



数字时代图书馆学情报学研究论丛

(第二辑)

基于业务规则的 档案信息资源管理

ARCHIVES INFORMATION RESOURCES MANAGEMENT
BASED ON BUSINESS RULES

王新才 江善东 等 著



WUHAN UNIVERSITY PRESS

武汉大学出版社



数字时代图书馆学情报学研究论丛

(第二辑)

本书为粤港关键领域重点突破项目“电子政务环境下基于业务规则的档案信息资源管理支撑平台”的成果之一

基于业务规则的 档案信息资源管理

ARCHIVES INFORMATION RESOURCES MANAGEMENT
BASED ON BUSINESS RULES

王新才 江善东 著

参与撰著人员：

李 雯 吴佳鑫 沈 阳 丁石征
赵欲勤 刘海超 王莉莉 邓春来



WUHAN UNIVERSITY PRESS

武汉大学出版社

图书在版编目(CIP)数据

基于业务规则的档案信息资源管理/王新才,江善东著. —武汉: 武汉大学出版社,2014. 6

数字时代图书馆学情报学研究论丛

ISBN 978-7-307-12703-6

I . 基… II . ①王… ②江… III . 档案工作—信息管理 IV . G27

中国版本图书馆 CIP 数据核字(2014)第 004341 号

责任编辑:詹 蜜

责任校对:鄢春梅

版式设计:马 佳

出版发行: **武汉大学出版社** (430072 武昌 珞珈山)

(电子邮件: cbs22@ whu. edu. cn 网址: www. wdp. com. cn)

印刷:武汉中远印务有限公司

开本: 720 × 1000 1/16 印张:15. 5 字数:219 千字 插页:2

版次:2014 年 6 月第 1 版 2014 年 6 月第 1 次印刷

ISBN 978-7-307-12703-6 定价:32. 00 元

版权所有,不得翻印;凡购我社的图书,如有质量问题,请与当地图书销售部门联系调换。

《数字时代图书馆学情报学研究论丛》

编 委 会

顾问：彭斐章(武汉大学资深教授,信息管理学院博士研究生导师)
孟广均(中国科学院国家科学图书馆教授,博士研究生导师)
吴慰慈(北京大学资深教授,信息管理系博士研究生导师)
胡述兆(台湾大学图书资讯学研究所教授,博士研究生导师)
梁战平(中国科技信息研究所教授,博士研究生导师)
倪 波(南京大学教授,博士研究生导师)
黄长著(中国社会科学院研究员,学部委员)
冯惠玲(中国人民大学副校长,信息资源管理学院博士研究生导师)

Raymond von Dran (Professor and Dean, School of Information
Studies, Syracuse University)

Harry Bruce (Professor and Dean, Information School, University of Washington)

主编：陈传夫 马费成 胡昌平

编委：(按姓氏笔画排序)

马费成	方 卿	王新才	邓仲华	司马朝军
刘家真	朱玉媛	朱静雯	余世英	吴 平
张玉峰	张李义	张美娟	李 纲	肖希明
邱均平	陆 伟	陈传夫	周 宁	周耀林
罗紫初	查先进	胡昌平	赵蓉英	唐晓波
徐丽芳	曹 之	黄先蓉	黄如花	焦玉英
董有明				

序

“图书馆学情报学”是我国的习惯用法，是涵盖图书馆学、情报学、档案学、出版发行学等学科的名称。在我国台湾被称为“图书馆与资讯科学”，英文为 Library and Information Science。美国也用 Library and Information Studies 来称谓这一学科。

1807 年，德国学者马丁·施莱廷格（Martin Schrettinger, 1772—1851）首次使用了“图书馆学”这一概念，1808 年他又在《试用图书馆学教科书大全》中建立了以图书馆整理为核心的学科体系，标志着图书馆学学科正式诞生。

自 1887 年美国学者杜威（Melvil Dewey, 1851—1931）在哥伦比亚大学创办世界第一所图书馆学校，1930 年在卡内基基金的资助下芝加哥大学设立第一所图书馆学博士班课程以来，图书馆学开始走进大学殿堂，成为高等教育中的一个专业。

图书馆学教育在美国的兴起带动了全球图书馆教育的发展。1919 年英国在伦敦大学建立了图书馆学院。目前，美国有 56 所美国图书馆学会（ALA）认可的图书馆学院，每年招收图书馆与情报学学生 26 000 人左右。

在施莱廷格后的两个世纪，图书馆学科不断变化。特别是在 20 世纪 50 年代以来的“冷战”期间，美苏军备竞赛，两大阵营形成。苏联卫星上天，美国实施阿波罗计划，科技文献激增。科学家对文献信息的获取变得困难。一门新型学科——情报学应运而生。1963 年美国文献工作学会正式更名为美国情报学会（ASIS）。大量增设图书馆学与情报学硕士点、博士点。图书馆学课程表中也增加

了大量的情报学课程。

20 世纪 70 年代，计算机技术在图书馆与信息工作中广泛应用，自动化、地区性图书馆网络形成，机读目录广泛应用，国际图联将世界书目控制列为核心计划。图书馆学（Library Science）发展为“图书馆与情报学”（Library and Information Science），后来又进一步演变为“图书馆与情报研究”（Library and Information Studies）。

20 世纪 80 年代高新技术迅速发展，信息时代到来。美国里根政府实施星球大战计划，欧洲实施尤里卡计划等。联机图书馆系统广泛建立，并扩展至世界主要发达国家。商业性联机数据库如 ORBIT，DIALOG 发展迅速，图书馆与情报职业面临挑战。为适应信息时代要求，国际上图书馆学情报学专业开始调整。国际上有较多大学将图书馆学院易名为图书馆与情报学院或信息研究学院，图书馆学、情报学在硕士、博士层次合二为一。

20 世纪 90 年代，全球进入后信息时代——数字时代到来。克林顿政府开始实施国家信息基础设施计划（NII）、全球信息基础设施计划（GII）。新一代互联网投入使用。欧美初步建成信息社会，全球进入无缝信息环境。世贸组织建立和一揽子贸易协定生效，使全球经济一体化并逐步进入知识经济时代。各国继续加强图书馆学、情报学学科调整。图书馆学、情报学学科内容向情报科学汇集。

进入 21 世纪以来，国际上信息管理学科变化很快。自雪城（SYRACUSE）大学将学院更名为信息研究学院（The School of Information Studies）后，在美国立即出现了 iSchool 的浪潮。伊利诺依大学、华盛顿大学、密歇根大学、匹兹堡大学、加州大学伯克利分校、北卡罗来纳大学等知名大学的图书馆与情报学院宣称自己为 iSchool。这些 iSchools 通过宪章组成 I-Schools 联盟（ISG）。目前共有 20 所美国的大学加入联盟（联盟宪章不允许超过 25 个）。iSchool 强调信息、技术与人的关系（relationship between information, technology and people）。iSchool 的标准包括：必须有杰出的研究和杰出的博士教育；必须能在科学、企业、教育与文化进步过程

中提供任何形式的信息所需的专门技术；必须能提供信息技术及其应用、信息使用与用户方面的专门知识。2004—2006 年的联盟领导委员会协调人是雪城大学信息研究学院的 Raymond von Dran 院长，2006—2007 年将由匹兹堡大学信息学院院长 Ron Larsen 担任。联盟成员的标准主要强调研究即实质性承担研究活动（三年中每年研究支出达到 100 万美元），同时，致力于培养未来的研究者（通常通过研究型的博士点），引领推动信息职业领域。

国际上图书馆与情报学科的发展表现出明显的特征：研究范围由传统的图书馆领域扩大到信息领域（information field），研究视野由实体的图书情报机构扩大到虚拟空间，研究对象由图书文献转向了信息内容。一系列相关学科如图书馆学、情报学、档案学、出版科学、信息管理与系统乃至数字商务汇集于信息科学（Information Sciences）下，从而使图书馆学情报学研究发生了根本的变化。

武汉大学图书馆学科起源于 1920 年美国学者韦棣华女士创办的武昌文华大学图书科，档案专业起源于 1940 年的文华图书馆学专科学校的档案管理科。1978 年武汉大学创办科技情报学专业，后改为情报学专业。1983 年创办图书发行学专业，2002 年创办电子商务专业。1984 年经教育部批准建立武汉大学图书情报学院。2001 年更名为信息管理学院。图书馆学和情报学两个二级学科被国务院学位委员会批准为国家重点学科。“图书馆、情报与档案管理”被国务院学位委员会批准为一级学科博士学位授权点。教育部批准“武汉大学信息资源研究中心”为国家人文社会科学重点研究基地。信息产业部批准成立“国家信息资源管理（武汉）研究基地”。新闻出版总署批准建立“新闻出版总署武汉大学高级出版人才培养基地”。“网络信息资源开发与数字图书馆建设”被国家计委、教育部等批准为“十五”211 重点学科建设项目。建立一级学科博士后流动站。武汉大学信息资源研究创新基地被列为国家“985 二期工程”建设项目。一批院内校级重点研究基地如武汉大学四库学研究所、武汉大学中国科技评价中心、武汉大学政府信息

研究中心、武汉大学数字图书馆研究所、武汉大学出版发行学研究所、武汉大学图书馆学情报学国际合作研究中心也在科研和人才培养中发挥着重要平台作用。

强调一级学科内学科群建设和学科协调发展是武汉大学图书馆与情报学科建设的基本目标。以图书馆学、情报学两个国家重点学科为龙头促进图书馆学、情报学、档案学、信息资源管理、出版发行学等学科的协调发展。

我们深刻认识到信息资源与自然资源、人力资源共同构成支撑现代经济社会发展的资源体系。信息资源是知识经济时代重要的国家战略资源，是实现经济和社会全面、可持续发展的基础条件。对信息资源的拥有、开发和利用水平，是衡量一个国家综合国力和国际竞争力的重要标志之一。消弭信息鸿沟、实现信息公平，是消除贫困、促进经济发展、构建和谐社会的重要条件之一。

信息资源管理人才培养是学院的基本任务。学院每年为国家培养本科生 260 名，硕士研究生 150 名，博士研究生 55 名左右。学院有一支知识结构和年龄结构合理的优秀学术队伍。这支队伍中有武汉大学人文社会科学资深教授 1 人，博士研究生导师 26 人，国务院政府特殊津贴专家 6 人，教育部新世纪优秀人才支持计划 3 人，武汉大学珞珈特聘教授 2 人。作为实现研究型学院建设目标的一部分，在教学的同时，广大教师承担了大量的科学研究任务。为了推动本学科领域的前进，分享他们的见解，在武汉大学出版社的大力支持下，并报有关部门批准，我们拟出版《数字时代图书馆学情报学研究论丛》（简称《论丛》）。

为了编辑这套丛书，武汉大学邀请了国内外知名学者担任《论丛》的学术顾问，组建了主要由信息管理学院的博士研究生导师担任委员的编辑委员会。

《论丛》拟用 4 年时间出版著作共 20 卷。20 卷著作将分为三个系列：（1）学科年度进展。主要约请信息管理学院图书馆学系、档案与电子政务学系、信息管理科学系、现代出版系、信息系统与

电子商务系的有关教师和校外专家共同编写本学科的年度研究进展，主要有《图书馆学研究进展》、《情报学研究进展》、《档案学研究进展》、《出版学研究进展》、《信息资源管理学研究进展》；（2）个人学术专著。涉及图书馆、情报与档案管理基本理论研究、信息组织与检索、信息资源管理、信息资源建设与信息服务、文献编纂与出版、数字图书馆与信息系统工程等研究方向；（3）研究报告系列。我院研究人员共承担教育部哲学社会科学研究重大攻关项目、国家社会科学基金重点项目、教育部人文社会科学重点研究基地重大招标项目、国家自然科学基金项目、国家社会科学基金项目多项。特别是211项目和985项目，围绕数字信息资源开发与管理、数字信息资源服务与保障、信息资源公共获取与知识产权协调管理、数字图书馆关键技术与系统、资源与服务整合、信息构建与知识管理等主题正在进行探索。在信息构建的理论与方法、信息系统与资源整合、元数据知识表达、网络计量与参考、信息服务集成机制、信息资源与服务集成技术、媒体及数字出版、数字内容分销、信息资源的长期保存、商务信息流等关键领域力图实现图书馆学科在数字图书馆领域、情报学科在数字资源管理领域、档案学在数字化政务信息管理领域、出版发行学在数字出版与数字化分销、信息系统科学在集成系统以及数字化商务信息流研究方面取得研究成果。本系列将对部分研究结果进行报告。

丛书的出版是学院广大教师和研究人员辛勤探索的结果，在此，谨向严谨治学、辛勤耕耘的各位著作者表示感谢！对武汉大学出版社的支持表示感谢。同时，还望广大读者不吝批评指正，共同推动图书馆学、情报学、档案学、出版发行学和信息资源管理学科的进步！

武汉大学信息管理学院 陈传夫
武汉大学信息资源研究中心 马费成

前 言

21 世纪以来，计算机技术、通信技术和网络技术的持续创新与进步，全方位推动着社会信息化的进程。在信息化大潮下，我国档案信息化建设在档案数字化、档案网站和档案数据库建设、电子文件与电子档案的一体化管理、面向社会的档案信息资源整合、共享与利用，档案信息化标准及法规建设、档案信息化人才培养等方面都取得了长足的进步。

然而，在这些成绩的背后，却依然存在许多亟待解决的问题，特别是现有的档案信息管理系统在开发和实施过程中，仍旧沿袭着传统归档管理模式，归档效率不高、归档质量管理薄弱，档案信息管理系统有效使用期短、运作效率低、维护成本高等问题的存在根本无法满足信息时代人们对于信息管理智能分析、辅助决策支持的要求，且严重制约了档案部门优质高效地开展业务工作，成为影响档案信息化进程的重要桎梏。

为此，档案工作规划者和系统设计者，必须构建一个具有智能分析、能够提供档案业务提示并对该业务系统进行自动控制、集中统一管理零散的档案业务规则、促进档案标准化建设，并能实时监控和综合评价档案管理工作等功能的高效档案信息管理系统。业务规则管理技术的出现，恰好满足了档案部门这一需求。

业务规则（Business Rule）最初兴起于 20 世纪 80 年代国外的人工智能研究领域，20 世纪 80 年代后期得到较快发展并逐步受到社会各界关注。当前，国外几乎所有大型通用软件提供商、大型行业软件提供商以及大型行业用户，如 ILOG、IBM 等公司在系统开发和运行过程中都使用了业务规则这一技术。而且基于业务规则的应用系统越来越多、越来越普遍，在银行、保险、医疗卫生、电子

政府等领域都有业务规则的“身影”。在我国，比较有代表性的有凯科斯特（北京）科技有限公司的 KCOM 商业工程。

业务规则组织（Business Rule Group, BRG）把“业务规则”界定为两个层面的含义：从企业业务的角度看，业务规则是支持企业决策，影响或控制企业业务行为的指示；从计算机信息系统的角度看，业务规则是一条语句，它定义或约束业务的某些方面。其目的是对业务结构做出断言，或者对业务行为施加控制和影响。也就是说，业务规则内涵丰富，既包括明文规定与业务相关的法律法规、操作规范、行业标准等，也包括业务指导长期实践的经验、知识总结、专门技能等没有形成文字的惯例。而这些在档案管理系统中随处可见。但我国档案管理系统中，业务规则主要存在于档案工作人员脑海中，或者相关的业务手册、管理规程等非结构化的文档中，即使某单位有了信息系统，规则逻辑也只是作为一个过程写进程序代码中，当业务和需求发生变动时无法做出及时响应，必须对原有程序代码进行修改，造成了工作上的诸多不便。

现代业务规则管理技术将传统业务系统引入业务规则层，彻底改变了以过程形式处理业务逻辑的方式，将业务规则的实现从具体的程序代码中抽取出来，以结构化的业务规则数据来表示组织的业务行为，使得业务规则与组织的数据信息一样成为组织的主要资产。从这个角度看，业务规则并不是新生事物，而只是一种新的数据类型。这为解决我国目前档案信息管理系统存在的众多问题提供了一条新的思路和途径。

从国内外业务规则研究和发展的实践，结合我国档案管理系统的发展现状、特征及需求思考，选取业务规则的视角，设计档案信息管理系统可谓另辟蹊径。基于业务规则的档案管理系统（从计算机信息系统的角度）具备除一般档案管理系统的档案管理、利用编研的功能外，拟从管理和技术两个角度来探讨档案信息资源开发与管理问题，重点研究一套基于信息资源规划、依托业务规则库的档案信息资源收集管理体系，同时也为档案管理部门建立起智能化的网上监督指导系统、绩效评估系统和上岗人员自学习系统。还可以基于档案业务规则库（ABRD）和推理引擎（RM）构建档案



信息雷达，自动从相关业务系统收集档案信息资源，进而大幅度提高档案管理系统的效能，增强档案管理系统的交互性能和可维护性，优化档案管理个性化服务的特征。因此，基于业务规则的档案管理系统将有利于组织和利用档案信息、提高档案管理工作质量、规范档案信息资源的采集与共享、提升档案管理工作者的业务水平、积累和传播档案管理知识、降低档案信息系统扩展和变更的成本、便于业务规则的修改和维护，促进档案管理信息系统向着规范化、精细化、集成化、智能化的方向迈进。

本书的编写起源于多年前东莞市万维网络科技信息有限公司与武汉大学信息管理学院档案系的合作。东莞万维，现在已发展成为广东万维博通信息技术有限公司，是国内为数不多的专业的电子政务解决方案提供商和档案信息资源综合服务商，长期以来一直致力于电子政务与档案信息资源管理软件研发、信息资源规划与建设，并于2009年创办了华南地区首家档案信息资源研究所——东莞市万维档案信息资源研究所。公司已经通过双软认定、ISO9001:2008认证和CMMI（能力成熟度模型集成）认证，是广东省首批档案中介备案企业，是东莞首批“国家级高新技术企业”，是粤港关键领域重点突破项目和国家创新基金项目的承担单位。本书即属粤港关键领域重点突破项目“电子政务环境下基于业务规则的档案信息资源管理支撑平台”的成果之一。这一项目中标于2008年12月，由东莞万维、武汉大学信息管理学院和东莞市档案局共同承担。2009年2月14日，三方就项目的主要研究开发内容、达到的技术指标、获得的研究开发成果及形式进行了详细讨论。

本书的编写也在这之后正式确定下来。由王新才和江善东统稿，具体撰写者有李雯（第1章、第2章、第4章、第6章），吴佳鑫（第3章），沈阳、赵欲勤（第5章），刘海超（第4章），丁石征（第2章），王莉莉（第6章、第7章），邓春来（第7章）及王新才（第1章、第2章）等人。特此说明。

目 录

1 绪论	1
1.1 档案信息化现状	1
1.2 档案信息化发展趋势	4
1.2.1 档案管理规范化	4
1.2.2 档案管理精细化	5
1.2.3 档案管理集成化	6
1.2.4 档案管理智能化	7
1.3 业务规则和业务规则管理	8
1.3.1 什么是业务规则	8
1.3.2 什么是业务规则管理	10
1.4 业务规则管理技术国内外研究和应用现状	11
1.4.1 国外	11
1.4.2 国内	15
1.5 基于业务规则档案信息管理系统开发的意义	16
2. 档案业务规则管理概述	20
2.1 档案业务规则的内涵	20
2.2 档案业务规则的特性和作用	21
2.2.1 档案业务规则的特性	21
2.2.2 档案业务规则的作用	22
2.3 档案业务规则来源	23
2.3.1 档案法规和标准	24
2.3.2 未成文的经验和惯例	33
2.3.3 管理信息系统里的文档	33

2.4	档案业务规则管理	34
2.4.1	传统档案工作业务规则管理	34
2.4.2	利用业务规则管理技术实施档案业务 规则管理	35
2.5	档案业务规则管理系统	38
2.5.1	业务规则管理系统的基本原理	38
2.5.2	档案业务规则管理系统的结构	39
3.	档案业务规则的表达和组织	43
3.1	业务规则的逻辑学基础	43
3.1.1	命题逻辑	43
3.1.2	谓词逻辑	44
3.1.3	模态逻辑	45
3.2	业务规则的分类	46
3.2.1	不同的业务规则分类方式	47
3.2.2	Ross 的业务规则分类	48
3.3	档案业务规则的组成结构	51
3.3.1	BRG 的业务规则组成结构	52
3.3.2	SBVR 的业务规则语义模型	60
3.4	档案业务规则的表达	63
3.4.1	业务规则表达的生命周期	63
3.4.2	业务规则模板表示方法	65
3.4.3	决策表表示业务规则	68
3.4.4	决策树表示业务规则	72
3.4.5	规则标记语言	73
3.4.6	其他的业务规则表示方法	77
4.	基于业务规则的档案信息资源规划	79
4.1	档案信息资源规划概述	79
4.1.1	信息资源规划和档案信息资源规划的定义	79
4.1.2	档案信息资源规划的实施步骤	82

4.2	档案业务规划	88
4.2.1	档案信息资源的定义和定位	88
4.2.2	档案业务分析	88
4.2.3	档案业务流程重组	92
4.3	档案业务规则的规划	95
4.3.1	档案业务规则源	95
4.3.2	档案固化规则与动态规则	96
4.4	基于业务规则的档案信息资源管理模型	97
4.4.1	现有档案信息资源管理系统 (AMIS) 的管理模式	97
4.4.2	基于业务规则技术的档案信息资源管理模型	98
4.4.3	基于业务规则的档案信息资源管理系统的特 点	99
4.5	基于业务规则的档案信息资源管理标准规范体系	101
4.5.1	档案业务标准	102
4.5.2	项目质量标准	104
4.5.3	软件工程标准	107
4.5.4	安全标准	108
4.6	基于业务规则的档案信息资源管理关键技术	111
4.6.1	面向服务的架构 SOA	111
4.6.2	人工智能技术	112
4.6.3	网络信息雷达	114
5.	档案业务规则管理系统	116
5.1	档案业务规则管理系统简介	116
5.2	系统开发方法和开发步骤	118
5.2.1	系统分析	118
5.2.2	系统设计	120
5.2.3	系统部署和实施	130
5.2.4	系统维护与评价	133
5.3	档案业务规则管理系统的功能	135

5.4	档案业务规则管理系统的结构	137
5.4.1	档案业务规则集成开发环境	139
5.4.2	档案业务对象模型	142
5.4.3	档案业务规则引擎	146
5.5	规则权限管理和系统安全管理	147
5.5.1	规则权限管理	147
5.5.2	系统安全管理	149
6.	业务规则在档案信息资源管理中的应用	152
6.1	档案信息鉴定	153
6.1.1	基于业务规则的档案鉴定标准	154
6.1.2	基于业务规则的档案价值鉴定主要内容	158
6.2	档案信息资源收集	160
6.2.1	档案移交（被动收集）	160
6.2.2	档案采集（主动收集）	163
6.3	档案信息资源整理	167
6.3.1	电子档案信息预处理	167
6.3.2	档案整编	168
6.3.3	档案著录	169
6.4	档案信息资源管理的监督指导和绩效评估	175
6.4.1	监督指导系统	175
6.4.2	绩效评估系统	177
6.5	档案从业人员网上自学习系统	179
6.5.1	档案管理理论学习子系统	179
6.5.2	电子档案管理模拟实验室	180
6.5.3	考核子系统	181
7.	基于业务规则的档案信息资源管理案例	183
7.1	案例单位简介——广东万维博通信息技术有限 公司概况	183
7.2	万维博通档案业务规则管理系统介绍	184

7.2.1 系统特点	184
7.2.2 软件架构	190
7.2.3 运行环境	196
7.3 规则库建立及应用	199
7.3.1 规则库创建和管理	199
7.3.2 规则引用	203
7.4 档案业务规则在档案整理中的作用	204
7.4.1 档案整编	205
7.4.2 检索利用	216
7.4.3 档案编研	217
7.5 系统分析与评价	217
参考文献.....	221

1. 绪 论

1.1 档案信息化现状

20 世纪 90 年代以来，信息技术不断创新，信息产业持续发展，信息网络广泛普及，信息化成为全球经济社会发展的显著特征，并逐步向一场全方位的社会变革演进。进入 21 世纪，信息化对经济社会发展的影响更加深刻，电子政务在提高行政效率、改善政府效能、促进信息资源共享、扩大民主参与等方面的作用也日益显著，“推行电子政务”和“加强信息资源的开发利用”已列入《2006—2020 年国家信息化发展战略》的九大战略重点行列内。

在当今信息化时代，信息资源与物资资源、能量资源，并列为现代经济发展的三大支柱，成为每个国家和地区都必须高度重视的核心战略资源，更是国家综合竞争力的重要因素，档案的原始记录这一本质属性使得档案信息在信息资源体系中占有重要地位并发挥着独特作用。《中共中央办公厅、国务院办公厅关于加强信息资源开发利用工作的若干意见》（中办发〔2004〕34 号）中明确了档案馆是我国政务信息公开的主要场所之一，并要求加快以传统载体保存的公文、档案、资料等信息资源的数字化进程。档案工作面临着前所未有的发展机遇。把档案信息化纳入国家信息化建设、电子政务建设的总格局中是档案工作适应社会信息化发展的必然趋势。

我国档案信息化始于 20 世纪 80 年代，当时只是利用计算机辅助档案管理，只能对档案工作中的部分环节的信息进行处理。从 21 世纪开始，计算机信息技术在档案管理工作中的应用已相

当普及，主要应用为：数字档案馆的建设、电子政务中的电子文件与电子档案的一体化管理、档案信息资源的整合与面向社会的应用等。近年来，各地档案行政管理部门相继对档案信息化建设进行了部署，取得了一定的成效。许多市和区、县级档案局（馆）以及大部分专门档案馆，已建立了局域网和档案信息网站，计算机辅助档案管理全面推行，档案数字化工作有了较快进展，有些市、区档案馆也在建设档案目录中心。同时，档案信息化的法规体系的建立有了一定基础，部分标准已经产生。总体来说，我国档案信息化现在正处于发展阶段，全国档案信息化包括硬件、软件环境，档案信息资源以及人才等方面都有所进步。

然而，在档案信息化如火如荼地开展过程中仍然存在许多亟待解决的问题，现有的档案信息资源管理系统还远远不能令人满意，不能为智能分析、决策支持提供基础服务，各级档案部门陷入档案业务工作低质低效的困境。在实际的档案信息资源管理系统开发和实施过程中，我们经常会遇到以下四个问题：

（1）档案归档管理模式没有根本性突破

我国档案正在逐步实现计算机系统的目录管理和全文管理，有的还实现了网络化电子文件归档。但是目前档案归档管理模式并没有根本性的突破，这主要表现在，对归档部门的业务指导方法不得力，兼职档案人员的管理不到位，归档文件的质量管理处于薄弱状态等。传统的归档模式对于如何提高归档效率，实现对文件形成与归档的过程、时间和质量的实时监控，以及建立对归档人员激励评价的科学管理机制，仍缺乏科学有效的解决方法。

（2）档案信息资源管理系统有效使用期短暂

由于现在我国档案事业处在一个快速发展的阶段，各种规章制度、管理章程、操作规范会发生变化，又由于档案种类和形式多样化，对于不同种类的档案，不同级别部门有着不同的业务管理标准。不少档案部门为了提高开发效率而把档案信息资源管理系统的开发整体外包给 IT 公司，业务规则以代码形式散落、驻留于计算

机系统内，非档案专业的技术人员对档案业务规则不熟悉，这些业务规则每改变一次，软件就需要修改一次。经常有档案管理系统刚上线，还没等验收，因为某些规则的调整，业务部门已经提出修改软件的要求。

(3) 档案信息资源管理系统低效运作

档案信息资源管理系统的直接用户是单位或部门的档案工作人员，他们熟知档案的管理流程、运作模式以及业务规则的制定和管理。但由于业务规则以代码形式散落、驻留于计算机系统中，系统缺乏灵活性和友好性，一切规则的调整必须经过档案工作人员与 IT 人员沟通后由 IT 人员直接参与实施，业务规则的增删和修改往往都需要经过一个非常长的系统修改时间，影响了管理系统的高效运作。

(4) 档案信息资源管理系统维护成本高昂

现在多数人把信息系统的上线运行，作为系统成功的标志，实际上这仅仅迈出了万里长征的第一步，从工作量上说，至多完成了 30%，更多的工作还是以后的软件升级、系统维护、数据资源开发等，这些工作使系统的维护成本极为昂贵。档案信息资源管理系统正常上线运行后，由于档案业务人员一般缺乏专业 IT 人员的技术能力，档案信息资源管理系统的维护依赖于相应的 IT 人员和公司，相比于信息系统技术“自给自足”的部门，档案部门的信息系统维护成本会更加昂贵。

由此可见，现有的档案信息资源管理系统还远远不能令人满意，更没有达到智能分析、辅助决策支持的要求，各类档案部门陷入档案业务工作低质低效的困境。如何有效地构建一个能够支持辅助决策、及时响应需求变化、可充分扩展并具长久生命周期的档案业务系统成为档案工作规划者和系统设计者所面临的现实问题。

高效档案管理要求信息系统应该具有智能分析功能，能够为档案业务工作者提供业务提示并对业务流程进行自动控制，实时监控和综合评价档案管理工作，有效地指导档案工作人员规范化工作，

从而改变传统档案归档管理模式。此外，还应提高档案业务人员在信息系统设计、实施中的参与度；对散落于档案管理系统中的业务规则实施集中统一管理，促进档案标准化管理；自动分析抽取出档案信息中对当前决策有参考价值的信息。这些都需要在档案信息化建设过程中解放思想、积极探索、利用先进的计算机和网络技术，结合人工智能和业务规则管理方法，促进高效档案管理和档案信息辅助决策的实现。

1.2 档案信息化发展趋势

在各种信息资源中，档案以其与人类社会各种活动的伴生关系及所载信息的原始性、真实性等特殊性质，占有独特的地位，它是各级领导决策参考的重要依据之一，档案信息化建设也成为电子政务建设的重要内容之一。发挥档案信息资源价值最好的途径就是在档案信息化建设中实现档案信息自动化、标准化、集成化和智能化的管理，尽量减少人工干预，建立高效的、具有智能分析功能且易于维护的档案信息资源管理系统。也就是说，规范化、精细化、集成化、智能化的档案信息管理成为现阶段档案管理信息系统建设的发展方向。其任务就是关注信息新技术的发展和应用情况，选取和利用适合档案信息化的相关理论和技术，如信息资源规划、本体理论、业务规则管理、文本挖掘等，解决现阶段档案信息系统建设中的问题，促进档案信息化的顺利进行。

1.2.1 档案管理规范化

加强档案的规范化管理，是提高机关档案管理工作的重要手段，也是衡量一个单位档案整体管理水平的重要指标。目前我国文件档案管理各自为政的局面日益突出，在网络工作环境下因缺少具有一致性和互操作性的统一标准而带来的信息资源共享与利用孤岛问题越来越大，信息化进程难以推进。制定档案信息管理标准体

系，实现档案规范化、标准化管理是档案信息化顺利开展的保障，是档案信息共享的基础。

在实际工作中，要贯彻落实《中华人民共和国档案法》和《中华人民共和国档案法实施办法》，将档案规范化管理列入机关管理工作的重要议事日程，使档案规范化管理与机关其他工作同部署、同实施、同考核，切实解决档案管理工作中的突出问题。明确档案工作的分管领导，建立健全档案人员岗位责任制，完善档案移交、接收、保管、鉴定、保密、销毁、利用、统计等一系列规章制度；建立查阅、借阅、档案利用效果登记簿；结合各门类档案，制定切实可行的归档范围和保管期限表等。档案业务工作要做到井然有序，做到档案查阅、借阅有登记，利用效果有记载。还要加强公文制作规范化管理，促使文书处理工作规范化，制定一系列技术标准和管理标准，如《档案著录规则》（DA/T18-1999）等，并将这些标准纳入业务规则库中，对其进行集中统一管理，形成实现数据共享、信息集成以及指导档案信息全程管理的规范化系统。此外，档案管理的主体是人，要重视档案工作人员的培训，提高档案人员的综合素质和相关业务能力。

1.2.2 档案管理精细化

档案是组织或个人在以往的社会实践活动中直接形成的、清晰的、确定的、具有完整记录作用的固化信息和重要信息资源。档案的原始凭证性要求从档案形成伊始就必须精细地管理才能保证档案的终身质量。现阶段还有很多企事业单位的档案工作并不规范，专职档案管理人员偏少，流动性也大，档案整理工作一般是雇佣临时员工或是业务外包，只有制定详细具体可操作性强的管理措施才能保证档案工作的质量。此外，随着档案信息化程度的提高，高效档案信息资源管理系统的设计和实现依托的就是现实档案业务流程，需要细化档案工作流程和制定完善的工作制度、标准。

科学发展观要求档案工作必须尊重科学，按客观规律办事，协调发展，实行科学的精细化管理。首先要安排部署到位、责任明确

到位、监督检查到位，执行的责任落实到具体的部门和个人，在多部门或多人共同执行时，必须明确谁是唯一责任人或第一责任人，谁是第二责任人，谁是相关责任人。档案馆在业务指导，制定分类方案、归档范围和划分期限上要细化，确保每一个文件和档案的进出和去留都有据可循；将所有的档案工作作为一种流程，进行流程化管理。档案的前身是文件、图纸、表格、照片、录音、录像等多种载体的文件和材料，每一份文件和材料生成后，从归档、移交到档案室再经过鉴定、电子档案的载体有效性测试和复制、整理、系统分类、著录、编目、数字化、信息化、装订、上架等一系列过程，以及档案安全管理、提供查询利用和档案编研等，都要制定明细表，标注执行的质量和数量标准、执行的重点和难点、进度要求和完成时限等；还要实行档案质量监督考核机制，明确检查、考核的程序、方法，质量标准以及检查和考核责任人，另外还要制定对执行者的责任追究和奖惩办法，作为档案管理监督工作的保障。

1.2.3 档案管理集成化

为了对档案信息资源进行高效管理，提供高质量的信息服务与决策支持，档案信息资源的集成管理是一个必然趋势。档案信息资源的集成管理不是对档案信息的简单集中与叠加，而是以集成思想为指导，通过集成的方式、方法，对档案以及与档案相关的各种管理要素进行优化整合，加强档案信息之间的有机联系，使之便于检索利用，保证其发挥最大效用，即实现利用的最大化。

档案集成管理的内容主要包括：①数据和信息集成，即将分布在某一信息系统中自治与异构的多处局部数据源中的信息和数据进行有效集成，将各个应用系统的数据分类，实现各信息子系统间的信息共享，通过多媒体数据集成将原有信息系统中信息不一致的各个自治的数据库进行集成，建立主题数据库。②应用系统功能集成，即在集成系统的整体功能目标的统一框架下，将各应用系统的功能开放协议、标准或规范集合在一起构成一种一体化的多功能系统，以便互为调用、互相通信，更好地发挥集成化信息系统的作用。

用，如文档管理系统与业务系统和数字信息资源管理系统集成等。

③过程集成，其包括文件与档案管理业务流程集成、文档综合管理与服务系统集成、档案部门对机关和事业单位职能活动前端性介入过程的集成管理（即实现档案文件的前端控制）。

④技术集成，即综合利用和合理配置实现档案集成管理所必须依靠的数据库技术、计算机网络技术以及硬件条件等。

1.2.4 档案管理智能化

自动化、智能化管理是计算机管理系统的优势和发展趋势所在，是一个需要始终加强的基本点。从某种意义上讲，智能化就是现代化的标志，是真正意义上的网络化和信息化。因此，文件和档案管理信息系统的设计也应努力向智能化靠拢，以提高档案管理效率和质量，满足现代化、信息化时代档案高效管理的要求。

然而，尽管现有的一些软件产品标榜自己已经实现档案管理自动化，实际情况却不能令人满意。许多“自动化”管理环节不是效果太差，就是功能上有所欠缺，档案业务人员使用起来不“顺手”，也无法完全寄予信任，需要“人工干预”的地方还有很多。以自动著录为例，现代文件著录是一种全面的过程化的著录，需要软件自动捕获或提取诸如文件内容、结构、背景等很多方面的信息，以维护电子文件的真实性、可读性、可靠性。然而目前软件大多仅仅著录题名、作者、时间等几个项目就算大功告成，仍然模拟原先的手工著录，这种“自动著录”显然不能满足电子环境下档案工作自动化、智能化管理的要求。

高效的档案管理要求信息系统面对各种复杂问题具备一定的判断推理功能，能够为档案业务工作者提供业务提示，如文件该归档了、档案该入库了、档案保管期限到了、提示档案工作者库房的容量受到威胁、利用率高的档案应采取及时的保护措施等；能够实现档案管理的自动化，如判定档案的密级、保管期限，统计各部门归档数量等；能够实时监控和综合评价档案管理工作，如档案收集是否齐全完整、内容是否真实有效、整理是否规范

等；能够支持档案工作者自学习，如档案上岗人员利用信息系统学习各种档案法律法规、相关制度规范，摄取长期实践经验、知识总结，档案馆利用信息系统对相关归档部门进行业务指导等；能够研究分析出档案信息中对当前决策有参考价值的信息，如自动分析全宗介绍、目录、著录、全文等信息推理出与政府中心工作有关的档案信息等。总之，从公文书中的自动催发、自动转发，到档案的自动归档、自动组卷、自动标引，甚至自动著录、自动鉴定和自动销毁，都是档案管理智能化的表现，是现代新型档案管理信息系统攻克的方向。

1.3 业务规则和业务规则管理

1.3.1 什么是业务规则

规则，英语为“rule”，其中文意思有规则、惯例、统治、章程、准则、标准、控制等^①。在日常生活中，像我们见到的管理制度、业务手册、工艺流程、操作规范、收费标准、促销策略以及没有形成文字的惯例等都可以称为业务规则（Business Rules，BR），它们体现了一个组织拥有的全部业务知识。如果没有业务规则，做一切事情都是临时做出决策，需要以一事一议的方式逐个比较可能的方案，这种做事方式，不仅非常慢，还会产生不一致、不协调的后果。可以说在当今的世界中，所有有组织的业务过程都有业务规则。

业务规则的含义丰富，目前还没有统一的标准定义。一个比较公认的定义是由业务规则组织（Business Rules Group，BRG）给出的，该定义同时包含两个层面的理解，从业务的角度来看，“业务

^① 杨静. 业务规则管理技术的研究与应用 [D]. 上海：华东师范大学软件学院，2007.

规则是支持决策，影响或控制组织业务行为的指示”；从计算机信息系统的角度来看，“业务规则是一条语句，它定义或约束业务的某些方面。其目的是对业务结构做出断言，或者对业务行为施加控制和影响”。^①

业务规则实质上也可以理解为一组条件和在此条件下的操作，是一组准确凝练的语句，用于描述、约束及控制企业的结构、运作和战略，是应用程序中的一段业务逻辑。该业务逻辑通常由业务人员、企业的管理人员和程序开发人员共同开发和修改。它的理论基础是：设置一个条件集合，当满足这个条件集合时，触发一个或者多个动作。下面列举一些业务规则的实例：

(1) 一条基于服务质量的定价规则

“当一条 CDR 的类型是 IP 电话时，根据服务质量给这条 CDR 定价，对服务质量进行潜在的评测。”

(2) 一条跨服务的折扣规则

“当客户既享有互联网服务又享有移动通信服务时，如果客户使用互联网服务超出指定的互联网服务使用下限时，并且客户使用移动通信服务超出指定的移动通信服务使用下限时，那么对指定客户实行指定的折扣。”

(3) 一条基于内容的定价规则

“如果是彩信服务，那么每条彩信收费 1 元。”

(4) 一条基于使用量的定价规则

“如果通话时间超过 10 分钟，那么 10 分钟以后的话费按半价计费。”

从以上分析，不难看出业务规则内涵丰富，既包括明文规定与业务相关的法律法规、操作规范、规章制度、管理章程、行业标准等，也包括业务指导长期实践的经验、组织内经验丰富的从业人员的知识总结、员工日常使用的专业知识和专门技能等没有形成文字的惯例，这在档案管理工作中也随处可见。但是，目前

^① D. Hay, A. K. Healy. Defining Business Rules—What are They Really? [R]. The Business Rules Group, Technical Report1.3, 2000 July.

国内档案管理系统中,业务规则大多数存在业务人员的大脑中,或者有关的业务手册、管理规程、操作手册等非结构化的文档中。即使某单位有了信息系统,规则逻辑也是作为一个过程写进了程序代码中,当业务和需求发生变化的时候,不能做出相应的响应变化,就必须修改原有的程序代码,给工作带来极大的不便。

1.3.2 什么是业务规则管理

业务规则管理方法源自人工智能技术,是改变业务系统设计和开发方法的主要范例。事实证明使用业务规则技术可以提高业务活动的效率,加强业务系统的有效性和灵活性^①。

现代业务规则管理技术较之传统业务系统引入业务规则层,彻底改变了以过程形式处理业务逻辑的方式,将业务规则的实现从具体的程序代码中抽取出来,以结构化的业务规则数据来表示组织的业务行为,使得业务规则与组织的数据信息一样成为组织的重要资产。与此同时,软件开发的习惯也开始因业务规则管理技术而改变。业务规则像数据一样独立于程序之外,业务人员可以使用行业术语而不是专业编程语言来编写规则,从而使组织的业务系统真正面向业务人员^②。

业务规则管理技术将使我们重新认识业务规则,重新认识业务规则在整个软件系统中的地位,以新的方式来表示、运行、管理业务规则,从而为组织和软件开发商在处理业务规则的问题上提供新的解决方式。

业务规则并不是新的事物,但是当我们不再把业务规则当做一

^① Andozej Maciol. An Application of Rule-based Tool in Attributive Logic for Business Rules Modeling [J]. *Expert Systems with Applications*, 2008 (34): 1825-1836.

^② 杨静. 业务规则管理技术的研究与应用 [D]. 上海: 华东师范大学软件学院, 2007.

个事务的处理过程，而是把它当做一种新的数据类型时，它就为我们提供了一种新方法来解决易变的业务规则的难题。这种根本的改变在于：把业务规则作为数据来进行管理。

将业务规则技术应用于档案管理系统的改进，建立一个基于业务规则的能够支持辅助决策、及时响应需求变化、可充分扩展并具有长久生命周期的档案信息资源管理系统具有深远意义。

1.4 业务规则管理技术国内外研究和应用现状

1.4.1 国外

业务规则管理技术的发展主要经历了三个时期，即启蒙时期、低谷期和繁荣发展时期。20 世纪 70 年代，斯坦福大学使用 LISP 语言开发的用于诊断血液疾病和推荐治疗方法的 MYCIN 系统^①是第一个基于业务规则的专家决策系统。该系统实现的主要理念是知识和控制的分离，将以规则表示的知识从用于评估、执行的控制逻辑（程序）中分离出来，这是规则管理技术的启蒙时期。20 世纪 80 年代，随着人工智能的研究热潮，业务规则的研究逐渐兴起，但由于最初的业务规则以传统方式嵌入应用程序代码中，紧密嵌入数据库结构中并以 DBMS 触发器和存储过程的形式进行编码，缺少管理业务规则的规范方法和标准的业务规则语言，非技术人员无法对业务规则进行创建、修改和管理，此阶段业务规则技术的应用困难重重。20 世纪 80 年代后期，随着面向对象技术的兴起，分类机

^① Bruce G. Buchanan, Edward H. Shortliffe. *Rule-based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project* [M/OL]. <http://www.aaai.org/AITopics/pmwiki/pmwiki.php/AITopics/RuleBasedExpertSystems>.

制、信息隐藏（封装）、消息通信机制等技术为人们解决复杂应用软件系统的问题提供了新的概念和模型，同时也为基于规则的程序提供了更好的集成和实现方式。此外，丰富的图形化的业务规则管理工具也是业务规则管理技术得到广泛应用的重要基础。近些年，业务规则越来越受到业界同仁的关注，出现了 G. Ronald Ross、Holger Herbst、Daniela Rosca 等许多著名研究者，以及业务规则论坛（Business Rules Forum）、业务工作组（Business Rules Group）、业务规则社区（Business Rules Community）等一批专门研究组织，此外国外几乎所有大型通用软件提供商、大型行业软件提供商以及大型行业用户，如 ILOG、IBM、Versata 和 Blaze 等公司，都在系统开发和系统运行过程中使用了这一技术。业务规则管理技术的研究和应用已经进入繁荣发展时期。

国外业务规则管理研究主要分为业务规则基本概念、业务规则管理基础技术和业务规则管理技术在各领域中的应用这三大方面。

“业务规则”一词被使用于不同领域，Daniela Rosca^① 等人认为业务规则是指“关于企业业务目标、政策或约束的陈述语句”；著名的 Holger Herbst 博士^②将业务规则定义为“如何做业务的陈述语句，也就是指关于组织内业务状态和过程的指导性、约束性语句”；Mitchell Krammer^③ 则认为业务规则是商业组织习惯和政策的实施纲领。P. Kardasis 和 P. Loucopoulos^④ 教授基于信息系统分析角度，将业务规则看成是“对组织工作方式和支持组织工作的信息系统功能的外部约束的反映”。总体来说，这些业务规则定义主要分

① D. Rosca, S. Greenspan, C. Wild. Application of a Decision Support Mechanism to the Business Rules Lifecycle [C]. 10th Knowledge-based Software Engineering Conference (KBSE95), Boston, MA, 1995.

② Holger Herbst. *Business Rule Oriented Conceptual Modelling* [M]. Heidelberg: Physica-Verlag, 1997.

③ M. Krammer. Business Rules: Automating Business Policies and Practices [J]. *Distributed Computing Monitor*, 1997 May.

④ P. Kardasis, P. Loucopoulos. Expressing and Organizing Business Rules [J]. *Information and Software Technology*, 2004 (46): 701-718.

为从企业业务角度和计算机信息系统角度两大类，只从其中的某一个角度来看待业务规则是片面的。

业务规则管理基础技术的研究内容主要包括规则开发环境、规则模型、规则语言、规则库和规则引擎等。如，Ronald G. Ross 提出业务规则 Ross 模型^①，依据该业务规则管理模型，业务规则主要分为三种类型：规则、事实和术语，业务规则可以是原子的，也可以推理而来。每条推理规则可包括两条或更多的原子或推理规则。该方法共有 32 种原子和 58 种推理类型，最独特之处是在图表中用严格的图形符号正式地表达业务规则。相比较而言，Ross 模型是现在提出的有关业务规则管理最成熟、最完善的方法论。H. Herbst 和 T. Myrach^{②③} 阐述了建立规则库的模型和实例，在其提出的 BROCOM^④ 方式中，规则语言是一种结构化英语，具有高表达力，规则是根据丰富的元模型进行组织的，可以根据多种不同标准进行检索。Daniela Rosca 等人提出 DSS^{⑤⑥} 方法，该方法采用 ECA (Event Condition Action) 范式来表达、构建业务规则，关注规则建立背后的逻辑，并建立起业务规则与企业底层模型实体之间

① R. G. Ross. *Business Rule Solutions* [M]. Houston: The Business Rule Book, Houston, 1997.

② H. Herbst, T. Myrach. A Repository System for Business Rules [C]. IFIP Conference Proceedings, 6th IFIP TC-2 Working Conference on Data Semantics: Database Applications Semantics, Georgia, 1995: 119-139.

③ Holger Herbst, Business Rules in System Analysis: A Meta-model and Repository System [J]. *Information Systems*, 1996, 21 (2): 147-166.

④ Holger Herbst. *Business Rule Oriented Conceptual Modelling* [M]. Heidelberg: Physica-Verlag, 1997.

⑤ D. Rosca, S. Greenspan, C. Wild, H. Reubenstein, K. Maly, M. Febowitz, Application of a Decision Support Mechanism to the Business Rules Lifecycle [C]. 10th Knowledge-based Software Engineering Conference (KBSE95), Boston, MA, 1995.

⑥ D. Rosca, C. Wild. Business Rules in the Real World: A Decision Support Approach [C]. The Software Engineering and Software Engineering Conference (SEKE96), Lake Tahoe, NE, 1996: 121-128.

的关联。后来 D. Rosca 与其合作者^{①②③}又描述了一种支持业务规则生命周期的方法。P. Kardasis 和 P. Loucopoulos 教授研究了表达和组织业务规则的方法和模式^④, 依据企业知识发展 (Enterprise Knowledge Development, EKD) 模式, 提出一种名叫“曼彻斯特业务规则管理”(Manchester Business Rules Management, MBRM) 的以业务规则为中心的开发模式。Ram Sudha 和 Khatri Vijay^⑤ 提出一种在概念数据库设计阶段组织成套业务规则 (Set-based Business Rules) 模型的综合框架方法。Marko Bajec 和 Marjan Krisper^⑥ 认为不应只强调业务规则方法能够使应用软件更加灵活, 而更要注重业务规则方法对于建立组织信息系统与组织业务联系的桥梁作用。他们提出的管理业务规则的方法适用于组织信息系统的各个方面, 同时, 也提出了支持该方法的业务规则工具。

现在基于业务规则的应用系统越来越多, 除了银行、保险、医疗等业务规则应用较为成熟的传统领域, 一些国外学者开始进一步

① D. Rosca, S. Greenspan, C. Wild. Enterprise Modeling and Decision-support for Automating the Business Rules Lifecycle [J]. *Automated Software Eng. agement* 2002, 9 (4): 361-404.

② D. Rosca, J. D. Attilio. Business Rules Specification, Enforcement and Distribution for Heterogeneous Environments [C]. 25th International Computer Software and Applications Conference (COMPSAC 2001), Invigorating Software Development, Chicago, IL, 2001: 3-9.

③ D. Rosca, M. Feblowitz, C. Wild. Decision Making Methodology in Support of the Business Rules Lifecycle [C]. IEEE International Symposium on Requirements Engineering (RE' 97) 3rd, Annapolis, MD, 1997: 236-246.

④ P. Kardasis, P. Loucopoulos. Expressing and Organizing Business Rules [J]. *Information and Software Technology*, 2004 (46): 701-718.

⑤ Ram Sudha, Khatri Vijay. A Comprehensive Framework for Modeling Set-based Business Rules during Conceptual Database Design [J]. *Information Systems*, 2005 (30): 89-118.

⑥ Marko Bajec, Marjan Krisper. A Methodology and Tool Support for Managing Business Rules in Organizations [J]. *Information Systems*, 2005 (30): 423-443.

拓展业务规则应用领域的探索性研究。如 Flavio Corradini 等人^①将业务规则技术引入电子政府的建设应用中,提出一种自适应集成业务规则方法 (Business Rules for Adaptive Integration, BRAIN),该方法能将业务流程与业务规则结合起来,有效整合各个单位,促进组织之间或组织内部的联合与合作。

1.4.2 国内

在国内,目前对业务规则管理的研究主要集中在应用领域,对业务规则管理理论的研究相对较少,主要参照和引用国外的标准和定义。目前业务规则管理理论与技术研究方面的论文还比较少,主要有《业务规则管理技术研究》、《业务规则管理技术的研究与应用》等。而针对业务规则管理在各个行业中的应用的论文则较多,主要集中在保险、银行、电信、证券、电力、远程教学等行业,如《基于业务规则的汽车保险报价系统设计与实现》、《规则引擎在电信结算摊分系统的应用及业务规则管理》、《业务规则管理的研究及其在变压器故障检测系统中的应用》、《基于业务规则的远程教学教务系统研究与实现》等。

在应用上,目前,业务规则管理已成为商业应用的技术新贵。国外几乎所有大型通用软件提供商、大型行业软件提供商以及大型行业用户都在系统开发和系统运行过程中使用了这一技术,如 ILOG、IBM、Versata 和 Blaze 等公司。国内有代表性的有凯科思特(北京)科技有限公司的 KCOM 商业工程。由于系统的适应能力强,反应变化的速度快,因此采用业务规则系统可以给现代企业带来极大的效益。在银行、保险、信贷、证券等领域,几乎每个流程、每个处理环节都需要大量地应用业务规则。比如银行有信贷规则,保险公司有核保规则和理赔规则,不同的领域还有自己的行业

^① Flavio Corradini, Alberto Polzonetti, Oliviero Riganelli. Business Rules in E-Government Applications [J/OL]. *Electronic Journal of E-Government*, 2009, 7 (1): 45-54, <http://www.ejeg.com>.

