

数字档案馆建设指南

国家档案局
2010年

目 录

1. 总体要求.....	1
1.1 概述.....	1
1.2 建设原则与要求	1
1.3 建设目标.....	2
1.4 建设内容.....	3
1.5 建设步骤.....	4
2. 管理系统功能要求.....	6
2.1 收集功能要求.....	7
2.2 管理功能要求.....	8
2.3 保存功能要求	9
2.3.1 长久保存策略.....	9
2.3.2 存储架构选择	9
2.4 利用功能要求.....	10
3. 应用系统开发和服务平台构建.....	11
3.1 运用系统开发.....	11
3.2 服务平台建设.....	12
3.3 软硬件设备配置.....	14
3.3.1 系统软件选择.....	14
3.3.2 硬件设备配置.....	15
4. 数字档案资源建设.....	16
4.1 电子文件接收.....	17
4.2 档案数字化.....	18
4.3 资源整理.....	19
4.4 建立数字档案资源库.....	20
5. 保障体系建设.....	21
5.1 安全保障体系建设.....	21
5.2 标准规范体系建设.....	23

1. 总体要求

1.1 概述

数字档案馆是指各级各类档案馆为适应信息社会日益增长的对档案信息资源管理、利用需求,运用现代信息技术对数字档案信息进行采集、加工、存储、管理,并通过各种网络平台提供公共档案信息服务和共享利用的档案信息集成管理系统。

建设数字档案馆,有利于提高档案馆工作效率和现代化水平,有利于确保数字档案永久存储与安全保管,有利于促进公共档案服务能力拓展和实现档案信息资源的社会共享。

建设数字档案馆,有利于促进国家信息资源总量增加、质量提高和结构优化,有利于提高各级政府公共服务能力,有利于促进社会主义文化的发展繁荣,有利于满足广大人民群众对现代信息服务的现实需求。

数字档案馆建设是一项系统工程,需要经过调研、立项、论证、软件开发、平台构建、资源准备、系统试运行、项目验收、运行维护等诸多环节,需要各业务部门、档案管理部门以及其他相关单位的共同参与,是一项长期的工作任务。

1.2 建设原则与要求

数字档案馆建设应当遵照“统筹规划、循序渐进,项目带动、重点突破,需求导向、保证安全,合理适用、稳步实施”的原则,积极推进,务求实效。

统筹规划、循序渐进。应当根据各级政府电子政务建设整体框架和基本要求,及时将数字档案馆建设与区域电子政务和信息化建设相衔接,综合考虑自身信息化发展水平、技术力量、资源规模、基础工作水平、资金投入等因素,确定总体布局和实施步骤。

项目带动、重点突破。实施项目带动战略,支撑和带动数字档案馆整体建设和发展。将数字档案馆建设分成若干子项目,以项目带动数字档案馆建设及推进各项工作的开展,有利于获得党委政府和有关部门的重视和投入,有利于集中人力物力有针对性地重点解决数字档案馆建设中的核心问题。

需求导向、保证安全。数字档案馆建设应以需求为导向,着眼于党政机关,社会公众在线利用数字档案信息资源的需求,推动档案信息资源的共享,应当按照信息安全等级保护的要求,采取相应的安全保障技术方法,配备必要的软硬件设施,建立健全安全管理制度完善灾难恢复服务机制,确保数字档案馆建设和运行的安全。

合理适用、稳步实施。要重视前期调研,进行可行性研究,结合本部门业务工作实际,提出创新性高,应用性强的功能需求,以确保系统开发和项目建设的合理性、适用性、在项目实施过程中,档案部门的业务人员应全程参与,及时研究和发现问题,纠正偏差。重大项目应当按照信息系统建设规范要求引入监理机制,对项目质量、进度、投资、安全等方面进行全程监理。

1.3 建设目标

数字档案馆建设的目标是:紧紧依靠国家和当地信息化基础设施建设环境,充分利用各种政务网平台,公众网平台以及各类网络资源,以先进的信息技术为手段,集成建设适应本部门本单位一定时期内数字档案管理需要的网络平台,开发应用符合功能要求的管理系统,推动馆藏档案资源数字化、增量档案电子化,逐步实现对数字档案信息资源的网络化管理以及多层次多渠道提供档案信息资源利用和社会共享服务。

其中,数字档案信息资源建设,提供档案信息利用和社会共享服务是一项长期任务应分阶段推进,有步骤实现。

1.4 建设内容

数字档案馆总体建设内容主要包括:

