行分析的目的主要在于对档案信息的表现形式有更加深入的了解,为实现档案信息检索奠定理论基础。

一、原始记录性

档案信息是人们在社会活动中直接形成的原始记录,所以“原始记录性”是档案信息区别于其他信息的本质特性之一。人们之所以认为档案信息是最具有权威的信息,就是因为它不是事后编写的,也不是事后追忆的,而是伴随着事情的发生发展而逐渐形成的,它是事情发展轨迹的真实记录。正是因为档案信息的原始记录性才使得档案管理的一系列工作具有特殊的意义,因为这种原始记录不同于一般书籍的内容或新闻稿件的内容,它需要专业的理念和技术来

① 党跃武、谭祥金:《信息管理导论(第二版)》,高等教育出版社2006年版,第4页。

② 《中华人民共和国档案法》(1987年9月5日第六届全国人民代表大会常务委员会第二十二次会议通过,根据1996年第八届全国人民代表大会常务委员会第二十次会议《关于修改〈中华人民共和国档案法〉的决定》修正)。

维护这种原始性,因此,这种原始性也使得对档案信息的检索不同于对一般信息的检索,它是一项具有特殊技术要求的工作。

二、广泛性

信息广泛存在于人类社会和物质世界的各个领域。无论是从档案的定义开始还是从现实生活中档案信息的实际情况看,档案信息的广泛性都表现得非常明确和直观。从时间上看,跨越了过去和现在;从内容上看,包括了国家、社会组织以及个人在政治、军事、经济、科学、技术、文化、宗教等活动中直接形成的有价值的信息;从形式上看,包括了各种文字、图表、声像等不同形式的记录方式;从载体上看,包括了甲骨、缣帛、纸张、胶卷、磁性载体、光盘等多种载体。因此,信息广泛性特点通过档案信息反映得最为全面和具体。

三、可识别性

人们不仅可以通过自己的感官识别信息,也可以通过各种检测手段来感知信息、接受信息,从而达到识别信息的目的。这也是信息之所以能够被利用的基本条件之一。档案信息可识别性是档案信息检索的重要条件基础,人们首先能够识别不同类型的档案信息,才能够进一步对档案信息进行检索利用。档案信息的可识别性在档案检索中表现得非常充分和全面。第一,档案信息检索可以对单独的信息载体形态进行识别,如可以对一份文件、一组案卷、一个全宗等信息个体在一定范围内加以识别;第二,档案信息检索可以把档案信息具体内容作为识别单元,如对概念、题名、主题、分类、责任人等内容要素加以识别;第三,档案信息检索可以对一定的信息利用单元进行识别,如对用户类型、利用方式、检索效率、工作量、人员构成、管理流程等加以识别。

四、可存储性和可传递性

信息总是可以通过各种载体加以存储。档案在几千年的发展过程中,其内容的存储方式发生了翻天覆地的变化。从远古时代的龟甲、兽骨、青铜、缣帛,到纸张、胶片、磁盘、光盘,档案信息的载体形式伴随着科技的进步不断发生变化,这种变化集中反映了档案信息可存储性的这一特征。与可存储性相伴随的是存储载体的可转换性,档案信息可以从一种载体形式转换为另一种载体形式,这种特性使得档案信息数字化处理和计算机检索成为可能。

信息的可存储性使信息的传递成为可能,存储信息一方面是为了保存信息,另一个更为重要的目的是为了传递信息,让更多的人知道信息的内容,因此传

递性往往与信息共享是密不可分的。在档案管理的发展进程中,人们由最初的重视档案的“存本”作用逐渐发展到重视档案信息的“资治”作用,档案信息的可传递性起到了非常重要的作用,正是因为可以传递,档案信息才能够在一定条件下发挥其应有的社会价值。

五、共享性

共享性是信息的重要特征之一。档案信息作为人类社会中一种特殊信息在一定条件下是能够被人们所共享的。《中华人民共和国档案法》规定,国家档案馆保管的档案,一般应当自形成之日起满30年向社会开放。经济、科学、技术、文化等类档案向社会开放的期限,可以少于30年,涉及国家安全或重大利益以及其他到期不宜开放的档案向社会开放的期限,可以多于30年。档案馆应当定期公布开放档案的目录,并为档案的利用创造条件,简化手续,提供方便。中华人民共和国公民和组织持有合法证明,可以利用已经开放的档案;机关、团体、企业事业单位和其他组织以及公民根据经济建设、国防建设、教学科研和其他各项工作的需要,可以按照有关规定,利用档案馆未开放的档案以及有关机关、团体、企业事业单位和其他组织保存的档案。可见,档案信息作为社会信息的一种,其共享性特征在这里得到了具体体现。

第三节 档案信息的类型

档案信息内容广泛,形式多样,按照不同的划分标准有多种不同的类型。

一、按照档案信息形成时间划分

按照档案信息形成的时间可以划分为:历史档案信息和现行档案信息。

历史档案信息是指档案信息所涉及的时间较早,离现在较久远,它主要能够帮助人们从事历史研究和进行文化的积累。现行档案信息是指档案信息所涉及的时间较晚,离现在的时间距离较近,它主要能发挥现实性查考作用。这种划分与档案定义中所说的“档案是具有保存价值的历史记录”是相一致的,它说明了档案信息不仅对现实工作和生活有查考作用,而且对帮助人们了解历史、研究历史,传承精神文化同样是一种不可缺少的重要信息。

二、按照档案信息形成属性划分

按照档案信息形成的公、私属性可以划分为:公务档案信息和私人档案