条规等。可以说，没有业务规则，就没有金融行业。业务策略或业务规则，不是一天创建起来的，而是经历了几十年的发展和积累。随着各种法规和行业条例的不断进步，金融行业的业务规则也越来越缜密，可以用规则引擎来实现可运行时修改的业务规则。这样可得到高度灵活性的软件，更好地动态适应业务变化。目前，国内基于 BRMS 的信息系统还比较少，业务规则管理在国内档案行业还没有相关的研究。

1.5 基于业务规则档案信息管理系统开发的意义

档案是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具有保存价值的各种文字、图表、声像等不同形式的历史记录。简单地说，档案是保存下来的记录，是历史的真实记忆，是传承文明的载体，是现实工作的必然依托。档案和档案信息是社会唯一的原生信息源。无论历史文化还是科学技术，一切其他社会信息的形成从本质上讲都是源于对档案这一原生信息源的收集保管和开发利用。因此，对档案信息的有效保管与利用，从古到今，从国外到国内都受到政府和群体的重视。档案是人类社会发展的重要的信息资源，因此，要开发、管理和利用好档案，使其为政治、经济、文化建设服务。

基于业务规则的档案管理系统除具备一般档案管理系统的档案管理、利用、编研的功能外，拟从管理和技术实现两个角度来探讨档案信息资源开发与管理问题，重点研究一套基于信息资源规划，依托业务规则库的档案信息资源收集管理体系，同时也为档案管理部门建立起智能化的网上监督指导系统、绩效评估系统和上岗人员自学习系统等。也可通过基于档案业务规则库（Archives Business Rules Database，简称 ABRD）和推理引擎（RM）构建的档案信息雷达自动从相关业务系统收集档案信息资源，有效提高档案管理系统

统的性能和效率，增强档案管理系统的交互性能和可维护性，特别是增强档案管理个性化服务的优势。

构建与开发基于业务规则的档案信息管理系统具有重要的现实意义，主要体现在以下 7 个方面。

其一，有利于档案信息的组织和利用。

随着信息化进程的加快，信息资源呈现出爆炸式的增长趋势，档案信息也不例外。档案信息有着信息量大、增长快速、涉及内容广泛、存取分散、表现形式各异等特点，从而带来获取、鉴定、保管、查找和利用等方面的困难，常常干扰和降低档案管理人员获取档案信息的全面性和档案鉴别的准确率，以及用户获取档案信息的准确度和相关度。档案业务规则管理系统重点研究一套基于信息资源规划、依托业务规则库的档案信息资源收集管理体系：通过建立档案业务规则库（ABRD），同时将主管部门与指导部门的规范、制度、标准等也整理到 ABRD，通过强大的推理引擎（RM），对已有的档案信息数据（ADI）进行分析与推理，做出相应的判断，能够帮助立档单位达到“档案收集齐、保管期限设置准、档案著录规范”的目标，能够实现档案信息资源组织和利用自动化。

其二，有利于提高档案管理质量。

利用档案业务规则库（ABRD），建立一种新型的档案监督指导管理体系与绩效考核体系。遵循归档管理的工作流程和内在规律，涵盖归档各项管理细节的质量、时间、数量、难度、行为等特征的全方位的质量控制和评价；归档标准层次化，可以细化和量化到每一项归档要求，如归档范围、归档时间、归档人、归档文件的质量、案卷整理的质量、归档数量、归档难度等，组成一个权责分明、相互协作的质量管理评价标准，有利于提高档案归档、保管、利用的质量。

其三，有利于档案信息资源的规范采集和信息共享。

通过 Internet/Intranet 网络互联技术将主管单位与下级单位或企业各分支机构的档案信息组建为一个集中的目录中心，实现档案实体与档案信息的分离管理，实现档案部门与归档部门的沟通，有利于对归档管理各环节过程的控制和管理。

通过全面梳理档案管理业务流程和相关信息资源，摸清信息或数据产生的业务环节，确定第一数据源的产生部门和环节，编制档案信息资源知识库，避免信息资源的重复采集，为信息资源的整合和共享提供依据。

其四，有利于提高档案管理工作者的业务能力。

推进档案信息管理进步的关键是档案管理人员的业务能力。档案工作是一门实践性极强的管理工作，要实行档案的科学化、标准化、规范化管理，档案管理人员除了具备良好的政治、道德素养外，还应具备良好的业务素质和能力素质，只有这样，才能在技术上做到精益求精，不断提高档案管理工作的服务质量。

档案管理人员不仅要熟悉档案业务工作的各个环节、步骤，而且还应掌握各门类档案、各种载体档案的管理方法，在务实基础上促进创新，努力成为一名掌握科学的、先进的工作方法，一专多能的档案工作人才。

档案业务规则管理系统利用档案业务规则库（ABRD），建立一种自学习系统，为档案上岗人员提供便利、切合实际的业务培训途径，并通过不断完善档案业务规则库，找到一条传承资深档案人士知识的快捷方法，从各个方面不断提高档案上岗人员的业务能力。

其五，有利于档案管理知识的积累和传递。

随着变化效率的加快，社会机构的人员必须不断学习，尽可能全面地熟悉政策和规程。但是由于人员变更造成的业务经验的流失是目前比较普遍的现象。将业务逻辑以规则的形式表示，不但解决了规则流失的问题，更提供了一个机构内部知识交流和学习的平台。

其六，有利于降低信息系统扩展和变更成本。

对于一个以服务满意度为使命的机构来说，政策的正确、及时的体现和处理对其使命的完成起着关键作用。政策和规则总是在不断变化，一旦把业务规则像数据一样从程序中剥离，无疑在技术上为业务应用的实施团队提供了更大的灵活性，加快 IT 和业务部门的决策速度，缩短新政策和规则的部署实施时间。

其七，有利于业务规则的修改和维护。

由于业务规则已经与程序分离开来，业务规则的修改不需要修改程序，同时我们基于业务人员熟悉的术语实现规则编辑功能，就可以实现业务人员自行维护规则。这样就更加快了变更的速度，消除了业务人员与 IT 部门之间交流的困难。

2. 档案业务规则管理概述

2.1 档案业务规则的内涵

业务规则是描述和约束业务的语句，用来刻画业务的结构或者控制和影响业务的行为^①。

档案业务规则主要是用来描述、约束和控制档案信息管理业务中各步骤及其产生的数据，其大量存在于相关档案信息管理的各类业务文档之中，还普遍存在于档案管理的各个环节之中及档案管理信息系统中。在档案管理系统中存在着如归档范围、保管期限、工作流程、分类标准等多种规则，还包括一些没有形成具体条例的惯例，均称为档案业务规则。

通过前文对“业务规则”概念的分析，档案业务规则的理解也主要存在业务应用和计算机信息系统两个角度。从业务应用角度来看，档案业务规则包括对档案、档案工作、档案管理流程等有描述、约束和控制作用的法律法规、操作规范、规章制度、管理章程、业务标准、档案业务相关人员的经验总结等；从计算机信息系统角度来看，档案业务规则是采用特定的业务规则语言对档案业务中某些定义和限制的规范化描述，用于维持档案业务结构、控制或影响档案业务的声明语句，具有定义业务术语、陈述事实、约束行为、进行推理等作用。

两者之间的关系是前者主要用来指导、约束人们日常从事档案

^① GUIDE. GUIDE Business Rule Project [R]. 1995, Nov. 6.



业务活动所遵循的规则，采用人类交流和思维所用的自然语言来表达；后者主要来源于前者，是从已成文的法律法规、操作规范、规章制度、管理章程、业务标准和档案业务相关人员的经验总结中提炼而来的。与前者不同的是，该角度研究的业务规则应采用结构化的语言，为计算机管理信息系统的开发服务，可采用 RuleSpeak 模板、决策表、决策树、规则标记语言等方式来表示。本书的主要研究对象是从计算机信息系统角度定义的档案业务规则，而法律法规、规章制度、管理章程、业务标准、操作规范和档案业务相关人员的经验总结等是业务规则的来源和基础。

2.2 档案业务规则的特性和作用

2.2.1 档案业务规则的特性

业务规则通常具有声明性、准确性、原子性、一致性及非冗余性。在基于业务规则的档案信息资源管理系统中，档案业务规则的特性可分为一般特性和应用特性。

(1) 档案业务规则的一般特性

①原子性。档案业务规则不可再分，每条规则只定义一种判断和操作，复杂的业务逻辑由多条规则协同处理。

②独立性。档案业务规则彼此之间独立，复杂的逻辑关系由规则引擎来处理。档案业务规则存储在规则库中，独立于数据和程序。

③简单性。档案业务规则用简单直接的类自然语言来描述，很容易被档案业务人员和技术人员所理解。

④动态性。档案业务人员可以实时地修改业务规则，快捷地更新系统，低成本地维护系统。

⑤逻辑性。档案业务规则至少包含条件和执行两个部分，条件

是对业务数据作用的判定，执行是对业务数据的处理。

(2) 档案业务规则的应用特性

①档案业务规则的非“固化性”。固化在程序代码中的策略和规则必然是僵硬的。随着时间的推移和业务环境的变化，档案适用业务规则和策略会发生变化，如果规则的每次改变都要求对系统程序进行“伤筋动骨”式的修改，那么系统的维护和升级必然代价昂贵，甚至难以维持。

②档案业务规则的“逻辑性”。档案业务规则具有逻辑性，每条约束行为的业务规则至少包含两个部分：条件部分和执行部分。规则的条件涉及对业务数据作用的判定，规则的执行涉及对业务数据的处理。

③档案业务规则的“非过程性”。每条规则只能定义对一种现象的判断和操作，复杂的业务逻辑应该由多条规则协同处理。规则的“非过程性”带来的好处是：每条规则的制定变得非常单纯，可以“就事论事”，将复杂的过程处理平摊成一个个有条件的执行单元，从而实现了从简单到复杂的知识积累过程。

④档案业务规则的“事件触发性”。档案业务规则会根据相应的条件被触发执行，触发规则执行的“事件”就是业务数据本身。比如一套入库分析的规则集合，一旦库存件入库记录信息进入系统处理，这组规则将会被激活，启动相应的分析过程。

⑤档案业务规则的“非技术性”。档案业务规则是属于档案业务人员的，档案业务人员应该使用行业语言而不是计算机技术语言（如程序语言、数据库语言、脚本语言等）编写规则。正是由于档案业务规则的非技术性特点，使得使用业务规则方法的系统可以被档案业务人员进行维护，而不仅仅像过去一样，只能通过技术人员进行维护工作。

2.2.2 档案业务规则的作用

档案业务规则具有重要的作用，主要有以下几个方面：



(1) 能够对档案业务术语进行定义

档案业务规则的最基本成分是术语。术语是与单位或部门密切相关的概念，它在一般情况下是档案业务人员所知道并在单位或部门共享的一个基本单词或词组。术语包含特定的档案业务含义。对于一种具体的使用背景环境，术语不应该有歧义产生。术语总是名词和经过修饰的名词。

术语所具有的基本特征有：

①基本性：即术语表示的是业务中最为基本的知识，它不能用其他的业务词汇导出或计算而得出来。

②原子性：即术语表示的是个体知识，它不能是术语的组合物体。

③可知性：即术语总是表示业务人员能够了解到的事情，而不是所发生的事情。

(2) 能够描述由相关术语组成的事实

事实是由相关术语的简单描述语句给出的。它表示业务的公共或共享动词或动词短语，构建了所需的概念之间逻辑连接的结构。

(3) 具有行为约束的作用

通过约束和条件来限定和控制行为，例如，一条鉴定机关档案保管期限的规则：如果存档的是本机关制定的法规政策性文件，那么其保管期限应列为永久保管。

(4) 根据已有规则可以进行合理的推理

档案业务规则可以从一种已有知识推导出另一种知识。

2.3 档案业务规则来源

总的来说，档案业务规则主要来源于明文规定与档案业务相关

的法律法规、业务标准、操作规范、规章制度、管理章程，档案业务相关人员的需求调查，以及旧档案信息资源管理系统的文档这三大方面，在实际操作中，需要通过有效的手段和方法获取一套完整的业务规则。

2.3.1 档案法规和标准

我国的档案法规随新中国档案事业的起步而逐步发展起来，1987年《中华人民共和国档案法》（以下简称《档案法》）颁布后，形成了以《档案法》为核心，以档案行政法规、地方档案法规为框架，部门档案规章和地方政府档案规章为主体，档案标准和规范性文件为基础，其他法律、法规中的档案管理规定为补充的多层次档案法规体系。据不完全统计，目前，我国现行档案行政法规有7部、地方法规30多部、档案规章120多部，内容几乎涉及档案工作的各个方面。

我国的档案管理模式从建立之日起就具备一个相对完善的管理体系，保障顺利、有序开展各级各类档案管理工作的规则很丰富、细致，不仅有最具权威的档案专门法律，还根据各地各类档案的特殊性制定了一系列的条例、章程、纲要、办法、方案、决定、规定、指示、细则和标准等，起到了规范管理档案和档案事务的作用，是档案业务规则的主要来源。根据各地各单位和部门档案种类、档案管理特色，以及档案信息资源管理系统要求，可以将这些法律、条例、规章、标准结构化处理后编译到档案业务规则库中。

(1) 档案法规体系

档案法律法规，是指国家制定的有关档案收集、整理、保护、利用等各种活动的法律规范总称。它包括国家权力机关根据宪法制定的档案法律、法令和其他各种法律；法令中有关档案的条款；也包括国家行政机关根据宪法、法律、法令在其职权范围内制定的关于档案管理的各种规范性文件，如条例、章程、纲要、办法、方案、决定、规定、指示、细则、通知等。档案法规的规范作用主要



体现在以下 5 个方面：指引作用、评价作用、预测作用、教育作用、强制作用。其中与业务有关的条款，如归档范围、保管期限、分类方法、归档文件材料的质量要求等，是开展档案业务活动的重要依据。我国档案法规体系是以《档案法》为核心，由符合立法规定的若干有关档案工作的法律、行政法规、地方性法规和规章所构成的相互联系、相互协调的统一体，分为档案法律、档案行政法规、地方性档案法规和档案规章 4 个层次。其中包括：《中华人民共和国档案法》1 部档案法律；《中华人民共和国档案法实施办法》（以下简称《档案法实施办法》）等 6 部档案行政法规（含现行有效的中央制定发布的规范性文件）；《档案馆工作通则》等 95 部国务院部门档案规章（含联合发布的规范性文件）^①。此外，地方性档案法规和地方政府档案规章分别由享有立法权的地方人大及其常委会和享有规章制度制定权的地方人民政府视需要自行决定。

①档案法律。档案法律是全国人民代表大会或其常务委员会根据宪法或依职权制定的法规性文件，主要有《档案法》以及宪法、刑法等法律中涉及档案的内容或条款，是档案法规体系中的最高层次，对所有相关法规、规章和规范性文件的制定与实施起着约束、指导作用。

《中华人民共和国档案法》由第六届全国人大常委会于 1987 年制定、第八届全国人大常委会于 1996 年修订。它是由全国人大常委会制定的规范性文件，是公民、法人及其他各种组织从事档案事务的行为准则，是管理我国档案事业最基本的法律依据。

②档案行政法规。它指由国家最高行政机关——国务院在法定职权范围内为实施宪法和法律的规定制定和发布的有关档案行政管理的规范性文件的总称，包括条例、规定、办法等，它们是法规体系的一个重要层面，法律效力仅次于国家法律。档案行政法规是一种具有普遍约束力的规范性文件，是针对档案事业中某一方面的行政管理工作制定的。主要有《档案法实施办法》、《机关档案工作条例》、《科学技术档案工作条例》、《全国档案馆网设置原则和布

① 国家档案局. 档案法规体系方案 [R]. 2009.

局方案》等。此外，根据宪法和立法规定，国务院还有权发布决定和命令，其中具有规范性内容的，与行政法规具有同等效力。

③地方性档案法规。地方性法规是各省、直辖市、自治区所在地的市和国务院批准的较大市的人民代表大会及其常务委员会在不与宪法和法律抵触的前提下，根据本地区的实际情况制定的规范性法律文件，它是与国家法律及行政法规的重要联结点，其法律效力等同于行政法规。我国地域辽阔，地方差异较大，国家的法律、法规不可能面面俱到地适用所有地区，《档案法》和《档案法实施办法》的规定比较原则，而各省、市、自治区的档案事业发展各有特殊性，因此需要针对地方特点对国家法律法规中的原则性要求和宏观性规定进行细化，将档案法律法规的规定与本地的具体情况和实际结合，制定出地方性法规。

1995年6月，《上海市档案条例》公布实施，此后各地地方性档案法规陆续颁布施行。目前全国有30个省、自治区、直辖市颁布了地方性档案法规。长春、沈阳、青岛、宁波、武汉、济南、深圳等17个较大的城市也颁布了地方性档案法规（如“省档案条例”、“省档案管理条例”等）20件。

地方性档案法规是国家法律和行政法规的补充和细化，比起国家法律和行政法规，地方性档案法规可操作性更强。对《档案法》和《档案法实施办法》没有规定而实践中又急需做出规定的事项进行规定，是档案地方立法的一个重要的内容。具体在制定档案地方性法规时，要着重解决本行政区域实际存在的需要由法律来调整的，而法律、行政法规的规定原则或不便于操作或没有规定的问题，而不是对法律、行政法规的简单翻版。

④档案规章。档案规章包括国务院部门档案规章和地方政府档案规章两类。其中国务院部门规章是依据法律和行政法规的规定，由国家档案局单独制定颁布或国家档案局与有关部委联合制定颁布的有关档案工作的规范性文件的总称，其地位和效力低于宪法、法律和行政法规。政府规章相对法律、法规而言数量较大，内容具体、明确，针对性强，是法规体系的主体。目前，由国家档案局发布的档案规章就有40多件，内容集中在专业档案和专门档案管理、



文件处置、档案保护、档案行政管理和档案执法等方面。

地方政府档案行政规章是由省级人民政府以及省、自治区所在地的市和国务院批准的较大的市的人民政府依法定权限和程序制定，包括人民政府批转的地方档案局制定的规范性文件。目前全国有《苏州市城市建设档案管理办法》等地方政府的档案规章百余件。

及时清理法规、规章，是保证法规、规章的权威性、有效性，是建立健全档案法规体系的重要步骤。它可以有效地避免法规、规章之间互相重复、矛盾等现象，更好地完善档案法规体系。一般地讲，法规清理是指行政机关对过去所发布的行政法规和规章进行审查，确定哪些法规、规章已不适应现实需要，应予废止；哪些法规、规章已部分不适应现实需要，应部分废止；哪些法规、规章有互相重复、矛盾之处，应废止重复者，消除矛盾处。例如，根据2011年1月30日发布和实施的《国务院办公厅关于做好规章清理工作有关问题的通知》（国办发〔2010〕28号）要求，我国家档案局对2010年以前由档案局发布和以档案局为主与有关部委（局）联合发布的档案规章、规范性文件进行了全面清理。在编制档案业务规则和组建规则库时应关注这些新发布的法规废除通知，避免使用失效规则。

（2）业务标准体系

档案业务标准体系包括国际标准、国家标准、行业标准、地方标准和企业标准，是档案法规体系的重要补充。档案国际标准主要是指国际标准化组织制定的标准。我国标准是指由国家质量监督检验检疫总局发布国家标准和国家档案局发布或批准的行业标准中带有强制性或指令性的规范，分为国家标准（GB）、行业标准、地方标准（DB）、企业标准（QB）四级。

①国际标准。国际标准是指国际标准化组织（简称ISO）、国际电工委员会（简称IEC）和国际电信联盟（简称ITU）制定的标准，以及国际标准化组织确认并公布的其他国际组织制定的标准。国际上制定有关档案标准的影响较大的标准化组织有国际标准化组

织、国际电工委员会、国际档案理事会（简称 ICA）、国际文件管理者联合会（简称 IARMA）和联合国档案与文件管理署（简称 UNARMS）等。

国际标准化组织是世界上最大的国际性标准化组织，国际电工委员会主要负责电工、电子领域的标准化活动，这两大标准化组织在制定档案信息标准方面，都集中了国际上一流的专家、学者，并广泛地吸取了各国在这方面的最佳实践经验，其所制定的标准集中反映了人类在这一领域的发展与应用，质量是最高的。国际档案理事会（ICA）、国际文件管理者联合会（ARMA）及联合国档案与文件管理署（UNARMS）都是专业性的文件/档案组织，其所制定的档案标准十分专业。^①

ISO 所制定的主要档案信息标准有：《ISO 15489：2001 信息与文件-文件管理》、《ISO 23081 信息与文件-文件管理过程-文件元数据》、《ISO 22310：2006 信息与文件-标准中文件管理需求规定导则指南》、《ISO/CD 26122 信息与文件-文件保管工作过程分析》、《ISO/TR 18492：2005 基于电子文件信息的长期保存》、《ISO 19005 文件管理-电子文件长期保存文档格式》、《ISO 22938（DIS）文件管理-电子内容/文件管理数据交换格式》、《ISO 12651：1999 电子文档管理-词表-电子成像》、《ISO/TR1654：1997 电子成像-为确保记录于 WORM 光盘文件证据性管理电子记录系统的建议》、《ISO/TR15801：2004 电子图像信息数字化存储的真实性和可靠性建议》、《ISO 14721：2003 空间数据与信息传输系统-开放档案信息系统（简称 OAIS）-参考模型》、《ISO/NP 26102 电子文件长期保存需求》等。

ISO 与 IEC 联合所制定的主要档案信息标准有：《ISO/IEC 82045 文件管理》（共包括 5 个部分）、《ISO/IEC WD 24706 信息技术-技术标准与规范文档元数据》等。

ICA 所制定的档案标准主要有：《国际档案术语词典》、《鉴定

^① 段荣婷. 论国际档案信息标准化现状、发展趋势及我国的对策 [J]. 档案学研究, 2008 (1).



标准：《档案鉴定业务规范》、《电子文件数据交换标准》、《电子文件管理：档案工作业务指南》、《电子文件：档案工作者手册》、《国际标准——档案著录规则（总则）》（简称 ISAD（G））、《国际标准——团体、个人和家族档案规范记录规则（简称 ISAAR（CPF））、《法人团体活动/职能国际标准》（简称 ISAF）等。

ARMA 所制定的档案标准主要有：《ANSI/ARMA 5-2003 重要文件：识别、管理和恢复重要业务文件标准》、《ANSI/ARMA 8-2005 文件和信息维护管理（PDF）》、《ANSI/ARMA 9-2005 管理电子邮件文件的功能需求》、《ANSI/ARMA TR-01-2002 文件中心运行标准》、《ANSI/AIIM/ARMA TR48-2004 电子文档管理系统和电子文件管理系统集成框架》、《电子政务与电子文件：电子文件管理》等。

UNARMS 所制定的档案标准主要有：《文件管理系统的设计与开发指南》、《文件管理项目评价技术》、《文件保管系统功能需求》、《文件保管元数据》等。

一项国际档案标准的制定是国际档案界经验与实践的总结，尤其在技术层面，应该加以积极地吸取与采用。根据 2001 年施行的《中华人民共和国采用国际标准管理办法》第 2 条的规定，“采用国际标准”是指“将国际标准的内容，经过分析研究和试验验证，等同或修改转化为中国标准（包括国家标准、行业标准、地方标准和企业标准），并按中国标准审批发布程序审批发布”。如 1999 年国际档案理事会颁布的《国际标准——档案著录规则（总则）》就可以在充分研究的基础上，加以吸取与采用。该标准中的多级著录规则是充分反映档案特点的著录规则，而在我国的档案著录规则中就没有，这就可以将其吸取过来为我所用。

②国家标准。国家标准是在全国范围内统一的技术要求，由国务院标准化行政主管部门编制计划，协调项目分工，组织制定（含修订），统一审批、编号、发布。国家标准分为强制性国标（GB）和推荐性国标（GB/T）。国家标准的编号由国家标准的代号、国家标准发布的顺序号和国家标准发布的年号（采用发布年份的后两位数字）构成。

1985年5月10日,国家标准局批准发布了我国档案工作的第一个国家标准《GB3792.5-85 档案著录规则》。之后,与之相配套的《中国档案分类法》的发布,标志着有中国特色的档案分类标准的基本框架已经确立。

现行使用的档案国家标准主要有:《GB/T 15418-2009 档案分类标引规则》、《GB/T 24573-2009 金库和档案室门耐火性能试验方法》、《GB/T 24422-2009 信息与文献 档案纸 耐久性和耐用性要求》、《GB/T 9705-2008 文书档案案卷格式》、《GB/T 11822-2008 科学技术档案案卷构成的一般要求》、《GB/T 13967-2008 全宗单》、《GB/T 20163-2006 中国档案机读目录格式》、《GB/T 20530-2006 文献档案资料数字化工作导则》、《GB/T 19667-2005 基于 XML 的电子公文格式规范》(包含2个部分)、《GB/T 11821-2002 照片档案管理规范》、《GB/T 50323-2001 城市建设档案著录规范(附条文说明)》、《GB/T 17678-1999 CAD 电子文件光盘存储、归档与档案管理要求》(包含2个部分)等。

③行业标准。由我国各主管部、委(局)批准发布,在该部门范围内统一使用的标准,称为行业标准。例如:机械、电子、建筑、化工、冶金、轻工、纺织、交通、能源、农业、林业、水利等,都制定有行业标准。

为了依法加强全国档案标准化管理,国家质量技术监督检验检疫总局将国家档案局列为档案工作行业标准的归口管理单位,负责标准的起草、颁布、组织实施和修订工作。标准代号为“DA”,档案工作的行业标准多以 DA/T (T 表示推荐性标准) 打头。现行有效的档案工作行业标准主要包括:《DA/T 1-2000 档案工作基本术语》、《DA/T 22-2000 归档文件整理规则》、《DA/T 23-2000 地质资料档案著录细则》、《DA/T 28-2002 国家重大建设项目文件归档要求与档案整理规范》、《DA/T 29-2002 档案缩微品制作记录格式和要求》、《DA/T 30-2002 满文档案著录名词与术语汉译规则》、《DA/T 31-2005 纸质档案数字化技术规范》、《DA/T 32-2005 公务电子邮件归档与管理规则》、《DA/T 33-2005 明清档案目录中心数据采集标准 明清档案机读目录数据交换格式》、《DA/T 36-2007 人



身保险业务档案管理规范》、《DA/T 3-1992 档案馆指南编制规范》、《DA/T 2-1992 科学技术研究课题档案管理规范》、《DA/T 8-1994 明清档案著录细则》、《DA/T 9-1994 明清档案档号编制规则》、《DA/T 13-1994 档号编制规则》、《DA/T 17-1995 有关历史档案资料著录、标引的系列标准》、《DA/T 18-1999 档案著录规则》、《DA/T 19-1999 档案主题标引规则》、《DA/T 20-1999 有关民国档案著录、标引的系列标准》，等等^①。

此外，船舶工业（CB）、测绘（CH）、城镇建设（CJ）、地质矿产（DZ）、核工业（EJ）、公共安全（GA）、国家军用（GJB）、航空（HB）、环境保护（HJ）、海洋（HY）、机械（JB）、林业（LY）、民用航空（MH）、农业（NY）、汽车（QC）、航天（QJ）、气象（QX）、电子（SJ）、检验检疫（SN）、兵工民品（WJ）等行业标准中还存在针对该行业档案工作制定的标准。例如：编号为CB 3495-1992的一系列船舶工业档案管理规则（共包括17个部分）；核工业行业标准里的相关标准有《EJ/T 1225-2008 核电文件档案管理要求》、《EJ/T 657.1-1992 中国档案分类法核工业档案分类与代码》、《EJ/T 657.2-1992 中国档案分类法核工业档案专用综合复分表》、《EJ/T 965-1995 核工业档案全宗单位代码》；国家军用标准中有一系列关于军队档案工作的标准，主要有《GJB 2418.2A-2009 军队档案主题词标引规则》、《GJB 3108-1997 军队文书档案目录数据库结构与著录格式》、《GJB 3865-1999 光盘存储档案信息技术与管理规范》、《GJB 4292-2001 军队装备科技档案目录数据库结构与著录格式》、《GJB 4339-2002 军队档案分类标引规则》、《GJB 6640-2008 军队档案工作基本术语》、《GJB 6651-2009 军队后勤科技装备管理档案目录数据库结构与著录格式》等；属于城镇建设行业标准的《CJJ/T 117-2007 建设电子文件与电子档案管理规范（附条文说明）》；属于地质矿产行业标准的《DZ/T 0431-2005 矿业权档案立卷归档规范》；属于公共安全行业标准的

^① 档案工作行业标准目录 [EB/OL]. 中国国家档案局官方网站, <http://www.saac.gov.cn/>.

《GA/T 749-2008 公安档案信息数据交换格式》；属于航空行业标准有《HB 7837-2008 航空产品数字化文件档案管理通用要求》、《HB 6588-1992 航空工业档案主题标引规则》、《HB 6446-1990 航空工业档案分类法》等；属于环境保护行业标准的有《HJ/T 295-2006 环境保护档案管理规范 环境监察》、《HJ/T 7-1994 中国档案分类法 环境保护档案分类表》、《HJ/T 78-2001 环境保护档案管理数据采集规范》等；属于海洋行业标准的有《HY/T 004-1991 全国海岛资源综合调查档案 分类法》、《HY/T 030.2-1993 中国极地考察档案著录细则》等。

④地方标准。地方标准又称为区域标准，对没有国家标准和行业标准而又需要在省、自治区、直辖市范围内统一的工业产品的安全、卫生要求，可以制定地方标准。有关档案工作的地方标准是各省、自治区、直辖市标准化行政主管部门在国家、行业档案标准的基础上，按照《档案法》的要求制定出相应的地方标准，并报国务院标准化行政主管部门和国务院有关行政主管部门备案。例如：江苏省地方标准《DB 32/505-2002 文书档案文件级目录数据库结构与数据交换格式》、湖南省地方标准《DB 43/T 263-2005 档案工作规范化管理》、山东省地方标准《DB 37/T 536-2005 档案目录数据采集规范》、杭州市地方标准规范《DB 3301/T 13-2008 杭州市上城区社区档案管理规范》、辽宁省地方档案《DB 21/T1018-1998 林业科技事业单位档案综合管理》等。2005年1月，经四川省人民政府审查同意，由四川省人民政府办公厅下发了《四川省机关档案工作规范化管理标准》、《四川省企业、科技事业单位档案工作规范化管理标准》和《四川省各级档案馆工作规范化管理标准》，是近年来以政府办公厅名义出台的第一个地方档案工作法规性标准。

⑤企业标准。《中华人民共和国标准化法》规定：企业生产的产品没有国家标准和行业标准的，应当制定企业标准，作为组织生产的依据。企业的产品标准须报当地标准化行政主管部门和有关行政主管部门备案。已有国家标准或者行业标准的，国家鼓励企业制定严于国家标准或者行业标准的企业标准，在企业



内部适用。

档案企业标准是对企业范围内关于档案工作需要协调、统一的技术要求、管理要求和工作要求所制定的标准。企业标准由企业制定，由企业法人代表或法人代表授权的主管领导批准、发布。企业标准一般以“Q”作为企业标准的开头。例如：《QB/YW 216-03-2007 会计档案管理标准》（云南电网公司企业标准）、《XTAP-QMS 110-2006 记录和档案管理程序》、《建筑施工企业档案工作规程》，等等。

2.3.2 未成文的经验和惯例

档案业务规则还包括指导长期实践的档案业务经验、组织内经验丰富的档案从业人员的知识总结、员工日常使用的档案专业知识和专门技能等没有形成文字的惯例。

具体开展档案管理工作时所依据的规则往往都是具体和细化的，相对来说，明文规定的法律、法规和规章、制度等是纲领性的，对实践具有指导作用。具体实施档案管理工作时还需要有经验的员工或决策者参与。例如：具体针对某个单位时，什么是重要文件、什么文件反映了该单位重要职能活动、档案价值的具体判定等。这些都是档案工作人员们长期从事档案管理活动的知识积累和总结。如不记录、固定下来，这些隐性知识会随着经验丰富的档案专职人员的离职或时间的推移而不复存在。

对于此类业务规则，可以利用访谈业务人员、开座谈会、问卷调查、跟班作业等方式进行调查总结，并将其文档化、结构化处理后编译到档案业务规则库中，完善档案业务规则体系。

2.3.3 管理信息系统里的文档

信息系统的文档，是系统建设过程的“痕迹”，是系统维护人员的指南，是开发人员与用户交流的工具。文档还可以作为监理和审计的对象，作为开发其他信息系统的参照。没有文档的软件，不

是合格的软件；没有文档的信息系统，不是完整的信息系统。一个良好的信息系统，需要做好系统文档的管理工作。档案管理信息系统依据标准设计，文档的编制在信息系统的开发工作中占有突出的地位和相当的工作量，规范了系统文件和档案的管理，支持档案采集、移交接收、归档、存储、借阅利用和编研发布等全过程的信息化处理。

档案管理信息系统开发、设计、建设过程中形成的用户手册、操作手册、系统开发计划书、系统分析说明书、系统设计说明书、程序设计报告、数据字典、详细设计、程序代码等包含着大量档案业务规则。将从中经过分析、筛选出的档案业务规则集中到档案业务规则库保存，不仅体现了业务规则集中式管理的策略，而且从旧系统中具有规范格式的信息系统文档中提炼出的结构化业务规则更接近编译到档案业务规则库前的句法要求，避免了重复劳动，减轻了编译工作量。

要注意的是，通过以上三个来源获取的业务规则会存在许多重复或与规则无关的信息，必须通过挖掘整理从中获取有用的业务规则，按规则类型将业务规则进行分门别类，然后再进行业务规则的编译活动。

2.4 档案业务规则管理

2.4.1 传统档案工作业务规则管理

从业务应用角度来看，档案业务规则主要存在于对档案、档案工作、档案管理流程等有描述、约束和控制作用的法律法规、操作规范、规章制度、管理章程、业务标准、档案业务相关人员的经验总结等。这些明文规定的、依照法定程序发布的与档案工作有关的法律法规和业务标准是传统业务规则的主要存在形式，主要以单份文件或法规汇编、标准汇编形式实行集中管理。法律法



规和标准的汇编是依据国务院或国务院各部门制定的汇编编辑出版管理规定，依照法定程序发布的法律、行政法规、国务院部门规章、地方性法规和地方政府规章，按照一定顺序或者分类汇编成册的公开出版物。这些非结构化文件里关于档案业务规则的定义和描述是以人工实施管理的纸质、实物等传统档案管理模式的依据，不能满足电子文件管理及其档案管理信息系统的设计、开发和实施的需求。

从计算机信息系统角度来看，档案业务规则是采用特定的业务规则语言对档案业务中某些定义和限制的规范化描述，用于维持档案业务结构、控制或影响档案业务的声明语句。这一类规则主要嵌入信息系统文档和程序中，其中嵌入文档中的规则是人可读的，程序里控制和影响档案业务的规则是机器用来执行的。这些档案业务规则与文档、程序里对非档案业务活动的定义和限制共同存在，缺少单独管理业务规则的规范方法和标准业务规则语言，不能及时响应系统业务逻辑的变化，从而影响了档案管理信息系统的有效性和灵活性，降低了业务活动的效率。

2.4.2 利用业务规则管理技术实施档案业务规则管理

(1) 业务规则管理技术简介

20 世纪 70 年代，斯坦福大学使用 LISP 开发的 MYCIN 是第一个基于规则的系统。该系统用于血液疾病的诊断，并推荐治疗方法，可以算作一个专家决策系统。该系统实现的主要理念是知识和控制的分离，将以规则表示的知识从用于评估、执行的控制逻辑（程序）中分离出来，这是规则管理技术的启蒙时期。

20 世纪 80 年代，基于规则的编程方法随着人工智能的研究热潮达到顶峰，那时候有不少的开发者和组织购买商业化的规则引擎，但当时规则引擎性能很差，并且没有与当时主流系统的集成能力，基于规则的编程方法没有得到很好的应用。按照 Gartner 的副

总裁和研究主管 Jim Sinur 的话说, 由于早期的许多业务规则产品都以人工智能领域和推理系统为基础, 所以这个阶段的业务规则技术非常复杂, 运营和维护费用高昂, 从而导致了这种技术的“垂死经历”。但是研究工作仍在继续, 基于规则的编程方法仍然在某些方面得到应用。

20 世纪 80 年代后期随着面向对象技术的兴起, 分类机制、信息隐藏(封装)、消息通信机制等技术为人们解决复杂应用软件系统提供了新的概念和模型, 同时也为基于规则程序提供了更好的集成和实现方式。可以说, 面向对象机制很好地解决了数据与数据操作的关系, 而业务规则管理技术为对象之间的消息通信提供了(以业务规则数据形态处理业务逻辑的)触发机制, 使对象模型和业务规则模型较完美地结合在一起。此外, 丰富的图形化的业务规则管理工具也是业务规则管理技术得到广泛应用的重要基础^①。

近年来, 业务规则管理技术正以越来越快的速度重新崛起。不少软件开发商已经开始利用业务规则管理技术来开发商用软件, 它们不仅能够为用户搭建规则库, 让用户随意添加自己的业务规则, 而且会在一些针对行业的应用中, 将自己的行业经验以业务规则的形式加进去, 为用户提供最佳实践经验。

如今的业务规则已具有明确的定义、规范的表达和系统的管理, 它们不再以代码形式嵌入计算机系统中, 而是采用易读、易懂的语言表达, 即使非技术人员也能够迅速掌握并灵活运用这种语言管理业务。这种清晰、准确的表达方式大大提高了组织内部的交流效率和决策速度。业务规则软件已被公认为提高业务灵活性和保持组织竞争力, 实现“一站式”服务的必备技术。

(2) 业务规则管理技术的基本思想

业务规则管理技术的基本思想是将系统处理的业务逻辑从程序

^① 梁冰. 业务规则成资产 [N]. 计算机世界, 2004-04-19.

代码中抽取出来，将其转变成简单的业务规则，以结构化的业务规则数据驱动规则引擎执行业务行为^①。

现代业务规则管理技术较之传统业务系统引入业务规则层，把业务规则作为数据来进行管理，使得业务规则与组织的数据信息一样成为组织的重要资产。以结构化的业务规则数据表示组织业务行为，业务人员可以使用行业术语而不是专业编程语言来编写规则，便于将业务规则创建、修改和维护的权力交给业务人员，从而使组织的业务系统真正面向业务人员。引入业务规则管理技术，可以将业务规则存储在完全独立于数据和程序的规则库中，业务人员可以对业务规则进行查询、添加、更新、统计、提交等操作，并且可以在线修改和测试业务规则。业务规则可以不断积累、调整 and 共享，并能对规则进行版本管理，设定规则的有效期，实现对业务行为的知识管理。系统的稳定性也因此得到了保障，系统的维护成本大大降低。

业务规则管理技术将使我们重新认识业务规则，重新认识业务规则在整个软件系统中的地位，以新的方式来表示、运行、管理业务规则，从而为组织和软件开发商在处理业务规则的问题上提供新的解决方式^②。

业务规则方法提供了一种依照业务规则概念进行分析问题和解决问题的方式，帮助人们发现规则，表现规则，管理规则，自动执行规则，建立规则运行机制，最终目的是实现业务规则管理系统^③。

① 吕良双，蔡欢．一种面向业务人员的基于对象的规则描述语言 [J]．计算机与数字工程，2008（5）．

② 杨静．业务规则管理技术的研究与应用 [D]．上海：华东师范大学软件学院，2007．

③ 来自 CSDN 博客 [EB/OL]．<http://blog.csdn.net/liuyingsheng/archive/2008/09/21/2957444.aspx>．

2.5 档案业务规则管理系统

2.5.1 业务规则管理系统的基本原理

业务规则管理系统（Business Rules Management System，简称 BRMS）是实现业务规则的建立、管理、维护、部署和运行的信息系统^①。BRMS 能够利用规则引擎把传统的应用程序或参数化实现的业务策略从系统中提取出来，被提取的业务规则将作为一种可管理的对象存储在规则库中。BRMS 是构建在操作系统、数据管理以及事务处理之上的高级管理层次，在应用系统中的地位与数据库管理系统（DBMS）类似，处于比较基础的位置，是其他高端应用的基础。图 2-1 是 IT 市场研究公司 GIGA Information Group 给出的 IT 架构中 BRMS 的位置图^②。

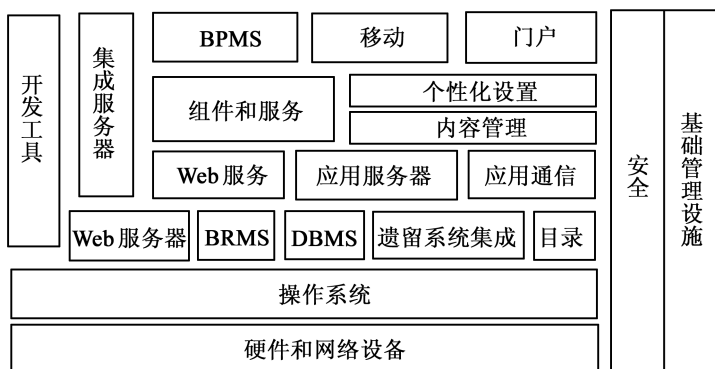


图 2-1 IT 架构中 BRMS 的位置图

① 徐黎，冯元勇，糜宏斌．基于受限中文的轻型业务规则管理系统的设计与实现 [J]．计算机科学，2005（2）．

② 杨静．业务规则管理技术的研究与应用 [D]．上海：华东师范大学软件学院，2007．

业务规则管理系统主要包括规则的形式化表示、规则编程语言、规则匹配与执行，其基本原理是：用一个或多个规则引擎替换原先以程序代码（或其他抽象代码）“固化”在应用系统中的业务逻辑，业务逻辑不再以程序代码的形式驻留在系统中，实现了业务逻辑和程序代码的分离。业务规则的创建、修改是通过规则操作工具界面，对规则库的修改而实现的，经编译后的业务规则保存在规则库中。规则库中的业务规则可以实时“在线”地被加载到规则引擎中去，供规则引擎调用执行，参见图 2-2 东莞市万维网络科技有限公司绘制的业务规则管理系统工作模型图。

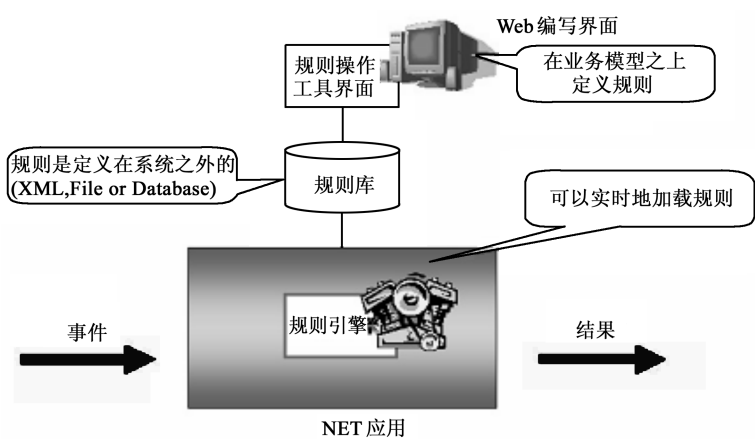


图 2-2 业务规则管理系统工作模型

2.5.2 档案业务规则管理系统的结构

档案业务规则管理系统（Archives Business Rules Management System，简称 ABRMS）主要包含业务规则集成开发环境、业务对象模型、业务规则语言、档案规则库、业务规则引擎、与外部系统部件的数据和指令接口良好的组件化结构等，见图 2-3。

下面具体阐述档案业务规则管理系统设计过程中最重要的 5 个

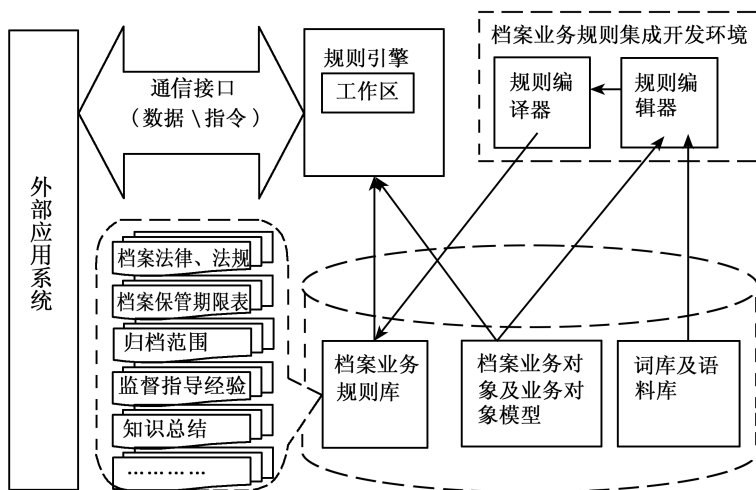


图 2-3 档案业务规则管理系统的组成

部分：

(1) 档案业务规则集成开发环境 (Archives Business Rules Integrated Development Environment)

在档案业务规则开发环境中集成了规则编辑器和编译工具,是用来管理、创建、修改、部署、编译和测试业务规则的图形化工具。利用图形界面用户可以方便地建立业务对象模型;可以对其中的规则库和规则进行编辑操作,如新建、删除、复制、粘贴等;可以检查和测试规则的有效性和正确性。此外,相应的编译工具还可以将用户输入的档案业务规则语句转换成规则引擎能够执行的规则代码^①。

① 师艳辉, 庄越挺, 施坚强. 基于 JAVA 的业务规则管理系统设计和开发 [J]. 计算机应用与软件, 2006 (1).

(2) 档案业务对象模型 (Archives Business Object Model)

档案业务对象模型是用来描述档案管理工作领域中所有的业务对象 (如文书档案、会计档案、设备仪器档案、电子公文等), 它记录了档案各个业务对象的名称、属性、方法、参数类型、返回类型, 以及对象之间的约束。业务对象模型为业务规则语言提供了绝大多数的词汇, 多由档案业务系统分析员设计, 由系统开发人员具体实现。

(3) 档案业务规则语言 (Archives Business Rules Languages)

面向不同的用户, 业务规则语言的具体表达形式不同。档案业务规则管理系统的三层规则语言模型^①为档案业务人员和 IT 技术人员设计了不同的业务规则语言, 它们最终都会被系统转换成规则引擎的执行代码, 见图 2-4。

由于熟知档案业务规则的档案业务人员和分析人员通常不具备计算机编程知识, 只能理解和表述自然语言形式的业务规则, 但是计算机处理中文自然语言又比较困难。为平衡业务和 IT 之间存在的矛盾, 可以采用受限中文业务规则语言作为一般档案业务人员使用的业务规则表述语言, 采用类 Java 语言文法作为系统 IT 技术人员使用的编程语言。

(4) 档案业务规则库 (Archives Business Rules Database)

档案业务规则库是存储档案业务规则及规则有关属性的地方, 用户录入的规则经编译后保存在规则库中, 供规则引擎调用执行。档案业务规则库将档案各项专业知识分门别类融合, 东莞万维在其设计的基于业务规则的档案信息资源管理支撑平台中将档案业务规则库分为以下三类: ①存储以《档案法》为主的相关法律、法规和制度等针对外部的业务规则库; ②存储如对综合档案以立档单位

^① 徐黎, 冯元勇, 糜宏斌. 基于受限中文的轻型业务规则管理系统的设计与实现 [J]. 计算机科学, 2005 (2).

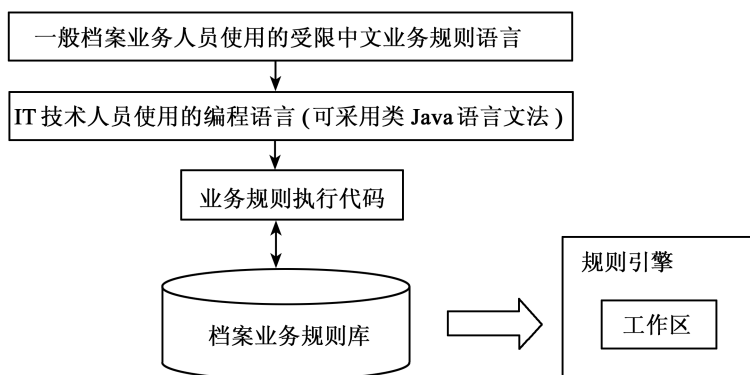


图 2-4 档案业务规则语言模型

为基础的《归档范围》、《保管期限表》，对工程档案以工程性质为基础的《建设工程竣工档案归档范围》等针对结构化的内部规则库；③存储档案指导长期实践的经验，单位经验丰富的档案从业人员的知识总结等针对非正式的内部规则库。

(5) 档案业务规则引擎 (Archives Business Rules Engine)

档案业务规则引擎是执行业务规则的软件组件，嵌入在程序中，是档案业务规则管理系统的核心元素。规则引擎是应用系统调用规则库的接口，以基于规则库中创建的规则来获取自动化的决策。系统开发人员在程序中使用规则引擎的基本步骤为：创建规则引擎对象；向引擎中加载规则集或更换规则集；向引擎提交需要被规则集处理的数据对象集合；命令引擎执行；导出引擎执行结果，从引擎中撤出处理过的数据。

3. 档案业务规则的表达和组织

3.1 业务规则的逻辑学基础

业务规则作为一种需要被表示和推理的对象，具有其相应的逻辑学基础，即规则的表示和推理要以一定的逻辑学理论作为根基。在业务规则管理中，其逻辑学基础主要包括命题逻辑、谓词逻辑和模态逻辑。

3.1.1 命题逻辑

命题逻辑是现代逻辑较简单、较基本的组成部分，它不考虑把命题分析成个体词、谓词和量词等非命题成分的组合，只研究由命题和命题连接词构成的复合命题，特别是研究命题连接词的逻辑性质和推理规律^①。

命题逻辑具有一定的命题形式。用特定的语词把命题连接起来可以构成复合命题；在复合命题中起连接作用的语词称为命题连接词；构成复合命题的命题称为支命题，支命题本身也可以是复合命题。命题逻辑研究复合命题的逻辑形式、推理形式和公理系统。传统逻辑关于假言推理、选言推理和二难推理等的理论，都属于命题逻辑的范围。复合命题的形式可以用公式明晰地表示，公式通常由

^① 黄华新，胡龙彪．逻辑学教程 [M]．杭州：浙江大学出版社，2000：5.

以下 3 类符号组成：

- ①表示任意命题的命题变元，它们是 p, q, r, p_1, q_1 等；
- ②5 个基本的命题连接词，即 $\neg$ 、 $\wedge$ 、 $\vee$ 、 $\rightarrow$ 、 $\leftrightarrow$ ；
- ③用来显示公式的结构层次的括弧 (,)。

5 个基本的命题连接词依次称为否定词、合取词、析取词、蕴涵词和等值词。在汉语中，它们通常分别用语词“并非”、“并且”、“或者（可兼的）”、“如果……则”以及“当且仅当”来表达。在这 5 个连接词中，否定词属一元连接词，其余 4 个都是连接两个命题以构成复合命题，称为二元连接词。复合命题的形式都可以用这 3 类符号构成的公式表示。如 $\neg p$ 表示否定命题的形式， $p \wedge q$ 、 $p \vee q$ 、 $p \rightarrow q$ 、 $p \leftrightarrow q$ ，分别依次表示合取、析取、蕴涵和等值命题的形式。它们是和 5 个基本的连接词相应的 5 个基本的复合命题形式，参见图 3-1。

p	$\neg p$	p	q	$p \wedge q$	$p \vee q$	$p \rightarrow q$	$p \leftrightarrow q$
1	0	1	1	1	1	1	1
0	1	1	0	0	1	0	0
		0	1	0	1	1	0
		0	0	0	0	1	1

图 3-1 命题逻辑的真值表

3.1.2 谓词逻辑

谓词逻辑是形式逻辑的最根本部分，也是最基本的逻辑系统或理论。在谓词逻辑中，除研究复合命题的命题形式、命题连接词的逻辑性质和规律外，还把命题分析成个体词、谓词和量词等非命题成分，研究由这些非命题成分组成的命题形式的逻辑性质和规律。谓词逻辑把命题逻辑作为子系统，但为了研究方便，同时也由于命



题逻辑具有某些重要的特殊性质，命题逻辑通常又作为一个独立的系统先研究，而在谓词逻辑部分则集中研究由非命题成分组成的命题形式和量词的逻辑性质与规律。

只包含个体谓词和个体量词的谓词逻辑称为一阶谓词逻辑，简称一阶逻辑，又称狭义谓词逻辑。此外，还包含高阶量词和高阶谓词的称为高阶逻辑。谓词逻辑也分为经典的谓词逻辑和非经典的谓词逻辑，后者包括作为子系统的非经典的命题逻辑。经典的一阶谓词逻辑是谓词逻辑的基本部分。

在谓词逻辑中，一个论域中的元素称为个体，个体词是表示个体的符号；表示某个论域中的一个特定个体的符号称为个体常项或个体常元，个体常项也就是它所表示或指称的那个个体的名字；不表示某一确定论域中的特定个体的个体词，称为个体变项或个体变元，用符号 x, y, z 和 $x_1, y_1, z_1, \dots$ 表示；个体变项取任一论域中的任一个体为值。谓词是表示个体的性质和个体之间关系的符号。个体的性质也称一元关系，表示个体的性质即一元关系的称为一元谓词。两个个体之间的关系称为二元关系， n 个个体之间的关系称为 n 元关系。表示二元关系的为二元谓词，表示 n 元关系的为 n 元谓词。如“是素数”就是一元谓词，“大于”是二元谓词，“在……之间”是三元谓词。表示某一论域中的特定的性质或关系的称为谓词常项或谓词常元，“是素数”等都是谓词常项。不表示某一确定论域中的特定性质或关系的称为谓词变项或谓词变元。

除了个体词和谓词，组成命题的成分还有量词。量词是命题中表示数量的词，它分为全称量词和存在量词。例如，在“所有阔叶植物是落叶植物”、“有的水生动物是肺呼吸的”这两个命题中的“所有”、“有的”都是量词，其中前者是全称量词，后者为存在量词。在汉语中，“所有”、“一切”、“凡”等表示全称量词，“有的”、“有”、“至少有一”等表示存在量词。

3.1.3 模态逻辑

模态逻辑是逻辑的一个分支，研究必然、可能及其相关概念的

逻辑性质^①。模态逻辑所研究的命题“必然 A”和“可能 A”与通常命题演算中的命题不同。后者是真值函项，前者不是。因为，当 A 真时，“必然 A”既可以是真也可以是假；当 A 假时，“可能 A”既可以是真也可以是假。模态命题演算是现代模态逻辑的基本内容之一。它是应用数理逻辑的方法研究模态命题逻辑的结果。最先开始这方面研究的是 19 世纪末的 H. 麦克考尔。在他的影响下，美国哲学家、逻辑学家 C. I. 刘易斯于 1914 年构造了一个模态命题演算。19 世纪 40 年代末，卡尔纳普开始从语义方面研究模态逻辑。19 世纪 50 年代末到 60 年代初，S. 坎格尔、J. 欣梯卡与 S. A. 克里普克等人发展了卡尔纳普的理论，提出了比较完整的模态逻辑的模型理论。19 世纪 60 年代以后模态逻辑有很大发展，出现了许多新的系统，特别出现了许多非标准的模态逻辑系统，如认知逻辑、道义逻辑、时态逻辑等。模态逻辑由于研究和阐明了必然、可能、应当、知道等本体论和认识论概念的逻辑性质，因而具有深刻的哲学意义。

在现有业务规则管理方法与实践，往往以命题逻辑、一阶谓词逻辑为基础，而对于高阶谓词逻辑和模态逻辑，则作为表达更高级别业务规则的理论基础。在 OMG 的“业务词汇与业务规则语义”（Semantics of Business Vocabulary and Business Rules）标准中就是以一阶谓词逻辑为基础，并提供模态逻辑（modal logic）的接口，用于对规则进行形式化表示。

3.2 业务规则的分类

当前业务规则没有统一的分类标准，不同的研究背景和业务规则的含义存在不同的分类。

在 1990 年，业界在业务规则管理方面达成了一致，即业务规则涉及术语、事实和规则。这被看做业务规则管理领域内的里程

^① 张君. 逻辑学教程 [M]. 长沙: 中南大学出版社, 2009: 8.

碑，在此基础上，形成了对于业务规则的各种分类视角。但不管哪种视角，业务规则的基本组件，即术语、事实和规则是被公认的^①。

3.2.1 不同的业务规则分类方式

BRG 将业务规则分成三类：

①派生规则：由业务中其他规则和事实推导出新的事实，具体为数学演算和推理。

②结构化断言：表示企业某些结构的事实的声明或者定义的概念，包含术语和由这些术语构成的事实。

③行为断言：限制或控制业务行为的约束和条件的断言，具体有权限、条件和完整性约束。

Ross 将业务规则分为九类：术语、事实、规则、约束、推导、推理、定时、顺序、启发式。

C. J. Date 将业务规则分为七类：约束、状态约束、转换约束、刺激/响应、推导、计算、推理；同时从数据结构的角度提出了约束分类：领域约束、列约束、表的约束（一个表中约束）和数据库约束（多个表中的约束）。

Versata 公司将业务规则分成五类：引用完整性规则、推导（属性计算）、验证（属性强制/可选值、最大值、最小值）、约束（同一个实体属性对属性约束）、行为/事件、表示规则。

Usoft 公司将业务规则分为五类，即约束规则：对存储信息的业务约束；行为规则：根据环境系统该怎么运行；系统自动化规则；推导规则：系统怎样展示给用户，任务怎么被组织；指令性规则：在确定的情形下，用户将怎样运行系统。

Barbara Von Halle 将业务规则分成术语、事实和规则三类。

^① BR Community. A Brief History of the Business Rule Approach (2nd ed) [J/OL]. *Business Rules Journal*, 2006, 7 (11), <http://www.BRCommunity.com/a2006/b317.html>.