(1)按照数字档案馆基础网络架构,主要技术路线与软硬件配置基本要求,集成建设适应馆藏档案基础数据和今后一定时期内数字档案增长规模的数据管理、满足数字档案馆各项管理与服务需求的基础设施。

(2)开发或应用具备“收集、管理、存储、利用”等功能要求的数字档案管理系统,实现档案管理各业务环节的自动化、网络化;

(3)全面推进馆藏数字档案基础数据库建设,优先建立馆藏档案的文件级目录数据库,逐步进行传统载体档案的数字化转换,积极推进电子文件的接收和管理,建立各类数字档案资源库群。

(4)应用先进技术和相关管理手段,保证数字档案信息资源的可靠可信和长期可用,减少数字档案对硬件的依赖性,从而使数字档案具备传统档案所具有的原始性、凭证性和长期可读性。

(5)运用多种技术手段,针对不同利用对象,通过不同渠道,实现档案信息资源分层共享,方便、快捷满足各类用户利用需求。

(6)配套建设数字档案馆保障体系,确保数字档案馆系统安全和数字档案信息安全。

1.5 建设步骤

数字档案馆建设一般分项目规划与立项、项目招投标、项目实施、运行维护等几个步骤进行。

(1)项目规划与立项

根据社会信息化和档案事业发展需要,研究提出数字档案馆建设项目。成立由档案馆主要领导、信息化人员、相关技术人员、档案管理人员等参加的项目筹备小组,对本地区信息化现状和国内外数字档案馆建设现状等,特别是各立档单位产生电子文件状况、党委政府各部门和公众对档案利用需求情况进行充分调研,研究提出项目可行性研究报告,向当地主管部门提出数字档案馆立项申请。在与相关主管部门充分沟通后,由主管部门组织专家对项目可行性研究报告进行论证。

(2)项目招投标

项目经立项批准后,应按照国家有关项目管理规定进行招标或邀标工作。充分征求档案馆各业务部门,相关专家的意见,对数字档案馆的功能需求进行细化,形成功能需求方案。功能需求方案应既符合档案管理的实际需要,又便于计算机编程人员理解,能够在系统开发中实现。

(3)项目实施

按照数字档案馆建设目标、总体内容和项目可行性研究报告立项批复的要求,制订详细的项目建设实施方案,明确项目建设的内容、进度和时间安排,有计划地推进系统开发与调试、软硬件集成、项目试运行与验收等项目实施工作。项目建设中应当加强项目监理,确保项目建设的进度和质量。

(4)运行维护

项目验收后,应适时调整、明确数字档案馆各功能模块的管理职责,加强应用培训,建立数字档案馆业务工作与技术工作相互衔接、相互配合,以业务管理工作为主线、以技术保障工作为支撑的管理机制。要重视人才培养,形成以档案馆人才自主管理为主体,社会外包服务保障为依托的运行与维护体系。

2. 管理系统功能要求

数字档案管理系统应当具备“收集、管理、保存、利用”四项基本业务功能,以及用户权限管理、系统日志管理、数据备份与恢复、系统及其数据安全维护等功能。数字档案管理系统还应当采取必要措施保证馆藏数字档案信息,特别是由电子文件归档形成的电子档案信息的可靠和可用。数字档案管理系统功能可以根据信息化发展和档案管理的要求而有所侧重并不断拓展。

2.1 收集功能要求

系统应当具备接收立档单位产生的电子文件及其元数据、对传统载体档案进行数字化和采集重要数字信息资源等功能。

(1)系统能够根据相关要求接收立档单位产生的各类电子文件及其元数据,并在建立一整套接收机制基础上,保证接收过程责权明确,杜绝安全隐患,从源头上保证数字档案的真实、完整、可用。

(2)系统应当提供选择在线接收和脱机接收方式。

(3)系统应当能够批量导入或导出数据,保证数据的可靠和可用。

(4)系统应能对在线或离线接收的档案数据进行真实性、完整性、可用性验证。

(5)系统应具备目录数据和内容数据等多种信息资源的采集功能。

2.2 管理功能要求

系统能够对所接收的各类数字档案信息进行整理、比对、分类、著录、挂接、鉴定、检索、统计等操作,使无序信息有序化,并实施有效控制。系统应能:

(1)按照设定的分类方案,将数字档案信息存储到系统中,或根据管理要求进行适当调整。

(2)过滤重份数据和重新分类、编号。

(3)对档案内容进行抽取和添加元数据等操作。目前档案管理都是基于目录数据库挂接方式来实现,将来不排除使用新的技术方法对数字档案进行有效管理。

(4)辅助人工完成档案的开放鉴定工作。

(5)对档案内容数据及其元数据等相关信息建立持久联系,形成长期保存档案数据包和利用数据包。

(6)对档案类型、数量大小等按照设定要求进行统计、显示或打印输出所需各类档案信息。

(7)辅助完成馆藏实体档案编目(著录标引)整理出入库房管理等工作。

(8)定制档案业务流程或进行流程再造。

2.3 保存功能要求

系统能够实现对数字档案长久的安全保存,包括两方面的要求,一是长久保存策略的确定,二是存储架构选择。

2.3.1 长久保存策略

长久保存策略包括存储格式的选择,检测、备份和迁移等技术方法的采用等。

(1)当选择符合国家标准格式、暂时未制定标准的,选择开放格式或主流格式。

(2)定期对载体及其软硬件环境进行读取、测试,发现问题,及时解决。

(3)根据数据重要程度以及管理和利用的需要,选择在线、近线、离线、异地、异质和分级存储等技术和方式。

(4)计算机软硬件以及技术或标准规范发生重大变化或发生重大事件时,为了保证数字档案信息可读,应采取迁移等手段对所存储的数据进行技术处理。

2.3.2 存储架构选择

根据档案数据量和管理目的不同而采用不同的存储技术及其相关设备。安全性和稳定性是选择存储设备的首要因素。在数字档案馆建设过程中,应根据数字档案馆的数据量和利用并发用户数的需求,以保证数字档案馆合理安全的存储容量和较快的网络传输速度,适当选择采用单一应用平台,配备数据库服务器、文件存储器、备份服务器、备份软件等构成的存储服务平台,以及采用 SAN、NAS、DAS、IP-SAN 或其他形式的存储技术方法。

2.4 利用功能要求

数字档案管理系统应当根据档案信息的利用需求和网络条件,分别通过公众网,政务网、局域网等建立利用窗口。系统应能实现档案查询、资源发布、信息共享、开发利用、工作交流、统计分析等功能。系统应当:

- (1)能够运用最新检索技术方法满足利用者在各种利用平台对档案数据进行快速、准确、全面的利用查询要求。
- (2)能够通过网络平台或特定载体发布档案信息和信息资源共享。
- (3)能够辅助进行档案信息智能编研、深度挖掘。
- (4)能够为档案管理者 and 利用者提供在线交流平台、远程指导、远程教育。
- (5)能够辅助开展数字档案的增值服务。
- (6)能够进行档案利用访问量统计、分布分析、舆情分析等相关工作。
- (7)能够对用户、数据项、功能组件进行利用权限的角色授权处理,能够进行门类设置、结构设定、字典定义等系统代码维护工作。

3. 应用系统开发和服务平台构建

3.1 应用系统开发

鼓励各地区统一开发或购置成熟的数字档案管理系统软件,减少重复开发,降低应用成本。

应用系统开发应当遵循整体性、开放性、稳定性等原则。鼓励软件开发公司应用先进技术手段对“收集、管理、保存、利用”各功能模块或子系统进行专业深度扩展、创新开发。

(1)整体性。系统应考虑所配备和购置的软硬件及其网络平台环境,选择恰当的开发工具和技术路线,正确处理各子系统或模块的关系,形成一个整体。

(2)开放性。系统应能够随着信息技术的不断发展和档案管理的最新要求而具有兼容性和拓展性。

(3)稳定性。系统开发应采用先进、成熟、稳定的技术,保证系统的稳定、可靠和安全。

3.2 服务平台建设

数字档案馆网络架构一般应面向不同对象,立足现有不同网络,构建三个服务平台,并提供相应层级数字档案信息资源利用共享服务。

数字档案馆应根据不同服务对象和不同档案开发范围建立相应的服务平台。一般主要包括:一是基于局域网面向档案馆工作人员和来馆利用档案人员的馆内档案利用服务平台;二是利用当地政务网建设的面向本级党政机关各立档单位的电子文件归档和档案信息共享平台;三是利用公众网建设的面向广大社会公众和进行馆际交流的公共档案信息服务平台。

(1)基于局域网的档案服务平台。

局域网档案服务平台是数字档案馆建设的基础平台。局域网档案服务平台应当具备馆

藏数字档案传途、交换、存储、安全防护的功能,承担档案馆“收集、管理、保存、利用”四项基本功能,满足日常数字档案馆业务管理和提供利用服务的需要。局域网通常还要承担辅助档案实体管理的功能。