业务规则分类依赖于使用业务规则的人员。为业务人员服务提出的业务规则是从业务层次来对业务规则进行分类，如 Barbara Von Halle 的分类就属于这类。为开发人员服务提出的业务规则是从平台无关的系统模型层次来对业务规则进行分类，如 Versata 公司的规则分类就属于这类，这种层次的业务规则分类可能包含区别验证规则的引用完整性规则的分类。编程人员对验证规则进行编码，而引用完整性约束规则在 DBMS 中实现。C. J. Date 的分类规则是为数据库开发人员服务。这类业务规则包含不同表的规则和列的规则。没有考虑业务规则的使用者，也不考虑业务规则是否简单或复杂的分类，如 BRG 和 Ross 的业务规则分类。这两种分类没有考虑编程时哪里需要它们，它们和哪个业务模型或者架构相关。

林金娇认为^①，根据用户群来分类业务规则，从企业协作系统开发的角度并结合企业协作系统的实现目标，将业务规则划分为实体规则和行为规则两类，如图 3-2 所示。

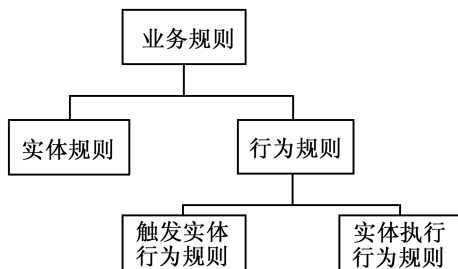


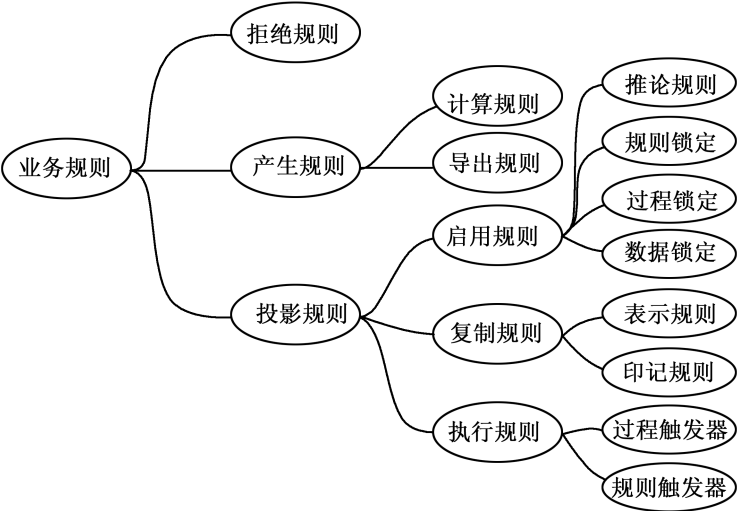
图 3-2 业务规则分类

3.2.2 Ross 的业务规则分类

Ross 作为业务规则管理领域的专家，他的分类方法具有代表

^① 林金娇. 业务规则与企业协作系统适应性的研究 [D]. 济南: 山东大学博士论文, 2008.

性，也成为被研究者引用最多的分类方法，其分类并没有违反业务规则所需的术语、事实、规则的框架，而是在此基础上进行了更深层次的细分。Ross 提出了业务规则的分类模式（BRS Rule Classification Scheme）。BRS 的规则分类采用了三级分类方式（见图 3-3），其将业务规则分为：拒绝规则、产生规则、投影规则。各个规则的定义如表 3-1 所示。



注：表中的“事件”是指对“状态”发生的任何改变，这些事件需要被检查以接受、拒绝、计算、推导、触发等。

BRS 规则分类及其定义		
功能分类/ 子分类	常用名称	定义
1. 拒绝规则	限制	如果导致规则冲突则一般会拒绝业务活动发生（事件）的所有规则。拒绝规则使业务保持正确的状态，不受与业务规则冲突的信息的影响

续表

功能分类/ 子分类	常用名称	定义
2. 产生规则		既不拒绝事件，也不投影事件，而只是根据某种数学函数计算或导出取值的所有规则
2.1 计算规则		根据明确规定的标准算术操作（如加减乘除、取平均等）计算数值的产生规则。计算规则会为计算过程提供精确的公式
2.2 导出规则		根据明确规定的逻辑操作（与、或、非、逻辑等于，等）计算真值的产生规则
3. 投影规则	触发/响应规则	当相关的事件出现时，一般会采取某种除拒绝之外的行动的所有规则。投影规则不会拒绝事件，而是投影事件，即引发一些新的事件。投影规则一般要规定自动化的系统行为，减轻业务人员的工作强度
3.1 启用规则	锁定	锁定某种事物打开或关闭的一种投影规则
3.1.1 推论规则		推导出在合适的环境中为真的事物的一种启用规则，如“在某个人的年龄和性别准则得到满足时，可以认为是妇女”
3.1.2 规则锁定	例外类型规则	在合适的环境中，打开或关闭另一个规则：使该规则能够或不能够触发的一种启用规则。例如，给出一个规则锁定，说明在紧急情况下应取消某个正常操作规则
3.1.3 过程锁定		在合适的环境中，打开或关闭一个操作、过程或规程，使其能够执行或不能够执行的一种启用规则。例如，可以给出一个过程锁定，说明当出现安全漏洞时不能执行某种敏感过程

续表

功能分类/ 子分类	常用名称	定义
3.1.4 数据锁定		在合适的环境中，创建或删除实际数据实例的一种启用规则。例如，可以给出一个数据锁定，说明当文献过时，清除其在数据库中的记录
3.2 复制规则		复制实际取值的一种投影规则
3.2.1 印记规则		确定某种持续存在的事物取值（如数据库中的内容）的一种复制规则。例如，印记规则可以用来在学生注册时，把学生在给定学期的所欠学费初始化为该学期的学费
3.2.2 表示规则		确定与数据表示方法有关的取值或参数（如通过屏幕、报表等表示）的一种规则，例如可以通过描述一个表示规则，说明如果文件已处理，在屏幕上用红色显示
3.3 执行规则	触发器	导致执行某个操作、过程或规程，或触发某个规则的一种投影规则
3.3.1 过程触发器		导致执行某个操作、过程或规程的一种投影规则。当保管期限到期时，触发向指定的业务人员发送通知
3.3.2 规则触发器		导致触发某个规则的一种投影行为。例如，当有关文件的数据显示在屏幕上时，可以指定一个规则触发器触发另一个规则以预测文件的保管期限。

3.3 档案业务规则的组成结构

根据 BRG（Business Rules Group）的业务规则研究报告，业务

规则被看做一个复杂的对象加以研究，它具有复杂的结构并且由大量相关元素共同构成。在其最终版本的报告中，其给出了业务规则的完整形式①。

3.3.1 BRG 的业务规则组成结构

(1) 业务规则的高层结构

业务规则所在的单位（组织和企业）首先具有与其业务运行相关的一些政策或规定性条款，这些条款被看做企业的政策（Policy），在这些政策中，蕴含了业务规则的表述（Business Rule Statement），在业务规则管理中，要将这些潜在的业务规则表述为业务规则（Business Rule），每个业务规则必须进行形式化规则表述（Formal Rule Statement），而形式化的表述方式可以有多种表达类型（Formal Express Type）（见图 3-4）。

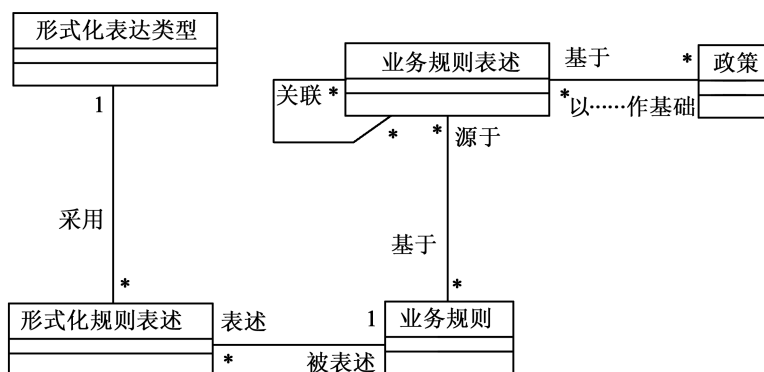


图 3-4 业务规则的高层概念结构

① Defining Business Rules—What are They Really? [EB/OL]. http://www.businessrulesgroup.org/first_paper/BRG-wha-tisBR_3ed.pdf.

在档案信息资源管理领域，档案管理单位包括一系列的管理办法和文件，这些管理办法和文件可以看做 Policy，而其中恰恰蕴含了档案业务规则的表述，这些表述必须表达成为形式化的档案业务规则，表达的途径可以采用各种形式化的档案业务规则表示方法。

根据上述定义，可以发现业务规则源于档案管理部门的内部政策性文件和规定，而为了进行有效的管理，必须识别这些政策文件中的业务规则表述，并采用某种形式化表达方法，进行形式化的表述。只有进行了形式化表述，才能进行规则的计算机化管理并最终与档案的业务流程逻辑相关联。

(2) 业务规则的分类与组成

在 BRG 的业务规则定义中，将业务规则分为三类：派生规则 (Derivation)、结构化断言 (Structural Assertion)、行为断言 (Action Assertion) (见图 3-5)。

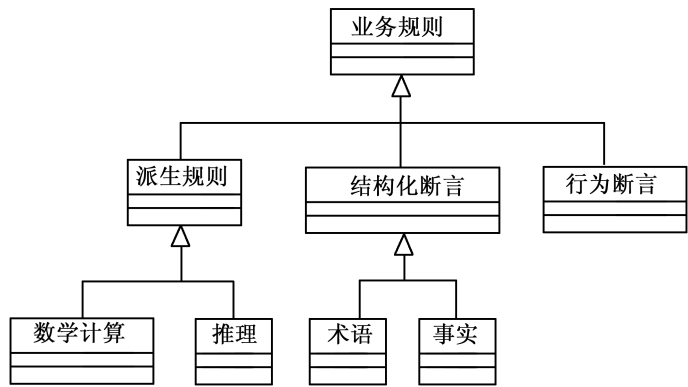


图 3-5 业务规则的分类

结构化断言：用于表达组织（企业、部门）业务的结构化特征，表达相关的业务概念或组织的各类事实，即用于表示与业务相关各类术语和事实。

行为断言：体现为某种条件约束，用于限制和控制业务活动的

执行。

派生规则：这类规则是从其他业务规则中派生（推导）出来的。

在档案业务规则管理中，结构化断言对应档案管理领域的各类专业术语、概念以及相关的事实，如《中华人民共和国档案行业标准：档案工作基本术语》中定义的大量术语，以及术语之间的有机联系。档案针对不同的档案业务，会有特定的术语和事实产生，而这些术语和事实都可以作为结构化断言对待。

行为断言则对应档案业务活动中，对活动的执行和开展起限制和约束作用的各类规则，如满足何种条件的档案文件需要永久保存，满足何种约束的档案文件可以直接公开等。这些规则往往蕴含在各类《档案工作手册》和《政府信息公开条例》之中。

派生规则是在已有规则或与业务活动相关的知识的基础上，通过谓词演算和逻辑推理形成的规则，在档案业务中，通过对各类相关的规则进行组织并采用推理和演绎，能够形成各类派生规则，扩大规则库中规则的量。

（3）术语

术语是表达特定业务中的概念的词语或词组。术语可以分为两类：一种是业务术语（Business Term），一种是通用术语（Common Term）（见图 3-6）。前者是在特定业务活动中使用的（特定业务活动被看做业务术语所使用的背景上下文），后者则不限于某种特定的业务，属于日常使用的各种通用概念。业务术语与通用术语往往结合起来构造具体表达相应的业务事实。在档案管理中，“全宗”、“科技档案”、“档案馆”等就是业务术语，需要进行清晰的界定以表达业务事实和业务规则。

除了按照术语的通用性划分类别外，还可以将术语分为型（Type）和值（Literal）两类，其中型是术语所对应的某类事物的抽象，而值则是对事物的具体实例，代表某个具体的业务对象（见图 3-7）。

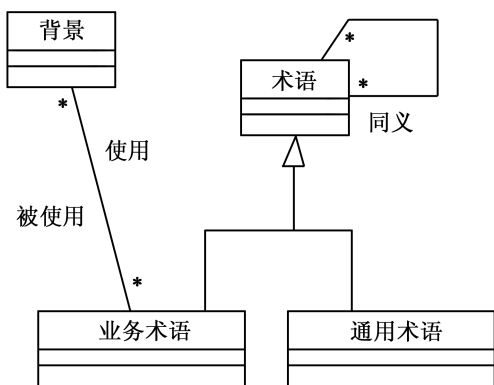


图 3-6 术语的分类（业务术语与通用术语）

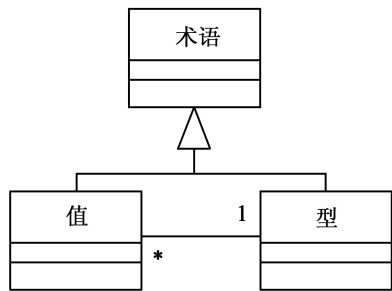


图 3-7 术语的分类（型和值）

（4）事实

事实表达了两个或两个以上的术语之间的关系，即一个事实可以包含多个术语，而一个术语也可以在多个事实中出现。术语在事实的表述中扮演一定的对象角色（Object Role），因此术语与对象角色相关联，代表它在表述事实过程中在语义层面上所起的作用。以“档案馆员著录档案文件”这样一条事实为例，它包含两个对象角色，这两个角色分别与“档案馆员”和“档案文件”对应。因此在形式上，事实通常用有序的文本序列（Text Ordering）加以

表达，即一条事实语句，在文本序列中，可以包含一个或多个词组 (Phrase)，这些词组对应于事实中所包含的多个对象角色，通过为每个词组设置标示符来为对象角色设置标示，以定义术语在事实中的位置（见图 3-8）。

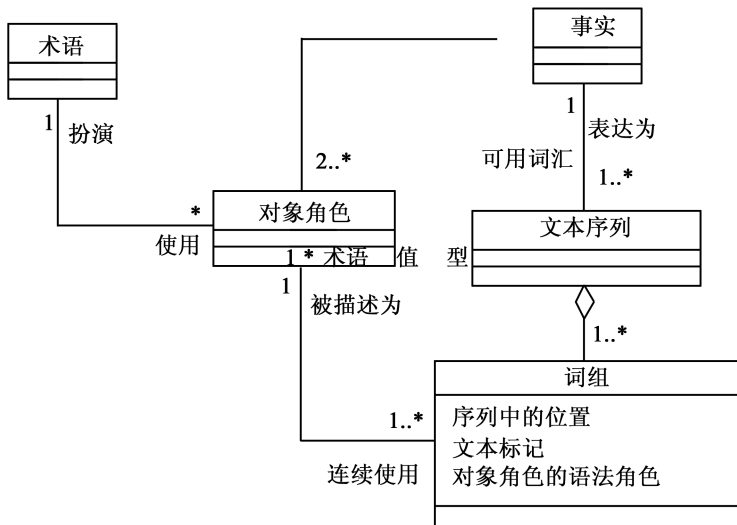


图 3-8 事实的组成结构

(5) 事实的分类

事实可以是基本事实 (Base Fact)，也可以是派生事实 (Derived Fact)。其中基本事实是明显的基础性的事实，而派生事实则是通过派生规则推理而来的，一个派生规则可以派生出多个派生事实（见图 3-9）。

事实还可以分成以下三类（见图 3-10）：

属性事实 (Attribute)：代表事实中的两个术语中，一个术语是另一个术语的属性，例如，标题是档案文件的一个属性；姓名是档案馆员的一个属性。

关系事实 (Participation)：表征事实中的术语之间存在一定的

结构化断言提供了业务活动的基本内容，但却并未添加业务活动发生需要符合的条件，而行为断言在结构化断言的基础上附加“必须”、“应用”、“不能”、“不应该”这样的限定条件（见图 3-11）。

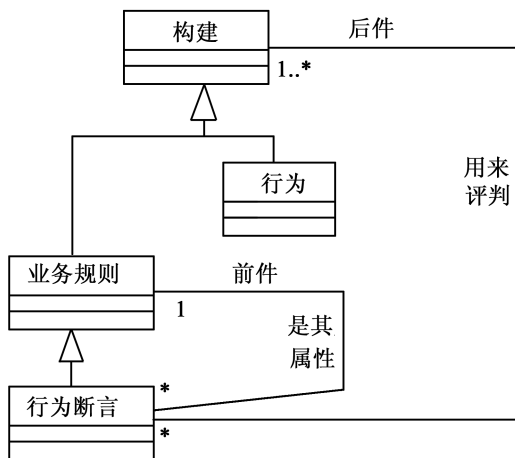


图 3-11 行为断言的组成结构

行为断言可以是一个业务规则的一部分，而行为断言包含一个前件（anchor object），在“如果……，那么……”这样的业务规则中，“如果”后面的部分即为前件，即代表业务规则中行为的发生所需要满足的前提条件或约束，前件是某个业务规则，通常情况下为结构化断言，当然也可以是其他的行为断言；“那么”后面的部分为后件（correspondent object），它往往对应一定的行为（Action）或某个业务规则，因此此处采用 Action 和业务规则的泛化概念构建（Construct）来表达后件。

以“一个档案文件必须有编号”这样一个行为断言为例，档案文件即为前件，对应一个术语，而编号则为后件。

行为断言可以分为如下三类：条件、完整性约束、授权（见图 3-12）。

①条件型行为断言表达：如果……为真，则执行相应的业务规则，它用于验证执行某种业务规则或行为的前提条件是否满足

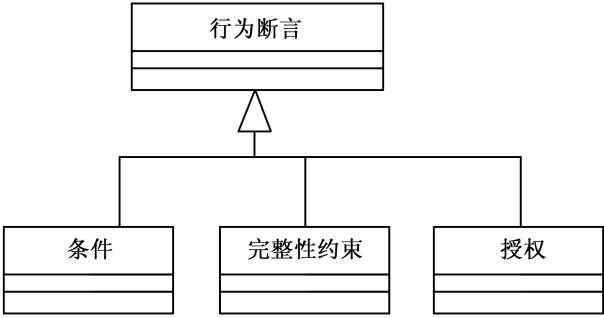


图 3-12 行为断言分类（条件、完整性约束、授权）

（为真）。例如“如果某个档案文件保管期限完毕，那么它可以被公开”或“如果用户没有提供有效证件，那么无法提供档案查询服务”，这两个行为断言都需要判断前提条件的真伪。

②完整性约束行为断言：是一个必须为真的断言，它设定了某个业务术语必须具备的特征，如“档案必须登记在册”。它体现了一种强制性的约束，违反这个约束是违背业务逻辑的。

③授权行为断言：体现了执行某个业务规则或业务行为所需的特定权限。如“只有档案馆员才可以查阅档案”，这条断言就将查阅档案的权限赋予给了档案馆员。

行为断言还可以分为以下两类：控制型行为断言（Controlling）、影响型行为断言（Influencing）（见图 3-13）。

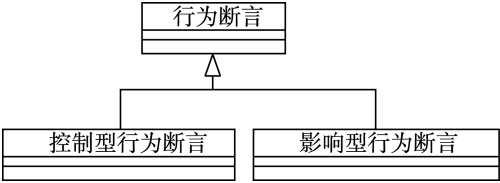


图 3-13 行为断言分类（控制、影响）

控制型行为断言对应那些采用“必须”和“必须不”修饰符定义的行为断言，这些断言必须被执行，因此可以直接内化到计算

机系统中，而有些断言，在定义过程中，往往采用“应该”、“不应该”这样的修饰符，即并没有完全强制性，因此往往不会内化到业务管理系统中，但可以在用户或业务活动违反该规则时给予提醒。

(7) 派生规则

派生规则必须在其他规则的基础上进行派生，派生的途径有两种，一种是通过事实、术语的数学计算或通过规则推理完成，通过派生规则的计算过程和推理过程，可以形成新的派生事实。因此派生规则可以用来丰富事实库（见图 3-14）。

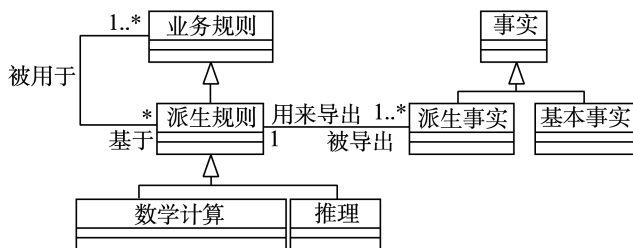


图 3-14 派生规则

3.3.2 SBVR 的业务规则语义模型

业务词汇与业务规则语义标准（Semantics of Business Vocabulary and Business Rules, SBVR），包括与业务词汇和业务规则相关的词汇。它采用 XML 作为业务规则及业务词汇的语义建模的标记语言，并具有相应的 XML 模式（Schema），可以用于进行业务词汇和业务规则的共享、交流和交换。

在 SBVR 以业务词汇的视角对业务规则的语义结构进行建模，从中派生出了面向对象模型。

(1) SBVR 模型的前提假设

模型具有以下几个假设：



①高质量的领域模型必须建立在对该领域业务熟悉的基础上，而熟悉之后将会形成成熟的该领域的共享业务词汇（Shared Business Vocabulary）。

②共享的领域词汇支持标准集成。

(2) SBVR 的设计目标

①SBVR 的模型定义中不包含任何关于信息系统的内容，因此更加概念化，能够体现业务人员的真正所想。在模型定义过程中使用的语言为最通用的日常语言，因此可以最大限度地保证语义的一致性。

②业务规则完全来自于业务词表。

③业务人员采用自己熟悉的语言交流业务词汇，即不论 SBVR 不对业务人员的语言作前提限定，英语、法语、汉语等任何一种语言在 SBVR 中，都可以用来交换和共享业务规则与业务词汇。

④SBVR 中业务规则的语义完整性是建立在谓词逻辑的基础上的。谓词逻辑可以更好地表达业务规则中的语义。

⑤业务人员通过 SBVR 可以表示各类业务词汇和业务规则，并将其提供给 IT 人员。

(3) SBVR 的五个组成方面

SBVR 标准在表达规则的语义结构时包含了 5 个方面（见图 3-15）：

①社区（Community）。社区是业务规则的基础，因为企业或单位的社区是即将建立和管理业务规则的场所。当然其他的社区，如其他的业务合作伙伴、标准化机构等，都需要加以识别。

社区的很重要的特点是在一个企业或组织中，还包含许多子社区，这些子社区之间需要共享概念、术语、事实和规则，而这些概念、术语、事实和规则又可能采用不同的词汇加以表达，可以是专业的术语，也可以是普通的自然语言。

②业务语义（Business Meaning）。一个社区具有共享的概念、事实和业务规则，所需要共享的是具体含义，而不是这些概念、事

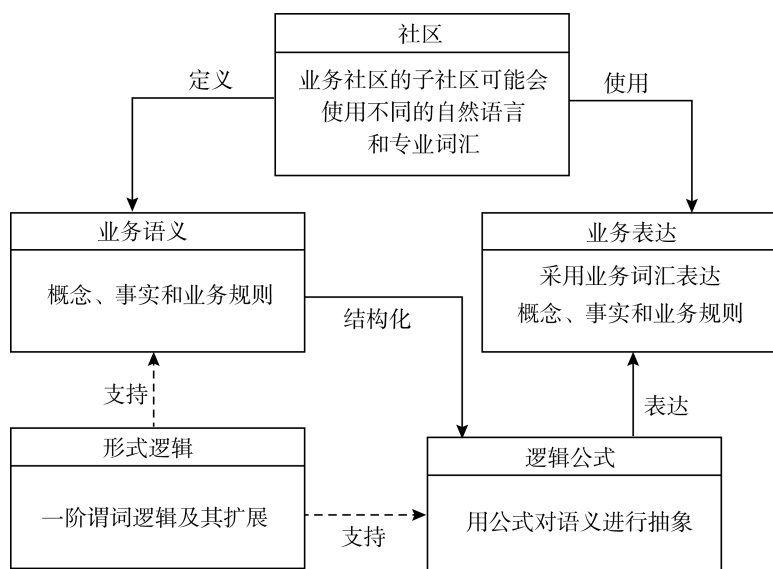


图 3-15 SBVR 的五个主要方面

实和规则的外在表现形式。尽管在传递知识时，首先需要对相关知识进行形式化表达，但 SBVR 模型将业务的含义与业务具体的表达方式隔离开了。业务语义的结构模型是将各种抽象的业务语义元素联系起来（如概念构成事实，事实又是规则的基础），并且使之不局限于任何语种的语言。

③逻辑公式（Logical Formulation）。逻辑公式提供了一种正式的、抽象的、独立于特定语言的语法，表示业务语义模型。它支持各种形式的表示方式，例如，表达式中的名词形式、动词形式，关系的双向关联等。

逻辑公式支持 SBVR 的两个重要特征。首先是将社区中的业务语义（共享概念、事实和规则的社区）映射为业务词汇，其次，将 SBVR 业务语义映射到 XMI，使各种支持 SBVR 的不同计算机软件工具可以交换、交流、分享。

④业务表达（Business Expression）。为了使社区中的成员共享概念、事实、业务规则，不可避免地要采用某种表达方式加以表



示。这些业务词汇可以采用自然语言、人工语言（如 UML）或者自然语言的子集加以标示。

SBVR 通过逻辑公式，可以将业务语义映射为某种具体的语言。SBVR 支持从其他标准、规范中提取词汇并应用到 SBVR 中，例如，SBVR 就采用了 ISO 组织中许多标准中的术语和词汇。

⑤形式逻辑（Formal Logic）。SBVR 模型的理论基础为形式逻辑，它是业务语义和逻辑公式的根基。在 SBVR 中，采用的最基础的逻辑理论是一阶谓词逻辑，同时也支持向高阶逻辑的扩展。除了一阶谓词逻辑外，还提供了到模态逻辑的扩展，用于表达规则中的可能性和不确定性。

3.4 档案业务规则的表达

3.4.1 业务规则表达的生命周期

根据业务规则的生命周期，从发现到实现阶段，业务规则表达存在 4 种类型表达形式如图 3-16 所示。该图是业务规则通过分析人员和业务人员进行对话讨论得到的，从开始对话讨论发现规则到整理成业务人员可以读懂的业务规则，再到技术人员能够读懂的业务规则，最后到机器可以读懂的业务规则。

在业务规则生命周期的不同阶段，业务规则的表达具有不同的特点。

（1）规则发现阶段的业务规则表达

该表达方式是业务规则分析人员使用的。如果是采用和业务人员进行会话交流来得到业务规则，则所得到的规则是一个一个的业务会话内容，是非形式化的。它是分析人员针对所需要完成的业务目标，通过和业务人员进行交流整理的与业务决策和业务建议或推荐业务相关的易变业务需求。具体为与业务环境相关的业务需求和

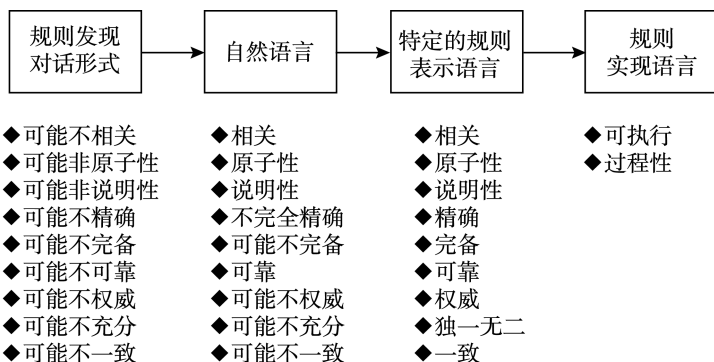


图 3-16 业务规则不同阶段采用的规则表示语言

政策性极强的业务需求。该阶段表达的业务规则之间可能是不相关，不精确，不完整，也不一致。

(2) 自然语言表示的业务规则

自然语言是针对业务人员提出的，分析人员通过和业务人员交流得到可以作为业务规则的大致内容的表达形式。业务人员通过自然语言也即业务人员的语言来对其进行描述。用该语言对业务规则进行描述，所得到的业务规则之间是相关的明确描述，但可能仍是不精确、不完整也不一致。而且用该语言表达的业务规则具有很强的业务味道，只考虑或捕捉了领域相关的内容，而不考虑其技术实现。

(3) 业务规则说明语言是针对业务人员和计算机技术人员

这种语言明确表达了业务规则完成了什么，但是没有说明怎么去完成这个目标。用这种形式表达的业务规则是唯一的、精确的、完整的、具有一致性。而且这个形式的业务规则是产生业务规则代码的基础，只可惜目前对这种语言没有统一的标准。

(4) 业务规则实现语言

业务规则实现语言是一种目标技术，该语言是可被计算机执

行，对应于某种过程语言。与前几个阶段所使用的语言相比，过程语言可以支持规则管理的自动化，因此往往需要将说明性的规则表达语言转换成为过程性的规则实现语言。

在描述业务规则中，主要有以下几种表示方法，这些方法更贴近业务人员、业务分析人员，并可以用来与规则管理系统的实现人员进行共享和知识沟通，这些业务规则描述方法有：业务规则模板、决策树、决策表、特定的规则描述语言、专家系统类的业务规则表示方法。

3.4.2 业务规则模板表示方法

最为著名的规则模板表示方法为 Ross 的 RuleSpeak 模板表示方法，其将每个模板与相应的业务规则类别对应，通过一定的约定和表示方法，可以将一个业务规则方便地转换成为业务规则模板。在业务规则模板中，业务规则由简单的规则从句格式组成，其格式为：<term1><operator><term2>。其中 term 是统一定义的名词或名词短语，包括概念、概念的属性、值和值的集合等。operator 是对各 term 类型有意义的任何操作。任何对业务有意义的术语和操作都将存在。结构化规则语言通过连接规则从句来表达一个业务规则。如果业务规则太复杂用结构化规则语言表达不了，可以借助于自然语言，科学注释等任何有意义的方法。

表 3-2 展示了 RuleSpeak 的业务规则模板表示方法，在该方法中，将每个具体类型的业务规则映射为具有特定格式的规则模板，模板中包括主语、规则关键词、事实、条件、数据项、介词、数学公式等（不必全部包含）。

- 其中符号“<>”表示所包含的句法项是强制的。
- 符号“[]”表示所包含的句法项是可选的。符号“/”表示需要选出所列出的两个或多个句法项中的一个。
- 条件永远包含一个逻辑表达式（必须永为真货假）。条件永远以一个或多个术语和事实（或数据项）为基础，并且包含逻辑操作符（与、或、非）。

- “<>”中的事实引用主语后面的事实描述的其他部分，例如“档案馆员管理档案文件”事实描述的<事实>是“管理档案文件”。此外在所有<事实>出现的地方，都允许嵌入条件。例如“管理近2年的档案文件”中嵌入了条件“近2年”。
- 在规则模板中，规则关键字应该（Should）代表的是一种非强制性的建议。

表 3-2 RuleSpeak 模板

分类	规则的主语 必须是……	模板
1. 拒绝规则 规则关键词：必须（MUST）、只有（ONLY）	术语或事实（也允许数据项）	<主语>[不]必须/应该<事实>[当<条件>] <主语>可以/应该<事实>只有如果/当<条件> <主语>可以/应该<事实>只有<介词><条件>
2. 允许描述 规则关键词：可以（MAY）不必（NEED NOT）	术语、事实、规则或过程（也允许数据项）	<主语>可以<事实/规则关键字>[如果/当<条件>] <主语>不必<事实/规则关键字>[如果/当<条件>]
2.1 计算规则 规则关键词：被计算为（BE COMPUTED）	计算术语（也允许数据项）	<主语>[不]必须/应该被计算为<数学公式>[如果/当<条件>] <主语>=<数学公式>[如果/当<条件>]
2.2 导出规则 规则关键词：被看做是（BE TAKEN TO MEAN）意味着（MEANS）	所导出的术语（也允许数据项）	<主语>[不]必须/应该被看做是<逻辑表达式>[如果/当<条件>] <主语>[不]意味着<逻辑表达式>[如果/当<条件>]

续表

分类	规则的主语 必须是……	模板
3. 投影规则		
3.1 启用规则		
3.1.1 推论规则 规则关键词:被认为 (BE CONSIDERED)	术语 (也允许数据项)	<主语>[不] 必须/应该被认为<术语>如果/当<条件> <主语>[不] 是<术语>如果/当<条件>
3.1.2 规则锁定 规则关键词:除了 (UNLESS) 除 非 (EXCEPT) 执行 (BE ENFORCED)	规则	<规则描述>, 除了/除非<条件> <规则名称>[不] 必须/应该被执行 如果/当<条件>
3.1.3 过程锁定 规则关键词:被启用 (BE ENABLED) 被禁用 (BE DISABLED)	过程或规程	<主语>[不] 必须/应该被启用/禁用 如果/当<条件>
3.1.4 数据锁定 规则关键词:被创建 (BE CREATED) 被删除 (BE DELETED)	数据项	<数据项>[不] 必须/应该被创建/删除 如果/当<条件>
3.2 复制规则		
3.2.1 印记规则 规则关键词:被设置为 (BE SET)	术语或事实 (也允许是数据项)	<术语>[不] 必须/应该被设置为<术语/取值>[如果/当<条件>]
3.2.2 表示规则 规则关键词:被显示 (BE DISPLAYED)	术语或事实 (也允许是数据项)	<主语>[不] 必须/应该<在<媒介>上>以<显示方式>被显示 [如果/当<条件>]
3.3 执行规则		

续表

分类	规则的主语 必须是……	模板
3.3.1 过程触发器 规则关键词:被执行 (BE EXECUTED)	过程或规程	<主语>必须/应该被执行 当<条件>
3.3.2 规则触发器 规则关键词:被触发 (BE FIRED)	规则	<规则名称>必须/应该被触发 当<条件>

3.4.3 决策表表示业务规则

决策表又称判断表，是一种呈表格状的图形工具，适用于判断条件较多，各条件又相互组合、有多种决策方案的情况^①。它是一种采用表的结构，精确而简洁地描述复杂逻辑的方式。它类似于高级程序语言中的 if-then-else 和 switch-case，将多个条件与这些条件满足后要执行动作相对应。但不同于传统程序语言中的控制语句，决策表能将多个独立的条件和多个动作直接的联系清晰地表示出来。

在业务规则表示过程中，每个规则都包含一个主语 (Subject)，并且总有一个或多个其他术语。对于决策表，这些其他术语就是“评价术语”，或称为“决策准则”，这些决策准则提供了决策表的行标签和列标签。决策准则的取值将决定主语的输出取值。

采用决策表的前提条件有：

- ①当有多个并行规则，这些规则共用一个主语，并且具有相同

^① Jiawei Han, 等. 数据挖掘：概念与技术（第3版）[M]. 北京：机械工业出版社，2012.

②每个评价术语的取值的个数是有限的。

在此处,评价术语相当于 IF<conditions>THEN<actions>语句中的条件项,而主语的输出则为 THEN 行为项,代表给定条件下的决策行为。

R1: 如果 档案类型为“党组会议纪要”, 则保管期限为永久;

R3: 如果档案类型为“交流材料”, 则保管期限为30年;

则会形成如表 3-3 所示的决策表。

决策表 1

档案类型	保管期限
党组会议纪要	永久
简报复印件	10 年
交流材料	30 年

R1：如果档案类型为“党组会议纪要”而且来源单位为本机关，则保管期限为永久；

R3: 如果档案类型为“交流材料”而且来源单位为本机关, 则保管期限为 30 年;

则会形成如表 3-4 的决策表:



表 3-4

决策表 2

档案类型	保管期限	
	本机关	上级机关
党组会议纪要	永久	永久
简报复印件	10 年	30 年
交流材料	10 年	30 年

在上一个决策表中，分别将评价术语分布于表的行和列，而右下角的则为主语的取值，代表在行列两个评价术语特定取值（条件与条件取值）的情况下，主语输出。

当评价术语的数目为 3 个时，可以采用行分裂和决策表分组的方式进行，如表 3-5 所示。

表 3-5

决策表 3

档案类型		保管期限	
		本机关	上级机关
会议文件	党组会议纪要	永久	永久
	代表发言材料	30 年	10 年
	会议决议	永久	30 年
政策文件	基础性文件	10 年	30 年
	技术性文件	30 年	10 年
	操作性文件	10 年	10 年
计划总结	年度总结	10 年	30 年
	季度总结	10 年	10 年

将上述决策表采用分组的方式可以形成如表 3-6，表 3-7，表 3-8 所示的决策表分组。

表 3-6 会议文件决策表

档案类型	保管期限	
	本机关	上级机关
党组会议纪要	永久	永久
代表发言材料	30 年	10 年
会议决议	永久	30 年

表 3-7 政策文件决策表

档案类型	保管期限	
	本机关	上级机关
基础性文件	10 年	30 年
技术性文件	30 年	10 年
操作性文件	10 年	10 年

表 3-8 计划总结决策表

档案类型	保管期限	
	本机关	上级机关
年度总结	10 年	30 年
季度总结	10 年	10 年

当决策表的评价术语数目多于 3 个时，整个决策表的结构和分组的数目增多，此时可以采用如表 3-9 所示的决策表格式。

表 3-9 通用决策表

评价术语	评价术语 1（条件 1）	V_{11}	V_{12}	V_{13}	V_{14}	V_{15}
	评价术语 2（条件 1）	V_{21}	V_{22}	V_{23}	V_{24}	V_{25}
		...	...	...	...	...
	评价术语 N （条件 N ）	V_{N1}	V_{N2}	V_{N3}	V_{N4}	V_{N5}

续表

主 语 行 为	主语取值 S_1	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}
	主语取值 S_2	S_{21}	S_{22}	S_{23}	S_{24}	S_{25}
	主语取值 S_3	S_{31}	S_{32}	S_{33}	S_{34}	S_{35}
	主语取值 S_4	S_{41}	S_{42}	S_{43}	S_{44}	S_{45}

在表 3-9 中, V_{ij} 代表第 i 个评价术语的第 j 个取值; 而 S_{ij} 则为
主语行为的取值。

对文件的保管期限采用上述表达方式表示, 形成表 3-10, 需要
说明的是, 在主语行为输出部分可以是“是/否”, 也可以是概率、
频率甚至是模糊值。

表 3-10 通用决策表范例

评 价 术 语	机关	本	上级	本	上级	本	上级	本	上级
	时间	年度	月度	月度	年度	年度	月度	月度	年度
	会议	纪要	纪要	纪要	纪要	决议	决议	决议	决议
主 语 行 为	10 年	—	—	—	—	—	—	—	是
	30 年	是	—	是	—	是	是	—	—
	永久	—	是	—	是	—	—	是	—

3.4.4 决策树表示业务规则

决策树是用树形图来表示处理逻辑的一种工具, 因此可以用于
业务规则的表示中。与决策表相比, 它可以更直观、清晰地表达业
务规则的结构, 特别适合于评价术语较少、逻辑组合不复杂的情
况。

决策树提供了一种展示类似在何种评判术语的取值下会得到什
么行为取值的这类规则的方法。一棵决策树, 由决策节点、分支叶

子组成。决策树中最上面的节点称为根节点，是整个决策树的开始。其对应一个评判术语，该节点的分支数目取决于评判术语的取值数目。

决策树的每个节点子节点的个数与决策树所用的算法有关。如 CART 算法得到的决策树每个节点有两个分支，这种树称为二叉树。允许节点含有多于两个子节点的树称为多叉树。

每个分支要么是一个新的决策节点，要么是树的结尾，称为叶子。在沿着决策树从上到下遍历的过程中，在每个节点都会遇到一个评判术语，对每个节点上评判术语的不同取值导致不同的分支，最后会到达一个叶子节点。这个过程就是利用决策树进行分类并获取主语行为的过程。

图 3-17 展示了档案文件保管期限的一棵决策树。

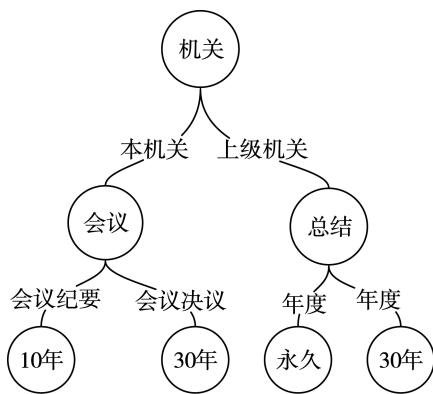


图 3-17 决策树

3.4.5 规则标记语言

规则模板、决策表和决策树表达业务规则的方式，直观清晰，适合于业务分析人员和用户使用，但是当向规则的形式化迈进并最终使之成为可执行的规则处理程序时，就必须采用更加复杂的规则

标记语言对规则本身进行标记。

规则标记过程就是定义规则的过程，使规则更加结构化，在业务规则标记方面，出现了许多的标记语言：包括 RuleML、SWRL、BRML 等，这里我们主要介绍 RuleML 和 SWRL。

3.4.5.1 RuleML

RuleML (Rule Markup Language) 的目标是为学界和工业界中致力于规则管理的研究者和从业者提供一套规范的规则标记语言，方便各方标记、交换和共享规则词汇和各种规则，因此表达业务规则是 RuleML 的一个目标，而更重要的是，通过 RuleML 标记的规则可以进行交流、交换和共享。

RuleML 采用 XML 作为规则标记的元语言，通过相应的 XML 模式，定义了 RuleML 标记规则的模式。RuleML 标记语言核心的语言基础是符合 Horn 逻辑的 Datalog 子语言，Datalog 是 SQL 语言与 Prolog 语言的交集。由于采用 XML 作为标记语言，RuleML 可以满足各类智能系统、规则系统以及 Agent 主体在数据交换过程中对通用格式的需求，可以较好地实现 Web 语义的规则表示。RuleML 对于规则表示的各个部分都进行了细致的刻画，它把一个规则划分成为若干个原子部分，支持反映性规则、完整性规则、派生规则和事实性规则四种规则。在进行规则标记时，采用 Head 标签表示事实的结论或 Then 部分 (IF...Then...形式)，Body 则表示规则体 IF 部分。RuleML 不但可以支持各种流行的规则系统，还可以用 var、ind 等标签表示不同的单个常量、逻辑变量和结构等，此外它还能很好地支持 Horn 逻辑语句，特别适合规则系统的表达。

下例为采用 RuleML 表达的规则实例，即如果档案来源是本机关，而且档案类型是会议纪要，那么档案保管期限为 30 年。

例 1：RuleML 范例

```
<Implies>
  <head>
    <Atom>
      <opr> <Rel>是</Rel> </opr>
```



```

    <ind>30 年</ind>
    <Var>档案保管期限</Var>
  </Atom>
</head>
<body>
  <And>
    <Atom>
      <opr> <Rel>是</Rel> </opr>
      <ind>会议纪要</ind>
      <Var>档案类型</Var>
    </Atom>
    <Atom>
      <opr> <Rel>keep</Rel> </opr>
      <ind>本机关</ind>
      <Var>档案来源</Var>
    </Atom>
  </And>
</body>
</Implies>

```

在上述 RuleML 标记中，<ind>标签代表常量，<Var>标签代表变量，对应规则中的术语，<And>标签代表两个<Atom>（原子事实）是与的关系，<body>对应规则的 IF 部分，而<Head>部分对应规则的 THEN 部分。

3.4.5.2 SWRL

SWRL（语义网规则语言：A Semantic Web Rule Language Combining OWL and RuleML）是用于描述语义网的规则语言，它结合了 OWL 和 RuleML，因此是在 RuleML 以及 OWL 的基础上发展起来的。SWRL 扩展了 OWL 的公理系统（Axioms），增添了类 Horn 的规则。因此可以支持类 Horn 规则构建 OWL 知识库。SWRL 同样采用 Head 元素表征规则的结论（后件），Body 元素表征规则的前件。前件和结论包含零个或多个原子事实（Atom），Atom 可

以是 $C(x)$, $P(x, y)$, $\text{sameAs}(x, y)$ 或 $\text{differentFrom}(x, y)$ 四种之一, 其中 C 是一个 OWL 描述符, P 是一个 OWL 属性, x 和 y 可以是变量、OWL 个体或 OWL 数据值。

下例表示了如何采用 SWRL 表示规则^①, 其中 `ruleml`、`swrlx` 为命名空间, 体现了 SWRL 对 RuleML 的借鉴。在该例中, `Body` 元素包含了三个 `Atom`, 对应规则的前提 (IF 部分), 即 `hasParent` (x_1 , x_2) (x_2 是 x_1 的父母), `hasSibling` (x_2 , x_3) (x_3 是 x_2 的兄妹), `hasSex` (x_3 , `#male`) (x_3 的性别为男); `Head` 元素包含一个 `Atom`, 对应规则的结论 (THEN 部分), `hasUncle` (x_1 , x_3) (x_1 是 x_3 的叔叔)。整个范例表达了: 如果 x_2 是 x_1 的父亲或母亲, 而且 x_3 是 x_2 的兄妹, 而且 x_3 为男性, 那么 x_3 是 x_1 的叔叔。

例 2: SWRL 示例

```
<ruleml:imp>
  <ruleml:_rlab ruleml:href="#example2"/>
  <ruleml:_body>
    <swrlx:individualPropertyAtom swrlx:property="hasParent">
      <ruleml:var>x1</ruleml:var>
      <ruleml:var>x2</ruleml:var>
    </swrlx:individualPropertyAtom>
    <swrlx:individualPropertyAtom swrlx:property="hasSibling">
      <ruleml:var>x2</ruleml:var>
      <ruleml:var>x3</ruleml:var>
    </swrlx:individualPropertyAtom>
    <swrlx:individualPropertyAtom swrlx:property="hasSex">
      <ruleml:var>x3</ruleml:var>
      <owlx:Individual owlx:name="#male" />
    </swrlx:individualPropertyAtom>
  </ruleml:_body>
```

^① SWRL 参见 <http://www.w3.org/Submission/SWRL/> [EB/OL]. 2011-05-22.

```
<ruleml:_head>
  <swrlx:individualPropertyAtom swrlx:property="hasUncle">
    <ruleml:var>x1</ruleml:var>
    <ruleml:var>x3</ruleml:var>
  </swrlx:individualPropertyAtom>
</ruleml:_head>
</ruleml:imp>
```

3.4.6 其他的业务规则表示方法

除了上文所述的规则模板、决策表、决策树等表示方法外，还包括许多其他的业务规则表示方法。这些方法包括：

①一阶谓词逻辑表示法。一阶谓词逻辑演算主要是用来更好地解决具有不确定性的系统。通过采用一阶谓词逻辑演算、临时逻辑、模糊逻辑、指示逻辑来进行推理。当前 BRMS 很少支持不确定逻辑推理，其内部的规则往往都是明确定义的。

②UML 表示法。采用 UML 可以表达业务规则中的事实模型，通过添加类、子类并建立各种不同类和对象的关系，可以有效地表达术语之间的各类事实；同时 UML2.0 标准也提供了表达业务规则的语言 OCL，OCL 是 UML 标准中的对象约束语言（Object Constraints Language），它通过向 UML 图添加注释，对 UML 图中的类或对象添加约束。但是 OCL 的注释只能针对某一对象或类，无法注释对象或类之间的关联关系。

③语义网络表示法。相对于产生式规则主要用以描述因果知识，语义网络则能够用来表达更加复杂的概念及其之间的相互关系，形成一个由节点和弧组成的语义网络描述图。在语义网络图中节点指示事物，节点间以有向弧连接，而弧上的标签则指示节点间关系。带标签的有向弧也称为关系弧。对于表示关系的标签中有两个比较重要，其一是 IS A 关系，它表示 A 概念是 B 概念的一个实例，比如 Window。这个概念，我们这样表示：Windows IS A Operate System，就能把它从一般的描述窗户的概念中区分出来，

专指一种操作系统。其二是 AKO (A Kind Of) 关系, 它表示 A 概念是 B 概念的一个子类, 人 AKO 动物, 就表示人是动物中的一个子类。

4. 基于业务规则的档案信息资源规划

档案信息化是一个以信息资源建设为核心的长期发展过程，信息资源规划是其基础和先导工程。将信息资源规划理论引入基于业务规则的档案信息资源管理，实施档案信息资源的全面规划，将会更好地整合现有的各类档案信息资源和业务规则，降低档案信息资源管理系统开发失败的风险。

基于业务规则的档案信息资源规划，首先需要定义档案信息资源，对档案管理业务各流程进行梳理，构建档案业务模型；在此基础上，确定档案业务规则来源，组织档案业务规则，构建档案业务规则知识库；最后，从全局出发构建基于业务规则的档案信息资源管理模型，此外，制定管理标准和选择关键技术也是规划阶段的必要环节。

4.1 档案信息资源规划概述

4.1.1 信息资源规划和档案信息资源规划的定义

(1) 信息资源规划的定义

不论是在一个具体的组织机构范围内还是在行业、地区、国家等更大的范围内，信息资源规划都是指对信息资源采集、描述、处理、存储、管理、定位、访问、重组与再加工等全过程的全面规划

工作^①。根据规划对象、规划目的以及规划方法等的不同可以将信息资源规划分为宏观和微观两种规划，二者在规划对象、目的、规划方法、规划重点、执行方式和存在形式等方面均存在不同之处。两者之间的不同点比较详见表 4-1。

表 4-1 宏观信息资源规划与微观信息资源规划的不同点^②

比较项目	宏观层面信息资源规划	微观层面信息资源规划
对象	所有信息资源	系统规划中的信息数据
目的	加强信息资源的综合开发与利用，提升信息资源在社会运转中的功能，减少社会运行的成本，保证社会向预期目标发展	在计算机系统内统一数据标准，消除信息孤岛，实现信息共享，保证企业利润最大化的实现
规划方法	从调查研究入手，提出社会发展中信息资源领域里的发展目标、主要任务与保障措施	作为一种工程技术，包括需求分析、系统建模、支持工具等内容
规划重点	通过若干项目整合全社会信息资源，包括结构化和非结构化的信息资源	信息数据的标准化与信息数据的有序流动，主要为结构化的信息资源
执行方式	自上而下，政府引导，社会参与	计算机信息系统自动执行
存在形式	可独立存在并发挥作用	必须依附于系统建设才能发挥作用

总的来说，宏观层面上的规划是经济社会发展规划的一部分，是一项信息化的专项规划；微观层面上的信息资源规划主要是运用

① 国务院信息化工作办公室政府信息资源开发利用政策研究课题组．信息资源规划与国家基础数据库的政策建议 [J]．中国信息界，2004（4）：7-12．

② 罗程．两种不同的信息资源规划 [J]．中国信息界，2007（2）：44-45．



信息工程及信息资源管理等理论方法,对边界清晰的组织机构在业务活动中所需要的信息资源进行全面和详细地规划。档案信息化建设是个复杂的系统工程,其中档案数字资源建设是核心^①,应当进行完整、合理的规划。笔者认为档案信息资源规划应同时采用两种信息资源规划方法,既要体现国家档案信息化的战略思想和战略目标,也要强调对信息系统基本性问题的设计规划,以提高规划的有效性和可实施性,也就是说规划要点是既全、又专、也要细。

(2) 档案信息资源规划的定义

档案信息资源规划是指依据档案法律法规及国家相关规划纲要,在国家档案行政管理部门的宏观指导下,引用先进的信息工程和数据管理理论方法,对档案信息资源的收集、整理、鉴定、保管、统计、编研、检索、利用的全面规划,其目的是解决档案信息资源共享和整合问题,弥合信息孤岛,实现档案信息集成化的应用开发。

宏观层面上的档案信息资源规划,是国家信息资源综合规划的一部分,服从于国民经济和社会发展规划,是一项信息化的专项规划,如《全国档案信息化建设实施纲要》、《××省档案信息化建设规划纲要》、《××市档案信息化建设规划纲要》等。这些具有宏观指导作用的相关政策性规划一般由政府规划部门或政府委托的单位编制,档案管理部门需要在其指导下开展本单位的档案信息化建设。微观层面的信息资源规划已有一套成熟的理论和实施方法,在国内外的企业中运用得较为普遍,其先进的信息工程和数据管理理论及方法是政府、企事业单位的档案部门开展信息资源规划过程中迫切需要引进的。相对来说,微观档案信息资源规划的研究对于现今档案信息化建设更具有现实意义和实用价值,因此,本书的研究重点是基于微观信息资源规划理论与

^① 吴光华. 略论国家档案馆信息资源建设规划 [J]. 上海档案, 2008 (9).

方法的档案信息资源规划。

4.1.2 档案信息资源规划的实施步骤

档案信息资源规划是一项涉及人、财、物、技术等因素的系统工程，需要依据一定的步骤有序开展。不管档案信息化建设的对象如何，其基本实施步骤主要为：首先是规划前期准备，包括组织协调、规则小组成立和确定总体目标等；其次是规划主体，主要包括需求分析、系统建模和制定标准规范体系；最后是新旧系统过渡、设计系统实施进度和底层工具，以及提交规划报告等善后工作。

(1) 规划前期准备工作

确立档案信息资源规划项目，首先要得到单位最高层管理人员的赞成，然后成立规划小组，制定项目总体目标、主要任务、保障措施等前期准备工作。

档案是国家机构、社会组织或个人在社会活动中直接形成的，供保存备查的原始记录。档案信息源自各单位日常生产、科研、行政和经营管理过程中，档案信息资源规划涉及各方面人、财、物的支配和调用，需要高层决策者及时有效地组织协调各部门、各系统之间的关系，同时单位领导的参与和重视还必然提高了该单位其他人员对档案管理工作和档案信息资源规划的重视，带动部门全体员工积极参与。

在实施档案信息资源规划前，还必须成立一个专门的规划小组，小组成员应该包括信息资源规划项目负责人、各职能域业务代表、系统分析员、信息技术人员，其主要职能是统一解决管理和关键技术问题。由于各企事业单位档案工作现状和档案人员素质良莠不齐，多数单位完全靠自己组织信息资源规划和实施档案信息化的能力还不够，实践证明，与具有较高信誉度的档案专业咨询公司合作，不仅能够节省人力物力，还能够降低档案管理信息系统开发失败的风险。



档案信息资源规划的总体目标是指对档案信息化建设需要达到的目的和主要任务的总体规定,即解决档案信息资源共享和整合问题,实现档案信息集成化的应用开发。各单位可根据本单位规模和性质进行调整和细化,如企事业单位档案室的档案主要是本单位内形成的,其用户主要是本单位的员工,而国家档案馆是面向全社会的,两种档案管理部门信息资源规划的具体任务会存在差异。此外,还要制定好确保规划内容得到有效实施的各种保障措施,概括来说,这些保障性措施大致包括法治保障、管理保障、技术保障、资金投入保障、人才保障、思想保障和安全保障这七类^①。

(2) 规划主体

信息资源规划主要包括各职能领域的信息需求、数据流分析、信息资源管理基础标准的制定,以及全域和各职能领域信息系统框架——功能模型、数据模型和系统体系结构模型的建立^②。实施过程中可以借助有效的软件支持工具,如国产网络版信息资源规划工具软件 IRP2000net、信息分类编码管理工具 IRP3-icc 等。尽管不同信息部门的信息资源规划范围和内容有所不同,但其基本方法和实施步骤大致相同。档案信息资源规划是依据档案法律法规及国家相关规划纲要,在国家档案行政管理部门宏观指导下,以信息系统工程方法论^③为指导,采用工程化方法^④(详见图 4-1),经过需求分析和系统建模两大阶段,遵循一定的标准规范来进行的,一般由档案业务人员和系统分析人员密切合

① 赵国俊. 档案信息资源开发利用专项规划及其内容体系 [J]. 北京档案, 2005 (9): 16-18.

② 高复先. “数字校园”建设的基础工程: 信息资源规划 [J]. 中国教育网络, 2006 (8): 63.

③ 高复先. 信息资源规划的理论指导 [J]. 中国教育网络, 2006 (9): 62-64.

④ 高复先. 信息资源规划的工程化方法 [J]. 中国信息界, 2005 (20): 21-24.

作共同完成。

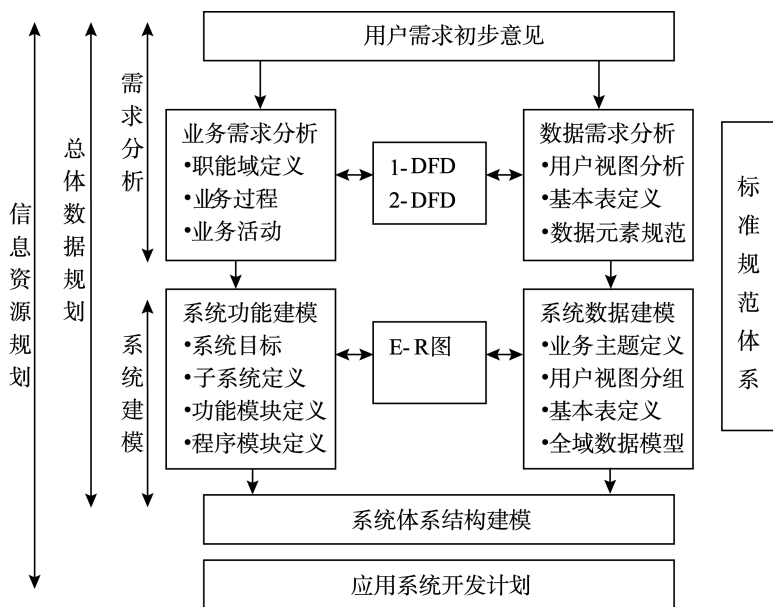


图 4-1 信息资源规划的工程化方法图①

1) 需求分析

根据信息工程化方法，信息资源规划正式实施的第一阶段是需求分析。需求分析的结果是理清本单位档案工作的业务流以及数据流。业务流程分析是按职能域进行的，一般要经过以下几个步骤：①先订立实施档案信息资源规划单位的职能域，确定规划工作范围；②分析每一个职能域的业务过程以及该职能域档案形成、流转、管理和开发利用的流程；③将每个档案业务过程分解为业务活动，对定义的业务过程和活动进行复查。经过以上业务梳理，形成以“职能域—业务过程—业务活动”三个层次来表达业务功能结

① 李姝琳，孙德杰，史震东．信息资源规划——IRP 技术研究 [J]．中国水运，2008（4）：230-231．



构的业务模型。数据流分析是依据业务模型，进行用户视图分析和基本表的定义，分析业务活动中档案流转方式、种类，同时确立档案信息管理各阶段所依据的规范、标准，如档案分类方案、著录标引标准、保管期限表等。

2) 系统建模

系统建模是需求分析的综合与定型，包括系统功能建模、数据建模和信息系统结构建模。在需求分析阶段，建立了由“职能域—业务过程—业务活动”三层结构组成的业务模型，但是并非所有的业务活动都能实现计算机化的管理，经分析可以发现有些业务活动可以由计算机自动完成；有些业务活动可以人-机交互完成；有些业务活动仍然需要由人工完成。我们将能由计算机自动进行处理和人-机交互进行的活动挑选出来，按“子系统—功能模块—程序模块”组织，就是系统功能模型。一般来说，业务模型与功能模型对应关系如图 4-2 所示。根据业务过程和业务活动分析，可以确定档案类别，一般来说档案分为文书档案、会计档案、人事档案、基建档案、设备仪器档案和特殊载体档案六大类，这些大类的名称即可作为主题数据库的名称，各主题数据库又由若干个基本表组成。在信息工程方法论中，信息系统体系结构是指系统数据模型和功能模型的关联结构，也就是确认主题数据库与业务过程的对应关系，可采用 C-U 矩阵来表示。

3) 标准规范体系

信息资源管理要从标准化做起，标准化是一项极其重要的基础性工作。总的来说，档案信息资源规划中主要涉及两类标准，即技术标准和管理标准。技术标准主要有数据元素标准、信息分类编码标准、用户视图标准、概念数据库标准以及逻辑数据库标准等，是实现数据共享、数据交换、数据库设计、应用软件设计和系统建设的保证。管理标准主要是档案信息资源管理的基础标准，包括档案分类标准、著录标引标准、鉴定标准、管理规范以及相关的信息政策和信息法律法规等，这些标准化的规则不仅是日常管理的规范依据，也是建立系统业务模型和功能模型的基础保证。标准规范体系

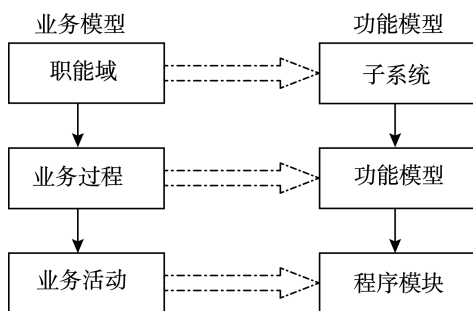


图 4-2 业务模型与功能模型对应关系

的建立就是要在规划过程中制定一系列相应的技术标准和管理标准，如《建设电子文件与电子档案管理规范（CJJ/T117-2007）》，《档案著录规则（DA/T18-1999）》等，并将这些标准纳入业务规则库中，对其进行集中统一管理，形成实现数据共享、信息集成以及指导档案信息全程管理的规范化系统。

（3）规划后期任务

1) 新旧系统过渡

对于信息化建设不发达的地区或单位，档案信息化工作正处于起步或者还没有开展，规划待开发的档案信息管理系统与现有的应用系统的衔接较为简单。对于已经建设和使用档案管理信息系统的单位，新系统在进行设计开发时，与旧系统相比，在功能和结构上都进行了相应的扩充和调整，使得新旧系统在功能和结构上存在一定的差异，应重视新旧系统过渡问题，做好准备工作，以保证新旧系统的良好衔接或平稳过渡。

一般来说新旧系统有三种关系，即用一个现存的软件包取代现有系统、新旧系统混合使用、重新开发一个新系统。三种不同的方式各有利弊，采用现存软件包，如直接购买、安装软件公司开发的通用数字化档案网上管理系统软件包，看起来比较方便，但如果该软件不符合本单位管理需要和特点，则需要修改以满足本单位业务



的需求，最终需要很大的附加投入；新旧系统混合使用，如沿用本单位原有的数字化档案采集与加工系统，可以减少浪费，但是恰到好处地设计新系统的互补功能本身就是个不容易的事，另外新旧系统接口问题也比较复杂；重新开发新系统一般来说功能更全、更适合本单位业务，但却是个高投资高风险的工程。档案管理信息系统的开发设计应结合单位性质和档案管理业务特点，基于需求分析，综合考虑各项指标确定新旧系统的衔接和转换方式，解决新旧系统过渡问题。

2) 设计系统实施进度和底层工具

管理信息系统建设不能随意进行，在执行过程中一定要按预先制定的时间计划和规章制度进行，设计好系统实施的优先顺序是项目成功与否的关键因素。因此，信息资源规划后期工作之一就是要按照科学的方法制定系统实施时间进度表，使得执行阶段能够按照规划制定的时间计划以及各项规章制度有条不紊的开展工作。

此外，信息工程方法论是一套包括自顶向下规划和自底向上设计的方法论，信息资源规划不仅应有顶层设计，还要细化到底层工具的设计。各单位在档案信息资源规划后期阶段，应结合规划的前期调研，分析本单位已使用的技术和未使用的先进技术，选择最适合本单位档案信息化建设的软件支持工具，建立一套与自顶向下规划工具相衔接的自底向上设计实现的技术体系。

3) 提交规划总结报告

在档案信息资源规划结束后，要认真对规划进行总结、记录，编写、提交规划总结报告，规划总结报告是信息资源规划最终成果的详细记录与展示，是规划成员集体智慧的结晶。总结报告内容丰富，既包含该信息资源规划项目实施的基本要素、项目总体目标、主要任务、措施保障、规划方法等总体规划的介绍，也包括需求分析、系统建模以及技术、管理等方面的细化规划陈述，报告中的规划成果一般来源于机内文档和纸面文档，要按照规划项目验收具体要求的格式书写内容和准备材料。

4.2 档案业务规划

4.2.1 档案信息资源的定义和定位

信息资源是对信息进行选取、组织、序化而形成的有用信息集合，信息资源是信息的一部分，是在当前生产力水平和研究水平下人类所开发与组织的信息，是可供人类利用的有用的、创造价值的信息。档案信息资源是信息资源的种概念，是信息资源的一种，档案的原始性、凭证性等价值树立了档案信息在信息家族中不可替代的地位。档案是社会活动的伴生物，只要有人类的社会活动，就会有档案的产生。在开发档案业务规则管理系统前，需要定义和定位系统范围内本组织形成的或对本组织有保存价值的档案信息及其来源，明确管理对象。

档案资源主要包括实体资源和信息资源，其中实体资源主要包括目前各档案机构保管的档案原件、电子文件的纸质文本、档案信息开发实物成果等；信息资源主要指的是实体档案内容所反映的信息和知识集合。基于业务规则的档案信息资源管理主要针对的是档案原件的数字化文本、电子文件以及其他开发成果的数字副本等所反映出的信息资源。

规划小组应召集本组织内各项工作的业务人员以及档案管理人员，共同梳理业务工作，结合上级部门的档案种类和归档要求，确定本组织归档范围和档案信息资源的种类。

4.2.2 档案业务分析

长期以来，在纸质办公环境中，对于档案管理来说，无论是各级机构组织的职责，还是档案归档、整理、保管和利用等管理程序方面，已经形成了一套完整的管理流程和办法。例如，档案工作实



行统一领导、分级管理的原则；档案实体管理内容包括档案的收集、整理、鉴定、保管和统计^①；档案管理应遵循来源原则、全宗原则、有机联系以及保管和利用方便原则；还有国家、地方、各单位制定的有关档案的《机构、人员编制规定》、《归档范围和保管期限表》、《归档文件整理规则》、《档案著录规则》、《照片档案整理规范》等规定和标准。这些是做好档案工作的根本保证，很大程度上也是电子网络环境下档案工作开展的基本制度和标准。然而随着网络信息技术的跨越式发展，不论是政府部门、企事业单位工作方式，还是人们的日常交流方式都发生了巨大变化，文档的形成和传递呈现数字化网络化趋势。现代信息技术，主要是计算机技术和网络技术在档案管理中的应用，使原有的管理模式发生了变化。自动化集成管理成为档案管理的发展方向，档案业务工作的变化主要体现在以下几个方面：

(1) 档案信息收集、管理方式的变化

在网络环境下，数据处理在文件处理阶段就可完成，档案的有关数据一经输入，可多次使用，公文处理与档案管理的有机结合，真正实现了文档一体化管理；电子文件的网上收集采取实时归档的方式，通过一个接口顺利地导入到档案管理系统中；将鉴定工作所涉及的相关要求在文件产生之前即在整个组织信息流程的计算机系统设计阶段就确定下来，有别于将对办理完毕的文件进行价值判断的过程单纯移至文件形成或流转之初的做法，在很大程度上增强了对未来利用需求预测的可操作性，使得通过计算机对档案鉴定实行批处理成为可能；可以自动完成各类档案资料的统计工作，绘出各类统计报表；可以实现库房档案管理工作，如库房档案总数、现存档案数量，借出、归还档案的情况，并可实现库房空气与温湿度、照明、防火、报警等系统的自动控制。

^① 陈智为，邓绍兴，刘越男. 档案管理学 [M]. 北京：中国人民大学出版社，2008.

(2) 档案信息组织检索的变革

第一, 档案信息组织与传播自动化。网络环境下将实现诸如自动分类、自动标引、自动编目及自动搜索网上信息源等档案信息组织与检索的自动化。第二, 档案信息检索简便化。利用者在网络环境下不再受时空的限制, 档案馆的利用者范围不再局限于到档案馆这个特定场所查阅档案的人群, 而是延伸到了社会的各个角落, 包括任何一个通过计算机访问档案馆网页的人。在网络环境下, 除作为凭证依据的档案原件或不能公开的档案需到档案馆借阅外, 其他内容的档案资源可随时随地利用计算机网络进行查阅, 必要时还可进行存储和复制。

(3) 档案信息利用方式的多元化

随着档案信息载体的多样化, 检索方式从传统的手工检索发展到数字检索、全文检索、超文本检索、多媒体检索等, 利用者对信息需求也从单一的纸质档案转向集档案利用服务、参考咨询服务等为一体的多元化服务。随着利用者需求的变化, 传统档案部门服务模式显然不能满足新环境下持续变化的利用者需求。新的环境和服务需求必然要求档案馆(室)的服务有新的组织与管理体制来适应。传统档案部门工作的重点是通过组织馆藏开展借阅利用服务, 而且馆藏建设也偏重于档案实体管理和实体信息的有序化。这种单向线性的实体管理方式, 以馆藏决定服务的内容, 没有考虑到用户需求动态实现的过程, 是一种被动服务。在档案信息流重组后, 档案部门应在关注用户问题解决的过程中提供主动服务。

(4) 档案编研工作的新特点

首先, 档案编研载体和传播媒介从有形实体向虚拟网络转变。网络环境打破了传统档案编研信息的传播必须借助于有形载体的概念。网络出版是一种借助于网络这种无形载体的新型的编研出版形式, 具有成本低、出版周期短、传播速度快、范围广等特点, 改变



了传统出版物只有少数人才能利用的弊端。多种形式的档案可以有机融合。在网络环境下,各种编排软件日趋完善,并较好地融合了网络技术、电子文档处理技术、多媒体技术以及数据库技术,可以实现文字、声音、影像、动画、图片等多种形式并用,可以充分利用互联网上的数字资源,充实汇编内容。

(5) 数字档案馆建设对档案业务流程的影响

数字档案馆有着与传统档案馆迥然不同的流程,馆藏结构由传统的物理现实馆藏进入虚拟现实混合的馆藏,出现了收集、征集、整理、分类、编目、编研一体化的趋势。其相应的业务机构可粗分为三部分:资源建设部门、用户服务部门和技术支持部门。数字档案馆资源建设部门应一并承担起信息资源获取、整序、维护的职责,实行收集、征集、整理、编目为一体的作业流程。而计算机网络的发展大大促进了档案馆信息资源特别是数字资源的使用,档案馆要实现数字资源和纸质文献的一体化服务,建立统一的档案检索方式,实现部分纸质档案的数字化,建立自己的特色数字资源。为了使用户能充分有效地利用数字资源,同样有必要对档案馆信息资源进行重组^①。

总之,信息技术重建了档案馆、档案室的运行环境和作业环境,传统档案业务管理显露出种种弊端。档案部门信息化建设(包括自动化和网络化建设),不仅提供了一个先进的服务平台,而且它还包含着减少重复的作业,资源共享的重要设计原则。档案部门信息系统的集成化管理将原来许多分离的工作实现了在功能上的合并和重组,原有的一些部门设置显得多余和重复。因此,档案业务分析及重构是基于业务规则的档案信息管理系统开发的前期准备,是档案信息管理面向网络化、智能化发展的必要准备。

^① 王雯. 档案馆业务流程重组的思考 [J]. 兰台世界, 2007 (9), 下半月.

4.2.3 档案业务流程重组

“业务流程重组”是由美国的 Michael Hammer 和 Jame Champy 于 20 世纪 90 年代提出的一种管理思想。该管理思想强调以业务流程为改造对象和中心、以关心客户的需求和满意度为目标、对现有的业务流程进行根本的再思考和彻底的再设计,利用先进的制造技术、信息技术以及现代的管理手段、最大限度地实现技术上的功能集成和管理上的职能集成,以打破传统的职能型组织结构,建立全新的过程型组织结构。业务流程重组源于企业业务管理领域,基于企业的奋斗目标和一切工作的归宿——怎样使顾客满意。可以说,档案信息资源的管理和保存目标以及一切档案工作的归宿就是使现存及潜在利用者满意,“业务流程重组”的核心思想同样适用于档案管理业务。通过上一节的档案业务分析不难看出,植根于手工作业环境下的传统档案组织机构、业务流程、管理体系越来越不适应信息时代档案馆发展与服务的要求,所以有必要借鉴企业业务流程重组的思想,重新考虑设计业务流程,从管理理念到组织机构,从业务内容到人员资源,从服务内容到服务方式进行全方位、深层次的变革。

在现代信息技术的支持下,对档案管理中的信息流进行整合,这种整合既不是对原有单向线性的档案实体管理方式细枝末节的修补,也不是利用计算机对原有手工管理程序的简单模拟或单纯的功能合并;而是在面向用户需求的集成管理思想下,在对组织信息流程全盘考虑的基础上对档案信息流的重新定位。档案业务流程重组主要包括组织结构再造和业务内容再造两个方面^①,此外还应重视管理理念和档案工作者的再造。

(1) 组织结构再造

长期以来,档案馆通过对档案实体的手工管理,建立了从收集

^① 陈霞.论网络环境下档案业务流程重组 [J]. 兰台世界,2005 (1).



到编研、利用等规范、机械的业务运作模式。我国档案馆一般都是按功能设置收集、整理保管、编研、接待利用等机构或岗位。随着现代信息技术在档案馆的逐步应用,档案馆的功能不断得以拓展,并形成了一些新的业务环节,而这些新增的环节无法得到原有组织机构的有效支持,造成档案馆运行不畅和功能失调。档案馆必须对其组织机构进行改革,把以传统业务流程为基础的组织结构转变为以现代业务流程为基础的组织结构,把直线型、职能化的组织结构向扁平化、智能化乃至学习型组织的转变,把传统的以收藏和手工进行信息服务为主的组织转变为以用户为主,提供智能化的信息检索和现代化服务手段的组织,重新调配档案馆的业务工作流程,建立起一套以信息网络技术为基础的,与知识经济时代相适应的管理体制。

(2) 业务内容再造

我国档案馆的内部机构大多按“收、管、用”分别设置,形成档案馆服务的整体。分工细、工作流程长,影响了信息的即时服务,难以提供高质量的服务。现代信息技术的发展使原来分立的工作实现了功能上的整合,弱化了分工之间的界限,实现了文档一体化及档案业务工作一体化。同时档案馆的业务内容也发生了重大变化,许多手工操作的业务被计算机取代,原有业务也被调整、淘汰,或推陈出新。档案馆应改变以先进的计算机技术、网络技术去适应或应对传统 workflows 的做法,根据自身的性质、特点、任务和规模大小,分析目前业务流程状况和理想流程状况,对原有业务流程进行重组、简化和糅合。重组后的业务流程应符合档案馆今后发展的需要,充分体现网络环境下以信息服务为中心,以用户需求为导向的思想。

(3) 管理理念和档案工作者再造

电子档案的前身——电子文件的内容和形式是相对独立的,不仅内容易于修改,而且失去了固定形式,从而使电子档案的原始记录性和法律凭证受到巨大的冲击。为确保具有档案价值的电子文件

真实可靠、且适合保存，有关保存和归档的要求应该在电子文件生命周期的开始阶段——设计阶段起做出，应确立前端控制和全程管理的现代文件、档案管理理念。档案工作者只有积极参与相关流程的设计中，使其方法适应于操作过程和过程控制职能的需要，并且利用他们的专业技能和专门知识来定义档案方面的基本概念，才能赋予电子档案的原始性以新的解释，确保电子档案的真实可靠、齐全完整及其可利用性与可理解性^①。档案管理流程重组满足了组织机构实施知识管理的需要，提高了管理效率。为适应这一发展，档案工作者必须改变自身传统的角色定位，提高自身整体素质，将工作重心逐渐从原先对档案实体的保管转向对电子文档处理全过程的监督；并运用自身文档管理方面的隐性知识向各业务部门提供档案管理方面的知识服务；以及通过组建知识库动态解决用户所面临的问题，重视对档案用户反馈信息的分析。档案部门是组织机构的有机组成部分，档案信息流的重组是整个组织信息流程重组的重要环节，因此档案工作者的角色定位很重要，档案工作者应积极参与组织信息管理流程的设计，融入组织的整体文化，协调与其他部门的关系，做好信息的交流与共享。

在信息流程重组后，档案工作者通过积极参与组织机构整个信息流程的设计，将自身的隐性知识（包括档案专业技能、档案知识、价值取向等）融入流程设计的各环节，以定义流程中与档案相关的基本概念，提出文档鉴定、流转、保存等方面的具体要求，确保了整个流程的科学性与高效。可见，重组精简了雷同的工作环节，改善了原有文档管理条块分割、效率低下的状态，使得信息流动更为顺畅，实现了档案工作者对文档处理全过程的监督，加强了组织内部的知识传递和交流；而且信息流重组的过程促进了档案工作者自身所蕴藏的隐性知识的显性化。档案工作者在参与流程再造的同时，通过与其他参与者的讨论与交流实现了各自隐性知识的交流与共享：一方面自身所蕴藏的隐性知识得到了显性化；另一方面

^① 薛匡勇. 新世纪档案职业的走向——中国首届档案博士论坛综述之一 [J]. 湖北档案, 2001 (12): 12-14.



通过交流和共享,他人外化的隐性知识、组织内外部的各种显性知识就内化为个人新的隐性知识,这种知识不断内化、外化的循环促进了知识创新的实现^①。

4.3 档案业务规则的规划

基于业务规则的档案信息智能化管理的核心技术是以结构化的业务规则数据驱动规则引擎执行业务行为,档案业务规则库是存储档案业务规则及规则有关属性的地方,用户录入的规则经编译后保存在规则库中,供规则引擎调用执行。因此,档案业务涉及的相关业务规则应集中统一采集和管理,并且规则的采全率和采准率高低直接关系到档案管理信息系统自动化管理效果的好坏和智能水平的高低。

4.3.1 档案业务规则源

总的来说,档案业务规则主要来源于明文规定与档案业务相关的法律法规、操作规范、规章制度、管理章程、业务标准,档案业务相关人员的需求调查,以及旧档案信息资源管理系统的文档这三个方面。

①我国的档案管理模式从建立之日起就具备一个完善的管理体系,保障顺利、有序开展各级各类档案管理工作,不仅有最具权威的档案专门法律,还根据各地各类档案的特殊性制定了一系列的条例、规章、标准等。如国家颁布的《中华人民共和国档案法》,各省人大批准的各地的档案管理条例,专门针对国家重大建设项目文件的《国家重大建设项目文件归档要求与档案整理规范》等。根据各地各单位和部门档案种类、档案管理特色,以及档案信息资源

^① 孟歆. 探析知识管理思想下的档案管理流程重组 [J]. 湖北档案, 2007 (9) .

管理系统要求，可以将这些法律、条例、规章、标准结构化处理后编译到档案业务规则库中。

②档案业务规则还包括档案业务指导长期实践的经验、组织内经验丰富的档案从业人员的知识总结、员工日常使用的档案专业知识和专门技能等没有形成文字的惯例。对于此类业务规则，可以利用访谈业务人员、开座谈会、问卷调查、跟班作业等方式进行调查总结，并将其文档化、结构化处理后编译到档案业务规则库中，完善档案业务规则体系。

③旧档案信息资源管理系统中的系统分析、需求说明书、数据字典、详细设计、程序代码、操作手册等也包含着大量档案业务规则。将从中经过分析、筛选出的档案业务规则集中到档案业务规则库保存，不仅体现了业务规则集中式管理的策略，而且从旧系统中提炼出的结构化业务规则更接近编译到档案业务规则库前的句法要求，避免了重复劳动，减轻了编译工作量。

要注意的是通过以上三个来源获取的业务规则会存在许多重复或与规则无关的信息，必须通过挖掘整理从中获取有用的业务规则，按规则类型将业务规则进行分门别类，然后再进行业务规则的编译活动。

4.3.2 档案固化规则与动态规则

档案业务规则是对档案业务中某些定义和限制的描述，用于维持档案业务结构、控制或影响档案业务的行为。档案业务规则主要来源于明文规定与档案业务相关的法律法规、操作规范、规章制度、管理章程、业务标准，档案业务相关人员的经验总结等。这些规则中有些是适用于各级各类档案部门的业务活动，是固定不变的，如《中华人民共和国档案法》、《中华人民共和国档案法实施办法》、《中华人民共和国保守国家秘密法》等法律法规和各类国家标准，各个档案部门信息管理系统所依托的业务规则库中都应包含这类规则。此外，各单位还自有一套规章制度、操作规范，根据自身情况所拥有的档案种类、归档范围、保管期限表、工作经验等



也是不同的，我们称这些规则为动态规则。

在档案业务规则库的建设过程中要注意区分档案固化规则和动态规则，依照这两种规则的特点实施不同的添加方案，并且在规则管理系统里对其的权限设置也不同。档案固化规则一定时期内是不会改变的，可以成套固定存储于规则库中，每个单位不必再花费时间去另建此类规则，而且由于此类规则具有权威性，不能随意修改。档案动态规则是需要本档案管理系统自己添加入库的，档案业务人员或者系统管理人员可以使用受限中文业务规则语言通过规则管理平台界面将本单位使用的规则添加进规则库中，此类规则的权限设置较为宽松，一般可随本单位档案业务实际情况新建、删除或修改规则。

4.4 基于业务规则的档案信息资源管理模型

4.4.1 现有档案信息资源管理系统（AMIS）的管理模式

现有的较科学的档案信息资源管理系统包括输入子系统、存储加工子系统、输出子系统、反馈子系统四大部分。输入子系统的主要功能是收集和鉴选档案，存储加工子系统的主要功能是对档案信息进行揭示、加工和组织，进行系统整理，便于检索和利用，输出子系统的主要功能是以借阅、检索、咨询等方式向社会提供档案信息服务，反馈子系统的主要功能是通过统计和用户研究等方式掌握系统运行及用户需求状况，使档案管理系统不断完善和优化。其简要结构图如图 4-3。

现有的档案信息资源管理系统（AMIS）的归档管理模式较传统管理模式并无根本性的突破，突出问题表现在：对归档部门的业务指导方法不得力，兼职档案人员的管理不到位；归档文件的质量管理处于薄弱状态，无法对以下问题进行全面系统的实时监控和综合评价，如各门类载体材料收集是否齐全完整、内容是否真实有

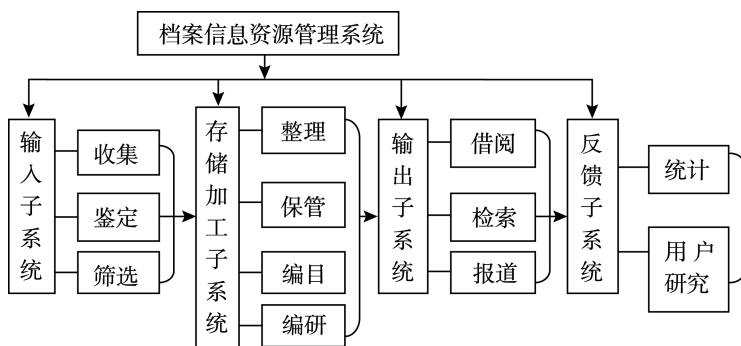


图 4-3 档案信息资源管理系统结构图

效，整理是否规范，归档是否及时，以及各部门的归档难度、归档数量等。即便是各类综合管理达标升级的考核程序，也只能对已归档的材料进行抽查。显然，传统的归档模式对于如何提高归档效率，实现对文件形成与归档的过程、时间和质量的实时监控，以及建立对归档人员激励评价的科学管理机制，仍缺乏科学有效的解决方法。

4.4.2 基于业务规则技术的档案信息资源管理模型

在现有档案信息资源管理系统存在一些突出问题的背景下，本研究将业务规则技术应用到新型档案管理信息系统中，在档案信息资源管理系统中加入“规则档案平台”。“规则档案平台”是一个在 Internet/Intranet 环境下基于微软 .NET 平台、三层结构、SOA 架构的 Web 应用，是一套引入人工智能专家系统的概念，构建的完整业务规则管理系统解决方案。SOA 系指面向服务的体系结构 (Service-Oriented Architecture)，它是一个组件模型，将应用程序的不同功能单元通过这些服务之间定义良好的接口和契约联系起来，使构建在各种这样的系统中的服务可以以一种统一和通用的方式进行交互。“规则档案平台”以档案业务规则库 (Archives Business Rules Database，简称 ABR) 为核心，通过规则推理引擎 RM，利用

档案管理过程中形成的不同业务数据，优化档案工作的多个环节，包括档案的整理分类、鉴定、保管和利用等，并根据单位内部 Intranet，搭建出不同的档案业务分支管理系统。推理引擎就是应用系统中用来完成推理功能的模块，一般包括调度器、执行器和一致性协调器三部分。调度器控制整个推理流程，使其按照一定的流程进行下去；执行器执行调度器所选定的动作，负责读取知识库中的知识和全局黑板中的信息；一致性协调器保持推理中间结果的一致性，用于反向推理中。

利用档案业务规则管理系统进行优化的档案轨迹的流程可简要描述为：首先，基于档案业务规则库与推理引擎建立档案信息雷达，利用雷达对海量的档案信息资源进行全面收集；其次，档案业务规则库对自然语言描述的归档范围进行分词，中文分词的方法主要包括基于字符串匹配的分词、基于统计规律的分词和基于语义理解的分词等，分词之后生成词群系列，对档案信息资源进行科学的分类整理；再次，ABR 对归档文件中的档案信息进行分词，提取关键词，并与已生成的若干词群一一进行匹配，从而能够确定该条档案信息的保管期限，完成鉴定；最后，完成档案信息资源的有效保存和合理利用。利用档案业务规则管理系统进行优化的档案工作的流程如图 4-4。

4.4.3 基于业务规则的档案信息资源管理系统的特点

首先，从业务的角度看，作为具体应用领域的业务软件，基于业务规则的档案信息资源管理系统具有一般的管理信息系统的特点。即基于业务规则的档案信息资源管理系统建立和实施，实现了档案管理部门的管理信息化和办公自动化；提高档案管理与科研利用的工作效率和质量；极大地方便了档案部门管理人员，也方便了其他部门和领域的档案利用人员。

其次，从 IT 的角度看，基于业务规则方法构建的档案信息资源管理系统具有良好的可扩展性和可维护性。

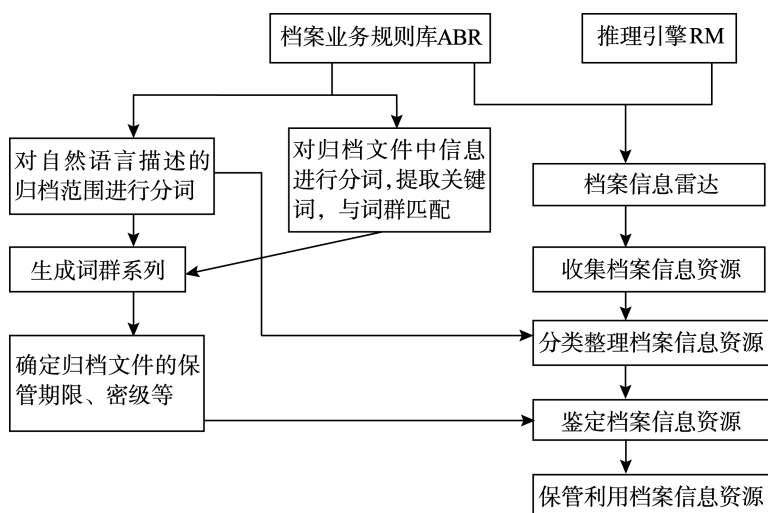


图 4-4 档案业务规则管理系统优化后档案工作流程图

(1) 良好的可扩展性

可扩展性的好坏体现了一个软件系统适应变化的能力。可扩展性的实现主要可以通过组件之间的解耦、公共功能的集中化和组件的重用来实现。基于业务规则方法构建的科技管理信息系统可以从以下方面获得良好的可扩展性：

①通过将各个模块的业务逻辑抽取出来形成独立的业务规则模块，可使模块之间减少不必要的逻辑关联，而只需要专注于解决本模块所要完成的计算任务。

②业务规则模块的形成，同时可以将业务逻辑进行集中处理，使原来散布于业务逻辑层的各个组件中的业务逻辑部分通过业务规则模块进行统一处理，提高了系统的扩展性。

③为构建科技管理信息系统形成的业务规则，作为科技管理部门的政策和管理策略，可以作为将来构建其他业务系统的策略储备进行良好的复用。由于业务规则只包含业务知识，它比代码级的模块复用更加直观和简单，而且减少了错误的发生。



(2) 合理的维护成本

①业务规则方法的主要优势之一就是它所具有的灵活处理动态可变业务逻辑的能力。通过将业务逻辑抽取成可以作为“数据”进行管理的业务规则模块,当出于政策和管理的要求对管理系统中的业务逻辑进行更改时,将不再反映为程序代码的重新编写、编译、测试和部署,而是直接通过修改“数据”形式的业务规则即可完成。这显然可以极大地降低维护成本。

②通过为科技管理部门制定符合本部门策略的领域语言(Domain Specific Language, DSL),业务规则甚至可以由业务人员自己根据需要制定即可,几乎不需要IT人员的参与。这可以使新的政策和策略近乎实时地在业务系统中得到反映,避免了由于业务系统修改造成的政策和策略的延迟执行问题。

4.5 基于业务规则的档案信息资源管理标准规范体系

实现档案信息资源共享,标准化正是减少直至消除档案信息化建设与运行过程中无序状态与重复加工等现象以达到规范化、系列化与统一化,从而促进档案信息的交流与共享,实现档案管理科学化与现代化的重要手段。档案信息资源管理标准体系则是指将规范档案信息资源管理及管理信息系统建设各个流程、环节的标准进行有机统一,从而形成的一个具有内在联系、功能得到最大限度发挥的标准有机整体。基于业务规则的档案信息资源管理标准体系包括确保档案信息管理智能化建设项目目标快速有效完成所需的现有国内外相关档案业务标准、项目质量标准、软件工程标准及安全标准。

4.5.1 档案业务标准

档案业务标准是档案信息资源共享体系建设的重要保障。1983年,国家档案局正式组建了全国档案工作标准化领导小组,负责组织和管理工作。1989年,国家档案局在综合科教司(1994年改为政策法规研究司)下设立了法规标准化处,具体承担全国档案工作标准化的行政管理工作,并开始筹备成立档案工作标准化技术委员会。1991年,全国档案工作标准化技术委员会(以下简称档标会)正式成立,法规标准化处同时是档标会的秘书处。

《全国档案信息化建设实施纲要》将档案信息化标准规范体系建设列入了档案信息化建设的六大内容之一,档案信息化基础设施建设、信息资源建设、应用系统建设和信息安全保障体系建设都需要相应标准进行规范。近年来,国家档案局和全国各档案部门根据档案信息化建设的需要,积极开展国标、行标以及地方标准的制定与修订工作。档案信息在规章方面,发布了《档案管理软件功能要求暂行规定》、《电子公文归档管理暂行办法》。在国家标准方面,《电子文件归档与管理规范》(GB/T 18894-2002)于2003年5月实施。在档案行业标准方面,《纸质档案数字化技术规范》(DA/T 31-2005)、《公务电子邮件归档与管理规则》(DA/T 32-2005)、《明清档案目录中心数据采集标准、明清档案电子档案著录标准机读目录数据交换格式》(DA/T33-2005)于2005年9月实施。《全国革命历史档案数据采集标准》、《民国档案目录中心数据采集标准》、《纸质档案数字化技术规范》、《公务类电子邮件归档管理规则》、《电子公文归档与管理办法》等标准规范,这些标准的制定和实施对档案信息化建设起到了积极的促进作用。

档案信息化标准规范的制定范围和内容要覆盖新的档案业务流程和具体工作内容,使其每项有标准可依^①。具体来说,档案工作

^① 路宜庆.论企业档案信息化标准规范体系的建设[J].山东档案,2008(2).



标准规范主要包括以下几点：

①档案分类规则及编号方案。结合电子文件与纸质文件并存的实际，以简化整理程序、提高检索质量、电子档案管理优先、满足网上原文查询需要的原则，编制适应档案信息化管理条件下新的档案分类规则及编号方案。编制的主要内容有：档案分类原则、类目设置、类目标识符号与标识方法、编号方案等。

②档案目录数据采集规范。由于电子文件的目录信息在线移交归档过程自动提取，档案信息载体的增多、档案类别的增多、档案利用方式的变化、档案信息管理程度的细化，需要对各类档案目录数据的采集进行规范。编制的内容主要可以分为文书档案、特殊载体档案、专门档案等几大类，每类档案的文件级和案卷级目录数据库都应明确字段名称、字段名、字段类型、字段长度、备注、著录项目细则等。为了规范材料移交单位（部门）向档案部门提供电子版档案移交目录，以减少档案目录信息采集的重复劳动，应规定各类档案移交电子目录格式，如：科技档案移交电子目录格式、基建类档案移交电子目录格式、声像类档案移交电子目录格式、OA公文在线归档移交目录提交格式等。

③OA 电子公文在线归档规程。其目的是规范在线归档过程中各程序的行为。编制的主要内容有：电子公文形成过程控制（包括起草格式、拟文单填选项目、痕迹留存、发文单填选项目等）、来文在进入办公自动化系统时的操作（包括收文单填选项目、纸质收文扫描模式及质量、图像存储要求等）、网上移交操作、网上接收归档操作等。

④档案数字化技术规范。其目的是规范纸质、磁性、实物等载体档案数字化过程中的行为。编制的主要内容有：档案数字化基本的要求、各类数字化加工的范围、档案数字化工作流程、纸质档案的数字化加工（包括文书、基建、照片等类档案）、音频档案数字化加工、视频档案数字化加工、实体档案数字化加工、数字化成果管理等。

⑤电子文件元数据标准。编制的主要内容有：电子文件元数据的概念、电子文件元数据表（包括元数据类型、元数据元素、是

否为必选著录项、说明等)。元数据类型包括标识性元数据、背景信息元数据、内容元数据、期限与条件元数据、保存历史元数据、结构信息元数据等。标识性元数据包括全宗号、分类号、文件编号等,背景信息元数据包括文件形成单位、文件责任者、文件形成时间、接收单位、接收人、接收时间、关联、语种、使用日志等,内容元数据包括文件标题、主题词、内容描述等,期限与条件元数据包括密级、用户权限、保管期限、处置授权机构、处置责任者、处置时间等,保存历史元数据包括保存日志,结构信息元数据包括载体类型、载体规格、载体数量、载体单位、稿本、文件类型、数据格式、存储位置、加密方式、应用环境、文件质量等。

⑥数字档案馆网信息发布规程。其目的是规范网站档案目录和原文、电子档务、现行文件、工作动态等信息的发布行为。编制的主要内容有:网站管理组织体系、发布信息内容范围、信息的发布审批、信息上传发布、系统维护与信息卸载、信息归档等。

⑦管理性标准。包括两个方面:一是对人的管理性标准,主要是指对与档案信息化建设相关的人员进行管理的标准,包括档案馆工作人员管理标准、软件设计人员管理标准、用户管理标准、用户角色控制标准、用户权限审批标准等;二是对物的管理性标准,主要是指对与档案信息化建设相关的物质进行管理的标准,主要包括档案信息化设备管理标准和档案信息化成果管理标准等^①。

4.5.2 项目质量标准

(1) 国际质量标准 ISO9001: 2000

20世纪80年代末期,随着世界经济领域的相互合作和竞争日益增强,产品质量成为各国和各公司竞争成功与否的关键因素,用户不仅对产品质量进行测试,而且也对产品生产企业的质量管理体

^① 赵传玉. 论档案信息化标准体系的构建 [J]. 湖北档案, 2009 (1-2).



系进行评估与检验，从根本上获得对企业的产品进行质量管理体系上的保证，使用户建立起对企业产品质量、质量的稳定性与生产水平的充分信心。随后，这一做法逐渐成为国际贸易和经济合作中的通常做法。为了进一步适应国际经济合作发展和进出口贸易的需要，保证公平公开的竞争，规范各国质量管理和质量保证行为并使之有序化、标准化，国际标准化组织（International Standard Organization, ISO）于1987年发布了ISO9000质量管理和质量保证系列标准，1990年以来又发布了一些与ISO9000系列标准配套的有关标准或标准草案。

ISO9000质量系列标准，作为一整套先进的国际质量标准，是贯穿单位全员、全过程的质量管理的体系标准，目前已经在世界许多国家和地区的企事业单位、机关、军队中被广泛采用，最新版的ISO9001:2000质量标准，适用性更广，而且管理理念也能较好地解决企事业单位工作行为的过程、行为结果的评价及行为实施后的纠正等问题。

ISO9000系列标准是一个文件化的质量体系，标准规定：质量体系中应制定文件化的质量文件和质量记录的管理程序，即对质量文件和记录进行标识、收集、编目、归档、存储、保管等工作，同时，还要把“记录”视为证实性文件，通过对这些证实性文件的审核，企业可以监视生产或者服务的全过程，从中发现问题，进而改进生产和服务过程，达到自我完善和自我改进并最终保证和提高产品质量或者服务水平的目的。ISO9000系列标准强调对产品质量形成过程的控制，需要相应的书面控制程序及执行程序的记录，对这些程序和记录的管理及存储，应当属于档案管理的业务范围。具体来讲，ISO9000质量体系标准的要求，涉及档案管理业务问题的主要有三个：文件及资料控制程序，产品标识、可追溯性和质量记录的控制。

根据ISO质量体系与档案管理的要求，企业应采取以下措施，以实现ISO9000质量体系与档案管理体系的有机结合。

①以ISO9000质量认证为动力，促进档案工作。在宣传贯彻质量认证、ISO9000系列标准活动中，将《档案法》及其他档案管理

法规制度的宣传贯彻有机地结合起来。应该认识到：ISO 质量体系是对企业质量体系的一个总体原则性的要求，这也是 ISO9000 质量体系灵活性的体现。明确 ISO9000 质量体系与档案工作的区别与联系，结合自身的特点，建立适合自己特点的合理的档案管理办法。

②按 ISO9000 系列标准要求，补充、修订、完善了部分档案管理制度，并纳入质量体系管理运行。我们既按 ISO9000 系列标准要求，新编制了质量体系文件，又修正、补充完善了原档案专业制度、标准，丰富和发展了档案法规体系，实现以档案工作为基础，保障 ISO9000 质量认证。

③充分发挥档案工作的优势，明确档案职能部门在质量管理体系中的作用和地位，加强培训工作，更好地贯彻 ISO9000 质量体系标准。

④ISO9000 质量管理的实质是将企业的质量目标通过全面、规范的可操作性、系统性的文件和指标体系形式确定下来，并能够运用这一体系进行质量管理问题的自我发现、自我改进、自我完善。将 ISO9000 质量管理的精神实质融会贯通，可以在档案管理中参考应用，使档案工作具有实现发展目标的针对性，发挥更多的作用，也使 ISO9000 质量管理在更广泛、更全面的意义上为档案的管理和发展服务。

(2) 能力成熟度模型集成 (CMMI)

很多不成熟的软件组织会面临许多问题，例如：需求经常得不到一致的理解，进度和预算经常得不到保障，项目的进展无法度量，产品的内容没有跟踪和控制，版本混乱，工程活动没有标准，实施不一致，开发团队没有经过培训，相互间不协调等。

能力成熟度模型集成 (Capability Maturity Model Integration, CMMI) 是由美国卡内基-梅隆大学软件工程研究所 (SEI) 发布的技术报告，它是软件过程改进方面得到国际认可的标准，为国内外众多大、中型软件/IT 公司广泛推行。CMMI 的提出主要是为了解决在项目开发中需要用到多个 CMM 模型的问题，为软件组织改进和提高过程能力提供指南，是评估组织当前开发管理状况的标尺，

它可以用于指导跨项目、部门或者整个组织的过程改进。它主要结合了 SW-CMM（用于软件的能力成熟度模型）、SECM（系统工程能力模型）、IPD-CMM（集成产品开发能力成熟度模型），将它们融合到一个统一的改进框架内，为组织提供了在企业范围内进行过程改进的模型。

CMMI 内容分为“Required”（必需的）、“Expected”（期望的）、“Informative”（提供信息的）三个级别，来衡量模型的质量和作用。其中，最重要的是“必需的”级别，是模型和过程改进的基础。第二级别“期望的”在过程改进中起到主要作用，但是某些情况不是必需的可能不会出现在成功的组织模型中。“提供信息的”构成了模型的主要部分，为过程改进提供了有用的指导，在许多情况下它们对需要的和期望的构件做了进一步说明。

4.5.3 软件工程标准

在开发一个软件时，需要有不同层次、不同分工的人员相互配合，在开发项目的各个部分以及各开发阶段之间也都存在着许多联系和衔接问题。如何把这些错综复杂的关系协调好，需要有一系列统一的约束和规定。在软件开发项目取得阶段成果或最后完成时，还需要进行阶段评审和验收测试。投入运行的软件，其维护工作中遇到的问题与开发工作有着密切的关系。软件的管理工作则渗透到软件生存期的每一个环节。基于业务规则的档案信息资源管理系统的设计和实施也是一种软件设计和开发行为，软件标准化管理是实现系统顺利运行的必要手段。所以，在档案信息系统开发前需要设定软件工程实施标准，从而为软件提供统一的行为规范和衡量准则，使得各工作都能有章可循。主要依据的软件工程标准有以下几种：

- ①GB/T 9385-2008《计算机软件需求规格说明规范》；
- ②GB/T 8566-2007《信息技术 软件生存周期过程》；
- ③GB/T 16260-2006《软件工程 产品质量系列标准》；
- ④GB/T 18492-2001《信息技术系统及软件完整性级别》；

- ⑤GB/T 20916-2007 《中文办公软件文档格式规范》;
- ⑥GB/T 16680-1996 《软件文档管理指南》;
- ⑦GB/T 9386-2008 《计算机软件测试文档编制规范》;
- ⑧GB/T 14394-2008 《计算机软件可靠性和可维护性管理》;
- ⑨GB/T 14394-2008 《计算机软件测试规范》。

4.5.4 安全标准

国际标准化组织 (ISO) 将信息安全 (Information Security) 定义为“为数据处理系统建立和采取的技术和管理的安全保护, 保护计算机硬件、软件和数据不因偶然和恶意的原因而遭到破坏、更改和显露”。信息安全不仅是保证信息的机密性、完整性、可用性、可控性和可靠性, 并且从主机的安全技术发展到网络体系结构的安全, 从单一层次的安全发展到多层次的立体安全。从管理方面来讲, 实现信息安全主要是健全组织内部的信息安全管理制度, 以防止因为内部人员的误操作或思想麻痹, 没有足够的信息安全知识而引起的严重后果。信息安全管理的原则之一就是规范化、系统化, 如何在信息安全管理实践中落实这一原则, 需要相应的信息安全管理标准。

(1) 信息安全国际标准: 《信息技术 安全技术信息安全管理准则法规》(ISO/IEC17799)

国际标准化组织的 ISO17799 (信息安全管理体系) 就是国际上通行的构建信息安全管理体系的标准。

BS7799 标准是英国标准协会 (BSI) 制定的国际上具有代表性的信息安全管理体系标准。该标准包括两个部分: BS7799-1: 1999 《信息安全管理实施细则》和 BS7799-2: 1999 《信息安全管理体系规范》, 其中, BS7799-1 标准目前已正式转换成 ISO 国际标准, 即 ISO17799 《信息安全管理体系实施指南》, 并于 2000 年 12 月 1 日颁布。该标准综合了信息安全管理方面的优秀控制措施, 为组织在信息安全方面提供建议性指南。因此该标准不是



认证标准，但组织在建立和实施信息安全管理体系时，可考虑采取该标准建议性的措施。BS7799-2 标准目前正在转换成 ISO 国际标准的过程中。BS7799-2 标准要求主要用于对组织进行信息安全管理体系的认证，因此组织在建立信息安全管理体系时，必须考虑满足 BS7799-2 的要求。ISO/IEC 17799 的目的并不是告诉用户有关“怎么做”的细节，它所阐述的主题是安全策略和优秀的、具有普遍意义的安全操作。该标准特别声明，它是“制定一个机构自己的标准时的出发点”，并不是说它所包含的所有方针和控制策略都是放之四海而皆准的，也不是其他未列出的便不再要求。标准主要讨论了如下的主题：建立机构的安全策略、机构的安全基础设施、资产分类和控制、人员安全、物理与环境安全、通信与操作管理、访问控制、系统开发和维护、业务连续性管理、遵循性。采用 ISO/IEC 17799 标准建立起来的信息安全管理体系（ISMS），是建立在系统、全面、科学的安全风险评估之上，是一个系统化、文件化、程序化、科学化的管理体系。它体现预防控制为主思想，强调遵守国家有关信息安全的法律、法规及其他要求，强调全过程和动态控制，本着成本费用与风险平衡的原则选择安全控制方式，保护组织所拥有的关键信息资产，确保信息的保密性、完整性、可用性，对网络环境下的信息安全管理无疑具有十分重要的意义。

（2）信息安全国家标准：《计算机信息系统安全保护等级划分准则》（GB17859-1999）^①

《计算机信息系统安全保护等级划分准则》（GB17859-1999）适用计算机信息系统安全保护技术能力等级的划分。计算机信息系统安全保护能力随着安全保护等级的增高，逐渐增强。本标准规定了计算机系统安全保护能力的五个等级，即用户自主保护级、系统审计保护级、安全标记保护级、结构化保护级和访问验

^① 中华人民共和国国家标准：计算机信息系统安全保护等级划分准则 GB 17859-1999 [R].

证保护级。

①第一级：用户自主保护级。本级的计算机信息系统可信计算基通过隔离用户与数据，使用户具备自主安全保护的能力。它具有多种形式的控制能力，对用户实施访问控制，即为用户提供可行的手段，保护用户和用户组信息，避免其他用户对数据的非法读写与破坏。

②第二级：系统审计保护级。与用户自主保护级相比，本级的计算机信息系统可信计算基实施了粒度更细的自主访问控制，它通过登录规程、审计安全性相关事件和隔离资源，使用户对自己的行为负责。

③第三级：安全标记保护级。本级的计算机信息系统可信计算基具有系统审计保护级所有功能。此外，还提供有关安全策略模型、数据标记以及主体对客体强制访问控制的非形式化描述；具有准确地标记输出信息的能力；消除通过测试发现的任何错误。

④第四级：结构化保护级。本级的计算机信息系统可信计算基建立于一个明确定义的形式化安全策略模型之上，它要求将第三级系统中的自主和强制访问控制扩展到所有主体与客体。此外，还要考虑隐蔽通道。本级的计算机信息系统可信计算基必须结构化为关键保护元素和非关键保护元素。计算机信息系统可信计算基的接口也必须明确定义，使其设计与实现能经受更充分的测试和更完整的复审。加强了鉴别机制；支持系统管理员和操作员的职能；提供可信设施管理；增强了配置管理控制。系统具有相当的抗渗透能力。

⑤第五级：访问验证保护级。本级的计算机信息系统可信计算基满足访问监控器需求。访问监控器仲裁主体对客体的全部访问。访问监控器本身是抗篡改的；必须足够小，能够分析和测试。为了满足访问监控器需求，计算机信息系统可信计算基在其构造时，排除那些对实施安全策略来说并非必要的代码；在设计和实现时，从系统工程角度将其复杂性降低到最低程度。支持安全管理员职能；扩充审计机制，当发生与安全相关的事件时发出信号；提供系统恢



复机制。系统具有很高的抗渗透能力。

4.6 基于业务规则的档案信息资源管理关键技术

4.6.1 面向服务的架构 SOA

面向服务的架构 (Service-Oriented Architecture, SOA), 其目标主要针对分布式系统的体系结构, 由一系列相互交互的服务组成, 每种服务都能提供对该定义功能集的访问。SOA 不仅仅只是一套构架, 其更像是一套设计思想、方法, 为解决客户所面临的业务敏捷性问题提供了一套新的解决方法。SOA 整体上设计和实现为一系列相互交互的服务, 这种将业务功能实现为服务的方法可以增强系统的灵活性, 系统通过增加新的服务来实现演化。SOA 定义了系统由哪些服务组成, 描述了服务之间的交互, 并将服务映射到一个或多个具体技术的实现。总体来讲, SOA 的优势在于它的高可复用性、灵活性, 以及更好的扩展性和可用性^①。

SOA 中定义了不同的角色, 这些角色包括: 服务使用者、服务提供者、服务注册中心。服务使用者是一个应用程序、一个软件模块或需要服务的另一个服务。它首先发起对注册中心中的服务查询, 其次通过查询结果来传输绑定服务并且执行服务功能。服务提供者是一个可通过网络寻址的可执行实体, 它接受和执行来自服务使用者的请求。它必须将自己的服务和接口契约发布到服务注册中心, 以便服务使用者可以发现和访问该服务。服务注册中心是服务发现的支持者, 它包含一个可用服务的存储库, 并允许服务使用者查找服务提供者接口。

^① 叶钰, 应时, 李伟斋, 张韬. 面向服务体系结构及其系统构建研究[J]. 计算机应用研究, 2005 (2): 32-34.