(2)基于政务网的档案服务平台。

政务网档案服务平台是数字档案馆连接本级各党政机关立档单位的主干平台。它依托本级政务网,能够接收各立档单位电子文件,能够为政务网用户提供在线档案查阅利用,档案业务指导或其他档案工作服务,实现党政机关的档案信息资源共享和资政服务工作。

鼓励具备条件的档案馆探索采用云计算等先进技术为各立档单位提供软件服务和存储服务。

(3)基于公众网的档案服务平台。

公众网档案服务平台,如公共档案信息服务平台,是档案馆实现公共档案服务和档案信息资源社会共享的有效途径之一。它依托公众网,通过档案网站建立满足公众查阅档案需求的利用窗口,同时,采集具有重要保存价值的各类数字信息,进行资源整合,实现公众档案信息资源的社会最广泛共享,该平台还可采取必要的安全措施,实现馆际档案信息交流。

档案服务平台建设当前应遵循的标准规范有:

(1)GB50174-2008 电子信息系统机房设计规范;

(2)GB/Z24294-2009 信息安全技术基于互联网电子政务信息安全实施指南。

3.3 软硬件设备配置

3.3.1 系统软件选型

数字档案馆系统软件一般包括操作系统,数据库系统等方面。

操作系统的选择要根据数字档案馆的规模和硬件配置等情况来选定。终端操作系统通常选择使用与其软硬件兼容,使用较为广泛的操作系统;服务器操作系统要综合考虑任务量、并发用户、安全性等因素选择操作系统。

数据库系统首先满足数据存储和管理的需求,并充分考虑技术和管理成本;其次应注意操作系统和应用软件对数据库系统的支持;再次要考虑其开放性和扩展性,为将来系统的升级、迁移免除后顾之忧。数据库系统必须考虑到稳定和高性能以及档案资源的安全性,所以在选型之前一定要认真规划,周密地分析与考虑,所选择数据库系统应能支持当前及可预见的将来采用的软件,并充分考虑今后几年数据量的要求。在前期规划时,不仅需要决定使用的类型和版本,还需要规划设计数据结构;不仅需要考虑当前存储和管理的数据信息类型,还要考虑将要接收进馆的数据信息类型。

其他软件配置。在数字档案馆系统运行中,除了操作系统和数据库系统之外,还需根据数据采集、管理、利用等方面需要,配置其他一些系统软件和工具软件,如数据仓库、数据交换、OCR、格式转换、非线性编辑、全文检索等方面的软件。

3.3.2 硬件设备配置

数字档案馆系统硬件设施主要包括服务器、终端、网络、存储及其他配套设备。系统中每个独立任务应有相应的服务器来承担处理任务。服务器在信息系统中的重要性不同,可以分成主服务器和辅助服务器两类。主服务器主要承担信息系统数据存储和管理以及应用系统的运行,主服务器的选配必须充分考虑系统性、安全性和高性能。辅助服务器包括辅助应用服务器、网络管理服务器、文件服务器等,承担系统中非核心功能的处理任务。一个数字档案馆系统配备服务器的数量,取决于系统的实际功能,不必强调所有任务都有独立的服务器,可根据财力、数据规模、用户数量及其访问量等,合理选定。

网络设备主要包括传输介质、集线器、交换机、路由器、光电转换等。双绞线和光纤是最常用的传输介质,物理上分散的多台计算机只有通过传输介质才能连接在一起形成网络;集线器可以放大传输信号,主要用于小型网;交换机是一种基于点对点发送数据的网络连接设备,具有网络流量控制等性能,是大型网络组网的必需设备;要使网络之间互连,则必须配备路由器;光电转换设备是实现双绞线和光纤两种不同传输介质的信号转换的设备。

存储设备主要包括磁盘、光盘、磁带等。应根据数字档案馆的存储架构而选择存储设备。

数字化设备包括扫描仪、数码照相机、数码摄像机、非线性编辑系统等。

终端设备包括输入、输出设备,通常是指独立工作的计算机、打印机、复印机等。数字档案馆各项业务的开展必须依托这些设备,通过终端访问网络应用系统,完成信息处理工作。一个数字档案馆终端设置应考虑三种用户,即社会用户、档案馆工作人员、档案馆聘请的数据加工和处理人员。