SOA 中定义了不同的操作, 这些操作包括: 发布、发现、绑定和调用。为了使服务可访问, 服务提供者需要发布特定的服务描述以使服务使用者可以发现和调用它。服务请求者为了定位服务, 首先要查询服务注册中心来找到满足其标准的服务, 在查询到服务描述之后, 服务使用者继续根据服务描述中的信息来调用服务。

SOA 把实现定义成两部分, 一部分集中于体系结构的功能性方面, 而另一部分集中于体系结构的服务质量方面。在服务质量方面, 服务使用者需要通过一定的策略来访问服务, 策略实际上它是一组条件和规则, 在这些条件和规则之下, 服务提供者使服务对于使用者可用。策略既有功能性方面, 也有与服务质量有关的方面, 因此, 在功能和服务质量两个区中都有策略功能。安全性是规则集, 可以应用于调用服务的服务使用者的身份验证、授权和访问控制。传输是属性集, 可以应用于一组服务, 以提供一致的结果, 它非常类似应用系统中的事物控制, 例如, 如果服务使用者要使用一组服务来完成某项业务功能, 则所使用的相关服务必须都完成, 或者没有一个完成。^①

SOA 参考架构由消费者 (Consumer)、业务过程 (Business Process)、服务 (Service)、服务组件 (Service Component) 和业务系统 (Operational System) 五部分组成, 它提供了一个从“服务”视角解决问题的思路和方法。

4.6.2 人工智能技术

人工智能 (Artificial Intelligence) 是一个信息知识处理系统, 涉及信息、数据和知识的获取、确认、表示、推理、解释和验证。通过研究人的思维过程, 并尽可能地利用机器体现或模拟人的智能行为, 如理解经验并从中学习、辨别模糊或互相矛盾的信息, 快速而成功地对新环境作出反应, 在解决问题时使用推理进行有效的推

^① 付玲. 基于 SOA 的吉林银行档案管理系统的设计与实现 [D]. 长春: 吉林大学硕士学位论文, 2009.



导、处理复杂的情况、应用知识控制环境等^①。美国麻省理工学院的温斯顿教授说：人工智能就是研究如何使计算机去做过去只有人才可以做的智能工作。人工智能在以电子计算机为工具的基础上，信息、数据和知识处理方面具有的优势已被人们所广泛认知，但并没有完全被开发利用。

目前，人工智能领域内最基本和最常用的方法有搜索算法、知识表示方法、不确定性推理方法、机器学习方法、知识处理系统、自然语言理解、高级搜索方法、专家知识库系统、人工神经网络以及模糊系统等。知识表示是人工智能的基本问题之一，常用的知识表示方法有：逻辑表示法、产生式表示法、语义网络表示法和框架表示法等。问题求解中的自动推理是知识的使用过程，推理过程一般可分为演绎推理和非演绎推理。机器学习是指在一定的知识表示意义下获取新知识的过程，按照学习机制的不同，主要有归纳学习、分析学习、连接机制学习和遗传学习等。知识处理系统主要由知识库和推理机组成。知识库存储系统所需要的知识，推理机规定使用知识的基本方法和策略，推理过程中为记录结果或通信需设数据库或采用黑板机制。如果在知识库中存储的是某一领域的专家知识，则这样的知识系统称为专家系统^②。

人工智能具有学习、推理、记忆、决策的能力，使计算机表现出与人类相似的智能行为。面对信息、数据和知识过载、非结构化和隐性知识的挑战，在档案管理系统中引入人工智能，将信息、数据和知识的获取、组织、创建、共享和应用的管理过程智能化，能够最快地把最精确、个性化的信息、数据和知识传递给最需要它的人，实现组织创造价值的目标，会极大地提高组织的竞争力，极大地提升档案管理水平^③。还可以加强人工智能专家系统在档案管理

① 邵军力，张景，魏长华．人工智能基础 [M]．北京：电子工业出版社，2000．

② 屈秀兰．应用人工智能技术开发利用企业档案资源的探讨 [J]．兰台世界，2008（2），下半月．

③ 屈秀兰．应用人工智能技术开发利用企业档案资源的探讨 [J]．兰台世界，2008（2），下半月．

中的应用,将语音识别技术应用于档案管理领域^①。“规则档案平台”采用 ILOG 公司的产品 ILOG Rules for .NET 作为人工智能的开发平台。

4.6.3 网络信息雷达

信息雷达是在已知数据样本的基础上,通过归纳学习、机器学习、统计分析等方法得到数据对象间的内在特性,据此采用信息过滤技术在网络中提取用户感兴趣的信息,获得更高层次的知识和规律的智能搜索引擎。其主要功能是实时监控和采集目标网站的最新事件和相关报道,对采集到的信息进行过滤和自动分类处理,将分类结果及时发布出来,实现统一的信息导航功能,同时提供包括全文检索、日期(范围)检索、标题检索、URL 检索等在内的全方位信息查询手段,最后根据用户个性化特征将信息主动推送给用户^②。

网络爬虫是与网络信息雷达技术原理相近的一种重要技术。其定义有广义和狭义之分。狭义上是指遵循标准的 http 协议利用超链接和 Web 文档检索的方法遍历万维网信息空间的软件程序;而广义的定义则是所有能遵循 http 协议检索 Web 文档的软件都称之为网络爬虫。Spider 是一个功能很强的自动提取网页的程序,它为搜索引擎从万维网上下载网页,是搜索引擎的重要组成,它通过请求站点上的 HTML 文档访问某一站点。它遍历 Web 空间,不断从一个站点移动到另一个站点,自动建立索引,并加入到网页数据库中。网络爬虫进入某个超级文本时,它利用 HTML 语言的标记结构来搜索信息及获取指向其他超级文本的 URL 地址,可以完全不

① 郭小磊. 浅谈人工智能技术在档案管理中的应用和加强应用的必要性 [J]. 才智, 2009 (18).

② 曹龄兮, 李建华, 娄悦. 基于多次权重鉴别的信息雷达及应用 [J]. 现代计算机, 2006 (2).



依赖用户干预实现网络上的自动爬行和搜索^①。

目前，爬虫在抓取网页时，一般采用两种策略：广度优先和深度优先。所谓广度优先是指爬虫沿着树的宽度方向遍历，直到抓完起始网页中链接的所有网页，然后再选择其中的一个链接网页，继续这个过程。此方法可以让爬虫并行运行，提高抓取速度。所谓深度优先是指爬虫沿着树的纵深方向遍历图中没有被访问过的节点。由于深度优先是一个递归的过程，爬虫程序在执行的时候就要大量消耗计算机内存资源，在很多情况下都会导致爬虫的陷入问题，甚至死机。另外，递归与多线程无法兼容（因为多线程允许一次运行多个任务，每个并行线程都有自己的堆栈。而当一个方法调用它自身时，它们需要使用同一个堆栈）。

^① 谢国强，蓝立新．基于 Web 的网络爬虫技术研究 [J]．科技文汇，2008（4）．

5. 档案业务规则管理系统

本章从档案业务规则管理系统概述入手，重点突出了档案业务规则管理系统的开发步骤，包括系统现状需求设计和可行性分析、系统逻辑性设计和机构设计、系统部署和实施、系统维护和评价等环节。接着，本章详细介绍了档案业务规则管理系统在档案收集、分类整理和著录、鉴定、保存以及推动档案工作自动化、智能化等方面的诸多功能，论述了档案业务规则管理系统结构中所包含的文件信息导入层、文件信息挖掘层和档案信息表现层这三大部分内容，深入剖析了档案业务规则管理系统的集成开发环境即档案业务规则编辑器、管理器和编译器，构建了一个关于档案业务规则的对象模型——DURA 模型，并阐述了该模型各个构成要素及相互之间的关系、档案业务规则引擎等方面的内容。最后，本章探讨了基于角色的档案业务规则权限管理，对规则权限管理过程中的角色定义、角色授权和角色分配等内容进行分析，并简要论述了档案业务规则管理系统的安全管理的必要性及其涉及的主要内容，以保证系统安全、稳定运转，推动档案工作发展。

5.1 档案业务规则管理系统简介

业务规则管理系统（Business Rules Management System，简称 BRMS）是实现业务规则的建立、管理、维护、部署和运行的信息系统。BRMS 能够利用规则引擎把传统的应用程序或参数化实现的

业务策略从系统中提取出来，被提取的业务规则将作为一种可管理的对象存储在规则库中。BRMS 是构建在操作系统、数据管理以及事务处理之上的高级管理层次，在应用系统中的地位与数据库管理系统（DBMS）类似，处于比较基础的位置，是其他高端应用的基础。

在现有档案信息资源管理系统存在一些突出问题的背景下，“电子政务环境下基于业务规则的档案信息资源管理支撑平台”（以下简称“规则档案平台”）应运而生。它是一个在 Internet/Intranet 环境下基于微软 .NET 平台、三层结构、SOA 架构的 Web 应用，是一套引入人工智能专家系统的概念而构建的完整业务规则管理系统解决方案。SOA 系指面向服务的体系结构（Service-Oriented Architecture），它是一个组件模型，将应用程序的不同功能单元通过这些服务之间定义良好的接口和契约联系起来，使构建在各种系统中的服务可以一种统一和通用的方式进行交互。“规则档案平台”以档案业务规则库（Archives Business Rules Database，简称 ABRD）为核心，通过规则推理引擎 RM，利用档案管理过程中形成的不同业务数据，优化档案工作的多个环节，包括档案的整理分类、鉴定、保管和利用等，并根据单位内部 Intranet，搭建出不同的档案业务分支管理系统。推理引擎就是应用系统中用来完成推理功能的模块，一般包括调度器、执行器和一致性协调器三部分。调度器控制整个推理流程，使其按照一定的流程进行下去；执行器执行调度器所选定的动作，负责读取知识库中的知识和全局黑板中的信息；一致性协调器保持推理中间结果的一致性，用于反向推理中。

档案业务规则管理系统就是“规则档案平台”中非常重要的组成部分。BRMS 是“规则档案平台”系统服务堆栈中的一部分，提高了 SOA 的便捷性，利用它建成的 ABR 及档案信息数据（Archives Information Data，简称 AID），将使技术人员可较快地搭建与档案配套的其他业务系统。

5.2 系统开发方法和开发步骤

5.2.1 系统分析

(1) 现状及需求分析

当前 AMIS 基本上是对已有档案信息的管理、利用、编研等环节进行管理,而对档案界反映比较强烈的采全率低、采准率低、及时性差的问题,即档案收集不齐、质量不高、期限不准、归档滞后等无法处理,更没有提升到软件平台概念的产品。由于档案的收集范围、保管期限对不同单位具有特定性,一般档案信息资源管理人员难于把握,使得大量的原生信息源未能收集保管,直接影响了档案的收集、著录以及今后的提供利用水平,正是由于原有系统的种种缺陷致使档案收集不齐,造成城市改造项目无法开工,道路开挖损坏电缆、光缆等问题,每年我国因此承受很大损失。

因此,国内急需研究网络档案信息资源采集、组织、检索、利用等工作的原则和实施办法,在借鉴、吸收国外成熟的电子政务信息资源管理理念与技术的基础上,解决现有网络文献归档管理系统中存在的问题,提出具有中国特色的网络文献归档信息资源管理体系,研发出一种操作性、实用性强的软件支撑平台,来有效解决现有的档案信息系统存在的问题,档案业务规则管理系统的开发任务迫在眉睫。

(2) 可行性分析

技术分析:在国内档案行业软件的研究和开发上,有南京大学开发的档案管理系统,清华紫光股份有限公司研发的“清华紫光电子档案管理系统”、杭州信雅达系统工程股份有限公司开发的“信雅达档案综合管理系统”、深圳世纪科怡公司开发的“世纪科

怡档案管理系统”、珠海泰坦软件系统有限公司开发的“DARMS 系列档案管理软件”。在人工智能技术方面，特别是业务规则知识库（BR）在国内档案行业目前还没有相关的研究。但在银行、证券、航空、电力行业已经有相关的应用。在政府信息资源目录体系方面，经过数年的探索和实践，我国政府信息资源目录体系建设取得了较大进展，已经进入原型试点阶段。《2006—2020 年国家信息化发展战略》将政府信息资源目录体系列入“电子政务行动计划”。在国家和社会的高度重视下，我国政府信息资源目录体系研究和实践工作稳步前进。在智能政务方面，2006 年清华大学管理科学与工程博士后流动站博士后樊博提出“政务智能”概念。政务智能是分析型的电子政务系统，是面向政务数据分析的决策支持系统（Decision Support System，DSS）。政务智能系统的目的是通过分析政务活动中积累的海量数据，得到有指导意义的知识和决策依据，使政务更好地管理和服务社会。20 世纪 90 年代提出的数据仓库（Data Warehousing，DW）、数据挖掘（Data Mining，DM）和在线分析处理（On-Line Analytical Processing，OLAP）已形成潮流，在美国数据仓库技术已成为继互联网后的第二位的热点。数据仓库、数据挖掘、OLAP 已经成为决策支持系统发展的新方向。因此，从目前的技术成熟程度来看，建立一个基于业务规则的档案管理系统是可行的。

经济分析：有形的物质资源和能量资源与无形的信息资源，并列为现代经济发展的三大支柱，只有三者具备，才能全面发展。全球经济进入 21 世纪，信息资源已成为每个国家和地区都必须高度重视的核心战略资源，更是国家综合竞争力的决定因素。随着经济发展和社会进步，档案业务规则管理系统的建立是有足够的人力、物力和财力保障的。

（3）系统总体逻辑模型图分析

档案业务规则管理系统的工作模型分析如下：规则是定义在系统之外的，通过 Web 编写界面，在业务模型之上来定义规则，然后将一系列的规则保存在规则库中。规则引擎是一种嵌入在应用程

序中的组件，它的任务是把当前提交给引擎的数据对象与加载在引擎中的业务规则进行测试和比对，激活那些符合当前数据状态下的业务规则，根据业务规则中声明的执行逻辑，触发应用程序中对应的操作。档案业务规则管理系统正是将这些法律、法规、标准及指导性意见整理成 ABRD，利用档案业务规则推理引擎 RM，贯穿整个档案轨迹的全过程，控制档案的收集、整理的质量与数量。档案业务规则管理系统工作模型如图 5-1 所示。

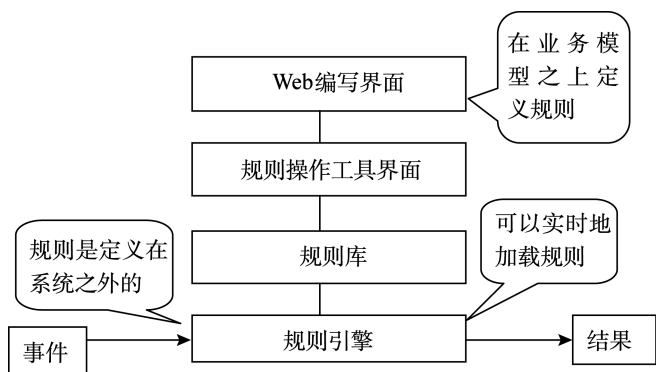


图 5-1 档案业务规则管理系统工作模型

5.2.2 系统设计

(1) 逻辑性设计

档案业务规则管理系统通过信息资源规划确定各立档单位的归档范围、保管期限表，并将其抽象为档案业务规则库，同时将档案主管部门与指导部门的规范、制度、标准等也整理到 ABRD，通过强大的推理引擎（RM），对已有的 AID 进行分析与推理，作出相应的判断，以帮助立档单位达到“档案收集齐、保管期限设置准、档案著录规范”的目标，同时也为档案管理部门建立起智能化的网上监督指导系统、绩效评估系统和上岗人员自学习系统等。也可

通过基于 ABRD 和 RM 构建的档案信息雷达自动从相关业务系统收集档案信息资源。

ABRD 是档案类别规则的格式化描述，它用来决定一个文档应该归档到哪个档案类别。本系统 ABRD 规则生成器以自然语言描述的档案分类规则为数据库源，通过中文分词系统提取规则描述中的关键词语，并过滤掉一些对档案分类无关的词语，从而生成档案类别的格式化描述。例如，某工业企业的档案利用 ABRD 将自然语言描述的规则存放在如下的 XLS 文档中，如表 5-1 所示。

表 5-1 某工业企业的档案分类描述(部分内容)

一级类目名称	二级类目名称	基本范围	保存期限
党群工作类	组织工作	组织建设,整党建党,党员和党员干部管理,党费管理等	
党群工作类	宣传工作	理论教育,各种工作活动宣传,政治思想工作与精神文明建设等	
党群工作类	统战工作	民主党派工作,无党派人事工作,港澳台工民主党派工作,无党派人事工作,港澳台工作,华侨工作,民族事务,宗教事务等	
党群工作类	纪检工作	党风治理,党纪检查,案件审理,信访工作等	
党群工作类	工会工作	职工代表大会,职工民主管理,劳动竞赛,劳动保障,女工工作,文化艺术和体育活动等	
党群工作类	共青团工作	组织建设,政治思想教育,团员大会,团员管理,团费管理,青少年工作等	
党群工作类	协会工作	各专业学会、协会工作,各群众团体活动等	
党群工作类	行政事务	企业综合性行政事务工作,厂务会议,厂长(经理)办公室工作,文秘工作,机要保密工作等	
政务管理类	公安保卫	社会治安,武装保卫,枪支弹药管理,民兵工作,消防,交通管理,刑事案件审理,人防工作等	
政务管理类	法纪监察	法律事务,政纪监察,违纪案件审理等	
政务管理类	审计工作	各专项审计工作活动等	
政务管理类	人事管理	干部管理,工人招聘、录用、调配工作,企业劳务出口工作等	

续表

一级类目名称	二级类目名称	基本范围	保存期限
党群工作类	组织工作	组织建设,整党建党,党员和党员干部管理,党费管理等	
政务管理类	教育工作	普通教育,中专和职业教育,高等教育,职工在职培训,幼儿教育等	
政务管理类	医疗卫生	卫生监督与管理,职工防病治疗,计划生育工作等	
政务管理类	后勤福利	职工生活福利,食堂,商店,幼儿园,农牧副业,职工住房,企业第三产业等	
政务管理类	外事工作	企业涉外活动	
经营管理类	计划工作	企业中、长期计划,年(季)度计划,各项专企业中、长期计划,年(季)度计划,各项专业发展计划,全面计划管理工作等	
经营管理类	统计工作	各种统计报表,企业综合性统计分析工作等	
经营管理类	财务管理	资金管理,价格管理,会计管理,资金流通等	
经营管理类	物资管理	物资供应,仓库管理,废旧物资回收与修旧利废等	
经营管理类	产品销售	市场分析,用户调查,产品销售,广告宣传,市场分析,用户调查,产品销售,广告宣传,售后服务工作等	
经营管理类	企业管理	企业普查,企业整顿和企业升级,经济责任制管理,企业管理现代化工作等	
生产管理类	生产调度	生产组织,调度指挥工作等	
生产管理类	质量管理	企业全面质量管理,产品质量检测和质量控制工作等	
生产管理类	劳动管理	劳动定额、定员,劳动调配,劳动工资,劳动保护等	

将一级目录和二级目录直接拷贝到 ABRD 中,对用自然语言描述的基本范围进行中文分词,然后把这些分词之后的词组系列保存到 ABRD 中,最后将保存期限表拷贝到 ABRD 中,ABRD 用 XML 文档表示。表 5-1 中的档案分类描述经 ABRD 规则生成器得到



如下的 XML 格式的 ABRD 文档部分内容，如图 5-2 所示。

```
<?xml version="1.0" encoding="GB2312" ?>
- <rules>
- <category1 id="001" name="党群工作类">
- <category2 id="001001">
  <name>组织工作</name>
  <description>组织建设,整党建党,党员和党员干部管理,党费管理等</description>
  <retentionPeriod>5年</retentionPeriod>
- <keywords>
  <item value="2">+组织;</item>
  <item value="2">+管理;</item>
  <item value="2">+整党;</item>
  <item value="2">+党群;</item>
  <item value="2">+工作;</item>
  <item value="4">+党员干部;</item>
  <item value="4">+组织建设;</item>
  <item value="4">+组织工作;</item>
  <item value="4">+干部管理;</item>
  <item value="2">+建设;</item>
  <item value="2">+建党;</item>
  <item value="2">+党员;</item>
  <item value="2">+党费;</item>
  <item value="2">+干部;</item>
</keywords>
</category2>
- <category2 id="001002">
  <name>宣传工作</name>
  <description>理论教育,各种工作活动宣传,政治思想工作与精神文明建设等。</description>
  <retentionPeriod>5年</retentionPeriod>
- <keywords>
  <item value="2">+政治;</item>
  <item value="4">+政治思想;</item>
  <item value="2">+党群;</item>
  <item value="4">+思想工作;</item>
  <item value="2">+理论;</item>
  <item value="2">+工作;</item>
  <item value="2">+建设;</item>
  <item value="2">+宣传;</item>
  <item value="2">+文明;</item>
  <item value="6">+政治思想工作;</item>
```

图 5-2 生成的档案业务规则库

档案业务规则库中的规则的来源主要有三大部分：第一部分是立档单位的文件的类目级别和保管期限表；第二部分是用自然语言描述的文件归档范围经过中文分词系统分词后得出的词组系列；第三部分包括相关法律法规、档案主管部门和指导部门的规范、制度、标准和时间经验，以及档案工作人员的知识总结等。其中，需要指出的是，立档单位在归档过程中使用的各种保管期限表是档案业务规则库中规则的重要来源，在此需做详细介绍。档案保管期限

表就是用表册形式列举档案的来源、内容和形式，并指明其保管期限的一种指导性文件。它是鉴定档案保存价值和确定保管期限的依据和标准，大致分为以下几类：

第一，通用档案保管期限表。它是由国家档案行政机关编制的，供全国各机关、团体、企事业单位鉴定档案使用，具有通用性，是全国各单位确定一般性档案保管期限的依据和标准，例如国家档案局于 2006 年 12 月第 8 号令中公布最新的《文书档案保管期限表》就属于这种类型，如表 5-2 所示。

表 5-2 文书档案保管期限表

1	本级党的代表大会、人民代表大会、政治协商会议，工会、共青团、妇联代表大会的文件材料	
1.1	请示、批复、通知、名单、议程、报告、领导人讲话、选举结果、讨论通过的文件、决议、纪要、公报、主席团会议记录等文件材料	永久
1.2	大会发言，人大代表建议和意见、人大议案及答复，政协委员提案及办理结果，简报，快报	永久
1.3	重要的贺信、贺电，筹备工作、选举过程中形成的文件，小组会议记录、会议服务机构的计划、总结等文件材料	30 年
1.4	讨论未通过的文件	10 年
2	本级党委、人民代表大会、政治协商会议、纪律检查委员会、共青团、工会、妇联的常委会、执委会、主席团、全体委员会会议，政府常务会、办公会议的文件材料	
2.1	公报、决议、决定、记录、纪要、议程、领导人讲话、讨论通过的文件、参加人员名册	永久
2.2	讨论未通过的文件	10 年
3	本机关党组（或实行党委制的党委）会议和行政办公会的纪要、会议记录	永久
4	本机关召开工作会议、专题会议的文件材料	
4.1	请示、批复、通知、名单、日程、报告、讲话、总结、决议、决定、纪要	永久
4.2	典型材料、代表发言材料、交流材料、简报	30 年
5	机关联合召开会议的文件材料	
5.1	本机关为主办的	

- 5.1.1 请示、批复、通知、名单、日程、报告、讲话、总结、决议、纪要 永久
- 5.1.2 典型材料、代表发言材料、交流材料、简报 30 年
- 5.2 本机关为协办的
 - 5.2.1 请示、批复、通知、名单、日程、报告、讲话、总结、决议、决定、纪要的复制件或副本 30 年
 - 5.2.2 典型材料、代表发言材料、交流材料、简报的复制件或副本 10 年
- 6 本机关承办国际性会议、大型展览会、博览会的文件材料
 - 6.1 请示、批复、申办和筹办组委会主要活动安排、议程、名单、主报告（原文及译文）、辅助报告（原文及译文），上级领导入贺辞、题词、讲话，会徽设计 永久
 - 6.2 代表发言材料、交流材料、简报、新闻报道 30 年
 - 6.3 委员会、分会会议和学术会的讨论记录，会议代表登记表、接待安排 10 年
- 7 上级机关、上级领导检查、视察本地区、本机关工作时形成的文件材料
 - 7.1 重要的 永久
 - 7.2 一般的 30 年
 - 7.3 本地区、本机关工作汇报材料 30 年
- 8 本机关业务文件材料
 - 8.1 本机关制定的方针政策性、法规性、普发性业务文件，中长期规划、纲要等文件材料 永久
 - 8.2 本机关的请示与上级机关的批复、批示
 - 8.2.1 重要业务问题的 永久
 - 8.2.2 一般业务问题的 30 年
 - 8.3 同级机关、下级机关的来函、请示与本机关的复函、批复等文件材料
 - 8.3.1 重要业务问题的 永久
 - 8.3.2 一般业务问题的 30 年
 - 8.4 本机关代上级机关起草并被采用的重要法规性文件、专项业务文件最后草稿 30 年
 - 8.5 机关联合行文的文件材料
 - 8.5.1 本机关为主办的
 - 8.5.1.1 重要业务问题的 永久

8.5.1.2	一般业务问题的	30 年
8.5.2	本机关为协办的	
8.5.2.1	重要业务问题的	30 年
8.5.2.2	一般业务问题的	10 年
8.6	本机关编辑、编写的文件材料	
8.6.1	大事记、组织沿革等	永久
8.6.2	简报、情况反映、工作信息等	10 年
8.7	行政管理、执法活动中形成的文件材料	
8.7.1	行政管理工作制度、程序、规定等文件材料	永久
8.7.2	执法检查情况汇总、通报, 整改通知等	永久
8.7.3	行政管理工作中形成的审批、审查、核准等文件材料	
8.7.3.1	固定资产投资、科技计划等项目审批 (核准)、管理、验收 (评估) 等文件材料	永久
8.7.3.2	不动产、自然资源的所有权、使用权确认的文件材料	永久
8.7.3.3	20 年 (含) 以上有效或未注明有效期的许可证、执照、资质证、资格证等的审批、管理文件材料	永久
8.7.3.4	20 年以下有效的许可证、执照、资质证、资格证等的审批、管理文件材料	30 年
8.7.4	行政管理工作中形成的备案文件材料	10 年
8.7.5	行政处罚、处分、复议、国家赔偿等工作中形成的文件材料	
8.7.5.1	重要的	永久
8.7.5.2	一般的	30 年
8.8	计划、总结、统计、调研等方面的文件材料	
8.8.1	年度和年度以上的计划、总结、统计材料	永久
8.8.2	年度以下的计划、总结、统计材料	10 年
8.8.3	重要职能活动的总结、重要专题的调研材料	永久
8.8.4	一般活动的总结、一般问题的调研材料	10 年
8.9	出国或出境访问考察、参加国际会议, 接待来访等外事活动形成的文件材料	
8.9.1	发表的公报, 签订的协议、协定、备忘录, 重要的会谈记录、纪要等	永久

8.9.2	出国审批手续、执行日程、考察报告、一般性会谈记录	30 年
9	本机关机构编制、干部人事、党、团、纪检、工会、保卫、信访工作文件材料	
9.1	机构设置、机构撤并、名称更改、组织简则、人员编制、印信启用和作废等文件材料	永久
9.2	人事工作制度、规定、办法等文件	30 年
9.3	人事任免文件	永久
9.4	先进单位、劳动模范、先进工作者的文件材料	
9.4.1	受县级（含）以上表彰、奖励的	永久
9.4.2	受县级以下表彰、奖励的	30 年
9.5	对本机关有关人员的处分材料	
9.5.1	受到警告（不含）以上处分的	永久
9.5.2	受到警告处分的	30 年
9.6	职工录用、转正、聘任、调资、定级、停薪留职、辞职、离退休、死亡、抚恤等文件材料	永久
9.7	人事考核、职称评审工作文件材料	永久
9.8	职工调动工作的行政、工资、党团组织关系的介绍信及存根	永久
9.9	职工名册	永久
9.10	党、团、工会工作活动中形成的文件材料	
9.10.1	工作报告、总结，换届选举结果	永久
9.10.2	重要专项活动的报告、总结等	永久
9.10.3	党团员、工会会员名册，批准加入党团、工会组织的文件材料	永久
9.10.4	情况反映、工作简报	10 年
9.11	纪检、监察工作中形成的综合性报告、调查材料	
9.11.1	重要的	永久
9.11.2	一般的	30 年
9.12	保卫部门的安全检查、调查记录	10 年
9.13	本机关处理人民来信来访的文件材料	
9.13.1	有领导重要批示和处理结果的	永久
9.13.2	其他有处理结果的	30 年
10	本机关事务管理文件材料	

10.1	房产、土地所有权和使用权的文件材料	永久
10.2	与有关单位签订的合同、协定、协议、议定书等文件材料	
10.2.1	重要的	永久
10.2.2	一般的	10 年
10.3	接待工作的计划、方案	
10.3.1	重要的	30 年
10.3.2	一般的	10 年
10.4	机关财务预算	30 年
10.5	机关物资（办公设备及用品、机动车等）采购计划、审批手续、招标投标、购置等文件材料，机动车调拨、保险、事故、转让等文件材料	30 年
10.6	国有资产管理（登记、统计、核查清算、交接等）文件材料	
10.6.1	重要的	永久
10.6.2	一般的	10 年
10.7	职工承租、购置本单位住房的合同、协议和有关手续	永久
10.8	职工住房分配、出售的规定、方案、细则，职工住房情况统计、调查表、职工住房申请	30 年
11	上级机关制发的文件材料	
11.1	上级机关制发的属于本机关主管业务的文件材料	
11.1.1	重要的	永久
11.1.2	一般的	10 年
11.2	上级机关制发的非本机关主管业务但要贯彻执行的文件材料	10 年
11.3	上级机关制发的关于本机关机构设置、领导人任免、人员编制等文件材料	永久
12	同级机关制发的非本机关主管业务但要贯彻执行的文件材料	10 年
13	下级机关报送的文件材料	
13.1	重大问题的专题报告	30 年
13.2	年度和年度以上的计划、总结、统计材料	10 年

第二，专门档案保管期限表。它是由国家档案行政机关会同有关主管部门编制的，是各机关、团体、企事业单位鉴定专门性档案时通用的依据。例如 1984 年财政部和国家档案局共同颁发

的《预算会计档案保管期限表》就是作为国家财政税收机关和使用国家预算的各种机关、团体、事业单位鉴定预算会计档案的统一标准。

第三，同系统档案保管期限表。它是由主管领导机关编制的，供同一系统内各单位鉴定档案时通用的依据和标准。例如《中国人民解放军文书档案保管期限参考表》就是这种类型。

第四，同类型单位档案保管期限表。它是由档案行政机关或主管领导机关编制的，是同一类型（如学校、医院等）各单位鉴定档案时通用的标准。

第五，单位档案保管期限表。它是各机关依据本单位档案的具体情况编制的，专供本单位鉴定档案时使用，它几乎包括了一个单位在工作活动中可能形成的所有文件及其保管期限，例如《国家计委机关文书档案保管期限表》就是这种类型。

以上种类的档案保管期限表之间具有相互制约关系，其中标准档案保管期限表对其他几种具有指导意义，而在利用档案业务规则管理系统对归档文件进行管理的过程中，上述几种类型的档案保管期限表是档案业务规则库中的规则的重要来源，立档单位在实际的归档和档案管理中，应将上述几类档案保管期限表拷贝至档案业务规则库中，以便于归档工作的开展。

档案信息数据 AID 包含全宗、目录/项目、案卷、卷内等信息也能够被保存至档案业务规则库 ABRD 中。而在此之后，档案业务规则推理引擎 RM 能够对保存在规则库中的 AID 进行分析和推理，从而准确判断其保管期限，与此同时，还能衍生出档案接收即时检查系统、档案著录规范检查系统、网上监督指导系统、绩效评估系统、上岗人员自学习系统等诸多子系统，共同推动档案工作过程的优化。而设计出的档案业务规则管理系统的整体流程如图 5-3 所示。

(2) 系统结构设计

对档案业务规则管理系统进行分析后，可得知档案业务规则管

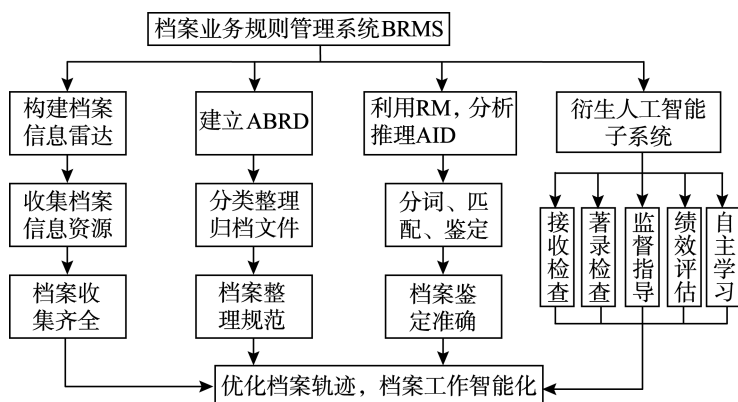


图 5-3 档案业务规则管理系统 BRMS 整体流程图

理系统主要的组成部分包括三大部分：用户工具，用户工具能够提供规则访问和规则管理。规则库，它指归档范围、保管期限表；相关专业档案的特殊规定；以《档案法》为主的相关法律法规和制度；档案从业人员的知识总结；档案局长期监督指导的实践经验等规则从各处进行提取，然后集中保存在规则库中。规则引擎，规则引擎能够提供规则的执行与规则的发布。档案业务规则管理系统三大部件之间是相互联系、相互作用的，共同为档案业务规则管理系统的正常运转服务，具体如图 5-4 所示。

5.2.3 系统部署和实施

(1) 系统的实施

一般的系统实施过程具有普遍适用性，档案业务规则管理系统的实施过程和一般系统实施过程基本上是一致的。具体的实施步骤为：第一步，根据档案业务规则管理系统设计的指标要求制作采购的标书，并邀请硬件供应商进行技术交流，选择资质好的厂商发放标书，经过评标、开标、商务谈判、获取并执行合同来获取档案业

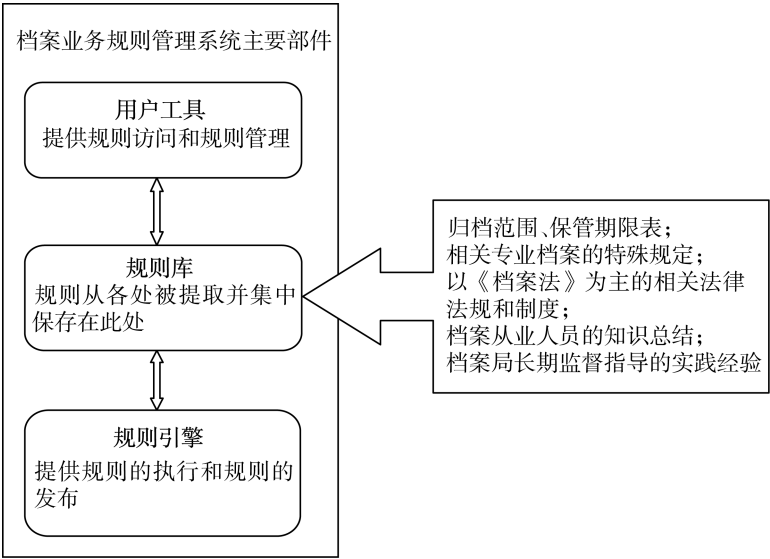


图 5-4 档案业务规则管理系统结构图

务规则管理系统实施所需的硬件设备。第二步，进行局域网、广域网的改造和建设，并根据档案业务规则管理系统设计的指标要求选择系统开发方法、软件工具和相关的数据库产品。第三步，便是终端用户准备。终端客户应具备的条件包括已有并可熟练使用档案管理系统（AMIS）；单位主管领导必须有强烈的档案意识；拥有专职档案从业人员，持证上岗，熟悉档案相关法律、法规、制度等；已制定完善的结合本单位实际的归档范围、保管期限表、分类编号方案等制度；配有高性能的档案专用服务器及备份设备；与下属单位或管辖单位网络已互通。第四步，进行档案业务规则管理系统相关管理人员的聘用和培训。参训人员应是拥有档案工作从业许可证件的专职人员，具有较深厚的档案理论知识功底、熟练的档案工作技能以及扎实的计算机软硬件操作技术，熟悉与档案工作、档案事业有关的法律法规及制度。参训人员只有拥有了这些必须掌握的知识 and 技能，才能在培训之后较为迅速地适应新的档案业务规则管理系统的实施环境，并熟练掌握新系统的操作方法，节约培训的时间和

成本，创造更好的效益。第五步，场所准备。它主要是指新系统的工作场所中的硬件设备、电力保障、办公家具的准备，数据准备是将原来的手工数据文件转化成计算机文件，将原来的不同格式的数据文件转换成新系统的格式。第六步，进行安装、测试、试运行、系统切换和用户验收。测试的方法主要包括模块测试（单调）、联合测试（联调）、系统测试（总调）。系统切换是新旧系统的交替过程，也就是旧的管理信息系统逐渐退出，由新的管理信息系统来代替。档案业务规则管理的具体的实施过程如图 5-5 所示。

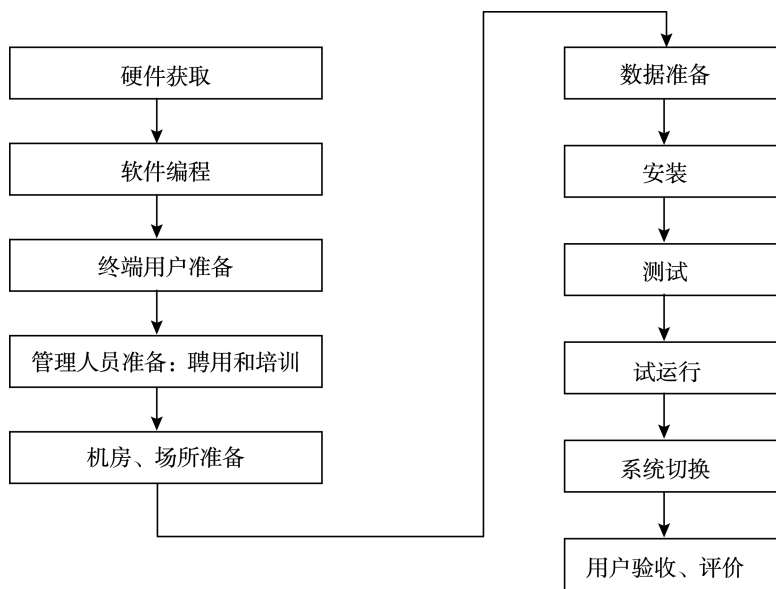


图 5-5 档案业务规则管理实施流程图

(2) 系统测试

压力测试：在系统工程中，压力测试（Stress Testing）是对系统不断施加压力的测试，是通过确定一个系统的瓶颈或者不能接收的性能点，来获得系统能提供的最大服务级别的测试。例如测试一个 Web 站点在大量的负荷下，何时系统的响应会退化或失败。在

档案业务规则管理系统的压力测试过程中，可以将某市级档案馆作为对象，在档案业务规则管理系统单机运行的情况下，系统管理人员、档案工作人员和档案用户同时使用是没有问题的，但是当该系统与互联网进行连接，网络用户便有可能大量涌入，该系统出现性能瓶颈，这时就需要增加若干服务器来满足大量的用户需求，避免系统运行迟缓甚至瘫痪。

功能性测试：包括白盒测试和黑盒测试。白盒测试称为结构测试，根据程序的内部结构和逻辑结构来设计测试系统，分析程序结构与过程执行路径，检验程序的每条通路是否按预期正常进行，力求提高测试覆盖率。黑盒测试是一种从用户观点出发的测试，它将档案业务规则管理系统看做黑盒子，不考虑其内部程序结构与处理过程，仅仅对程序接口进行测试，即检查适当的输入是否能够产生适当的输出。例如，若输入的档案信息的保管期限为“0”，这时系统就应当主动提示用户，避免出现保管期限为“0”的情况，或者系统默认某一时间段为档案的保管期限。功能性测试正是要通过对档案业务规则管理系统的各项预定功能进行检查和试验，及时发现问题并采用各种方法进行纠正，力求使系统功能更加完善，在正式运行阶段充分发挥作用。

5.2.4 系统维护与评价

(1) 系统维护

为了满足用户不断增加的新需求，有效应对运行环境变化、技术更新换代、硬件设备老化、政策调整等问题，需要对系统进行长期维护，以保证档案业务规则管理系统正常运行。在系统维护的过程中，要保证维护队伍的稳定性，保证维护的经费充足，节约软件维护的成本。档案业务规则管理系统维护的重点包括业务规则和词群表的维护、系统功能的整体维护、系统的日常营运维护三大部分。第一，业务规则和词群表的维护。由于 ABRD 中的规则和规则所对应的词群会随着系统的运转和海量的档案信息数据的录入而

不断增加，因此规则表和词群表应当根据最新录入的档案信息数据实时完成自动更新，并不断提高自身的容错性，以满足用户需求。第二，系统功能的整体维护。由于规则库中业务规则和词群表的不断增加，系统自身的功能也需要进行升级，以提高系统的易用性和规模的可伸缩性。第三，日常营运的维护。因为档案业务规则管理系统与互联网进行连接，极有可能遭受外来攻击，因此在系统的日常营运中，要进行系统的访问和管理权限的设置，同时要要进行系统网络安全维护，避免遭受网络攻击。

（2）系统评价

系统评价贯穿系统建设的全过程，尤其在设计、实施阶段更为重要。它分析系统的处理，以确保系统能按期望运行。档案业务规则管理系统可从性能评价、功能评价、经济评价三个角度来建立评价指标体系。第一，性能评价。例如，根据测试可看出档案业务规则管理系统的性能较好，可以支撑局域网约 200 人同时使用，达到了预期目标。第二，功能评价。档案业务规则管理系统的功能较为完备，其中，档案要素的录入、分词、词群自动匹配、系统智能自学习等功能是档案业务规则管理系统区别于其他档案管理系统的特殊之处，这些功能可提高档案工作的效率和档案工作人员的水平，从而有效推动实际档案工作向自动化、智能化方向发展。第三，经济性评价。例如，档案业务规则管理系统对于市级档案馆，开发和维护的费用在 5 万 ~ 10 万元，但由于该系统的实施能够在实际档案工作中节省大量的人力、物力和财力资源，因而可以提高工作效率，节约工作成本。总之，从这三个方面以及技术角度来对档案业务规则管理系统进行评价，可看出该系统创新地采用人工智能技术，贯穿整个档案轨迹的全过程，控制档案的收集、整理的质量与数量；构建档案信息雷达，通过数据挖掘技术将用户单位的各种业务系统的数据收集到“规则档案平台”；通过 Internet/Intranet 网络互联技术，将主管单位与下级单位或企业各分支机构的档案信息组建为一个集中的目录中心，实现档案实体与档案信息的分离管理；建立一种新型的档案移交、接收模式，做到档案实体正式移交前进

行相应的非现场预检查，保证档案接收质量并减少后续的著录量与数字化加工量；建立一种新型的档案监督指导管理体系与绩效考核体系，推动档案事业的规范健康发展；建立一种档案自学习系统，提供档案上岗人员接受培训的新途径，并找到传承资深档案人士知识的快捷方法；将在文件流转过程中采用智能派送、痕迹化管理；将数据固化技术应用于电子文件的归档，采用 SOA 的框架结构，集成各种高级应用。

该系统从技术上实现了将种类繁多的电子政务信息资源转换为单一的档案信息资源，并为档案信息资源的开发与利用提供可行的技术手段，提出我国电子政务建设的档案信息资源开发与管理的框架内容，为国家电子政务建设提供有价值的实践方案。因此，该系统具有可操作性，并具有较强的现实意义和实践价值。

5.3 档案业务规则管理系统的功能

一个完整的 BRMS 应该提供规则定制和设计（Rules Design and Authoring）、规则分析（Rules Analysis）、规则部署（Rules Deployment）、规则管理（Rules Management）等功能。

第一，本系统能通过信息资源规划确定各立档单位的归档范围、保管期限表，并将其抽象为档案业务规则库 ABRD。将档案业务规则进行抽象化的过程可简要描述为：以自然语言描述的档案类目级别和归档基本范围等一系列规则是 ABRD 最初的数据库源，这些规则存放在 XLS 格式的文档之中；这些规则中的一级、二级类目可直接拷贝到 ABRD 之中，而以自然语言描述的归档基本范围则需要通过中文分词系统采用基于字符串匹配、基于统计规律、基于语意理解等方法对其进行分词，生成词群系列，并保存至 ABRD；同时也将《档案保管期限表》、档案主管部门与指导部门的规范、制度、标准等也整理到 ABRD，这些 XLS 格式的文档经过 ABRD 处理之后成为 XML 格式的 ABRD 文档。这一抽象化过程可用简图表示（见图 5-6）：

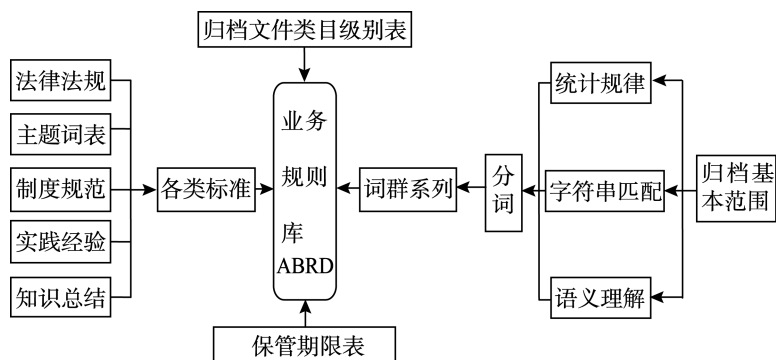


图 5-6 档案业务规则抽象化流程图

第二，本系统能利用强大的推理引擎 RM，对保存在 ABRD 中的全宗、目录/项目、案卷、卷内信息等档案信息数据 AID 进行分析与推理，并由 ABRD 采集若干档案要素，利用中文分词系统对其进行分词，与保存在 ABRD 中的一系列规则所对应的词群进行匹配，提高决策水平，推理得出相关结论：数据是否符合要求，档案著录是否规范，档案收集范围是否广泛，档案保管期限定义是否准确等，以帮助立档单位达到“档案收集齐全、保管期限设置准确、档案著录规范”的目标。还可通过基于档案业务规则库和推理引擎构建的档案信息雷达自动从相关业务系统收集档案信息资源。

第三，本系统可通过档案业务规则库在档案轨迹的各个环节衍生出相应的人工智能子系统，因而能有效地完善档案管理各项工作，推动整体的档案管理工作向自动化、智能化、集成化方向发展。

①人工智能子系统包括：档案接收即时检查系统。接收档案范围确定后，该系统可即时根据档案业务规则库对各立档单位的档案信息，通过知识规则推理引擎，对录入或导入的档案信息进行检查，系统自动提出整改意见或验收意见。②档案著录规范检查系统。该系统对责任者、题名、保管期限等，根据已存在的



档案业务规则库自动进行检查。③网上监督指导系统（W_NAA）。档案局/馆、档案主管单位利用基于 Internet/Intranet 的档案信息扫描技术，结合档案业务规则库，自动监督各立档单位的档案整理情况，发现问题自动提醒，并及时反馈，信息转入 Workflow，档案监督指导人员根据实际情况作出处理。它利用人工智能的原理，借助计算机模拟人类的思维过程，对档案信息数据进行分析与推理，作出相应的判断，提出建议及线索，以供档案管理部门进一步进行重点检查与监督指导。其目的就在于，提高监督指导效率，降低监督指导风险，进而保证档案收集入库的质量。④立档单位绩效评估系统（W_NAE）。绩效评测内容与目标建立后，系统可定期根据档案业务规则库对各立档单位的档案信息，通过知识规则推理引擎，自动生成绩效预评估报告，经过 Workflow 流转后产生最终的评估报告。⑤档案上岗人员自学习系统（W_NAL）。一般的档案人员起码需一年时间才能熟练掌握档案的归档范围与保管期限，通过 W_NAL 使上岗人员尽快了解档案的相关法律、法规和制度，充分利用档案业务规则库与规则推理引擎的作用，设置相关的 Workflow 使上岗人员在短时间内学习、测试对档案的熟练程度。

5.4 档案业务规则管理系统的结构

档案业务规则管理系统由文件信息导入层、文件信息挖掘层、档案信息表现层三大部分构成。第一，文件信息导入层包括两大部分，一是自然语言描述的归档文件类目级别，保管期限表，相关法律法规，档案部门的工作规范、制度和标准，工作人员的知识总结等；二是自然语言描述的归档基本范围经中文分词系统分词以后形成词群系列。这两大部分经过结构化、格式化和集成化，拷贝到档案业务规则库 ABRD 中，进入下一个环节。第二，文件信息挖掘层是档案业务规则管理系统中的核心环节。档案信息雷达从业务规则网络、档案网络、语义网络和要素网络这四大网络中收集归档文

件信息，经过聚类 and 分类处理，再到内容分析和业务规则引擎 RM 处理，在此之后由中文分词系统进行分词，剔除无用词，提取关键词，并将其与档案业务规则库 ABRD 中生成的一系列业务规则所对应的词群进行比对和匹配，从而确定具体的档案信息的保管期限、密级等，完成档案信息的著录。第三，档案信息表现层是档案业务规则管理系统的最终环节。在上一个环节中由档案业务规则库和规则引擎完成归档文件信息的分类、鉴定和著录等工作后，生成了完整的可查询的档案信息，同时还衍生了诸多子系统，包括档案接收即时检查系统、档案著录规范检查系统、网上监督和指导系统、绩效评估系统、上岗人员自学习系统等。档案业务规则管理系统的结构如图 5-7 所示。

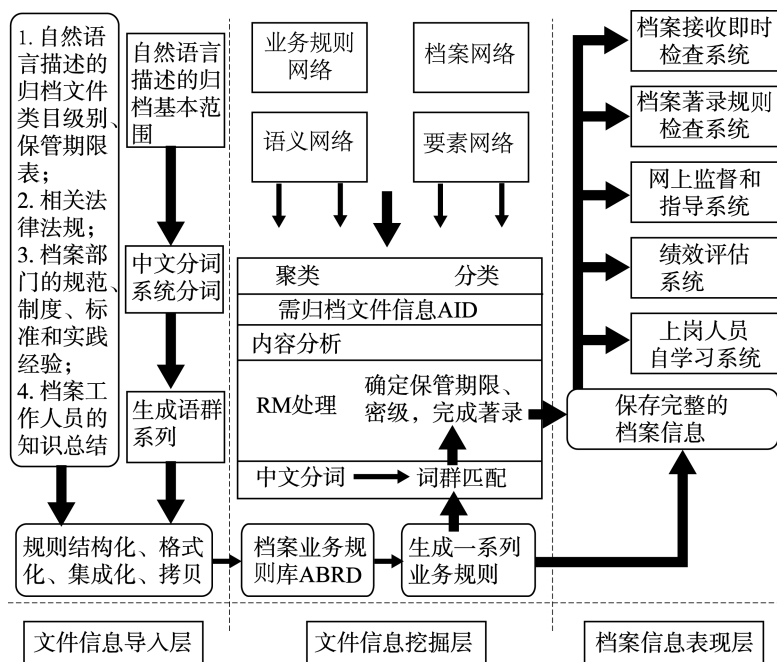


图 5-7 档案业务规则管理系统的结构图



5.4.1 档案业务规则集成开发环境

一般规则集成开发环境只有规则编辑器，而高级的规则集成开发环境可以实现对规则和规则库的管理，如规则的检索、版本控制、权限管理等；甚至可以实现对多个规则引擎的“在线”调试，对规则集合进行冲突检查等。而档案业务规则管理系统的集成开发环境主要包括规则编辑器、规则管理器、规则编译器三大部分。

(1) 档案业务规则编辑器

档案业务规则编辑器主要是实现对档案业务规则的创建和增删改两大功能。第一，业务规则的创建。档案业务规则的创建过程为：一方面，自然语言描述的以 XLS 和 Word 格式保存的归档文件的类目级别和保管期限表，档案工作法律法规，档案主管部门和指导部门的相关规范、制度、标准和实践经验，档案工作人员的知识总结等，可直接被拷贝到档案业务规则库中 ABRD 进行保存；另一方面，自然语言描述的以 XLS 格式保存的归档基本范围在中文分词系统的作用下进行分词，生成词群系列，然后拷贝到档案业务规则库 ABRD 中保存。这两大部分在 ABRD 中经过格式化、结构化、集成化处理后，成为以 XML 格式保存的 ABRD 文档，从而构建出一系列档案业务规则，每一条业务规则都对应着若干语义词群。第二，业务规则的增加、删除和改动。因为归档文件的数量是随着日常的业务活动和档案工作的开展不断增加的，因而档案业务规则库也必须实时更新，才能满足文件归档的需要。在 ABRD 中添加规则，可能由于多名操作者身份和认知不同，各个操作者对某一具体归档文件的规则设置问题的看法存在差异，操作者事先没有进行良好的沟通等原因，导致重复添加同一条业务规则的问题。为了避免档案业务规则库中出现规则重叠，不必要地占用规则库中有限的存储空间，档案业务规则管理系统须设置查重功能，在出现录入规则重复时，系统应自动提示录入者已出现该问题，需进行纠正。在 ABRD 中删除和改动某些业务规则，必须考虑到库内数据

的一致性。因为每一份归档文件都是经过 RM 处理和中文分词系统分词后提取归档信息中的关键词，并将其与 ABRD 中的业务规则所对应的词群进行匹配，然后确定其保管期限和密级等，完成该份归档文件的鉴定和著录工作，根据库内每一条归档文件信息都有与之对应的业务规则，该条业务规则是检验该份档案文件各类目是否设置准确的标准和依据，二者之间存在紧密的对应关系。现在如果要对该条业务规则进行删除或改动，必然会影响到依据此条规则进行归档的若干份档案文件，造成 ABRD 中归档文件信息数据的混乱。因而，当某条规则不再具有利用价值需要删除时，或者某条规则有些内容不再适应工作要求，需要进行修改时，可设置为：无操作，显示一条错误信息，告知用户不允许执行该删除或修改操作；级联，删除或修改某条业务规则，则系统提示用户该条业务规则所涉及的所有档案信息数据也会被删除或修改；设置空，如果规则库中的所有规则都可接受空值，则将该值设置为空；设置默认值，如果规则库中的所有规则均已定义默认值，则将该值设置为默认值。

(2) 档案业务规则管理器

档案业务规则管理器主要是实现档案业务规则的查询与检索、规则库的版本控制、档案业务规则的冲突管理等功能。第一，档案业务规则的查询与检索主要包括归档文件的类目级别的查询、归档范围的查询、规则库中词群系列的检索等。第二，规则库的版本控制。所谓版本控制，是指通过文档控制（Documentation Control）记录程序各个模块的改动，并为每次改动编上序号。一种简单的版本控制形式，例如，赋给一个系统初版以版本等级“A”。当做了第一次改变后，版本等级改为“B”，以此类推。简单说来，档案业务规则库的版本控制可采用管理员多人修改规则、协同工作，并在修改的部分保留标记，待审核员对修改处进行审核通过之后，最终确定规则修改是否合理并将其进行保存。第三，档案业务规则的冲突管理。在归档信息与规则库中的规则所对应的词群进行匹配时有可能出现三种情况：一是该条档案信息找到某条规则与之对

应，二是该条档案信息找不到任何规则与之对应，三是一条档案信息找到若干条规则与之对应，即它既可能属于 A 规则，也可能属于 B 规则。第一种情况属于正常状态，可利用档案业务规则库来对具体的档案信息进行处理。后两种情况称为规则歧义，业务规则的冲突管理就是要解决这一问题。如果出现某一条归档文件信息无法找到规则库中任何一条规则与之匹配，可以将此条档案信息闲置后放入“未找到规则与之对应”的模块中，系统自动对操作员进行提示。有些档案信息的规则也可以由操作员自己依据某一准确、具体的标准来进行定义，但是须统一交由审核员审核之后再做定夺。如果出现某条档案信息同时可与若干条规则进行匹配的情况，则需要操作员根据归档文件的具体内容，分别比较它与若干条规则之间的匹配程度，根据对应的准确程度，慎重判断该条档案信息究竟属于哪一规则，再交由审核员进行审核通过方可。

(3) 档案业务规则编译器

档案业务规则编译器进行规则的编译主要有如下 8 个步骤：第一步，词法分析。词法分析器根据档案业务规则和中文分词系统识别出 AID 档案信息数据所在源程序中的各个“记号”，每个记号代表一类“单词”。源程序中常见的记号可以归为几大类：关键字、标识符、字面量和特殊符号。词法分析器的输入的是源程序，输出的是识别的记号流。词法分析器的任务是把档案源文件的字符流转换成记号流，查看连续的字符，然后把它们识别为“单词”。第二步，语法分析。语法分析器根据档案业务的语法规则识别出记号流中的结构（短语、句子），并构造一棵能够正确反映该结构的语法树。第三步，语义分析。语义分析器根据档案业务语义规则对语法树中的语法单元进行静态语义检查，如类型检查和转换等，其目的在于保证语法正确的结构在语义上也是合法的。第四步，中间代码生成。中间代码生成器根据档案业务规则的语义分析器的输出生成中间代码。中间代码可以有若干种形式，最常用的一种中间代码是三地址码，它的一种实现方式是四元式。三地址码的优点是便于阅读、便于优化。第五步，中间代码优化。由于档案业务规则编译器

将源程序翻译成中间代码的工作是机械的、按固定模式进行的，因此生成的中间代码往往在时间和空间上有很大浪费。当需要生成高效目标代码时，就必须进行优化。第六步，目标代码生成。在生成目标代码时要考虑以下几个问题：计算机的系统结构、指令系统、寄存器的分配以及内存的组织等。编译器生成的目标程序代码可以有多种形式：汇编语言、可重定位二进制代码、内存形式。第七步，符号表管理。符号表的作用是记录源程序中符号的必要信息，并加以合理组织，从而在编译器的各个阶段能对它们进行快速、准确地查找和操作。符号表中的某些内容甚至要保留到程序的运行阶段。第八步，出错处理。档案工作人员编写的规则中往往会有一些错误，可分为静态错误和动态错误两类。所谓动态错误，是指源程序中的逻辑错误，它们发生在程序运行的时候，也被称为动态语义错误。静态错误又可分为语法错误和静态语义错误。语法错误是指有关语言结构上的错误，静态语义错误是指分析源程序时可以发现的语言意义上的错误。因而，档案业务规则编译器需要对这些错误进行处理，以保证规则库和系统的正常运转。档案业务规则库将各条业务规则分类并集成为多个模板，每个分类规则模板都是各种不同规则的集合。当在规则库中添加业务规则时，不用逐条规则进行添加，仅需要选择相应的分类规则模板即可。同时，增加和修改的规则可以另存为模板，模板也可以导出，进行备份。这样增加了可操作性，简化了操作，更方便档案业务规则库管理人员使用，以及与其他工作人员共享、交换模板。

5.4.2 档案业务对象模型

DURA 模型系指基于业务规则的智能归档管理模型，它是在档案业务规则管理系统（Business Rules Management System，简称 BRMS）中利用档案业务规则库（Archive Business Rules Database，简称 ABR）建立起来的档案业务对象模型。其构成要素包括：D（Document 文件）、U（User 用户）、R（Rules 业务规则）、A（Archive 档案）。DURA 模型就是 U（用户）利用 R（业务规则），

将电子政务活动中产生的 D (文件) 完整、及时、准确地转化为 A (档案), 向社会公众及时提供所需的电子政务档案信息资源和优质规范的档案服务。

文件 (Document): 包括内容、背景和结构信息, 可对文件进行如下的形式化定义: 设 content 为 Document 的内容, 即 Document 所包含的表达作者意图的信息, 在此处仅指文件的正文部分即 when/where/who/what/why/how; 设 context 为 Document 的背景, 即能证明文件形成过程和文件之间相互关系的信息; 设 structure 为 Document 的结构, 即文件内容信息的组织表达方式, 包括逻辑结构和物理结构。Document 可定义为一个由 content、context、structure 组成的三元组:

Document = <context, { content }, { structure } >.

Content = <when, where, who, what, why, how>.

Structure = <physical carrier, logical element>.

用户 (User): 在 DURA 模型中扮演的是操作者、管理者和观察者的多重角色。第一, 定义用户角色并授权。将用户角色设定为 A、B、C、D、E、F 六大类。A 类: leader of the institution。制定本单位基于 BRMS 的电子政务档案管理整体规划, 对系统进行宏观控制, 对本单位应用规则库开展电子政务档案归档的整体流程进行审核。B 类: head of the related department。对本业务部门的具体工作进行整体部署, 并对该项业务工作进行监督和审核, 对规则库中与本业务部门有关的规则进行控制和审核。C 类: staff of the related department。处理部门负责人所分配的细化的工作, 完成档案的著录, 提供本部门业务工作中的各类规则给数据库人员。D 类: end user。可利用 BRMS 查询检索该单位电子政务档案信息资源, 提出需求和反馈意见。E 类: system administrator。对系统进行管理和维护, 完成基于角色的规则权限的设置和用户分配。F 类: ABR administrator。对 ABR 进行创建、后台管理、日常维护、有权根据各业务部门的意见对与该部门密切相关的规则适时进行增删改。第二, 用户角色分配, 即根据政府机关和单位的人员管理体制和层级, 给 User 分配相应的角色, 赋予其对应的权限。User 可定义为

由以上六种角色组成的多元组:

User=<leader of the institution, head of the related department, staff of the related department, end user, system administrator, ABR administrator>.

规则 (Rules): 在 DURA 模型中其主要内容包括自然语言描述的归档文件类目级别, 保管期限表, 相关法律法规, 档案部门的工作规范、制度和标准, 工作人员的知识总结, 归档基本范围。这一部分将在后文中作重点论述。对规则进行形式化定义: 设自然语言描述的 category 为 Rules 的类目级别; 设 records retention schedule 为 Rules 的保管期限表, 包括通用、专用、同系统、同类型单位和单位档案保管期限表; 设 clumps 为 Rule 中经过分词的归档范围词群; 设 standards 为 Rules 中用自然语言描述的相关法律法规、制度标准、主题词表、实践经验和知识总结等。将 Rules 定义为一个四元组:

Rules = < category, records retention schedule, filing range clumps, {standards}>.

Standards = < archives laws, regulations, archives thesaurus, knowledge and practical experience >.

档案 (Archive): 在 DURA 模型中特指在电子政务活动处理完毕确认值得保存以供政府部门自身和社会公众查考利用, 而通过物理或逻辑归档方式移交给档案部门的文件集合。可对档案进行形式化定义: 设 original materials 为 Archive 的原始记录, 设 description elements 为档案的著录要素, 而著录要素包括 substance elements 内容要素和 form elements 形式要素。内容要素主要包括主题词、分类号、摘要; 形式要素主要包括形成时间、标题、责任人、档号、保存期限、密级、载体、文种等。可将 Archive 定义为一个由内容要素和形式要素组成的二元组:

Archive=<original materials, { description elements }>.

Description elements= < { substance elements } { form elements } >

Substance elements=<descriptor, classification code, abstract>.

Form elements= <date, title, responsible person, archival code, retention period, degree of secrets, carrier, record type>.

在 DURA 模型中，电子政务活动中文件是档案的来源之一，档案是文件的主要归宿，文件向档案的转化是一个单向传递过程，而业务规则正是文件向档案转化过程中归档的主要依据和标准，它决定文件是否可转化为档案，归档文件的保管期限和密级的设置等。用户在文件转化为档案的过程中扮演的是参与者、控制者和观察者的多重角色。他能定义、修改、删除、更新档案业务规则库的规则，对整个业务规则库进行管理，控制整个归档流程，是这个对象模型中非常重要和活跃的因素。

- D→A(descriptor): By Rules (archives thesaurus)
- D → A (classification code): By Rules (the Chinese archives classification)
- D→A(abstract, date, title, responsible person, carrier, record type) : By Users' analysis of the document, knowledge and practical experience
- D→A(archival code): By Rules(archives regulations)
- D→ A (degree of secrets): By Rules (archives laws) and Users' analysis of the document
- D→A(retention period): By Rules(records retention schedule)

DURA 模型的四个要素联系紧密，相互作用。其实现流程可用图 5-8 来表示。

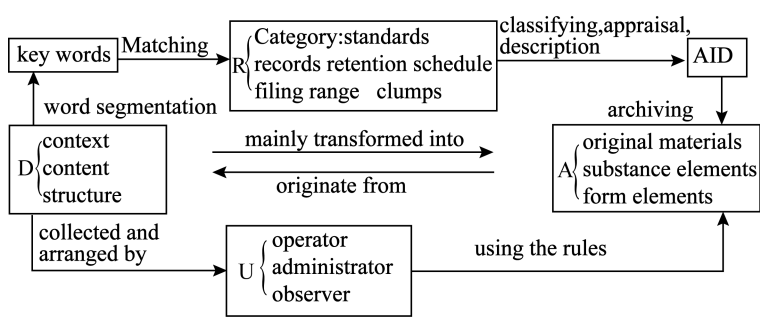


图 5-8 DURA 模型的实现流程图

DURA 模型能准确分析从文件到档案的转换过程中的迁移条

件,实现政府机关、业务部门和档案部门功能的无缝链接,大幅降低工作量,提高归档工作准确度。DURA 模型中采用语义匹配和内容挖掘技术,在电子政务建设中有现实意义。

但是由于现在 DURA 模型对电子政务档案信息自然语义的理解速度和准确度不高,BRMS 系统在对数字档案的分类上以及档案关键元素的提取上还存在较大的误差。如何实现系统对迁移条件的自动理解,实现业务规则动态更新,提高归档信息关键词提取的准确度、档案自动分类的精度和语义匹配的准确率,如何完成业务规则的半自动辅助构建是下一步的研究重点。

5.4.3 档案业务规则引擎

规则引擎是一种嵌入在应用程序中的组件,它的任务是把当前提交给引擎的数据对象与加载在引擎中的业务规则进行测试和比对,激活那些符合当前数据状态下的业务规则,根据业务规则中声明的执行逻辑,触发应用程序中对应的操作。

档案业务规则管理系统正是将这些法律、法规、标准及指导性意见整理成档案业务规则库 (ABR),利用档案业务规则推理引擎 (RM),贯穿整个档案轨迹的全过程,控制档案的收集、整理的质量与数量。档案业务规则引擎和规则库的关系可用计算机杀毒软件作为对比例子,互联网上新型计算机病毒总是在不断增加,并大肆入侵电脑,杀毒软件在对电脑全面杀毒的同时也在不断升级,并把新型病毒的详细资料复制到病毒库内进行保存,以便研究针对此种病毒的新的解决措施,病毒库内的病毒的资料也在不断丰富。而档案业务规则引擎就相当于杀毒软件,档案业务规则库就相当于病毒库。档案业务规则引擎在对档案信息数据进行处理时,将保管期限表,文件类目级别,归档基本范围,相关法律法规、制度、标准等进行分词处理,生成词群系列之后,拷贝到档案业务规则库内,以便下一步档案工作的开展。

5.5 规则权限管理和系统安全管理

5.5.1 规则权限管理

(1) 现存问题

随着海量档案信息资源呈指数增长，业务规则库中的规则不断变化，运用档案业务规则管理系统来进行归档管理的工作者以及利用档案的用户数量也在不断增加。如何既保证档案业务规则管理系统的稳定性和安全性，又能为访问者提供便利快捷的服务，成为一道难题。这就必须对档案业务规则库中的规则进行权限设置和管理：哪些访问者能够访问规则库中的规则，访问者访问的是“什么规则”，访问者对他所访问的规则拥有什么样的“权限”。可以用“who 对 what 是否能进行 how 的操作”来表述规则权限的需求。这就涉及网络安全的重要内容：权限管理与访问控制。访问控制是信息安全保障机制的核心内容。访问控制是为了限制访问主体对访问客体的访问权限，从而使系统在合法范围内使用。传统的档案管理系统通常是验证用户的用户名和口令来确定用户的身份，同时，系统通过维护一张访问控制 ACL 列表确定每个用户对特定系统对象的操作权限，这种方法因为简便易行而得到了普遍的应用。但是对于一些大型组织机构来说，其应用系统较多，用户数量巨大，采用这种方式需要不同的应用系统分别针对保护的资源进行权限的管理和控制，因而产生了一系列问题：第一，权限管理混乱。对一个档案部门而言，数据和人力资源都是统一的。但是由于原有的档案管理系统设计的原因，可能同时对相同的人员采用不同的管理方式，对机构内的共享数据采用了不同的权限分配策略，这显然不合理，也不利于对规则库中的规则和档案信息资源的管理。第二，系统安全性较低。不同的权限管理策略产生的安全强度是不同的。这

就可能造成机构信息安全管理漏洞，因此入侵者就有可能瞄准那些权限管理相对不安全的系统进行集中攻击，这就给规则库中的规则和档案信息资源的安全性带来极大的危害。第三，权限管理依赖于访问控制应用。权限的赋予和撤销往往都是在访问控制应用中产生的，不同的访问控制应用之间尽管有相同的用户和授权策略却往往不能互相使用对方的权限。每个应用都要维护自己的用户信息和授权方法，权限无法在分布的应用和远程应用中使用。

（2）基于角色的规则权限管理

针对上述问题，可应用基于角色的档案业务规则管理系统权限设置和管理的模式，既保证档案业务规则的安全性和稳定性，又使得以各种不同角色存在的访问者能够充分发挥系统已为各自设定好的权限，在自己的权限范围内对规则进行相应处理，方便访问者对档案业务规则管理系统的管理和利用。基于角色的规则权限设置和管理具有如下特点：第一，基于角色的策略将单独对特定用户授权分成两个动作：给用户赋予角色，给予角色对对象的权限；角色在实际应用也要分层，权限是通过继承方式组织的。第二，角色允许单独用户对数据传播最少的特权。被授予超级用户权限的用户，不一定使用超级权限，除非有特殊的需求，而只使用普通角色所拥有的权限，这样可以降低入侵者冒充合法用户对规则造成破坏。

基于角色的规则权限设置和管理的步骤为：第一步，定义角色。以某一普通市级档案馆为例，可根据一般的档案馆工作体制和人员配置情况，将权限设置和管理的角色暂定义为 A、B、C、D、E、F 六大类。第二步，角色授权，即根据授权策略给这些角色分配相应的权限。A 类：制定基于档案业务规则管理系统的档案馆发展规划，对整个档案业务规则管理系统进行宏观控制，对档案业务规则库中的规则可进行访问，并对应用规则库开展工作的整体流程进行审核；B 类：对本业务部门的具体工作进行整体部署，并对档案馆中此项业务工作的整体流程进行监督和审核，对规则库中与本业务部门有关的规则进行控制和审核，对与其他环节有关的规则无权审核；C 类：处理部门负责人分配给自己的具体细化的工作，分

工合作，共同履行本业务部门的职责，不断提供本部门业务工作中的各种具体的法规、制度、标准、知识总结等给数据库操作人员，使规则不断丰富，对规则库中与自己所做工作有关的规则有权访问，但无权完成其他操作；D类：可利用档案业务规则管理系统查询检索档案馆馆藏资源，提出需求和反馈意见；E类：对档案业务管理系统进行管理和维护，完成基于角色的规则权限的设置和用户分配；F类：对档案业务规则库进行创建、后台管理、日常维护，有权根据各业务部门的意见对与该部门密切相关的规则适时进行增加、更新、修改、删除等。第三步，角色分配，即根据档案馆人员管理体制和层级，对每一级别的人员分配相应的角色，也就赋予他们以这些角色所对应的权限。例如在上文所举例子中，应该将A类角色分配给馆长级别，以便于馆长对档案馆的业务规则管理系统和档案馆工作进行全面宏观调控；B类角色应当分配给档案馆各业务部门的主管负责人，如档案资源收集部、档案分类整理部、档案鉴定委员会、档案保管部、档案利用部等，各部门负责人对本部门的业务工作进行整体控制和审核；C类角色应分配给各个业务部门的工作成员，以完成本部门负责人所部署的具体的任务；D类角色应赋予档案馆注册用户和档案资源利用者，使他们能利用档案业务规则管理系统检索到自己需要查询了解的档案信息，然后在档案馆获取相应的档案信息资源；E类角色应分配给系统管理员，使他们能对整个档案业务规则管理系统进行管理和维护，添加删除普通用户，修改自身信息，进行业务控制和流程跟踪；F类角色应赋予档案业务规则库的操作人员，他们负责对档案业务规则库进行具体的后台数据管理和安全维护。基于角色的规则权限设置如图5-9所示。

5.5.2 系统安全管理

档案业务规则管理系统在构建、管理和日常维护过程中，始终存在安全隐患和风险。如果不重视系统的安全管理，安全防护能力无法提高，那么所谓的系统的结构合理、作用全面、功能强大等优势也只能是极易遭受安全威胁的“一纸空谈”，在实际的档案工作

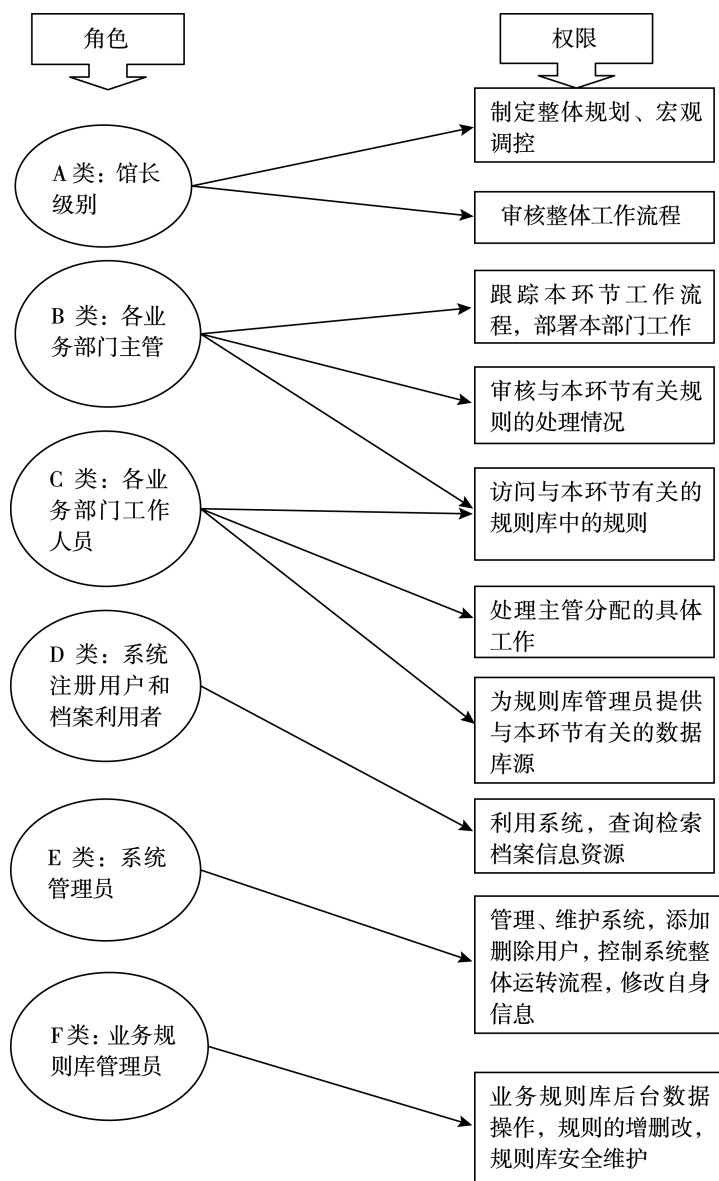


图 5-9 基于权限的档案业务规则的角色设置图

中根本无法充分发挥其作用。

档案业务规则管理系统的安全管理主要涉及三个方面：第一，规则库中的规则安全。随着档案信息资源数量的不断增加，与之进行匹配的规则库中的一系列规则也必须随之变化，比如规则的更新、增加、删除、修改等操作。但是，规则库中的规则也并非可以随意更改，因为它的变动涉及与之对应的若干词群系列，这些词群又与若干档案信息进行过匹配，“牵一发而动全身”，规则的随意改变最终影响档案收集、分类整理、鉴定、著录、保管、利用和编研等诸多环节工作的正常开展，因而必须对规则的安全性和相对稳定性给予足够的重视，只有拥有相应规则权限的角色才能对规则库中的规则进行访问或操作。第二，归档过程的安全性。利用档案业务规则管理系统实现已处理完毕的文件信息的归档，这一过程充满风险和各种不确定性因素，任何一个细节问题都有可能导致整个归档过程出现偏差，因而应当设置系统校验机制，对利用档案业务规则管理系统进行档案的收集、分类整理、鉴定、著录、保管、利用等环节进行实时跟踪和审核，以保证归档过程的安全性。第三，档案数据和系统安全。在利用档案业务规则管理系统进行归档工作之后，保存下来的档案信息数据的安全性就成为需长期重视的问题，同时，一旦档案业务规则管理系统与互联网进行连接，极有可能遭受外来攻击和计算机病毒的威胁，所以应当指定档案信息和档案业务规则管理系统安全策略，并采用访问控制、身份认证、信息加密、数字签名等多种计算机信息技术，以保证档案数据和系统的安全。

6. 业务规则在档案信息 资源管理中的应用

国家、省、市等档案主管部门根据档案的类别出台了一系列相关的法律、法规及指导意见，并会同国家标准部门颁发了一批有关档案的国家及行业标准，各立档单位也纷纷结合本单位职能与业务情况，制定了专门的归档范围与保管期限表。业务规则是业务过程要求与指标的最好体现形式，而这些法律、法规、标准、制度是它的基础与主要来源。在业务过程中，将档案业务规则与 workflow 相结合可以大大提高档案业务流程的抽象性，在业务系统中载入这些业务规则，更可以使开发的应用系统具有高度的可配置性。

一般而言，档案业务工作主要包括档案的鉴定、收集、整理、保管、统计、编研、利用等。对于档案行政管理部门而言，还包括对辖下各单位档案信息资源管理的监督指导、业务培训与达标评估。虽然，目前借助计算机来管理档案已经日益普及，但在常规的业务工作中，工作人员仍需借助于档案管理体系中的法律法规、标准规范、指导性意见以及成熟的应用经验等，对各环节的业务操作进行人为分析和判断，计算机几乎没有参与辅助决策。

如何让计算机识别档案业务开展所依托的法律法规、标准规范、指导性意见，智能化或半智能化参与档案业务人工决策，将档案工作人员从繁重的档案收集、整理、鉴定等工作中解放出来将成为档案领域一个全新的研究方向。基于业务规则技术，可以将档案管理体系中的法律法规、标准规范、指导性意见以及成熟的应用经验，通过业务规则管理系统整理成计算机可识别的档案业务规则库（ABR），利用规则推理引擎（RM）进行分析判断。

本章的研究对象具体指基于业务规则的电子档案管理。电子档



案又称数字档案，是在数字设备及环境中生成的，以数码形式存储于磁带、磁盘、光盘等载体中，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档案。电子档案与其他类型档案的区别在于记录方式和载体形式不同，它们在支撑业务活动和传承社会记忆方面的功能则是一样的。业务规则技术主要运用于电子档案鉴定、收集、整理三大环节。此外，借助业务规则库及其引擎推理技术还可以实现自动监督、评测档案信息资源管理效果和培训档案工作人员业务知识、技能的功能，此类档案信息资源管理包括监督指导和绩效评估系统，以及档案业务知识与技能网上学习和测试系统。

6.1 档案信息鉴定

所谓档案鉴定工作，就是档案馆（室）按照一定的原则、标准和方法，判定档案的保管期限的一项工作。简单地说，就是判定档案价值、决定档案存毁（生死存亡）的工作。档案价值鉴定工作的基本内容主要包括制定档案鉴定的有关标准、判定档案材料的价值、剔除无保存价值和保管期满的档案并按规定进行销毁等相应处理三个方面。

档案鉴定从来都是档案管理活动中最具有挑战性的环节，这种地位和特性在数字网络环境下不仅没有被削弱，反而得到了增强。现存档案信息价值鉴定的问题主要表现在：档案数量庞大，鉴定工作量大；档案价值判断主观性太强，缺乏规范化标准化，档案保管期限划分不科学；档案鉴定工作人员尤其是鉴定专家的宝贵经验多留存于大脑中，没有固定保存下来，无法有效地指导或被其他档案人员学习并用于开展鉴定工作。此外，每年大量的档案接收进库和到期待鉴定档案的有增无减，从长远来看必将造成档案信息库存压力的不断增大。与传统档案鉴定相比，基于业务规则的档案信息价值鉴定创新之处主要表现在对档案保管期限的划分、档案密级的鉴定以及档案期满鉴定环节的自动化控制。

6.1.1 基于业务规则的档案鉴定标准

档案的价值是客观存在的,传统的档案价值鉴定是人们对档案价值的认识和评价,带有很强的主观性。为了保证鉴定工作的质量和效率,可利用先进的人工智能技术,将固化于档案鉴定专家头脑中的隐形知识和档案鉴定的固化规则集中起来保存于规则库中,再通过规则引擎进行智能判断,从而真正实现档案价值判定的客观性、可靠性和准确性。为了达到计算机自动判断档案价值,首先必须要明确档案价值鉴定的标准,对于电子档案来说,价值鉴定标准主要包括来源标准、内容标准、形式标准、社会需求标准和技术鉴定标准。

(1) 来源标准^①

档案来源是指档案的形成者及文件的责任者(作者)和立档单位,来源鉴定标准是一种基本的档案鉴定标准。通过来源判断档案价值,一般反映在根据文件的发文字(号)、收发文单位、作者、主/协办单位规定、抄送范围等来辅助判断。

①分析文件的价值,应站在本机关的角度。一个机关的档案是否齐全完整,能不能反映本机关的历史面貌,在很大程度上是看本机关制发的重要文件是否保存了下来。除了保存本机关制发的文件,文件的保存还应包括外机关制发的文件,这一部分的文件判断较为复杂,主要看它在本机关的承办处理情况以及与本机关的关系:是针对本机关需要直接承办并产生的文件还是非针对本机关的文件;是直属上下级的文件,还是同级、非直属上下级的文件。例如,上级领导机关、业务指导机关印发的关于本机关主要业务的意见、规定、标准等的保管期限一般为30年,如果是上级机关制发的普发性不需本机关办理的文件材料,或任免、奖惩非本机关工作

^① 陈智为,邓绍兴,刘越男.档案管理学(第3版)[M].北京:中国人民大学出版社,2008.



人员的文件材料，只用于供工作参考的抄件等都不属于归档范围。另外，凡是记述和反映本机关主要职能、中心任务、基本情况方面的档案材料，对本机关、国家建设和历史研究有长远利用价值的档案，应作为永久保存；反之，可作为定期保存。这些都可以从分析文件的发文单位、收文单位以及文件抄送范围等判断出此文件对于本机关的重要程度，机关可依据“××局（单位）主要任务、内设机构和人员编制规定”等的规章制度判断机关之间的关系以及它们之间的业务往来内容及密切程度。

②分析文件的价值，应看立档单位在社会上的地位和作用，采用职能鉴定法。一般来说，职能和地位更重要的立档单位所形成的档案全宗及其所属档案文件，从总体上看具有较高价值。如党中央、国务院及各部委、中央军委及各总部、人大常委会等高级领导机关、地方上的党政领导机关所形成的档案，在政治、经济、科研等方面的价值要大一点。从人物全宗来说，党和国家领导人、著名科学家、社会活动家等知名人士形成的档案价值也相对较高。将以上鉴定标准分类、分条保存于档案业务规则库中，当立档单位符合上述要求时，就应适当增加保管年限，甚至考虑永久保存。

③在本机关制发的文件中，具体的撰写者，制发机构也对档案价值产生影响。机关领导人、决策机构、综合性办公机构、主要业务职能机构、人事机构、外事机构制发的文件大多比较直接地反映本机关主要职能活动和基本情况。当文件制发机构或撰写者符合以上情况的具有长久保存价值的文件比例较高。

（2）内容鉴定标准

档案内容鉴定标准是判定档案价值最重要的方面。一般通过文件标题、关键词、主题词、摘要或全文来分析判断档案内容价值。

①判定文件内容的重要性。归档文件内容为“重要的”，其保管期限相对较长。“重要的”主要是指四种情况：一是反映本机关的主要职能，其中以反映本机关职能的典型性问题尤为重要；二是反映本机关基本历史面貌，如反映机关辖内工作全貌、本机关工作变化、本机关工作发展等；三是反映本机关经济关系，如本机关的

房产、物产及其他债权债务文件材料；四是涉及个人利益，如本机关人事任免、调动、工资等文件材料。

②分析文件内容的独特性。凡是具有本机关、本系统、本地区特色的档案，以及特殊事件、特殊产品、特殊人物、特殊成果和具有开创意义的新人、新事、新方针、新政策方面的档案及各种特色档案，全应在归档范围内且保管年限较长。在实际鉴定过程中，可以根据规则库中定义的特色档案规则对文件内容进行逐件匹配，从而分析出文件内容的独特性，初拟出保管期限。

③分析归档文件利用时效性。档案价值的时效性，表现为档案可以在不同时期满足人们不同需要的阶段性。如条约、契约、协定、协议书、合同、法律、指示、方针政策、规章制度、规划、决算等文件材料，在一定时间和条件下，具有行政上和法律上的效力，应予以短期保存。当这类档案有效期满后，通常再自动增加相当于有效期的时间，以备查考。

(3) 形式特征标准

电子档案的形式特征主要是指文件的名称、作者（责任者）、文种、文本等（其中文件名称和作者/责任者已在前面详述过），这些形式特征在某种情况下，对档案价值产生影响，但其影响仅是辅助性的，重要的判断依据还是看文件主要内容。

①文件的文种。例如：具有权威性和重要性的决定、决议、命令、指示、条例、纪要、报告等比日常处理事务的通知、简报、来往信函等的保管期限要长；会计档案中的决算报告、账簿比凭证重要；教学档案中教学计划、教学大纲比教学日历、课程表价值要大。

②文件的文本。同一文件在撰稿、研制过程中可以形成各种稿本，电子文件也不例外。正本具有标准的格式，有机关的电子印章或负责人的数字签名等，其元数据记录的签署过程是较完整的，具有法定的效能和凭证作用，价值最大。相比之下，草稿、草案的可靠性要差一些。



(4) 社会需求标准^①

电子文件可以通过网络为全社会提供利用。在实际工作中,各档案馆(室)可根据电子档案借阅和利用统计情况,分析各时期、各类电子文件利用重点,预测社会需求的发展趋势,正确指导电子档案价值鉴定工作,纠正和补充由于个人的知识差异和工作经验造成的对电子档案价值鉴定的偏差和误判。基于业务规则的档案信息资源管理系统要求可自动通过分析电子文件的利用率来分析档案的价值,及时提醒调整不适合的鉴定规则,确保有价值的文件得到妥善保存。

(5) 技术鉴定标准

前文已经说过基于业务规则的档案信息资源管理主要针对的是形成于计算机和信息网络的电子档案,这里需要说明一点,就是电子文件的鉴定与传统纸质文件鉴定有一个很大的区别,即电子文件不仅有价值鉴定,还需要进行技术鉴定。电子文件的价值鉴定确定的是电子文件的保管期限,电子文件的技术鉴定是判断电子文件能否有效地按保管期限进行保存。这主要通过电子文件元数据介绍的技术状况和达标的技术要求相对照,如果电子文件所依赖的技术没有达标,无法支撑到需要保存的年限,技术鉴定就要提出迁移方案,保障电子文件在保存期内可以继续利用。

元数据是管理电子文件的命脉^②。元数据记录的完整性检查是电子文件技术鉴定工作的重要内容。档案元数据是一个描述过程的变量,它描述对象客体本身、对象所处环境及关系。不管是广义还是狭义的档案元数据,它所描述的内容有以下三个方面:一是内容信息,指那些反映了内容特征的信息,如标题、档号、分类号、主题词等;二是背景信息,指有关形成文件、维护文件的管理背景,

① 冯立. 电子档案价值鉴定及制约因素 [J]. 兰台世界, 2008 (1).

② 谭琤培, 章丹. 档案元数据在电子文件鉴定中的运用——元数据研究之三 [J]. 浙江档案, 2002 (6).

它包括形成文件的机构及其职能、形成文件的业务活动、信息系统的背景及文件使用情况等，如应用软件、文件形成机构、文件使用权限、文件的保存与处置决定等；三是结构信息，概括地说就是数据间的关系，包括电子文件的物理结构和电子文件的逻辑结构，如段落层次、文体、发（收）文者、发（收）文者地址、称谓、为显示而链接的数据间联系等。由于电子文件本身的特殊性，决定了其对元数据有极强的依赖性，因此，档案元数据的内容对电子文件鉴定具有特殊的意义。

6.1.2 基于业务规则的档案价值鉴定主要内容

（1）归档鉴定

①保管期限的划定。根据前端控制理论，对于电子文件来说，原来纸质文件鉴定中的某些标准要素和决策控制手段要提前到电子文件生成阶段，并主要依托于档案规则库内有关档案价值鉴定的大量规则及规则引擎的智能分析功能。文件一旦生成，系统即将文件元数据与鉴定规则进行对照判断，通过推理引擎，判断出其保管期限。由于档案属性特征的各个方面是互有联系、不可分割的，规则引擎在分析档案价值时，会综合地分析和考察文件各个方面的特征和作用。总的来说是以电子文件内容为中心，全面分析文件所属立档单位、文件的责任者、产生时间、名称、可靠程度、有效性等诸多因素，全面分析文件特点，科学判定档案价值。

②公开或密级的判定。档案公开和密级的判定是指对档案文献公开与否、保密等级及权限范围的划定。由于一个国家、一个组织或某个人物的有些活动在一定的时限和范围内具有不可公开性，从而决定了一部分档案内容具有保密性。此外，由于电子文件的特性和电子档案利用工作的需要，同一文件或档案在一定时限内的公开范围和使用权限是不同的。目前，我国对于电子文件密级的鉴定有宏观的原则性要求，但是，在实际工作中如何判定电子档案的密级和使用权限却没有具体的标准，这就给实际档案管理部带来很大的



困难与无序,出现“凡档案必保密”或秘密文件泄密现象。因此,制定细分规则并利用规则进行电子档案密级及使用权限的判定具有很强的现实意义。

基于业务规则的电子档案密级鉴定及权限设置,是在借鉴美国密级电子文件管理系统经验的基础上^①,依托规则库中关于档案秘密及权限规则,根据电子档案属性特征在电子文件形成之时就可先自动判定公开、权限或秘密等级,然后再针对不同密级档案进行加密处理,并且还可以基于事件、时间期限等进行自动分析解密。

(2) 期满鉴定

对于保管期限到期的档案,系统应自动提示用户以便下一步处置,如更改其保管期限,或作其他处理工作,如暂缓一段时间销毁等。依据档案的保管期限,系统自动监控到期档案,提示档案人员对其实施期满鉴定。此外,借助规则引擎的强大分析功能,可以依据最新实施的保管期限表实时改正以前错定的保管期限,重新给定档案的保留期限。

经过定性和定量化的双重把关鉴定,对失去保存价值的档案应进行销毁或处理。依照规则编制档案销毁清册,其登记项目有顺序号、案卷或文件的题名、起止日期、号码(案卷目录号、案卷号或文件字号)、数量、原保管期限、销毁原因、备注等。

(3) 进馆鉴定

具有永久保存价值的电子档案在形成机构中保存一段时间后,应向相应的国家或地方档案馆移交。对于这些电子档案,系统应比对相应进馆规则,从保存价值、完整性、可读性等方面进行复核后再接收进馆。

^① 张正强,崔屏.美国对密级电子文件管理的要求及我们对我国现状的思考[J].兰台世界,2004(9):16-17.

6.2 档案信息资源收集

各立档单位在工作和生产活动中形成的文件，是由各单位或单位内部的各个机构分散形成的，数量很多，而查找利用档案，则要求一定的集中。为了解决文件形成后的分散状态和利用要求集中的矛盾，就需要将分散的文件经过挑选，按照一定的制度集中保存起来，这就形成了档案的收集工作。

档案信息资源收集是档案工作的起点和基础，这项工作做得好与坏会直接影响到档案的数量与质量，因此，要把档案这项工作做好，首先要做好档案的收集工作。档案收集工作可以分为被动收集与主动收集两个方面。被动收集是指单位档案室按照归档制度对各部门应归档的文件材料进行接收或者是档案馆对辖下各单位长期永久的档案进行接收；主动接收是单位档案室或者各级各类档案馆根据归档要求或者业务活动如编研、利用需要，主动向有关各方索取或征集档案信息资源的活动。

6.2.1 档案移交（被动收集）

档案被各部门接收进档案室或者被接收进馆时，应该收集齐全，为后代积累完整的档案史料。归档制度与保管期限表是档案收集工作的一个重要依据，也是档案收集齐全、完整、准确的一项重要保证。然而，在传统的档案收集过程中，由于制度存在于人脑中，常因档案人员的业务水平有限造成对应归档的文件材料范围划定不准，档案管理部门在复核接受的档案资料时，时常发现有文件漏归、错归或者归档不齐全，档案收集质量难以保证。此外，庞大的档案数量和不断重复的接收鉴定工作也使得档案人员工作量急剧增加，档案接收工作始终处于被动局面。

采用人工智能的方法，将“××单位/机关文件材料归档与不归档的范围”等有形的业务规范、制度等进行细化和量化、分析、



排列、组合,将其抽象为归档范围规则库,并将辅助判断的各类因素如机构职能、人员组成,以及存于个人头脑中的规则与经验等加入规则库中,实现归档标准层次化,并利用规则推理引擎进行全自动或者半自动化分析、判断,对保证档案移交完整、准确,减轻档案人员工作负担,具有十分重要的意义。

本单位档案收集工作易受档案形成者个人主观意识影响,因此,在档案收集工作前,档案人员就需要对单位各部门进行综合研究、业务梳理和统筹规划,并根据各部门业务与职能制定明确的归档范围。因此,各部门的归档范围可作为档案接收自动检测的主要规则来源,同时,各部门的业务实际、人员状况、特色工作都可以成为该部门归档情况的规则源之一。而档案馆在接收管理辖下各单位的档案时,则以整个单位的职能和业务来考虑归档范围。在前面我们已经讲述了归档范围的鉴定,这里我们主要考虑归档档案的完整。

(1) 档案完整性控制规则

档案从立体结构来看,包括档案个体、同案档案、全宗档案三级结构,因此,在进行归档范围鉴定时,可以从档案个体、同案档案、全宗档案三级结构来检测,并根据各级结构检测结果提出修正建议,从而有效确保档案文件收集完整、准确。

①档案个体完整。档案文件个体是档案库藏的细胞^①。只有实现单份档案文件的完整,才能为整个档案的完整和优化奠定良好的基础。要保证单份档案文件个体完整,可以采用页码检测法、关联分析法等检测等。如利用页码可以检测电子档案或电子文件是否缺张少页,利用关联分析可以检测到该电子档案或者电子文件是否遗漏了附件。

②同案档案完整。同一项社会活动的档案应该是一个有机整体,收集来的档案在种类、内容方面应当齐全完整,如经受理的一

^① 陈智为,邓绍兴,刘越男.档案管理学(第3版)[M].北京:中国人民大学出版社,2008:78.

项业务、一个工程项目、一种产品等相关的档案文件要齐全。对此,系统可采用同案类比法、关联匹配情况等进行判断。如同类的会议产生的一套档案组成结构类似,可采用同案类比法即可对比相似会议材料是否有缺漏、可采用发文文号连续性或者发文总数判断发文文件是否完整。

③全宗档案完整。一般而言,各部门属于归档范围内的档案文件个体、同类案件完整后,全宗文件的收集基本可以保证齐全,但为保险起见,档案人员通常需要采用规则分析法或人工干预的方法进行二次检测,确保整个全宗档案接收的完整、准确,如可采用关键字分析法、规则匹配度等来进行二次验证。

档案馆在对辖下现行机关、企事业单位和撤销单位的永久、长期档案进行收集时,同样可以依赖于该单位的组织机构、职能、特色业务、所属系统、成员,以及业务规则库设定的规则等来分析档案个体、同案档案、全宗档案的完整与否,如可根据其机构规模判定其职工名册、干部名册是否完整、工作总结与工作计划是否归档齐全等。要保证收集的档案能够反映一个地区、一个部门、一个系统、一个专业、一个单位的历史脉络^①。

(2) 档案接收即时检查系统

利用研究得到的档案鉴定标准和档案完整性控制原则,可以进一步对电子档案或电子文件移交准确性鉴定与完整性保障进行相关研究,开发一套档案接收即时检查系统。将档案鉴定标准与档案完整控制原则中涉及的各项标准来源、标准要素、决策点、关联因素等进行整理和抽提,形成一套较为完备的业务规则库,利用相应的决策、控制办法绑定至该系统,从而可以建立一套基于业务规则库的档案接收即时检查系统,通过知识规则推理引擎,系统可以根据档案业务规则库对各立档案单位的档案信息,对录入或导入的档案信息进行检查,系统自动提出整改意见或验收意见。

^① 陈智为,邓绍兴,刘越男.档案管理学(第3版)[M].北京:中国人民大学出版社,2008:78.



在收到移交的档案后，系统可根据归档文件对标题、文种、关键字、责任单位来自动匹配归档范围规则库，从而智能化或者半智能化实现归档范围监测，来判断接收档案的真实性、完整性和准确性，并根据归档监测结果反馈监测报告，供科室人员进行修正与补充，从而保证归档数量充分、范围规范、信息完整。档案收集工作内容主要包括档案室对本单位各部门档案的接收与档案馆对辖下各级单位档案的接收。

6.2.2 档案采集（主动收集）

大量电子文件的产生使得现有的档案收集工作变得更加棘手。人们逐渐认识到，要确保档案信息收集完整、齐全、及时，档案收集应积极转变方式，实现机制创新，要变被动、滞后服务为主动、超前服务，在保证正常接收进馆档案的进度、质量和规模的基础上，坚持走出去主动收集、征集档案信息资源。

档案信息雷达借助于档案业务规则库对单位业务系统、外界网络、计算机的物理位置进行实时监控，及时把最新的已处理完的业务信息、物理位置信息和所需要的网络信息采集到指定数据库进行归档或者再利用。档案信息雷达在档案采集中的应用实现了电子文件与电子档案信息资源的自动搜索采集，大大增强了收集工作的主动性和收集效率，并且有效确保了收集的准确性。

（1）档案信息雷达的技术规则

档案信息雷达技术规则是档案归档鉴定的一个逆向规则。档案鉴定是根据归档范围等标准对已经收集到的信息进行范围鉴定、验证的过程；而档案信息雷达是根据归档范围等标准或者潜在需要主动获取数据的过程。

利用档案信息雷达技术实现对信息的定向、自动、智能收集，极大地提高了组织对外部信息的获取效率，降低信息采集成本，全方位掌控环境脉动，有效提高各个组织的快捷反应效能。此外，利用档案信息雷达实现对网络信息的定向采集，弥补了互联网搜索引

擎在专业需求方面的欠缺，通过聚焦定制采集源保证网络信息的实时性和全面性，并可以实现本地化存储，智能处理及再发布功能。其基本目标是满足机构用户及时跟踪和监控互联网信息，从而建立即时信息资料库乃至提供增值信息服务的专业需求。

如何通过业务规则制定，采集到所需的有效信息是档案信息雷达技术的关键所在（见图 6-1）。我们根据档案信息采集的目的和内容，简单分析如下：

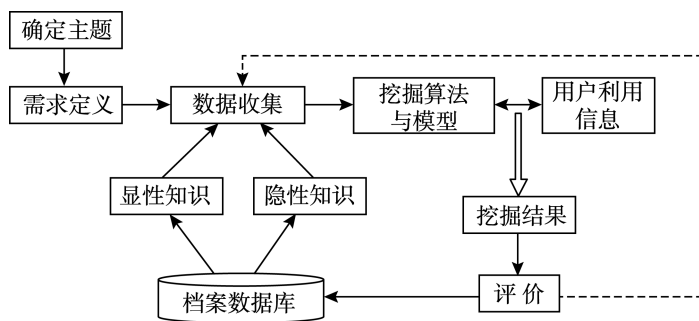


图 6-1 档案信息雷达技术采集信息流程

①采集与本单位密切相关的固化信息。它指的是与本单位相关的显性知识，即记录于一定物质载体上的知识，包括已数字化的馆藏资源、编研成果，业务系统中存在的办结的电子文件，与档案工作相关的各种法律法规、规章制度、行业标准等。一般而言，一个单位应该归档的档案可以根据该单位的职能、业务情况来判断，可以包括对本机关内部的、上级的、下级的、合作伙伴的、竞争对手的，所属的行业、专业或系统范围（如学校属于教育系统）进行判断，具体到档案的鉴定标准上则可以根据文件的标题、文种、主题词、责任单位、收文单位、抄送单位等进行判断。

②采集单位内部存在的智力信息。它指的是与本单位相关的隐性知识，是管理人员、技术人员、业务人员、政策法规研究人员等头脑中所储备的大量非编码智力资源，包括各种管理方法、计算机处理技术、处理问题的能力等。由于人是知识管理的核心，是知识



管理中最活跃的最主动的因素，这部分知识的挖掘将逐步成为档案采集的重点。目前已有不少专家针对该问题进行深入研究。

③采集与用户利用需求密切相关的。采集信息的目的主要是为了更好地提供利用与服务，因此，在实施采集时，要针对不同的部门和个人调查了解，将能够帮助用户解决现实问题，满足学术、科研、生产等需求的信息率先采集或者抓取过来。另外，也可根据档案利用人员提出的利用要求、访问内容、利用频率、访问时间等不断完善和增加新的利用需求，并将其制定为可为系统识别的业务规则。

(2) 档案信息雷达的应用

档案信息雷达在档案信息资源采集方面主要用于业务信息系统、网络信息抓取、物理位置数据抓取三方面。

①业务系统信息抓取。业务系统中含有大量重要的档案信息资源，但对于这些业务系统来说，只有解决了异构数据库的数据交换问题，才能顺利实现归档。目前，许多单位对业务系统要不然只进不归档，要不然则采用有限用户自主上传或者管理员上传方式来实现业务信息的归档，这种归档方式过多依赖于人工操作，且资源来源范围有限，难以实现业务信息的及时归档和档案信息资源库的动态持续更新，也制约档案信息资源的更新、利用效率，使利用者不能及时得到所需资源，从而影响利用效果。

业务系统数据抓取是基于导出接口，实现与第三方应用系统对接，并利用接口间设计的各项通信协议与绑定的各项规则，对业务系统进行定期扫描，从而可以将最新办结的业务信息系统自动采集并被自动导入档案管理系统中，确保实时归档（见图 6-2）。

②网络信息抓取。网络信息抓取是针对不同行业用户的实际应用，以 Internet 信息抓取为目的的跨平台软件，可按需且方便地指定需要监控采集的目标站点或频道，并设定监控更新的时间周期，包括定点执行、更新隔离，主要用于指定网络信息的抓取，实现在用户自定义的规则下，自动从互联网上的 HTTP 服务器上抓取指定信息的功能，并可以用于内部使用或外网发布，从而快速实现信息

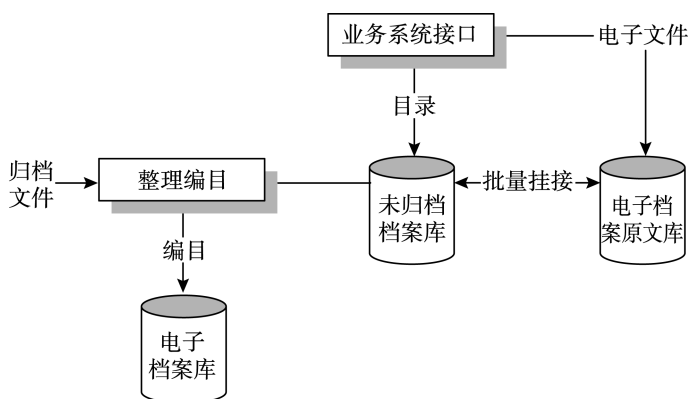


图 6-2 业务信息系统数据抓取

的全面共享。一般网络信息抓取要求可以支持各种标准格式信息资源的采集，如 HTML 页面、文本信息、图片等。

网络信息抓取软件实际是属于网络爬虫一类，只不过不像爬虫那样是通过分析网页来进行数据索引采集的，采集软件是通过用户定义的规则来进行数据采集的。单位的历史网站网页作为单位最重要的档案信息资源，应用档案信息雷达技术采集网站档案已经成为一项必不可少的档案工作。

③物理位置档案信息资源采集。物理位置档案信息采集是指档案信息雷达通过扫描组织或个人电脑，根据设定的规则，自动抽取电脑内需要归档的组织或个人信息。这种主动采集的方法适用于档案室主动收集各科室的定期更新的资料，同时也适用于企业档案信息及时归档管理，为了使软件好用，在以上的内容基础上，采集软件还会增加一些更加丰富的功能，譬如定时任务，增量采集等。

物理位置档案信息资源采集工作，对档案归档完整、齐全、及时具有重要意义，尤其是在电子文件真实性、凭证性得到认可和有效保障的电子信息时代。



6.3 档案信息资源整理

6.3.1 电子档案信息预处理

由于通过档案信息雷达采集下来的数据格式各异，甚至还带有很多干扰的乱码，这就需要对其进行加工。最理想的方式是采集后通过用户设定的规则直接将数据加工成需要的格式，一般这种用户规则已经集成在档案信息雷达中，数据采集后后台可自动按规则实现过滤重组、自动去重等预处理。

(1) 信息过滤与重组

由于采集时会产生大量的“噪声”数据，我们要做的首先就是数据的清理和预处理，对收集到的数据进行加工处理和组织重构，如检查数据的完整性及数据的一致性，去除噪声或删除广告、版权、栏目等无效数据，填补丢失的域，去除空白数据域，考虑时间顺序和数据变化，找到数据的特征，用维变换或转换方法减少无效变量的数目，数据类型转换等，精确获取内容主体，构建相关主题的数据仓库，为下一步的数据挖掘过程提供基础平台，做好前期准备。

(2) 基于内容的自动去重

档案向来以“孤本”著称，所以人们认为其珍贵。若过多保存重份文件，就会给保管、保护工作带来困难，造成人力、物力的浪费。因此，预处理时，我们应当注意对全宗内、各全宗间的交叉、重复文件进行清理和剔除，力求做到一个全宗内不出现重份文件，档案馆内各全宗间少出现重份文件。

采用的去重技术，不是利用简单的规则判断，而是利用内容的相似性进行排重判断，相似阈值可调，准确性高，不会因为标题或

内容的少许变化而产生漏判,即使把标题进行了改头换面,系统也会正确判定。被检出的重复档案文件不会被立即清除,只是作为相似或者推荐确认进行显示,提供给用户参考。

6.3.2 档案整编

收集起来的档案内容十分庞杂,数量很大,处于相对零乱状态,不好管理和无法提供利用,需要加以系统化。为了解决收集来的档案的零乱状态与条理系统便于查找利用的矛盾,必须把这一类档案进一步分门别类,这就形成了档案的整理工作。

然而,对于日益增加的电子文件整理,如何采用计算机来辅助进行电子文件、电子档案的整理、整编已成为不少专家研究的热点。从计算机专业层面来讲,档案的整编最核心的技术便是聚类。聚类是一种常见的数据分析工具,其目的是把大量数据点的集合分成若干类,使得每个类中的数据之间最大限度地相似,而不同类中的数据最大程度的不同。因此在数据进行聚类这一过程中没有指导,是一种无监督分类。聚类分析是用数学方法研究和处理所给对象的分类以及各类之间的亲疏程度,是在对数据不作任何假设的条件下进行分析的技术。

档案的计算机整编类似于对象的自动聚类。自动聚类由计算机系统按照被考察对象的内部或外部特征,根据一定的要求(如类别的数量限制,同类对象的亲近程度等),将相近、相似或相同特征的对象聚合在一起的过程。常用的自动聚类方法有:关联词法、文献—文献相似矩阵法、聚丛法和因子分析法等。自动归类是指计算机系统按照一定的分类标准或分类参考,将被考察对象划归到不同类目的过程。目前常用的自动归类方法有语义分析法、语法分析法和统计法等。

然而,由于计算机自动分析主题等研究还没有取得实质性的进展,所以,现在自动分类大部分都建立在题名或文摘中关键词的基础上,它的缺点是不能准确地按文献主题分类。但据有关资料表明,专家的偏爱也常使其分类的质量与普通标引员的分类质量相差



无几,而自动分类现有的水平与之也差不多,然而其速度与规定性则是手工分类无法比拟的。因此,它越来越受到人们的重视,成为档案整编、情报检索中一个重要的研究与发展方向。特别是它与聚类检索的结合,将使其有更强的生命力。

6.3.3 档案著录

(1) 自动著录

档案著录是指在编档案目录时,对档案内容和形式特征进行分析、选择和记录的过程。通过它可以具体描述每份文件或每份案卷的特征,揭示其主题内容、科学价值,指明出处,区别相互之间的异同,从而便于档案人员和利用者了解档案,更好地管理和利用档案。

档案信息的元数据在档案管理、利用等方面都起着重要的作用。元数据的捕获对于电子文件的生成以及作为含有证据意义的原始记录的归档保存都具有重要意义。档案著录可以实现对档案文件自身元数据的详细管理,支持一定范围内的用户检索和利用,帮助利用者在案卷和文件夹中定位档案文件。

①元数据的采集内容。通过借鉴国外较成熟的元数据标准,如《通用国际档案著录标准》(ISDA),《档案编码描述格式》(EAD),《永久保护电子文件真实性国际项目》(Inter PARES)等,参考国内的《中国档案机读目录》、《档案著录规则》等规范,根据档案的形式和管理特点,来研究制定符合本单位实际的档案信息元数据内容。

元数据的分类方法较多,但档案信息元数据的主要内容分为标识信息、背景信息和结构信息3个方面。

标识信息描述分为案卷和文件两个层次,案卷级描述主要有全宗号、档案馆档号、立档单位档号、案卷名称、文件份数、文件页数、起止时间、归档时间、保管期限、密级、控制使用符等;文件级描述主要有文件编号、发文字号、文件题名、责任者、形成时

间、密级、主题词、文种、文件附件、内容描述等。背景信息描述主要有机构名称、部门名称、操作者、加密方法、版本编号、版本说明、文件格式、文件大小、存储位置、关联类型、关联描述、保管日志、使用日志、扫描单位、扫描人、扫描时间等。结构信息描述主要有载体类型、载体规格、操作系统名称及版本、数据库系统、信息处置平台、硬件配置、网络配置等。

②元数据的采集途径。根据数字档案信息的采集过程，可以明确元数据的采集主要来源于自动获取和人工录入两个方面^①，自动获取包括取自系统、取自环境、取自文件、取自工作流等。

取自系统是指将元数据捕获机制嵌入文件采集系统中，依赖文件系统的元数据就能够自动地被系统生成。例如，所有产生于一个特定系统的文件会有相同的技术描述。

取自文档是指元数据可以从文档自身获取，这是由应用程序决定的。例如，电子邮件的标题就包含了诸多信息，如发件人、收件人、发送时间和日期以及文件的主题。

取自环境是指元数据可以从计算机环境中获取。特别是文件的创建者和创建时间（对文件的证据地位非常重要）能够通过智能卡或者用户的注册时间，通过产生文件的计算机获取。例如，取自技术环境的元数据，可以反映出文件是单机上生成、局域网上生成还是互联网上生成，不同技术环境就会有不同的元数据；取自行政环境的元数据，可以反映出文件生成时所遵循的法律、法规、制度和标准等。

取自工作流是指工作流就是一系列的步骤：自动产生于工作流的文件拥有一个上下关系，如文件与工作流中其他文件的关系。工作流中一个阶段的元数据可以被工作流所携带并自动录入以后的文件系统中，如文件停留在某处并由相应人员进行处理时，系统会记录下工作人员对工作文件的操作过程。

③元数据自动著录系统。元数据的添加可以通过系统自动生成

^① 屠跃民，李婉月．关于数字档案信息采集的思考 [J]．档案与建设，2006（9）：17-20.

(通过环境、文件、工作流和系统等渠道) 和人工干预两种途径获取。为减少人工著录工作量, 加快档案著录工作进度, 档案自动著录逐渐成为档案研究工作的热点之一。档案自动著录是指利用定义的规则, 自动采用计算机技术抽取档案文件对应信息, 完成对档案文件(案卷) 题名、文号、关键字、成文日期等元数据的著录的过程。

借鉴国内外自动著录在图书、情报等的应用, 档案自动著录系统应该包括档案文本输入、著录词典、抽词子系统、规则库、综合与转换系统、输出系统等。

➤ 档案文本输入子系统

对档案文本进行自动著录前, 必须使之变为机器可读形式, 如 txt 或者双层 PDF 格式, 并按一定格式加以组织, 这就是输入子系统的任务。通常, 输入子系统处理的结果是将文本以数据库或文档中的记录形式存储在磁介质上, 每个记录中包括供著录用的若干字段或子字段(如题目、文摘、文本段落等)。

➤ 词典子系统

自动著录词典是存储在计算机系统内的一部或多部词表。因著录的目的要求不同, 词表的结构和组织也有很大差别。在赋词著录系统中, 词表通常是手工著录用的受控词表(如主题词表)。这种词表中的词之间具有整体—局部、种—属、同义和相关关系。在抽词著录系统中, 全关键词著录词表是一种禁用词表(亦称非用词表), 即词表中收入的词都不作为检索标志; 而主关键词著录词表中的每个词具有若干信息, 如词类、组配等。在著录过程中, 通过查找机器词典确定关键词与非用词、词组构成与切分。

➤ 抽词子系统

通过查找机器词典中存储的信息, 对输入的文本逐字逐句进行扫描, 抽出供综合与转换子系统处理的检索标志。对于拼音文字来说, 词间有空格分隔, 可按空格进行逐词扫描, 作抽词处理。对于汉语这样的拼音文字来说, 词间没有空格, 不能进行逐词扫描。因此, 汉语抽词子系统的首要任务, 是将连续书写的汉字文本切分为有空格分隔的词。汉语自动切分, 又称汉语自动分词, 主要是通过扫描取文本的一部分同词典比较进行分词。分词的方法主要有最长

匹配法、最短匹配法、设立切分标志法、逆向扫描二字前进法和非用字后缀表法等。

➤ 规则库子系统

规则库子系统是人工著录中选词知识形式化的规则集合。总结人工著录经验,人们发现,人工著录包括主题分析、选择表述主题的关键词和把关键词转换为规范主题词3个阶段。这些经验通过主题句法、频率法、概率法、加权法、语法分析法和语义分析法等方法加以形式化,用语义网、框架等知识表达方式构造成知识库。

➤ 综合与转换子系统

依据知识库提供的知识,对抽词子系统抽出的全部词语进行统计、分析和综合,选取表达主题的关键词,然后,按词典子系统的规范化规则,将选出的关键词转换为规范化词语。

➤ 输出子系统

将选取的关键词存入到文本记录的有关字段,并将记录输出到要求的介质上。在数字化加工的过程中,可通过OCR自动处理技术,获取文字信息,以支持全文检索功能的实现。通过区域OCR识别技术,通过元数据提取,档案元数据可自动著录,大大减轻了档案著录工作量。

虽然,近年来自动化著录技术得到了长足的发展并取得了一系列应用成果,但由于档案文本中大量的隐形信息难以通过计算机进行准确分析,因此还不能完全依赖于自动著录。利用人工干预强化著录结果,在著录后进行人工著录审校,同时将审校中的问题记录下来,进行分析、论证,并不断改善、修正、充实业务规则库,将会对档案著录质量会产生重要的影响。

(2) 著录规范监测

档案著录质量是档案数据库和数字档案的命脉,标准化是实现档案工作现代化的基础之一。随着档案信息化建设的不断深入,人们对档案数据库和数字档案的质量要求越来越高,当前正在大幅开展的数字档案馆(室)建设更呼吁科学、系统的质检体系,对于大量的著录工作,具体存在哪里常见的著录问题,应该如何去避免



或者通过质检得到修正,实现档案的标准化、规范化著录,是人们最关注的问题之一。

①常见著录问题。档案著录内容通常包括著录原则、著录项目、标志符号、著录格式、著录用文字、著录项目细则等,手动著录受人员水平、熟练程度影响,在数字档案馆建设中存在一定质量偏差,主要包括以下几种:

➤ 著录项目不全、不准

著录内容不准确、不齐、不全,用词不规范。如文件库的档号、年度、分类号、保管期限、成文日期、文件标题、责任者等必填字段出现空值;

➤ 著录格式或者内容错误

书写形式、标点、格式等不规范,出现非法数据等。如成文日期规定是8位阿拉伯数字,却写成文本;数据不规范,如档号位数不统一,001-2005Y-001,全宗号001改成01就属于不规范数据;四是档号为唯一值,却出现重复值等。

➤ 标志符号不规范或者符号遗漏

找不到符号输入方法省略标志符号或者对标志符号识别错误。

➤ 著录用文字不规范

著录用文字不准确,滥用省略、随意简称等。如“中国共产党中央委员会”简称为“中共中央”、“广东省人民政治协商会议”简称为“广东省政协”等。

➤ 数据表之间的关系没有提示出来

如文件库与案卷库之间的关联,一个案卷具体有哪些文件,有多少件,有多少页,保管期限如何,这些都应自动生成。某一个专题,如省一级馆验收,既有大量的纸质档案验收材料,又有会计档案(预算、凭证、账册和报表)、照片档案(验收时的照片)、声像档案(验收时形成的音频、视频档案)、磁性载体档案(刻录成光盘的数字照片和数字声像)等,以前的数据库往往没有将这些关系提示出来。

➤ 维和度量不规范

比照多维数据仓库的要求,很多档案数据库的维和度量缺少或

不规范。以文件库为例,全宗号、单位类别、分类号、年度、保管期限就是5个维,而档号、页数等就是度量,页数属于数值度量。但实际上,文件库中往往没有设置这样的元数据,数据类型也不正确,如将页数设成文本而非数值。

② 档案著录规范控制与监测。我们发现很多数据不规范和内容不准确是由于软件设计时对输入没有严格的约束引起的。因此,好的档案管理软件必须能规范用户的数据录入行为。

在设计每种库的元数据时,必须对主键、外键、必填字段、维、度量、有效文本、数据类型等作明确的规范^①。为保证著录规范,可采取档案著录规范协助人员进行著录监测。

➤ 输入的软件控制

主要包括以下三个方面的要求:一要保证域的完整性,即输入的数据必须是合法数据和规范数据,同时要充分考虑是否允许空值。合法数据如成文日期为8位阿拉伯数字,20060518,如果输入7位或9位数字就会报错,提示用户输入合法的数据。规范的数据就是要保证数据的一致性,如档号统一由全宗号(3位)—分类号(1位)—年度(4位)期限—件号(4位)构成,如001-1-2005Y-0001,如果年度写成04就是不规范数据。以案卷库为例,档号、案卷题名、责任者、件数、总页数、起止日期、保管期限是必填字段,如空值,就不允许用户继续录入。二是实体完整性,每个数据表都有一个主关键字,在档案数据库中一般将档号作为主关键字,是唯一值,程序应严格控制档号不重复。三是引用完整性,案卷库与文件库的关联,可让用户清楚地看到一个案卷都由哪些文件组成;文件库与专题库、媒体库、员工库之间的关联,可让用户看到同一专题在不同种类数据库之间的联系,方便用户了解事件的真相和完整面貌。

➤ 校对的软件设计

包括双机录入校对、档号重复项校对、数量校对、内容校对、

^① 陈金冠,创新质检方法 规范数字档案——数字档案馆(室)质检方法研究[J].浙江档案,2006(9):34-36.



规范性校对等。这些均可通过程序，自动找错，实行智能化校对。在微软的 Access 数据库中，可用查询设计中的查找重复项向导来寻找重复的档号；用查找不匹配项的向导实现双机录入校对的功能；用数据透视表和数据透视图来查看数据与实体的一致性，包括噪声数据；用随机查询的功能进行档案内容与实体的校对等。

➤ 质检的软件设计

著录规范质检系统对责任者、题名、保管期限等，根据已存在的档案业务规则库（ABR）自动进行检查。利用国家档案局颁布的行业标准《档案著录规则》，可以将这些规则分解、抽提为业务规则库，并将各项定义的标准、规范、文字、符号的规则与各著录项绑定，对表的定义中，主要突出表的名称与含义、字段名称、所表示的中文意思（中文含义）、字段类型、字段长度、数值型字段中小数点以后的位数（小数）、是否允许为空值（允许空值）、缺省值、是否作为表的主关键字段（主键）和是否作为外部关键字段（外键）等字段信息作为描述，相关联表与必要注释等内容说明，在表格中无值的单元格表明当前项目不具备该属性。

据调查，目前还很少发现成熟、专业的数字档案质检软件。好的质检软件应具备智能化、网络化、框架式、流程定制、自动统计等特点和功能。智能化体现在处处为用户着想，如数据和图像同时显示、流程单可网上填写和打印、数量自动统计、流程可根据需要自定义、多人操作和协同等方面。

档案著录规范监测可以采用著录同步监测提醒、著录后统一监测与核对两种方法辅助档案人员进行档案著录规范检查。

6.4 档案信息资源管理的监督指导和绩效评估

6.4.1 监督指导系统

经过十多年的档案信息化发展，我国主要发达地区对档案信息

化的建设达到了一定的高度、深度、广度与力度，通过集成不同的现代技术，出现了不同层次的档案管理系统（AMIS），涌现出一批数字档案馆、电子档案馆，基本实现馆藏数字化、利用网络化，逐步实现了计算机系统目录管理和全文管理条件下的纸质、声像、实物和电子文件的归档，有的还实现了网络化电子文件归档。

虽然，当前存在的 AMIS 基本实现了对已有档案信息的管理、利用、编研等环节进行管理，但是，归档管理模式并没有根本性突破，主要表现在，对归档部门的业务指导方法不得力，兼职档案人员的管理不到位；归档文件的质量管理处于薄弱状态，无法对以下问题进行全面系统的实时监控和综合评价，如各门类载体材料收集是否齐全完整、内容是否真实有效，整理是否规范，归档是否及时，以及各部门的归档难度、归档数量等。即便是各类综合档案达标升级的考核程序，也只能对已归档的材料进行抽查。而对于如何提高监督指导效率，进而提升归档效率，实现对文件形成与归档的过程、时间和质量的实时监控，以及建立对归档人员激励评价的科学管理机制，仍缺乏科学有效的解决方法。

为了解决长久以来监督指导工作不便、低效的问题，相关专家开始从技术和管理两个角度研究档案信息资源开发与管理问题，拟利用我国完善的标准规范体系和先进的业务规则技术，建立一种新型的档案监督指导管理体系。将档案主管部门与指导部门的规范、制度、标准，以及立档单位的归档范围、保管期限表等抽象为档案业务规则库（Archives Business Rules Database，简称 ABRD），提供遵循归档管理的工作流程和内在规律，涵盖归档各项管理细节的质量、时间、数量、难度、行为等特征的全方位的质量控制和评价，利用人工智能的原理，通过强大的推理引擎（RM），借助计算机模拟人类的思维过程，对管辖范围内各单位已有的档案信息数据（ADI）进行智能分析与推理，提出建议及线索，以供档案管理部门进行进一步的重点检查与监督指导，从而帮助立档单位达到“档案收集齐、保管期限设置准、档案著录规范”的目标。

此外，为了更加及时地对立档单位的工作进行指导，并提出修改建议，可以利用基于 Internet/Intranet 的档案信息扫描技术，结



合档案业务规则库（ABRD），自动监督各立档单位的档案整理情况，发现问题自动提醒，并及时反馈，信息转入 Workflow，档案监督指导人员根据实际情况作出处理。

基于业务规则库的档案信息资源管理监督指导系统建立了一种新型的档案移交、接收模式，并做到档案实体正式移交前进行相应的非现场预检查，保证档案接收的质量并减少后续的著录量与数字化加工量；同时，通过网络来实现业务指导，打破了时空限制，节约了原来上门指导、电话交流所消耗的大量时间和精力，降低了信息传输的时间成本、人力成本，一方面使监督指导人员能够通过管理系统全面、及时地掌握各部门公文管理、预归档及归档情况，通过管理系统中的“短消息服务”，将预归档中存在的问题及时告知有关部门，确保归档后的文件不再返工；另一方面通过网上的监督和指导，有效地解决了上门指导来回往返效率低的问题。

6.4.2 绩效评估系统

档案管理达标升级是指在一定阶段内，以有关部门制定的一套既有定性要求又有定量考核的科学实用的标准体系，作为单位档案管理工作目标，并组织实施的一种管理办法。它是对档案工作进行整体建设，全面提高档案管理水平，逐步推进档案管理标准化、规范化、制度化、现代化的一项重要措施。

档案工作是政府、企事业单位的一项重要基础工作，实施档案管理达标升级是提高档案管理水平的有效而重要的措施。在我国，已经具备一个较为完善的达标升级的考核体系，然而，档案达标工作在实际工作中仍然是困难重重。首先，在目标管理等工作中，常常存在对档案部门的正常业务检查、指导不重视，对日常有关档案工作标准不贯彻，上级检查或进行目标管理时临时突击应付等，从而影响了档案工作的规模发展，影响了档案质量的规范和提高；其次，由于人员水平和业务素质的影响，对于主动申请目标管理的单位，档案行政管理部门往往需要在达标前进行若干次上门指导、辅助，提出改进意见，不断规范化、完善化；最后，由于达标管理未

真正得到重视，档案目标管理中存在许多较突出的问题，如档案构成材料的不规范，影响鉴定工作；编研资料少、利用率低；很多单位在目标管理时虽然编写了组织沿革、大事纪要、基础数字汇编等材料，但由于只为应付档案目标管理考评，编写质量不高，利用率低，经费投入不足，影响档案安全；缺少单独库房及相关装备和必要的调控设备，没有达到档案密闭、防光、防尘、防潮的要求，没有认真填写、查、借、阅记录，温湿度记录等；档案目标管理后放松了档案管理现象也十分突出。

为了加强档案目标管理工作的执行力度，确保档案行政管理部门能够不失时机地推进认定工作，同时为了使各单位能进一步明确达标管理要求，保证各单位档案管理做到规范化、标准化、现代化，可以尝试根据《市、县级国家综合档案馆测评办法》、《机关档案工作业务建设规范》、《城市建设档案馆目标管理考评办法》等测评办法，建立目标等级与考核要求库，并在此基础上建立一套新型的在线绩效评估系统，实现归档标准层次化，可以细化和量化到每一项归档要求，如归档范围、归档时间、归档人、归档文件的质量、案卷整理的质量、归档数量、归档难度等，组成一个权责分明、相互协作的质量管理评价标准。

在进行档案目标管理工作时，单位档案管理人员只需定期按照要求填写各类目标管理成果，如档案种类、构成材料、数量等，在线提交或者填写组织沿革、大事记、基础数据汇编等，同时可以按要求的填写或者选择档案库房管理已经具备的设施、设备及“八防”中具体已达到的要求与指数，并在线提交、导入或者填写档案管理制度以及相关的记录文件。

绩效评测内容与目标建立后，系统可定期根据档案业务规则库（ABR）对各立档单位的档案信息（AID），通过知识规则推理引擎（RM），自动生成绩效预评估报告，经过 Workflow 流转后产生最终的评估报告。将测评达标事后评价变为过程监控，实现重点工作的过程管理，掌握实时工作动态，提前预警问题。达标升级前，各单位都将积极开展一系列建设与整改活动，健全和完善各项档案管理制度。该系统则可实现模拟专家验收，发现问题及时整改，通



过自查，利于单位找到原因，也增强了信心。此外，采用科学的档案目标管理绩效考核体系，能够在同一平台上科学全面的对各党政机关进行综合评价，提供科学的考评方法，规避和解决考核过程中经常出现的各种导致评价不准的因素。

6.5 档案从业人员网上自学习系统

随着现代科学技术的发展和计算机、网络通信技术的广泛应用，档案种类和数量大大增加，档案管理工作领域在不断拓展，其信息化水平也在不断得提升。新环境下，档案信息化的顺利开展和档案高水平管理，要求档案从业人员所要掌握的技能水平也越来越高。培养一批专业化、复合型的档案管理人才成为迎接信息化、网络化社会挑战和机遇的必要准备。

档案网上自学习系统是档案从业人员进行自我学习档案管理知识和技能的最佳方式，其主要通过提供学习、试验和测试三个方面的功能，让档案人员自主学习并检测学习成果，一目了然自己的薄弱环节。业务规则技术引用到网上自学习系统的开发和设计中具有意义，本系统主要包括档案管理理论学习子系统、网上归档实验室和考核系统三个模块。

6.5.1 档案管理理论学习子系统

我国的档案管理模式从建立之日起就具备一个完善的管理体系，保障顺利、有序开展各级各类档案管理工作的规则很丰富、细致，不仅有最具权威的档案专门法律，还根据各地各类档案的特殊性制定了一系列的条例、规章、标准等。如国家颁布的《中华人民共和国档案法》，各省人大批准的各地《档案管理条例》，专门针对国家重大建设项目文件的《国家重大建设项目文件归档要求与档案整理规范》、机关档案管理必不可少的《机关文件材料归档范围和文书档案保管期限规定》等。这些都是档案人员需要掌握

的基础知识和开展档案工作的基本依据。

目前大多数档案人员都是通过培训和看资料（文件）来学习这些上岗知识。档案从业人员的这些传统学习方式存在着很大的弊端：首先，短期培训的时间和讲授内容都是有限的，这对于全面掌握档案业务知识是远远不够的；其次，档案人员自学习下发的文件，不仅效率低，而且还会抓不住要领，导致事倍功半的效果。构成档案业务规则库主体内容的法律法规、规章制度、标准规范等是网上自学习系统的强大资源后盾。

档案业务的实践经验，如组织内经验丰富的档案从业人员的知识总结、没有记录下来的惯例等，也是档案工作人员做好档案业务的重要知识。目前，这种知识的学习主要是通过员工之间的口口相传方式，不仅不规范，而且这些宝贵的经验知识会随着时间的流逝和人员的更换逐渐消失。将这些没有形成文字的经验总结、知识总结，通过访谈业务人员、开座谈会、问卷调查、跟班作业等方式进行调查总结，并将其文档化、结构化处理后编译到档案业务规则库中。这些隐形知识总结是自学习系统资源的补充，使得档案人员能学习到更加全面、准确的档案业务知识。

6.5.2 电子档案管理模拟实验室

随着现代办公信息系统和信息网络的普及，通过计算机网络产生和传递的电子文件量逐渐增多，鉴于记录与保存历史的责任和义务，以及利用的方便，电子档案的管理任务将会越来越重，这就要求现代档案管理人员要熟悉电子文件特性，能熟练操作电子文件（电子档案）管理信息系统。然而，现在的培训内容普遍缺乏实践操作训练，档案管理人员往往是只知理论不知实务，尤其是对于刚接触电子档案管理的新手，工作效率和工作质量都明显较低。在没有专业人员监督指导的情况下，新手在操作电子档案管理系统时出错了也没有提示，将会造成工作失误。这些电子档案管理仿真模拟训练的缺乏所造成的问题，在未来电子档案普及乃至成为档案管理主流的时候将会进一步暴露出来，造成工作上的很大不便。



基于业务规则的网上档案管理实验室模拟了电子档案从收集到服务的一整套工作流程,该实验室为学员提供了百分之百的仿真环境,其功能达到了实际工作中的一切要求。此外,利用业务规则引擎的强大功能,将学员每一步操作比对规则库中的相关专业知识,自动判断其操作是否符合档案管理规范,如果不符合相应规则将会以警示标志方式自动报错,且下一步的操作也将会自动失效。学员还可以选择帮助功能,这时计算机从规则库中调出本操作环节的正确规则,参考这些规则,学员可以轻松地学习本部分知识或者是进一步加深记忆。

6.5.3 考核子系统

社会的进步和发展离不开人才,同样档案事业的发展和档案管理水平的提升也离不开档案人才。合格档案业务人员的评判标准和严格的档案人员上岗考核是档案管理人才建设的保障,是规范档案工作、提高档案管理水平,乃至促进档案事业发展的基础,不容忽视。档案法律法规、档案基本理论知识、档案业务知识、档案信息化知识及计算机操作技能等都是每一位合格的档案业务人员必要的知识储备。而传统的考核体系包括的内容不全面,不能包括档案业务的所有知识;传统考核方式主要是考卷形式,形式单一,不能满足选拔信息化时代下现代档案管理人才的要求。

基于业务规则的档案考核系统,是集水平测试、模拟考试、专项测试、试卷编辑、考试管理等模块于一体的现代考试系统。该系统主要依托档案知识库内的各种规则和知识,考核题库来自于档案知识库中,其主要特点如下:

①根据知识库中规则种类,可以自动设定考试主题,可仅局限于档案业务的某一环节,如鉴定知识、档案整理知识,也可自动组配考察知识范围,而且具有综合考核功能,综合考核功能是考察考生能力最有效的一种考试方式,其中每一份考卷试题均涉及档案法律法规、档案基本理论、档案业务、档案信息化等知识,再加上以电子档案管理实践操作为考核内容的实验试题,不仅能考察考生的

理论知识，还能检验考生的实际操作技能，全面检验了考生的综合能力。

②本系统的第二大特点是考核系统可根据考生的答题情况，自动分析、总结考生成绩，给出分析报告。从分析报告中可以看出该考生的薄弱点，依据规则，拟给出下一步学习重点和学习计划。

7. 基于业务规则的档案 信息资源管理案例

7.1 案例单位简介——广东万维博通信息技术 有限公司概况

广东万维博通信息技术有限公司（原东莞市万维网络科技信息有限公司，以下简称“万维博通”）是一家长期专注于电子政务与档案信息资源管理的国家级高新技术企业，是广东省信息产业厅认定的“双软企业”，广东省科技厅认定的民营科技企业。公司已通过 ISO9001:2000 认证，同时，万维博通还是东莞市首家通过 CMMI（能力成熟度模型集成）国际认证的软件企业，是粤港关键领域重点突破项目的承担单位。目前，已有超过 2 000 家客户正在使用万维博通不同系列的自主知识产权软件产品。

万维博通自成立起就致力于档案信息资源行业软件的研究与开发，使软件不断往深度与广度发展，拥有多项行业专利、软件著作权，其产品“万维数字化档案网上管理系统”被广东省财政厅等六部门联合认定为省首批自主创新产品，享受政府优先采购政策。公司在松山湖科技园区设有“东莞市万维档案信息资源研究所”，并与武汉大学建立了“产学研”合作关系（见图 7-1）。

万维博通还是一家专业的档案系统集成公司，是东莞市首家备案的“广东省档案中介服务机构”。除提供档案信息资源管理系统（AMIS）外，万维博通还提供智能库房设计、档案管理体系建设、数字化加工设备配置、档案整理与数字化加工等服务，是国内较早

开展档案整理及数字化加工业务的企业。公司拥有一批精通档案业务、持证上岗的工作人员，有过百家档案整理、数字化加工、目标管理的成功案例。

通过近十年的技术沉淀与经验渗透，万维博通对电子政务与档案信息资源的建设有着一种独特的领悟，公司已从软件开发企业转型为行业管理咨询与实施机构。

7.2 万维博通档案业务规则管理系统介绍

7.2.1 系统特点

(1) 安装与操作简单方便

只需点击 .msi 安装文件或将安装光碟放入光驱，根据提示点击鼠标操作即可完成安装。安装完毕，在办公电脑打开 IE 浏览器，在浏览地址栏输入系统地址即可使用（见图 7-1）。

(2) 全面、智能、准确的档案业务规则控制

系统对档案业务规则控制，从档案著录到档案归档检测全程控制。业务规则存储在规则库中，独立于数据和程序。业务人员可以对业务规则进行增加、更新、删除等管理操作。档案业务规则在日常的档案业务整理过程中，不断累积、调整和共享，档案业务规则日趋完善，对档案业务规范更全面、合理。档案系统利用高效的规则引擎，调用庞大的规则库进行规范档案整理，实现档案规范化管理。

(3) 完善的审批流程

档案业务规范性，需要完善的审批监督机制。万维档案系统提供档案著录、自我检查、审批登记、审批办理、办事督办、审批检索、审批回馈等完善的审批流程。系统还提供工作提醒，确保档案



图 7-1 万维数字化档案网上管理系统界面

业务规范审批及时、全面。

(4) 继承一般档案管理软件的所有功能

可对各大门类档案系统进行独有详尽管理，如文书档案（包括立卷归档和单件归档）、特殊载体档案（包括照片档案、底片档案、音像档案及实物档案等）、专门档案（即用户根据单位实际情况，建立档案门类）另外，还包括会计档案、已改人员档案、诉讼档案等。

(5) 多全宗管理，支持网上档案移交

充分考虑档案管理的需求，全面使用 B/S 结构，使用 MS SQL Server 数据库，采用 ASP.NET C# 语言编写，主管单位与下属单位按全宗号可在同一系统整理档案，支持在线档案数据移交。

(6) 提供多种安全机制

①系统在数据库级、表单级、用户级采用多种存取控制和授权功能；

②在单位内部每个注册用户必须输入用户名和密码才可以登录系统；

③在政务专网为了防止其他单位人员进入本单位，可以借助密钥机制检测用户的合法性（类似于电子银行的模式）；

④支持 SSL2、数字证书（CA）等安全措施；

⑤完备的系统日志功能，为系统管理员提供监测系统运行的有力工具。

(7) 扫描操作简单轻松

①高速扫描仪最低每分钟双面 25 张的速度；

②选用万维公司“一键扫描”插件配合指定的高速扫描仪，更可大大削减操作时间，提高工作效率。

(8) 支持电子文件管理

用户可以通过扫描将纸质文件转换成电子原文上传到系统中，支持多种数字化文件格式，如图片的 JPG、音频的 MP3、声像的 MPEG、WMV、RM 及 OCR/ICR 后的原文，以便档案利用时更方便快捷了解全文具体信息。

(9) 超大容量管理

提供超大数据容量，超过 1 000 万数据条目的处理能力；文件服务器容量足够大时，影像文档容量不受限制。

(10) 支持远程档案查询利用

在单位内部有必要的安全措施条件下，利用 VPN 技术，可实现远程办公，如某位领导在外地出差，找一台可以上网的电脑即可实现远程档案查询，随时了解单位的情况。

(11) 强大的辅助管理，全宗卷、图书、资料、法律法规管理等功能

提供类似与立卷归档模式的资料管理方式。

(12) 提供专门档案自定义类别的录入整理方式

系统提供了专门档案自定义类别录入整理，各单位可根据专门档案的具体情况，决定采用何种整理方法进行整理，如教学档案可采用文书档案立卷归档的方法整理（见图 7-2）。



图 7-2 数字文书档案立卷归档方法

(13) 支持大事记、组织沿革管理

对于大事记及组织沿革的管理，系统提供两种管理方式，一种是以上传文件进行管理，另一种是逐项事件著录式管理，最后可输出统一格式的报表。

(14) 能自动生成档案编研光盘

系统对文件汇编和专题目录都提供详细管理，可自动产生编研目录册及电子文件集，方便制成编研光盘或图书。

(15) 档案检索利用效率大大提高

提供强大的检索功能，如综合检索可以实现跨档案类别的查询，自定义查询可以保存用户设定的查询条件，方便以后直接使用，档案分类查询提供更为快速简便的查询体验。WISE 检索技术实现跨档案类别快速检索，查询功能全面优化，保证查全率、查准率（见图 7-3）。

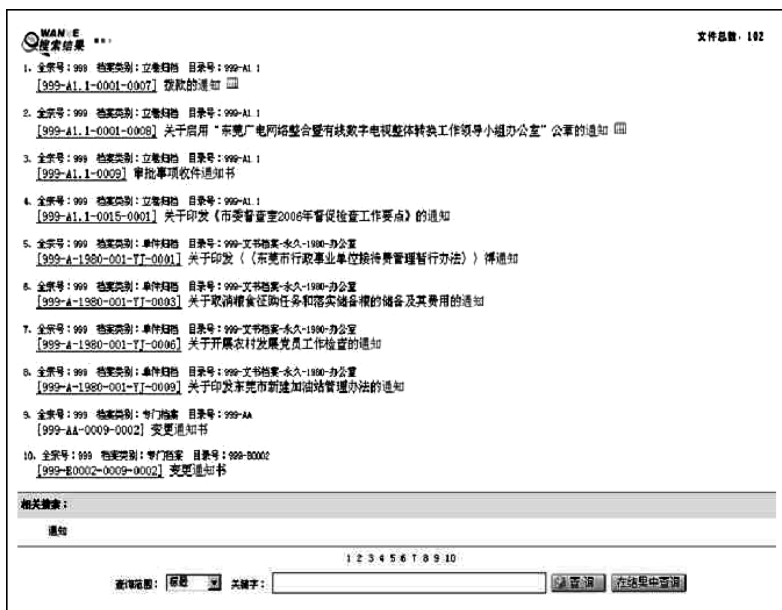


图 7-3 WISE 检索

(16) 完善的系统日志管理

提供了简单实用的日志管理功能，方便系统管理员了解进入系



统的各用户执行的操作，并以此作为参考和判断系统出现问题的依据。

(17) 支持从收、发文归档

可以通过手动归档与查询归档两种方式实现从收、发文归档。用户知道要归档的收文或发文编号时，适合用手动归档方式，通过输入收、发文编号后进行归档。当用户不清楚要归档的文号时，可以使用查询归档方式，通过多个查询条件的组合选择要归档的文件，然后归档。

(18) 建有档案软件外部接口

充分考虑外部数据接口（如业务软件管理、办事管理系统等），使其数据可直接归档，减少档案人员的工作量。

(19) 档案著录与数字化加工可分开进行

当用户需进行大批量档案整理时，可将档案著录与数字化加工分开进行操作，从而提高档案整理的效率。只要在扫描文件时按照万维博通公司提供的文件扫描命名规则进行扫描，在档案著录与数字化加工完成后，利用系统提供的电子文件加载功能，即可让档案记录与电子文件自动对应。

(20) 提供了更全面稳妥的备份方案

①管理员手工对数据进行备份，提供一键备份，操作简单轻松。

②利用备份程序设定执行时间，自动备份每天数据。即使服务器出现故障，也能迅速恢复数据。

(21) 标准规范性

系统遵照执行国家颁布的相关标准规范，符合国家有关文件与档案管理的规范要求、法律法规，系统开发从需求分析、设计编码、系统测试到正式运行的全过程遵循软件工程规范要求。

(22) 开放性

系统在体系结构、硬件产品、软件产品、数据交换协议等方面，充分考虑其开放性，使系统整体具有较好的互操作性、可移植性。

(23) 系统操作便捷性

系统设计人性化、专业化，符合用户日常习惯，易于操作使用、界面友好，用户不需培训或只需少量的培训即可使用。

(24) 可扩展性

系统软硬件配置要具备可伸缩及动态平滑扩展能力，通过系统框架和相应服务单元的配置，适应业务量的变化，以获得良好的性价比。系统架构基于开放体系结构之上，易于扩展，扩建新的业务功能。

(25) 技术的先进性和成熟性

满足系统在很长的生命周期内有持续的可维护性和先进性，充分选用符合国际标准的、代表未来发展主流方向的计算机产品、网络产品和系统软件平台来进行项目建设，使网络及应用系统符合主流技术，以利于将来维护与设计。

(26) 安全可靠

系统在数据库级、表单级、用户级采用多种存取控制和授权功能，保证各个环节数据的安全性，引入成熟的电子密钥、VPN 技术进一步加强了用户访问的安全控制。另外系统采用完善的备份方式和灾难恢复计划，确保系统的健康性和可靠性。

7.2.2 软件架构

万维博通软件开发，严格遵循 CMMI 的开发流程。建立基本项



目管理过程,跟踪费用、进度、功能等特性,制定过程体系,能重复使用已完成、稳定的功能模块,加速软件开发进度,增强软件的健壮性。

①采用 SOA 的框架结构,集成各种高级应用。

②系统安全:本系统各个功能模块访问权限独立分开,通过 MAC 绑定、IP 过滤、USB Key 电子密钥、日志痕迹管理等技术,确认合法用户的真实性。

③数据库安全:利用 Microsoft SQL Server 2005 可方便地建立数据备份及恢复策略,提高数据的可靠性,通过专用的备份软件及存储设备定期数据与电子文件备份,通过采购的灾难恢复设备,实现快速恢复。

④通过 Internet/Intranet 网络互联技术将主管单位与下级单位或企业各分支机构的档案信息组建为一个集中的目录中心,实现档案实体与档案信息的分离管理,实现档案部门与归档部门的沟通,有利于对归档管理各环节过程的控制和管理。

⑤构建档案中间接口,通过数据挖掘技术将用户单位的各种业务系统(如办事审批、发证管理)和 OA 系统的数据自动收集到档案信息平台。

⑥建立一种新型的档案移交、接收模式,做到档案实体正式移交前进行相应的非现场预检查,保证档案接收质量并减少后续的著录量与数字化加工量。

万维系统采用 B/S 与 C/S 混合结构,充分利用客户机的处理能力,提高了系统的运行效率,系统拓扑结构图如图 7-4 所示。

万维博通秉承学习、创新理念,对新兴技术保持敏感触觉,积极学习,把新技术融入软件开发中,为软件开发拓展思路、增添软件的活力。多年来从 ASP 页面开发,到 ASP.NET1.0、ASP.NET2.0……科技不断发展,万维探索、创新之路也在延伸、拓宽。

(1) ASP.NET 框架

ASP.NET 为构造新一代动态网站和基于网络的分布式应用提供了强有力的支持。与之前的 Web 开发模型相比,更简易、安全

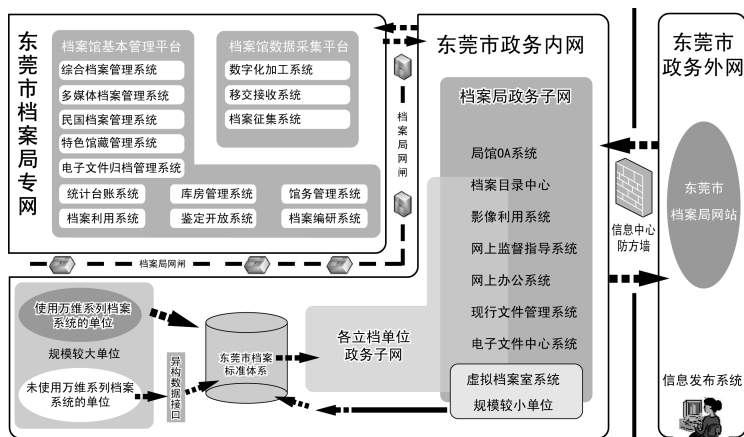


图 7-4 东莞市电子档案馆建设总体框架

及更具可管理性。且与基于过程的 ASP 页面技术相比，面向对象技术在 ASP.NET 完美体现。ASP.NET 技术结合 MVC，很好地实现了用户界面承担向用户显示问题模型和与用户进行操作 I/O 交互作用。采用 MVC 开发模式，把交互系统的组成部分分解成模型、控制器和视图三种部件。

视图部件把表示模型数据及逻辑关系和状态的信息以特定的形式展示给用户。

控制器部件是处理用户及软件的交互操作，是工作模型中任何变化的传播，确保用户界面与模型之间的对应联系。控制器接收用户的输入，将输入反馈给模型，进而实现对模型的计算控制，是模型和视图的协调工作的部件。

模型部件封装了问题的核心数据、逻辑和功能计算关系，它独立与具体的界面表达和 I/O 操作。

MVC 把模型、视图及控制器分离，使得一个模型可以具有多个显示视图。假如用户通过某个视图的控制器改变了模型的数据，所有其他依赖与这些数据的视图都会相应反映。因此，任何数据变化，控制器都会将变化通知所有视图。模型、视图、控制器间的工

程关系如图 7-5 所示：

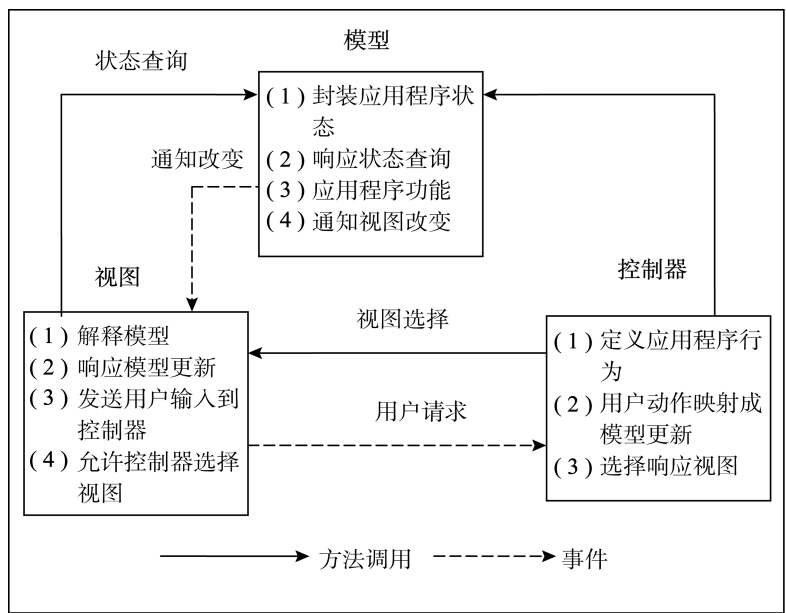


图 7-5 模型、视图、控制器间的工程关系

(2) Web Service 模式

Web Service 是一种应用组件，逻辑性的为其他程序提供数据与服务，各应用程序通过网络协议和规定的一些标准数据格式（如 HTTP，XML，SOAP）来访问 Web Service。Web Service 是 SOA 的实现技术。通过 Web Service 内部执行得到所需要的结果。Web Service 完全基于 XML（可扩展标记语言）、XSD（XML Schema）等独立于平台、独立于软件供应商的标准，是创建可交互操作、分布是应用程序的新平台，从而实现跨平台的可交互操作性。采用 Web Service，使万维档案系统具有以下特点：

①异步处理。采用 Web Service 模式，客户端不再提交页面，而是发送 Web Service 请求，并对收到的应当请求进行处理。由于

Web Service 是异步的方式，所以客户端在发送请求之后，无须等待，采用这种方式，服务器返回的数据量降低到最小。

②跨防火墙通信。同一单位使用多套万维博通软件，或者同一套软件有不同单位同时使用，而且分布在不同的区域，客户端和服务端之间的通信是一个很棘手的问题。因为客户端和服务端之间通常会有防火墙或者代理服务器。在这种情况下，使用 DCOM 就不是件容易的事，通常也是不便于把客户端程序发布到数量如此庞大的每一个用户手中。传统的做法是，选择用浏览器作为客户端，编写大量的 ASP 页面，把应用程序的中间层暴露给最终的用户，这样做的结果是，开发和维护的难度大。

调用 Web Service，可直接使用 Microsoft SOAP Toolkit 或者 .NET 的 SOAP 客户端，也就可以使用直接开发的 SOAP 客户端，然后把它和程序连接起来，不仅缩短开发周期，也降低了代码的复杂度，进而增强了程序的可维护性。

③软件局部更新灵活性。用户需求变动或者软件功能升级，都会涉及软件重新打包，再安装到客户的服务器上。如果只是很小的功能变动（比如在打印报表中增加一个字段，或者在打印功能增加一张报表打印），而对整个档案系统进行重新编译、部署，显然是不太现实的。为了增强档案系统的更新灵活性，将打印的后台处理放置在 Web Services，如需更新该功能，只需在客户端对前台代码进行稍做调整，或者只是对打印模块进行编译、更新即可。另外，采用这种方式，可减少系统频繁编译带来的版本控制问题。

(3) 采用 Ajax 无刷新提交

万维博通非常重视档案系统运行效率及良好的用户体验，采用 Ajax 技术也是出于这方面考虑。Ajax 技术发展至今，已相当成熟及稳定，也深受 IT 行业的宠爱。

①提交数据，页面无刷新，是 Ajax 技术的最大特点，也是用户体验最直接的反映。这一特点主要得益于 XMLHttpRequest 对象。这样就可以像桌面应用程序一样只同服务器进行数据层面的交换，而不用每次都刷新界面。



②Ajax 技术使用异步处理方式与服务器进行通信，不需要中断用户的其他操作，响应更迅速，提高程序运行效率。

③运用 Ajax 技术，可利用客户端空闲资源分担服务器的部分工作，从而减轻服务器及带宽的负担。“按需取数据”是 Ajax 技术的处理原则，减少冗余请求。

④Ajax 技术是基于标准化并被广泛支持的技术，无需下载插件或小程序。

(4) 采用自定义控件进行软件开发

万维博通在软件开发中，积极探索、勇于创新。为提升软件开发效率，缩短软件开发周期，增强软件的健壮性。通过底层开发，基于系统控件，开发出新的用户控件。通过拖曳相应控件到页面，在前台页面配置新添加控件属性，无须更改后台程序，即可达到信息增加。另外，数据读取、显示、查询、更新等操作，无须考虑在后台处理控件与数据库中的字段对应关系，只需调用底层方法，即可实现相应操作。采用定义控件创建的页面，代码风格统一、简洁、可重用性高、消除因页面编辑项过多，字段对应有误的情况。如果用户需要在某个页面增加一个字段信息，只需在前台增加相应的自定义控件，利用自定义控件新增的 Tag 属性，对应用户所需新增的数据库字段即可，无须对整个档案系统重新编译、更新（见图 7-6）。

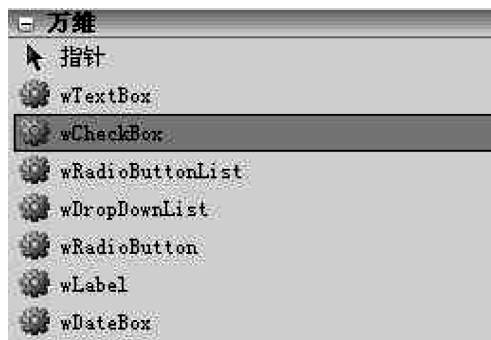


图 7-6 自定义控件

7.2.3 运行环境

(1) 系统环境要求

①服务器硬件环境要求：

配件名称	CPU	内存	硬盘	网卡
最低配制要求	P3 800MHz	256M	10G	10/100M 自适应网卡
建议配制	P4 1.8G 以上	512M 以上	40G 以上	

②服务器软件环境要求：

软件名称	操作系统	数据库	浏览器	框架（Framework）
配置方案	Windows 2003 企业版	MS SQL Server 2000 企业版	IE6.0	V2.0

③客户端硬件环境要求：

配件名称	CPU	内存	硬盘	网卡
最低配制要求	P3 600MHz	64M	3G	10/100M 自适应网卡

④客户端软件环境要求：

软件名称	操作系统	浏览器	应用软件
可选方案	Win98、Win Me、 Win 2000、Win XP	IE5.0 以上	MS-Word、MS-Excel、AcroRead 5.0

(2) 确保系统正常运行要点（以最佳方案来讲解）

①操作系统 Win2000 已经打上补丁程序 SP4，且安装了 IIS

(Internet Information Server)。

②数据库方面：一定要求安装上 MS SQL Server 2000，还要求安装 MS Access，MS Access 主要用于系统产生的数据据包。

③浏览器方面：IE 浏览器版本要在 5.0 以上，打上 IE 补丁程序 SP1。

④框架 (Framework)：Windows 2000 一定要安装 V1.1 的 Framework。Windows 2003 一般不需要安装框架。

⑤应用软件方面：服务器客户端必须安装 Adobe AcroReader，版本最好安装 5.0，建议安装上 MS-Word、MS-Excel。

⑥浏览器的一些设置 (见图 7-7)：在 IE 浏览器属性常规选项卡中，点击“设置”，在设置对话框中选中“每次访问此页时检查”，点击“确定”确认刚才所做的选择。

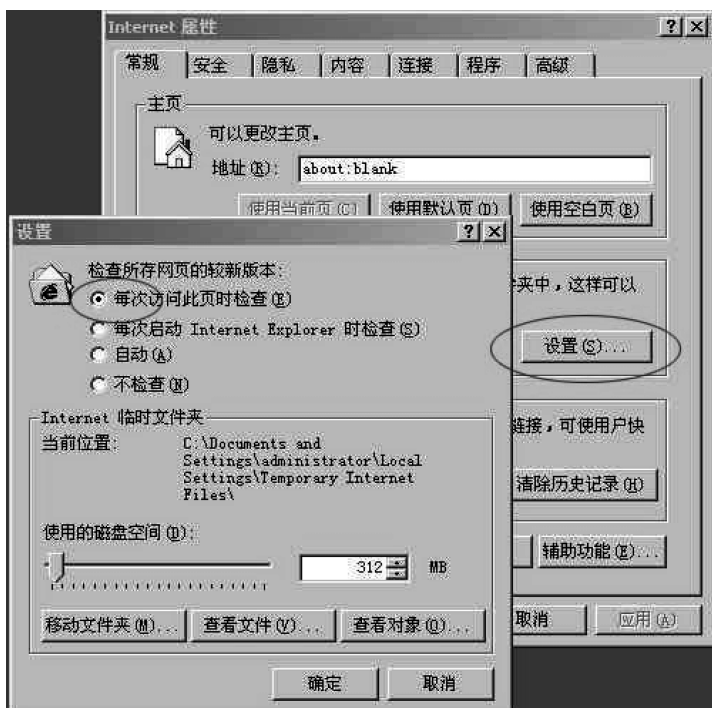


图 7-7 浏览器的设置

⑦服务器客户机上最好不要安装 3721、上网助手等，IE 浏览器附加的一些小软件，以及在一些杀毒软件中设置了“禁止弹出窗口”的，要求改为启用。

(3) 运行的网络环境

现有的业务应用主要基于局域网、政务网、互联网三个网络平台，图 7-8 是“三网”拓扑结构图。

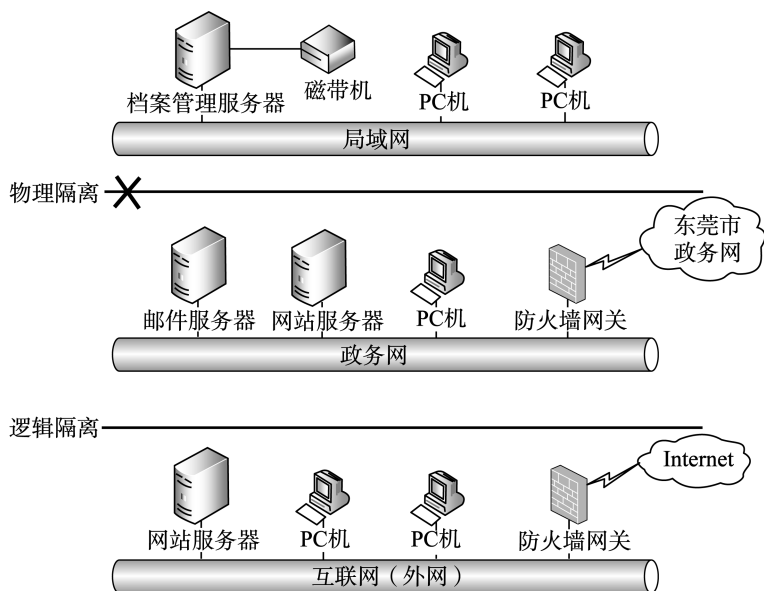


图 7-8 “三网”拓扑结构图

局域网：它是市档案局的一个内部业务网络，目前与政务网、互联网物理隔离，是档案局内部人员收集、管理、利用档案的主要网络。

政务网：政务网与互联网逻辑隔离，是与市级机关单位互联互通的专线网络。档案局局域网与政务网之间通过网闸互联。

互联网：互联网作为一个开放的网络，面向社会和广大民众，主要提供各类开放信息的共享查询和利用。



7.3 规则库建立及应用

电子政务环境下基于业务规则的档案管理平台是广东万维博通信息技术有限公司对业务规则在档案管理领域应用的软件方向。

建立档案业务规则库（ABRD），是架构档案业务规则管理系统的基础。万维博通档案系统为了兼顾档案业务规则的普遍性和特殊性，为用户提供建立档案业务规则模板的平台，使用户所建立的规则更适合自身业务的范畴，符合业务普遍性。任何一个业务规则模板，都不是万能的，不可能符合任何情况。例如：基建工程项目中的建造公路，涉及穿过山川时需要开凿隧道，横跨江河时需要搭建桥墩，不同情况下需要整理的基建档案资料有所不同，这就要求业务规则根据实际情况灵活变通，即业务规则的特殊性。任何业务规则的特殊性都是基于普遍性，无须为具体到每一种情况设定一个规则模板，只需在现有的具有普遍性的模板基础上做必要的增加、修改或删除即可。万维博通档案系统实现业务规则的普遍性和特殊性，充分体现了业务规则管理技术分层和复用。

7.3.1 规则库创建和管理

规则库用来存储规则元数据及规则集合，需要完善的管理机制进行管理。规则的选取、表述、切分等都会对规则的使用产生巨大的影响，所以只有在规则库中使用恰当的规则，才能给规则的处理、运算、逻辑表述等带来可能，给应用程序处理提高效率带来可能。

在万维博通档案系统的基本设定中，设立了业务规则模板设定接口。考虑到档案业务规则，会在日常的业务运作中，不断更新、新增，导致业务规则库出现指数形式的增长。万维博通档案系统，采用树型结构分节点管理方式管理业务规则库。采用树型结构进行分类管理，降低规则模板的内部耦合性。根节点为规则库的最基本、最底层的规则模板分类，如基建类、会计类等，在此基础上，产生多个分支（子节点）。在子节点层，可设定模板名称和其子节

点。模板名称设定后，可进行具体的规则模板设置。在设置规则模板时，结合保管期限表，设定案卷和卷内的档案业务规则。规则模板建立结构如图 7-9 所示：

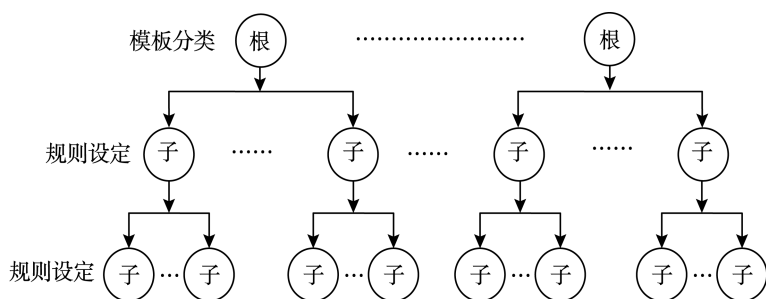


图 7-9 规则模板建立结构

档案系统规则分类运行界面截图如图 7-10 所示：

■ 系统设定 > 归档规则模板类别设定			添加	查询	刷新	打印
序号	类别名称	操作				
1	基建类	[修改] [删除] [设定模板]				
2	会计类	[修改] [删除] [设定模板]				
3	科技类	[修改] [删除] [设定模板]				
4	法规类	[修改] [删除] [设定模板]				

图 7-10 档案系统规则分类运行界面

点击图 7-11 中的选定区域，设定该档案类别下的档案规则。

■ 系统设定 > 归档规则模板设定 > 基建类				添加	查询	刷新	打印	返回
序号	模板名称	备注	操作					
1.00	地下管道工程	通信、通水、通气	[设定子规则] [修改] [删除] [设定归档规则]					
2.00	建筑工程	城市建筑、水利建筑等	[设定子规则] [修改] [删除] [设定归档规则]					
3.00	道路工程	普通、高速公路、铁路	[设定子规则] [修改] [删除] [设定归档规则]					

图 7-11 设定档案类别下的档案规则界面

规则树遵循树型结构的特性，子节点继承父节点的所有属性，并可衍生出自己的子节点。点击图 7-12 中的“设定子规则”，是对

地下管道工程的规则进行细化定义，可定义“地下通信管道档案规则”、“地下排水管道档案规则”等。点击“设定规定规则”，是对当前规则进行具体设定操作。

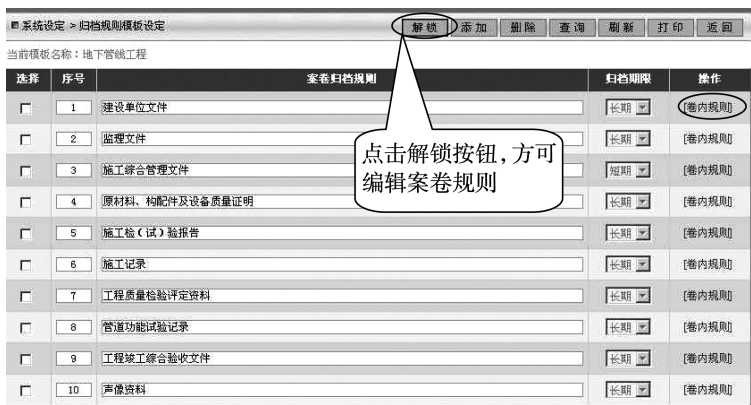


图 7-12 归档规则模块设定界面

为了规则的稳定性，防止误操作导致规则变动，采用规则套锁机制。有规则管理权限的用户，进入规则管理界面，需要点击解锁按钮方可对规则进行编辑操作（同时解锁按钮变成锁定按钮），编辑操作结束后，用户应进行锁定操作将规则锁定。分支节点上的规则设定，可分为案卷和卷内规则，也可只设定案卷规则。设定规则时，结合保管期限，确保每条规则都与保管期限有映射关系，但这种映射关系不是固定的。判定一份文件归档成档案资料的归档期限，还需根据责任者或文号、收发文单位间的上下级关系等综合因素。点击图 7-13 中的“卷内规则”进行卷内规则设定。

规则库是档案业务规范与否的参照，是档案业务规范化的方向标，规则库的优劣，很大程度影响档案业务的规范性。规则与档案标题完全匹配，只是档案整理理想化的结果。因此规则库的基本元素“关键字”，判定档案条目是否符合规则的关键。万维博通档案系统采用“复方分词法”，以最合理的方式将规则分成若干词组，作为档案规范匹配的关键字。

当前模板名称：地下管线工程

卷内归档规则设定——当前案卷规则为：建设单位文件

添加

查询

刷新

打印

序号	卷内归档规则	匹配关键字	备注	最小匹配次数	操作
1	建设工程计划、规划、设计、用地、开工批准文件及其他审批文件	建设工程计划;规划;设计;用地;开工批准文件;审批文件;		3	[修改] [删除]
2	工程勘测、设计、监理、施工合同书	工程勘测;设计;监理;施工;合同书;		2	[修改] [删除]
3	工程决算和工程审计文件	工程决算;工程审计;文件;		3	[修改] [删除]
4	按规划局要求由市勘测院测量的竣工测量	规划局;勘测院;测量;竣工;测量;		4	[更新] [取消]

关键字提取图标

跳到:

1/1

总数: 4

◀

▶

⏪

⏩

图 7-13 当前案卷规则

所谓“复方分词法”，是指综合运用多种分词算法，吸纳不同分词算法的优点弥补其他算法的不足，使分词达到更高的准确度。

①基于字符串匹配分词法（也被称做机械分词法），按照双向匹配（即正向最大匹配法和逆向最大匹配法）切分方式对规则进行分词。此种分词算法，作为初步分词，将较明显的关键字划分出来。

②基于理解的分词法，利用语法树，对句子进行语法、语义分析，减少歧义性，将符合语法规则的词组划分出来，作为规则的关键字。

③基于统计的分词方法，根据统计相邻字出现的概率或频率，计算相邻字的互现信息。采用串频统计和串匹配相结合的方式，利用无词典分词，识别生词和消除歧义，使分词更快、更准。

输入规则信息后，点击提取关键字图标，系统通过“复方分词算法”将当前规则进行分词操作，再结合《受控词表》，区分哪些分词可作为规则匹配词，哪些作为非用词（如“的”、“关于”、“以上”、“其中”等词，不作为规则匹配的关键字），并把最终筛选的分词结果显示在关键字信息框中。设定最小匹配数，限定规则匹配通过标准，如设定最小匹配数为4，在档案标题与规则匹配时，关键字匹配大于等于4，该标题匹配通过，反之匹配不通过。

业务规则主要效用就是行为约束和知识推理，规则引擎可以进行逻辑判断，建立数据对象和规则之间的映射关系，并能动态的组

织与数据对象相关联的、满足条件的业务规则，自动实现这些规则的一致性、完整性、相容性等逻辑关系。同时也可以已有的业务规则的基础上，根据实际情况，衍生出新的更贴切实际的业务规则。随着业务拓展，业务规则库也会日渐丰富、健全，只要能充分、智能地利用这规则库的规则，对相关业务资料的整理更具合理性、规范性将起着不可估量的作用。

7.3.2 规则引用

以基建工程档案整理为例，在万维博通软件系统中，建设单位用户可以根据各自的业务范畴设定相应的规则模板。在项目分工阶段，为项目的每个工程或标段设定档案业务规则。建设单位用户，可以在分工页面，选择规则模板名称。点击模板名称输入框，系统以树型结构显示当前用户所属建设单位设立的规则模板。选择任一规则模板名称后，页面显示相应的规则信息，相当于规则的一份拷贝，用户可根据实际情况，对当前案卷及卷内的规则进行添加增、删除、修改操作，而不影响母版规则，实现规则的有效复用，充分满足不同用户对业务规则的特殊化需求（见图 7-14）。

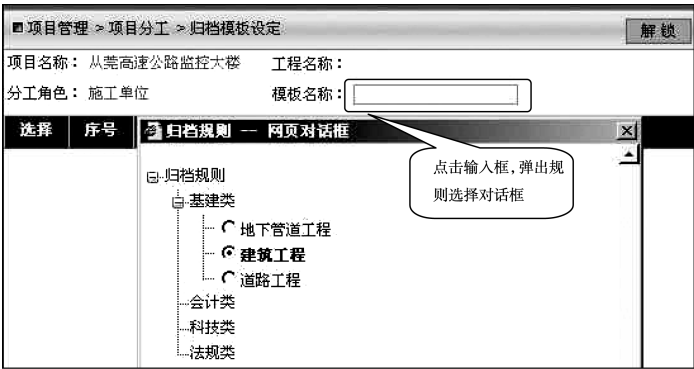


图 7-14 归档模板名称输入框

选择模板名称后，显示对应的案卷规则，如图 7-15 所示：

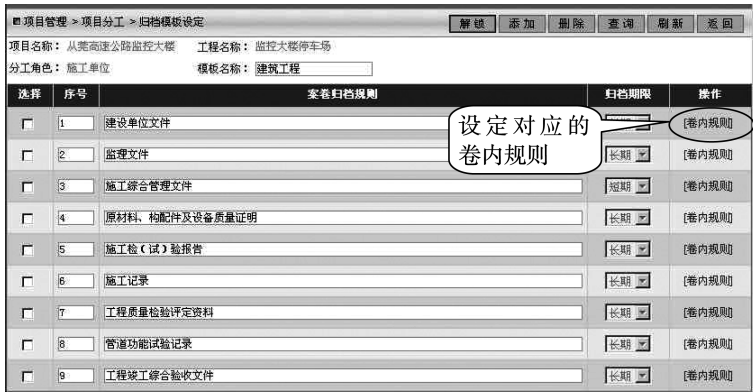


图 7-15 显示对应的案卷规则

点击图 7-16 中的卷内规则按钮，弹出相应卷内规则信息。



图 7-16 卷内规则信息

7.4 档案业务规则在档案整理中的作用

一个单位的档案采用传统的整理方式，就需要高素质的档案人员，做到仔细、认真，并且必须非常了解单位的业务运作。因此，



目前整个档案行业基本陷入一种困局，各个单位的档案管理很难做到收集完整、著录规范。在档案整理的过程中，每个环节需要大量的人工干预或检查，才能达到理想的整理效果，也就从管理和技术实现两个角度来探讨档案信息资源开发与管理问题，万维博通重点研究一套基于信息资源规划，依托业务规则库的档案信息资源收集管理体系：通过信息资源规划确定各立档单位的归档范围、保管期限表，并将其抽象为档案知识规则库（ABRD），同时也将档案主管部门与指导部门的规范、制度、标准等也整理到ABRD，通过强大的推理引擎（RM），对已有的档案信息数据（ADI）进行分析与推理，作出相应的判断，以帮助立档单位达到“档案收集齐、保管期限设置准、档案著录规范”的目标。

档案局/馆、档案主管单位利用人工智能技术，结合档案业务规则库（ABRD），自动监督各立档单位的档案整理情况，可对档案信息（包含全宗、目录/项目、案卷、卷内等信息）根据档案业务规则库（ABRD）推理出档案信息资源收集是否完整，保管期限是否把握准，著录是否规范等。发现问题自动提醒，并及时反馈，档案监督指导人员根据实际情况作出处理。

出于上述问题考虑，万维博通研发的“万维数字化档案网上管理系统”，从根本上解决因档案人员对业务不熟悉或者粗心大意，错误著录、漏著录导致档案资料整理不规范的情况。

7.4.1 档案整编

根据当前登录用户的管理全宗范围，整理档案资料。根据用户的不同档案业务需求，分成几大门类：文书档案、特殊载体档案、专门档案、会计档案、已改人员档案、诉讼档案等。用户可根据待整理的档案资料特性，选择最恰当档案整理方式，达到档案整理的合理性、规范化，使得整理的档案资料覆盖面广、检索档案时能快速准确地获得所需档案资料，获取最新整理的档案资料，充分体现档案信息化管理优势。

文书档案，细分为立卷归档和单件归档。立卷归档，先设定案

卷目录，然后在目录下进行案卷著录和卷内著录操作。单件归档，与立卷归档整理方式唯一不同的是，立卷归档是采用二级目录管理，而单件归档是一级目录管理，并根据全宗号、类别号、年度号、机构号以及保管期限进行区分整理，以确保档案整理有条不紊。文书档案针对档案整理的文字性质的档案资料，也是档案的最主要的组成部分之一。对文书档案的整理，根据档案业务规则，进行规范管理，显得尤为重要。另外，由于文书档案信息具有不同性质（有人事、财务、法制、行政、业务等）和不同的密性，有的甚至关系到国家和社会稳定，所以要具有一定的加密功能，“万维数字化档案网上管理系统”为文书档案著录时设立密级，只有等于或高于文件密级权限的用户才能查看到对应的档案资料，并对密级权限分配加强授权管理。

特殊载体档案，包括：照片档案、底片档案、音像档案、实物档案，根据不同载体类型档案，在对应的门类进行档案著录和管理。

专门档案，系统提供专门档案自定义类别著录整理，各单位根据专门档案的具体情况，决定采取何种整理方法整理，如人事档案可采取文书档案立卷归档方式整理，医疗档案可采取文书档案单件归档方式整理，科技档案则可以采用科技式整理。

已改人员档案，专用于整理已改人士的档案资料，有针对性的著录相关资料。

万维档案系统，为确保上述门类的档案整理的规范性，进行全面、综合控制，如著录时规范性控制、著录后规范性检测。

(1) 档案著录控制

档案整理要符合档案业务规则，档案著录控制是关键。系统对档案资料的规范控制中，著录控制是从档案整理的规范源头加以控制。根据业务规则控制档案著录，是把档案整理的不规范最大限度地扼杀在摇篮中，使档案资料核查和返工的工作量降低到最低水平，理想状态是达到档案零返工。

档案著录规范性控制，从档案目录设定已开始加以控制。输入



相同的目录号，系统提示目录号重复，避免不同的档案资料，整理到同一目录下。整个档案目录已作入库操作，不允许对档案资料进行编辑操作，只能查阅档案资料。如此控制，正是实现档案资料的统一管理，使实体档案与电子档案保持一致。

案卷、卷内著录控制，是按照档案整理规范对某些著录项进行控制，如页数、份数等著录项，要求输入整型格式。页次，合法格式为“001-012”或者“015”两种形式，其他形式的著录，都属非法著录，需要提示错误并不予著录。另外，档案整理过程中，档案日期存在较多不确定性，因此日期存在多种合法情况。档案日期中包含年、月、日三部分，只能确定其中一个或两部分，抑或年月日都不能确定等多种情况。因此：2005□□□□、□□□□1005、□□□□□□20、□□□□05□□等，都属合法日期。系统都可以通过前台脚本进行整型、页次格式、日期格式进行判断，在不增加服务器负担的情况下，对档案著录进行规范化控制。档案密级，根据著录用户的密级等级，筛选密级选项绑定到密级下拉框中，可有效地控制档案文件方法权限。如档案著录用户用于04的密级权限，密级下拉框只绑定01~04的密级，而05或以上的密级不可见。责任者著录，采用智能感知和选择方式进行著录。智能感知，即输入“市府”二字，输入框下面则把包含“市府”二字的责任者罗列显示，著录用户可以选择某一责任者，即可完成责任者著录。选择方式，用户点击责任者输入框，以层的形式显示已设定的责任者信息供著录用户选择。以这两种著录方式，著录操作更便捷，责任者著录更规范。

整理工程类档案的案卷资料，系统会自动产生虚拟案卷号（档案验收合格后，建设单位再统一分号），避免多人同时整理档案时，产生案卷号冲突现象。档案著录人员，点击归档资料分类编辑框，系统自动调用业务规则库，读取归档资料分类，并以树型结构显示。用户选择所选归档资料分类，系统自动填充归档资料分类和保管期限信息到对的编辑框内，使归档资料分类信息层次更分明、操作更简洁。

重复著录，是档案整理中不可避免的情况。多个档案整理人员

进行著录操作，由于协调不足，导致档案著录人员 A 已著录某一记录，档案著录人员 B，在不知情的情况，再次著录相同的记录。或者同一个档案著录人员，著录某一记录后，由于某种原因，对已著录的记录印象模糊，再次著录相同的记录。这时，重复性判断显得尤为重要。系统主要对著录项中的文号（如：东劳〔2009〕22 号）和标题进行重复性判断。

（2）档案规则智能检查

档案规则检查，是对已著录的档案资料进行规范性检查，著录后对不规范的档案资料进行排查及纠正措施。

作为基于档案业务规则的档案系统，不能根据单一的规则库进行断定一份文件任何归档为档案资料。同一标题的文件资料，承办单位、责任单位或文号等综合环境的不同，最终决定此文件归档成档案的保管期限就有所不同。万维博通档案系统，综合考虑责任者或文号，及收文与发文单位间的上下级关系，再结合规则库的关键词匹配进行综合分析，最终给出综合性强的分析结果，供档案管理人员参考。

对于不同的档案类型，整理方法有所不同，档案规则检查也相应有所不同。文书档案是以一份文件为一卷，而工程档案则是以一个专业阶段来组卷（多份文件），即文书档案独立性较强，可进行单独一卷文书档案资料进行规则性检查，而工程类档案则可进行整体性规则检查。

工程档案，由负责不同标段的归档单位进行档案资料整理，最终著录到万维博通档案系统中。如果档案资料不齐全，或者不符合档案整理要求，项目将不予验收，因此档案整理的规范性，关系到工程能否进行验证的关键。

工程档案整理完毕后，归档单位用户可先自己检查档案整理的完整性、规范性情况。只要进入工程目录，选定某一工程记录，点击检查按钮。系统将遍历所选工程下的所有档案资料（包括案卷和卷内档案资料，或者只检查其中之一），然后启动档案业务规则引擎（RM），规则引擎再调用档案业务规则库（ABRD）对应的规

则，对指定工程的档案资料进行规则匹配，并显示遗漏或者不规范的档案资料。

具体操作流程如图 7-17 所示：

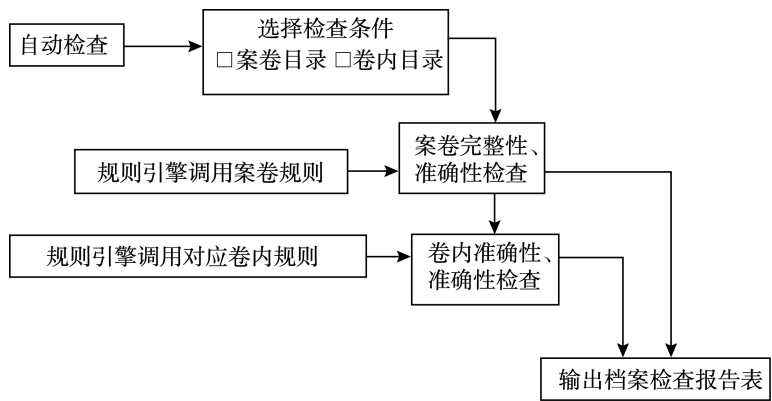


图 7-17 档案规则智能检查

档案业务规则引擎根据案卷和卷内，调用相对应的规则进行关键字匹配，匹配数大于等于规则最小匹配数，则认为该案卷或卷内符合业务规则，否则根据规则匹配失败情况，显示缺少的档案文件资料，档案业务规则智能检测结果如图 7-18 所示。

经档案资料智能检查操作后，能确保所选工程对应的档案资料的完整性、准确性、错误纠正或缺失不齐及时性。

基建档案整理，建设单位根据不同工程，设定相应业务规则，也是判断归档单位提交档案资料是否合格的标准。智能检查，调用档案业务规则库中对应的规则，对待检查的档案资料进行匹配检查，把不符合规则的档案逐一排除，直到指定工程所有档案条目都匹配完毕，如果存在未匹配通过的规则项，表示该工程档案存在资料不齐全情况，系统会把缺少的档案条目罗列出来，档案整理人员只要根据档案检查报告表补充对应的档案资料即可。如果规则中的所有项，都能在被检查的档案资料中找到匹配项，说明此次档案整理完整性得到保证。

经检查缺少以下案卷		
序号	名称	
1	建设单位文件	
2	监理文件	
3	施工综合管理文件	
4	原材料、构配件及设备质量证明	
5	施工检（试）验报告	
6	施工记录	
7	工程质量检验评定资料	
8	管道功能试验记录	
9	工程竣工综合验收文件	
10	声像资料	
11	竣工图	
已有案卷检查结果如下		
所属案卷	序号	缺少卷内
0002_ 从莞高速南城段设计图纸 2	1	主要原材料及构配件出厂证明及复试报告目录
	2	厂家资质、生产许可证、检验报告
	3	出厂合格证、出厂质量证明书
	4	主要原材料及构配件出厂证明及复试报告目录
	5	厂家资质、生产许可证、检验报告
	6	出厂合格证、出厂质量证明书
	7	见证取样记录及现场抽检试验报告

图 7-18 档案业务规则智能检测结果

为了达到档案检查准确性，档案规则引擎按照关键字匹配得分制。把待检查的案卷或卷内标题按规则拆分成词群，再把这该词群与未检查通过的规则的关键字进行逐一匹配，匹配一个关键字得一分，得分累计制，最终得分为该档案条目检查结果，是否达到或超过规定分值，为判断该条目检查通过与否的判定提供全面的数据参考。

档案智能检查从档案整理开始，到档案整理结束，整个过程中的任何时刻都可进行。随时进行档案检查，及时发现档案整理过程中的错误或者待补充的档案条目。

文书档案，也是人为著录进行档案整理，不可避免地出现档案著录偏差或者遗漏现象。在整理过程中，除了对档案人员著录过程

中进行规则限制，使档案整理尽可能的规范，还需配套的检查机制。

“万维数字化档案网上管理系统”对文书档案著录控制，充分体现 DURA 模型中提到将电子政务活动中产生的 D（文件）完整、及时、准确地转化为 A（档案）的思想。档案著录人员，在卷内著录页面，著录案卷题名后，可点击“规则分析”按钮，即可检查输入的案卷题名是否符合档案业务规则，如图 7-19 所示：

文书档案 立卷归档>卷内文件

上一条(N) 下一条(N) 继承登记(I) 新增(M) 返回(R)

序号: 0004

文号: 东府办 (2006) 3 号

责任人: 市府办

文件标题: 关于印发《东莞市哲学社会科学优秀成果奖励试行办法》的通知

主办单位: 万维博通

全宗号: 999 目录号: A1.1 案卷号: 0001

规则分析

万维档案业务规则库 — 网页对话框

规则分析结果 关闭

文件标题	关于印发《东莞市哲学社会科学优秀成果奖励试行办法》的通知 市府办		
建议归档期限	永久	查看分析情况	
编号	规则明细	匹配结果	得分
3.1	本机关制定的方针政策性、法规性、普发性业务文件，中长期规划、纲要等文件材料	永久	2
3.2	本机关为主办的涉及政策法规性、重大业务建设的联合行文的文件材料	永久	2
3.4	本机关为协办的涉及政策法规性、重大业务建设内容的	长期	2
6.1	本机关受县(含)以上表彰、奖励的先进单位、劳动模范、先进工作者、优秀党员、先进党支部的文件材料	永久	2
6.2	本机关受县以下表彰、奖励的或本机关内部表彰的先进单位、劳动模范、先进工作者、优秀党员、先进党支部的文件材料	长期	2
7.5	上级颁发的关于财务、统计、基建、房管、车辆、固定资产管理等政策、法规性文件	短期	2
得分结果	永久 19分，长期 12分，短期 8分。		

图 7-19 万维档案业务规则检查

档案著录人员，可以根据规则分析结果，判定当前著录的卷内标题是否规范，因此档案著录人员可以及时核实标题的规范性。另

外，系统根据“规则综合得分归档制”，建议此卷内文件归档到得分最高的保管期限的案卷目录下。比如当前卷内标题根据档案业务规则得分情况如下：永久 19 分，长期 12 分，短期 8 分，因此系统建议将该卷内文件归档到保管期限为永久的案卷目录下。“规则综合得分归档制”可帮助卷内文件的保管期限在模棱两可的情况下，得到有效的定位，从而使得档案整理达到合理性的最大化。

对于档案整理经验相对丰富，档案著录过程中，出错情况较小的档案著录人员，无须对每一个卷内标题进行逐一检查，以确保档案的规范性。此时，档案著录人员可以采取批量分析的方式，对档案进行检查，如图 7-20 所示：

文书档案 立卷归档>卷内文件							
批量分析 添加 刷新 查询 打印 分件 归档 调序 返回							
序号	文号	文件标题	责任者	页次	文件日期	附注	操作
0001	东物价办(2006)1号	东莞2006物价工作	东物价局	001-002	20060010	物价涨幅控制	[修改] [删除]
0002	东统字(2006)8号	关于做好2006年统计从业资格认定及相关工作的通知	市统计局	003-012	20060202		[修改] [删除]
0003	东常(2006)8号	东莞市人民代表大会常务委员会任命名单	市委办	012-015	20061212		[修改] [删除]

图 7-20 批量分析检查档案

批量分析结果如图 7-21 所示：

文书档案 立卷归档>卷内文件							
批量分析 添加 刷新 查询 打印 分件 归档 调序 返回							
序号	文号	文件标题	责任者	页次	文件日期	附注	分析结果 操作
0001	东物价办(2006)1号	东莞2006物价工作	东物价局	001-002	20060010	物价涨幅控制	※ [修改] [删除]
0002	东统字(2006)8号	关于做好2006年统计从业资格认定及相关工作的通知	市统计局	003-012	20060202		√ [修改] [删除]
0003	东常(2006)8号	东莞市人民代表大会常务委员会任命名单	市委办	012-015	20061212		√ [修改] [删除]

图 7-21 批量分析结果

批量分析功能，犹如浩瀚档案海洋中一座指明方向的灯塔，使迷失或偏离整理方向的档案，重新定位、归档。批量分析，是“规则分析”的扩充版。通过批量分析操作，无须步步为营地检测卷内文件的归档是否符合档案整理的业务规则。只需点击批量分析按钮，系统根据业务规则库，采用高效、合理的方式分析并呈现出当前卷内列表中，不符合档案业务规则的卷内文件。档案整理人



员,可根据分析结果,对个别分析未通过的卷内文件,有针对性地核实、分析并重新整理、归档。批量分析功能,深受广大档案整理人员青睐,离不开它的高效、实用,减少无效检查,最大程度降低档案人员对档案规范性整理的负担。

拥有如此高效的批量分析功能,并不意味着可以完全取代卷内文件的规则分析。犹如拥有强大的雷达系统,不能完全掩盖训练有素的侦察兵的光辉。雷达的作用,是最大范围的搜索异常信息,无法对捕获的异常信息作出精准的分析,要想得到更具体、更全面的情况,还需侦察兵的实地侦察、分析。同样,批量分析,较适合经验丰富或相对丰富的档案整理人员,在整理档案过程中,出错较少的情况。而经验较少,或工作较粗心的档案整理人员,使用规则分析,可从根本上减少不规范的情况出现。如果整理过程中,出现较多不规范,批量分析后,还需对大量的不规范进行逐一分析和重新归档,将使得档案规范性整理的效率大打折扣。

(3) 批量导入控制

档案著录,是档案整理过程中一个重要环节,也是档案整理人员最头痛的一个环节。面对海量的档案,逐一著录,需要消耗档案整理人员大量的时间和精力。尽管“万维数字化档案网上管理系统”提供友善、方便的著录操作,面对一次只能著录一份档案资料信息,频繁的著录和保存、新增操作,随着著录时间和工作量的增加,档案整理人员不可避免地产生急躁情绪。如果采用人海策略,分工著录的方式进行档案著录,需要为每个参与著录的人员分配账号或权限,产生额外的工作或不必要的麻烦,并且多用户同时著录,也会增加服务器负担。

“万维数字化档案网上管理系统”为了更好地服务于广大档案整理工作者,结合 SOA (Service-Oriented Architecture, 面向服务体现结构) 组件模型,通过 SOA 的实现技术 Web Services, 基于 UDDI 协议,采用 WS-BPEL (业务流程执行语言) 定义 Excel 数据导入接口。

通过该接口,档案著录人员,可暂时性地脱离系统,进行档案

著录。档案著录人员，只需从系统中下载“数据导入 Excel 模板”，并在下载的 Excel 文档中按照提示著录档案信息即可。采用这种方式，著录人员人数将不会受到任何限制，所有档案著录人员使用同一模板的 Excel 文档，著录自己范围内的档案信息，然后将所有著录人员著录在 Excel 中的档案资料整合成一份 Excel 文档中，最后通过接口，将整理好的 Excel 文档中的档案信息导入数据库中，完成无限制多人合作整理档案。

采用 Excel 数据导入接口著录档案，充分体现了 SOA（面向服务体现结构）的特点：

①服务是可重用的，服务所包含的逻辑不仅可以被用于最初的设计目的，也可被用于其他目的，即服务可以出现在不同的应用中。Excel 数据导入接口，不仅可以运用于案卷、卷内数据著录，还可运用于文件汇编、法规管理、图书管理等档案资料导入。只需调整 Excel 模板格式，及 Web Services 重载相关方法即可实现，功能的重用，进而实现功能范围的扩展。

②服务具有合同性，服务参与者（服务者和请求者）之间必须有相关约定，才能使服务顺利进行。服务请求者在请求服务之前，必须理解服务合同，在服务过程中，严格遵循服务合同，才能最终待定的效果。在下载 Excel 模板之前，系统将给出相关著录规则，档案著录人员需要按照规则，产生的 Excel 数据才能被系统识别并成功导入数据库中。

③服务具有松耦合性，这一特性使得服务变化相对比较容易，即增强服务的灵巧性、机动性，A 服务的变动，不会对 B 服务造成太大的影响。只要服务合同不变，服务逻辑实现的改变对服务之间的交互不会造成太大影响。下载 Excel 模板的实现方式可以有不同的方式，但都不影响数据的著录和导入服务。

④服务可组合性，下载 Excel 模板和导入 Excel 数据很好地体现了服务的可组合性。

⑤服务自治性，是指服务对其所包含的逻辑控制，即服务逻辑必须在限定的范围内进行服务。对案卷、卷内数据的导入服务，不能直接运用于其他导入，如大事记等，导入的数据内容不同，对于



数据库的字段也不同，不能跨越范围进行导入操作。

⑥服务无状态性，服务本身不保留状态，服务请求者提供服务处理所需的数据，服务则只包含处理逻辑，处理完毕，再把处理结果返回给请求者，服务不保留任何信息。系统不用为每个用户保存状态而导致系统性能降低或消耗服务器资源。档案人员，导入 Excel 时，从客户端将 Excel 上传到服务器上，系统从服务器读取相应的 Excel 文档，并将其数据导入到 SQL 数据库中，最后将 Excel 文档从服务器上删除并提示导入成功。

⑦服务是可被发现性，为了能够被发现，服务必须提供足够详细的描述，而不是要求请求者先阅读相关文档后，才进行服务申请。在确定下载 Excel 模板之前，服务提示请求者，该模板的用途、使用方法、相关规定等信息，即服务的合同，是请求者充分了解该服务是否符合自身需求。

⑧服务逻辑隐蔽性，服务请求者无须了解该服务的具体实现，只需了解服务合同即可。Excel 数据导入的具体逻辑实现，通过 Web Services 封装，对于请求者而言，服务相当于“黑盒”，只需知道怎么用，无须了解如何实现。

为了确保 Excel 导入的数据，服务档案业务规则，“万维数字化档案网上管理系统”采用双重规范——著录控制和导入验证。

①著录控制，放置在服务器上，供档案整理人员下载的 Excel 模板，已经进行“工作表保护”和“工作簿保护”处理的 Excel 文档，固定表头和著录项列数。避免档案人员在填写档案信息过程中，人为地修改表头信息或者增加档案信息项。未获得“工作表保护”密码的情况下，只能在没固化的单元格进行填写档案信息，确保 Excel 文档的基本格式的规范性，能被导入接口识别。

②导入验证，在导入 Excel 数据到 SQL 数据库之前，先对 Excel 数据进行遍历检验。主要对整数类型数据项（如页数、份数等）进行整型验证，对页次格式进行规范性验证（0001-0012 或 0015 两种为合法格式，其他为非法页次格式，如 0001 ~ 0012 等）。另外还有日期、案卷号、卷内号、案卷标题、卷内标题等信息的验证。导入验证采取“零差错导入”，即 Excel 数据中存在任何不符

合档案业务规则的信息，整份 Excel 文档的数据将不予导入，并把不符合业务规则的信息，如在 Excel 中的具体位置（所处行和列）及对应的范例以表单的形式显示出来。只有将所有不规范的信息规范化后，Excel 文档的数据方可导入数据库中，确保导入的档案信息符合档案业务规则。

考虑到多人同时导入，将消耗服务器大量资源，最差情况可能会导致页面响应超时。为了避免此种情况的出现，案卷、卷内数据格式化导入接口，采用“限额服务”和“队列模式”。一次性允许 20 个用户同时作导入操作，超过 20 个用户，则采用先来先服务模式进行数据导入，每当一个用户完成导入操作，将空闲出一个导入名额。类似银行的取号排队服务，导入操作前，如需等候，系统提示当前有几位用户正在排队等候，如果用户确认等候，将进入等候队列。

7.4.2 检索利用

档案业务规则的查询与检索，除了归档范围查询、规则库中词群序列的检索，还包含归档文件的类目级别检索。“万维数字化档案网上管理系统”的检索利用模块包括：综合检索，对所有门类的归档文件进行检索。综合检索，可根据全宗号、归档门类进行筛选检索。其中，全宗号筛选范围，是在登录用户的权限范围内进行选择；科技式快速查询，针对科技式门类的档案进行单条件或多条件方式的检索，针对性更强、更快捷；文书档案检索，包括立卷归档的案卷和卷内检索，以及单件归档检索；特殊载体档案检索，可对照片档案、底片档案、音像档案、实物档案分别检索；专门档案检索，针对档案整理单位自定义的档案门类档案进行对应检索。此外，还对会计档案、已改人员档案、诉讼档案、核算档案、图书资料进行对应检索。

为了精确获得所需的检索结果，系统还设置了检索显示列。用户可通过设定显示项，检索结果只显示符合检索条件的记录的相关数据项，如在确定检索前，设置检索显示项“档号”、“年度”、



“案卷标题”、“页数”、“期限”、“立卷人”、“立卷时间”等。检索结果则只显示以上数据项，把不需要查看的数据项隐藏，使得检索结果显示更简洁。

7.4.3 档案编研

为了记载、管理党政机关、企业单位或社会团体重要活动所产生的应用文件，系统专门设置管理大事记的管理目录。大事记按区域范围，对世界大事记、全国大事记、地区大事记、部门大事记及单位大事记等进行分类管理。另外，大事记还根据制文机构性质、时间跨度加以区分管理。

为预防档案整理人员经验不足或者疏忽操作，使相关资料著录信息有所偏差，系统根据大事记的标题、大事记的制文机构性质、时间跨度，结合保管期限表，启动业务规则引擎，调用业务规则库，把当前大事记档案资料对应的保管期限筛选显示。运用这种著录机制，降低档案整理人员的工作量，同时也确保档案资料的准确性，从总体上提升的档案整理的质量。

对各机关、企事业单位各大门类档案包括文书档案、科技档案、会计档案、声像档案、专门档案、特殊载体档案等实现独有详尽的管理。系统主要实现档案接收、档案著录、整理编目与档案数字化、在线检索利用、档案信息编研、档案资料征集、数据统计、馆务管理、系统维护等功能。

7.5 系统分析与评价

(1) 社会效益

档案资源是人类社会发展的历史记录，是国家的一种原生的、独有的、不可替代的信息资源，是社会的重要财富，更是各立档单位的宝贵财富。有形的物质资源、能量资源和无形的信息资源，并

列为现代经济发展的三大支柱，这三者的关系好比人的体质（物质）、体力（能量）和智力（信息），只有三者具备，才能全面发展。档案信息资源是整个信息资源的精髓。随着全国电子政务建设的发展、信息技术的广泛应用、办公自动化的推行，传统的档案工作方式和手段已不适应时代发展要求。该平台正是在这样一个背景之下立项开发的，它以档案局/馆与各立档单位作为一个整体对象来研发的系统工程，社会意义明显，主要体现在：

节约性：从东莞市档案资源管理的全局视角出发，统筹规划、分步实施，避免重复建设，节约人、才、物资源，提高 IT 投资的利用效率。

和谐性：它正对目前当前各立档单位档案信息化参差不齐的现状，为协调发展提供保证机制。对落后的单位可提升其建设与使用的速度，对信息化程度高的单位能够促进业务流程的规范化和档案信息的共享化。

共享性：它使电子档案馆的 IT 资源能够得到最大限度的共享，包括网络基础设施、服务器、存储设备等硬件设备，档案应用软件，开放的档案资源，编研利用方法等的共享使用。

便利性：它为实物档案馆提供网络化管理数字档案资源、开展现代档案工作的基础支撑平台，虚拟档案室的应用，职能比较单一的立档单位不需要自筹资金开展档案信息化建设工作。

发展性：通过电子档案馆管理平台的应用，能够满足各类档案资源的收、管、用、编、研、展的业务需求，做到系统功能不断随着业务的发展需要而实现扩充、完善和升级。

集中性：实现了东莞市区域性的档案信息资源的集中管理与分布存储，为大力开展档案信息资源的开发与利用提供了基础保障。

通过开发目录中心系统，能够实现解决电子政务建设中各委办局单位电子文件归档管理问题，实现在日常工作中产生的大量电子文件的有效积累和长久保存，防止国家档案信息资源的流失，维护城市历史记忆的完整与安全；通过对馆藏档案进行深入研究和开发，推出各类档案文化产品和档案资源服务项目，为弘扬东莞城市文化和塑造城市精神服务；通过档案信息资源总库及其应用系统建



设,全面整合全市各类档案信息资源,将市档案局馆与各镇、各部门档案机构连通,构建起档案基本信息、相关信息和其他信息之大成的信息资源总库,使其成为东莞市基础信息资源库的一个重要组成部分,为领导、机关和社会大众提供信息查考、信息共享和决策分析;通过在政务网、互联网上构建和深化档案信息化应用项目,形成包括文书档案、照片档案和声像档案的开放信息分布式数据库,实现全市统一的一站式档案共享利用平台,以更简便有效的服务手段和更丰富、更形式多样的档案信息为各级机关和社会大众提供服务。

该平台的建设,从宏观上来讲,将汇总各级档案部门的信息资源,建成东莞市档案信息资源总库,为各级机关和企事业单位提供信息服务;从微观上来讲,档案信息化建设将深入基层、深入群众,通过将名人档案、婚姻档案、房产档案、新农村建设档案等和人民群众生活密切相关的档案资源数字化并提供全方位、多角度、深层次的利用服务,是档案部门为普通老百姓做的一件实事,也是档案部门为建设社会主义和谐社会做出的巨大贡献。

(2) 经济效益

该平台集局/馆采集、保存、利用、管理、宣传等业务于一体,基本实现档案馆业务的全程系统化,规范管理程序,实现部门、各岗位、各员工的协同工作,降低管理成本。电子档案馆通过市政网、电子密钥、VPN 的应用,通过符合 Web Service 标准的接口建立和数据交换系统,将市档案局与全市各立档部门构建为一个有机整体。基本实现降低管理成本,减少工作量,提高工作效率,降低工作强度的目标,更远程实现掌握动态、把握结构、采集数据、有效控制管理手段。

该平台的建设不仅将产生巨大的社会效益,而且经济效益也是可观的。各类文书、财会、科技等档案的应用将为东莞市经济发展推波助力,为社会带来难以估量的经济效益。照片、声像等多媒体档案的利用将在经济开发和发展中,展现东莞市新的面貌和活力,为招商引资、产品推广、技术应用起到直接的推动作用。

跨行业、跨部门的档案资源整合将可以提供婚姻登记、国土房产、城建等类型的专题档案信息服务，还可以定期开展知名品牌、东莞名人等专题档案展示活动，并可以结合市级、国家级乃至世界级大型活动，进行档案资源的再挖掘、再结合和再利用，创造可观的经济效益。

本平台的建设另一个重大意义是，通过横跨区域和部门、纵贯档案专业条线的设计理念，对全市档案部门的信息化建设进行统一规划，避免各级档案部门、各机关档案室系统信息化项目的重复建设，大大减少政府投资，并能为全市党政机关提供档案资料信息资源的一体化管理和高质量服务。

参 考 文 献

(一) 著作

- [1] 安小米. 基于 ISO15489 的文件档案管理核心标准及相关标准解析 [M]. 北京: 中国标准出版社, 2013.
- [2] 陈兆祜. 档案管理学基础 [M]. 北京: 中国人民大学出版社, 2005.
- [3] 陈智为, 邓绍兴, 刘越男. 档案管理学 [M]. 北京: 中国人民大学出版社, 2008.
- [4] 丁海斌, 卞昭玲主编. 电子文件管理基础教程 [M]. 沈阳: 辽宁大学出版社, 2011.
- [5] 段荣婷. 中国电子文件知识组织 XML 集成置标标准化研究 [M]. 上海: 上海交通大学出版社, 2012.
- [6] 冯惠玲. 电子文件风险管理 [M]. 北京: 中国人民大学出版社, 2008.
- [7] 冯惠玲, 等. 中国电子文件管理: 问题与对策 [M]. 北京: 中国人民大学出版社, 2009.
- [8] 国际档案理事会. 电子办公环境中文件管理原则与功能要求 [M]. 王健译. 北京: 中国人民大学出版社, 2012.
- [9] 傅荣校. 档案管理现代化: 档案管理中技术革命进程的动态审视 [M]. 杭州: 浙江大学出版社, 2002.
- [10] 侯延香, 王霞. 信息采集 [M]. 北京: 知识产权出版社, 2012.
- [11] 金波, 丁华东编著. 电子文件管理学 [M]. 上海: 上海大学出版社, 2007.

- [12] 刘越男著. 建立新秩序——电子文件管理流程研究 [M]. 北京: 中国人民大学出版社, 2005.
- [13] 卢加元. 信息系统风险管理 [M]. 北京: 清华大学出版社, 2011.
- [14] 覃兆别. 企业档案的价值与管理规范 [M]. 上海: 世界图书出版公司, 2010.
- [15] 邵军力, 等. 人工智能基础 [M]. 北京: 电子工业出版社, 2000.
- [16] 王传宇, 张斌主编. 科技档案管理学 [M]. 北京: 中国人民大学出版社, 2009.
- [17] 王伟军, 甘春梅. Web2.0 信息资源管理 [M]. 北京: 科学出版社, 2011.
- [18] 王新才主编. 电子政务信息资源管理及其技术实现——2007 信息化与信息资源管理学术研讨会论文集 [M]. 武汉: 湖北人民出版社, 2008.
- [19] 王英玮主编. 专门档案管理 [M]. 北京: 中国人民大学出版社, 2010.
- [20] 吴志荣, 胡振华. 虚拟与现实: 数字信息交流研究 [M]. 上海: 上海辞书出版社, 2009.
- [21] 肖秋惠编著. 档案管理概论 [M]. 武汉: 武汉大学出版社, 2009.
- [22] 叶萍. 管理学基础 [M]. 北京: 电子工业出版社, 2007.
- [23] Ross G. R. 业务规则方法原理 [M]. 韩柯, 孟海军, 译. 北京: 机械工业出版社, 2004.
- [24] Ross G. R. *Business Rule Concepts: Getting to the Point of Knowledge* (4th Edition) [M]. Houston: Business Rule Solutions, 2013.
- [25] Ross G. R. *Principles of the Business Rule Approach* [M]. Boston: Addison-Wesley Professional, 2003.
- [26] Duranti L. *International Research on Permanent Authentic Records in Electronic Systems (InterPARES): Experiential, Interactive*

- and Dynamic Records* [M]. Padova: Associazione Nazionale Archivistica Italiana, 2008.
- [27] Shepherd E., Yeo G. *Managing Records: A Handbook of Principles and Practice* [M]. London: Facet Pub, 2003.
- [28] Herbst H. *Business Rule-oriented Conceptual Modeling* [M]. Heidelberg: Physica-Verlag GmbH, 1997.

(二) 期刊

- [1] 陈金冠. 创新质检方法 规范数字档案——数字档案馆(室)质检方法研究 [J]. 浙江档案, 2006 (9): 34-36.
- [2] 曹龄兮, 李建华, 娄悦. 基于多次权重鉴别的信息雷达及应用 [J]. 现代计算机, 2006 (2): 12-15.
- [3] 陈霞. 论网络环境下档案业务流程重组 [J]. 兰台世界, 2005 (1): 9-10.
- [4] 蔡学美. 档案管理视角下的电子文件元数据 [J]. 中国档案, 2014 (4): 63-65.
- [5] 段荣婷. 论国际档案信息标准化现状、发展趋势及我国的对策 [J]. 档案学研究, 2008 (1): 38-46.
- [6] 冯立. 电子档案价值鉴定及制约因素 [J]. 兰台世界, 2008 (1): 37-38.
- [7] 高复先. “数字校园”建设的基础工程: 信息资源规划 [J]. 中国教育网络, 2006 (8): 63.
- [8] 高复先. 信息资源规划的理论指导 [J]. 中国教育网络, 2006 (9): 62-64.
- [9] 高复先. 信息资源规划的工程化方法 [J]. 中国信息界, 2005 (20): 21-24.
- [10] 政府信息资源开发利用政策研究课题组. 信息资源规划与国家基础数据库的政策建议 [J]. 中国信息界, 2004 (7): 7-12.
- [11] 郭小磊. 浅谈人工智能技术在档案管理中的应用和加强应用的必要性 [J]. 才智, 2009 (18): 259-260.

- [12] 顾志峰, 董文生, 吴凤品, 仲跻炜. 业务规则管理系统在金融信息系统中的应用 [J]. 信息技术与标准化, 2013 (3): 54-57.
- [13] 黄玉明. 电子文件和相关数字资源封装案例比较 [J]. 档案学研究, 2012 (5): 55-58.
- [14] 罗程. 两种不同的信息资源规划 [J]. 中国信息界, 2007 (Z1): 44-45.
- [15] 吕良双, 蔡欢. 一种面向业务人员的基于对象的规则描述语言 [J]. 计算机与数字工程, 2008 (5): 33-37.
- [16] 李姝琳, 孙德杰, 史震东, 李霜鸿. 信息资源规划——IRP 技术研究 [J]. 中国水运 (下半月), 2008 (4): 230-231.
- [17] 李雯. 电子文件价值智能鉴定探析 [J]. 档案学研究, 2011 (6): 41-45.
- [18] 刘新安, 翟瑶, 黄小萍, 柳美丽. 非元数据档案整理技术的实践 [J]. 档案学通讯, 2008 (1): 19-23.
- [19] 路宜庆. 论企业档案信息化标准规范体系的建设 [J]. 山东档案, 2008 (2): 25-26.
- [20] 孟歆. 探析知识管理思想下的档案管理流程重组 [J]. 湖北档案, 2007 (9): 19-20.
- [21] 屈秀兰. 应用人工智能技术开发利用企业档案资源的探讨 [J]. 兰台世界, 2008 (4): 19-20.
- [22] 司莉. 网络信息资源组织研究进展 [J]. 情报科学, 2003 (6): 653-658.
- [23] 谭琤培, 章丹. 档案元数据在电子文件鉴定中的运用——元数据研究之三 [J]. 浙江档案, 2002 (6): 6-8.
- [24] 屠跃民, 李婉月. 关于数字档案信息采集的思考 [J]. 档案与建设, 2006 (9): 17-20.
- [25] 吴光华. 略论国家档案馆信息资源建设规划 [J]. 上海档案, 2008, (9).
- [26] 谢国强, 蓝立新. 基于 Web 的网络爬虫技术研究 [J]. 科教文汇 (中旬刊), 2008 (4): 198.



- [27] 吴品才. 从元数据与文件的关系准确认识电子档案的形成规律 [J]. 浙江档案, 2004 (7): 14-16.
- [28] 王雯. 档案馆业务流程重组的思考 [J]. 兰台世界, 2007 (18): 9-10.
- [29] 王新才, 李雯, 陈划生. 论档案信息资源规划的实施 [J]. 信息资源管理学报, 2011 (1): 91-94.
- [30] 王新才, 李雯, 丁时征. 业务规则管理在档案信息资源管理系统中的应用 [J]. 档案学通讯, 2010 (4): 79-84.
- [31] 王心裁, 吕元智. 超媒体数据库技术与档案信息组织 [J]. 中国图书馆学报, 2003 (1): 67-70.
- [32] 王英玮. 论档案的有机联系性 [J]. 档案学研究, 2003 (6): 2-4.
- [33] 薛匡勇. 新世纪档案职业的走向——中国首届档案学博士论坛综述之一 [J]. 湖北档案, 2001 (12): 12-14.
- [34] 徐黎, 冯元勇, 糜宏斌, 孙玉芳. 基于受限中文的轻型业务规则管理系统的设计与实现 [J]. 计算机科学, 2005 (2): 87-89+99.
- [35] 徐拥军, 周艳华, 李刚. 基于知识服务的档案管理模式的理论探索 [J]. 档案学通讯, 2011 (2): 24-28.
- [36] 叶钰, 应时, 李伟斋, 张韬. 面向服务体系结构及其系统构建研究 [J]. 计算机应用研究, 2005 (2): 32-34.
- [37] 赵传玉. 论档案信息化标准体系的构建 [J]. 湖北档案, 2009 (Z1): 37-39.
- [38] 赵国俊. 档案信息资源开发利用专项规划及其内容体系 [J]. 北京档案, 2005 (9): 16-18.
- [39] 钟晓, 马少平, 张钹, 俞瑞钊. 数据挖掘综述 [J]. 模式识别与人工智能, 2001 (1): 48-55.
- [40] 张正强, 崔屏. 美国对密级电子文件管理的要求及我们对我国现状的思考 [J]. 兰台世界, 2004 (9): 16-17.
- [41] Bajec M., Krisper M. A Methodology and Tool Support for Managing Business Rules in Organizations [J]. *Information*

- Systems*, 2005, 30 (6): 423-443.
- [42] Corradini F. , Polzonetti A. , Riganelli O. Business Rules in E-Government Applications [J]. *Electronic Journal of E-Government*, 2009, 7 (1): 45-54.
- [43] Cook T. Evidence, Memory, Identity, and Community: Four Shifting Archival Paradigms [J]. *Archival Science*, 2013, 13 (2-3): 95-120.
- [44] Duranti L. Concepts and Principles for the Management of Electronic Records, or Records Management Theory is Archival Diplomatics [J]. *Records Management Journal*, 2010, 20 (1): 78-95.
- [45] Herbst H. Business Rules in Systems Analysis: A Meta-model and Repository System [J]. *Information Systems*, 1996, 21 (2): 147-166.
- [46] Krammer M. I. Business Rules: Automating Business Policies and Practices [J]. *Distributed Computing Monitor May*, 1997.
- [47] Kardasis P. , Loucopoulos P. Expressing and Organising Business Rules [J]. *Information and Software Technology*, 2004, 46 (11): 701-718.
- [48] Maciol A. An Application of Rule-based Tool in Attributive Logic for Business Rules Modeling [J]. *Expert Systems with Applications*, 2008, 34 (3): 1825-1836.
- [49] Rosca D. , Greenspan S. , Wild C. Enterprise Modeling and Decision-support for Automating the Business Rules Lifecycle [J]. *Automated Software Engineering*, 2002, 9 (4): 361-404.
- [50] Ram S, Khatri V. A Comprehensive Framework for Modeling Set-based Business Rules during Conceptual Database Design [J]. *Information Systems*, 2005, 30 (2): 89-118.

(三) 学位论文

- [1] 付玲. 基于 SOA 的吉林银行档案管理系统的设计与实现 [D]. 吉林: 吉林大学, 2009.
- [2] 李泽锋. 基于 OAIS 的可信电子文件管理系统体系架构研究 [D]. 北京: 中国人民大学, 2009.
- [3] 林金娇. 业务规则与企业协作系统适应性的研究 [D]. 济南: 山东大学, 2008.
- [4] 马林青. 基于职能的文件分类体系研究 [D]. 北京: 中国人民大学, 2011.
- [5] 杨静. 业务规则管理技术的研究与应用 [D]. 上海: 华东师范大学, 2007.

(四) 会议论文

- [1] Herbst H. , Myrach T. A Repository System for Business Rules [C]. IFIP Conference Proceedings, 6th IFIP TC-2 Working Conference on Data Semantics: Database Applications Semantics, Georgia, 1995: 119-139.
- [2] Rosca D. , D'Attilio J. Business Rules Specification, Enforcement and Distribution for Heterogeneous Environments [C]. Computer Software and Applications Conference, COMPSAC 2001, 25th Annual International IEEE, 2001: 3-9.
- [3] Rosca D. , Greenspan S. , Feblowitz M. , et al. A Decision Making Methodology in Support of the Business Rules Lifecycle [C]. Requirements Engineering, Proceedings of the Third IEEE International Symposium on IEEE, 1997: 236-246.
- [4] Rosca D. , Greenspan S. , Wild C. , et al. Application of a Decision Support Mechanism to the Business Rules Lifecycle [C]. Knowledge-based Software Engineering Conference, Proceedings 10th. IEEE, 1995: 114-121.
- [5] Rosca D. , Wild C. Business Rules in the Real World: A Decision

Support Approach [C]. The Software Engineering and Software Engineering Conference (SEKE96), Lake Tahoe, NE, 1996: 121-128.

(五) 报告

- [1] Hay D. , Healy K. A. . Defining Business Rules—What Are They Really? [R]. The Business Rules Group, 2000.
- [2] GUIDE. GUIDE Business Rule Project, Final Report [R]. Nov. 6, 1995.

(六) 电子文献

- [1] 档案工作行业标准目录 [EB/OL]. [http: /www. saac. gov. cn/](http://www.saac.gov.cn/), 2010-01-11.
- [2] Buchanan B G, Shortliffe E H. Rule Based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project [M/OL]. Addison-Wesley Longman Publishing Co. , Inc. , 1984, [http: //www. aaai. org/AITopics/pmwiki/pmwiki. php/AITopics/RuleBasedExpertSystems](http://www.aaai.org/AITopics/pmwiki/pmwiki.php/AITopics/RuleBasedExpertSystems).
- [3] A Brief History of the Business Rule Approach [J/OL] Business Rules Journal, 2006, 7 (11) . [http: //www. BRCommunity. com/a2006/b317. html](http://www.BRCommunity.com/a2006/b317.html).
- [4] Horrocks I, Patel-Schneider P F, Boley H, et al . SWRL: A Semantic Web Rule Language Combining OWL and RuleML [EB/OL]. [http: //www. w3. org/Submission/SWRL/](http://www.w3.org/Submission/SWRL/), 2004-05-21.

(七) 标准

- [1] GB/T 17859-1999 计算机信息系统安全保护登记划分准则 [S]. 北京: 中国标准出版社, 1999.
- [2] GB/T18894-2002 电子文件归档与管理规范.
- [3] GB/Z26822-2011 文档管理电子信息存储真实性可靠性建议.
- [4] GB/T17678. 1-1999CAD 电子文件光盘存储、归档与档案管理

要求第一部分：电子文件归档与档案管理.

[5] GB/T17678.2-1999CAD 电子文件光盘存储、归档与档案管理
要求第二部分：光盘信息组织结构.

[6] GB/Z23283-2009 基于文件的电子信息的长期保存.



数字时代图书馆学情报学研究论丛

(第二辑)

书 目

竞争情报与企业危机管理	查先进 陈明红 杨凤
现代服务管理理论与实践	李枫林
数字信息生命周期管理	马费成
信息资源集成	唐晓波
多语种叙词本体	邓仲华 赵又霖 黎春兰 汤平
企业竞争情报智能采集	张玉峰
数字化信息服务	胡昌平 邓胜利
泛在环境下的信息采集理论与应用	张李义
数字化信息时代的图书馆管理	肖希明
基于业务规则的档案信息资源管理	王新才 江善东
人事档案管理原理与方法	朱玉媛 周耀林
图书馆与信息领域知识产权问题研究	陈传夫
数字信息资源的开放存取	黄如花
XML信息检索	陆 伟
现代企业信息系统的演进	张晓娟
抗战时期中国出版业研究	吴永贵
信息相似性研究	沈 阳
国故新证	司马朝军
商务信息检索理论与技术	李湘东