4. 数字档案资源建设

数字档案资源建设是数字档案馆建设的核心内容,也是一项经常性的业务工作。数字档案资源建设包括电子文件接收、传统载体档案数字化转换、资源整理加工、建立各类资源库等内容。

4.1 电子文件接收

应当根据档案接收范围,建立电子文件接收进馆制度和机制,配备必要的技术手段,从源头上保证数字档案信息的真实、完整、可用。

为了保证有价值的数字档案资源接收进馆,从数字档案馆建设开始,就应按照数字档案馆功能要求,研究确定电子文件接收范围,标准和方法。在范围上,除了参照纸质档案接收要求,还应针对电子文件多样性、形象性的特点,接收反映重大事件、重要活动、重大变化的电子形式的记录;在种类上,除了文本文件,还包括数码照片、图形图像、多媒体、数据库、网页等各种形式的电子文件。应当明确电子文件移交的方式和时限。移交方式可以是在线也可以离线移交,但不论何种移交方式,以安全、高效、准确为原则。与纸质文件对应电子文件的移交和接收之前,应要求和督促立档单位建立正确无误的对应关系,并保证其内容的一致性。应当加强对电子文件形成及其整理、归档、移交的监督指导,包括确定归档范围规范、进行质量检查、开展技术服务等。

电子文件的接收与管理应当遵循以下标准规范:

- (1)GB/T18894-2002 电子文件归档与管理规范;
- (2)DA/T32-2005 公务电子邮件归档与管理规则;
- (3)DA/T46-2009 文书类电子文件元数据方案;
- (4)DA/T47-2009 版式电子文件长期保存格式需求;
- (5)DA/T48-2009 基于XML的电子文件封装规范。

4.2 档案数字化

传统载体档案数字化是现阶段数字档案资源建设的一个重要途径。档案数字化工作是一项系统工程,涉及到档案保管、保护、整理、鉴定、转换、存储、利用等多个环节,应当统筹规划,分步实施。应当通过数字化工作,对馆藏档案进行全面的梳理。要高度重视档案整理、鉴定、保护等基础工作,为数字化工作提供高质量的来源。如果对馆藏档案无法一次性全面数字化,可以按照特殊载体优先、重要程度优先、共享性强优先等原则分步实施。

数字化加工一般采取自主加工和委托加工两种方式进行。自主加工是档案馆自行配备

数字化加工设备,自行组织人力开展数字化加工。这一模式适用于少量重要、核心档案的数字化加工。委托加工是将应数字化的档案,委托专业公司实施加工。这一模式效率相对较高,投资相对节省,普遍适用于各类档案的数字化加工。档案馆所委托加工的公司必须是具有相关保密资质的专业公司。

档案数字化工作当前应遵循的标准规范有:

- (1)DA/T31-2005 纸质档案数字化技术规范;
- (2)DA/T43-2009 缩微胶片数字化技术规范。

4.3 资源整理

在对信息资源进行有序、有效管理之前,必须进行整理,包括分类、价值鉴定、开放审查等。应当将数字档案信息进行有序排列,当前国家尚未制订针对数字档案信息分类的规则,在数字档案馆建设中,可以参照传统载体档案的分类、排序方式,结合数字档案信息的特点,确定数字档案信息的合理分类方案,实现对数字档案信息的有效控制。

电子文件或数字化档案信息发布利用之前,必须进行开放利用的鉴定工作,并通过技术检查,譬如清晰度、准确性、完整性,以便让利用者有效检索、阅读和理解数字档案馆信息。档案信息网上提供利用,要根据数字档案不同网络的传播范围、用户范围、使用方式等进行处理。对涉及国家秘密、知识产权或个人隐私及其他敏感信息的档案利用,应当按照国家法律法规要求,进行划控处理。涉密信息只能在涉密网发布;内部信息只能在内网(包括政务网和档案馆局域网等)使用;开放信息可以在公众网发布。开放鉴定是数字档案馆建设的重要工作内容之一。

4.4 建立数字档案资源库

运用计算机及其相关技术设备管理数字档案信息,当前一般采取数据库技术方法进行。档案资源库一般包括目录数据库或元数据库、内容数据集等。

目录数据库是数字档案资源管理的基础,它是将反映数字档案特征的规范数据,依照一定的字段要求存入计算机中,通过系统的排序等处理,形成由计算机检索的目录数据体系。目录数据库建立的方式主要有两种基本途径,一是通过传统载体档案数字化采集的档案目录数据库,一般是通过人工录入建库方式建立;二是通过接收电子文件方式形成的数字档案,一般通过档案管理系统自动采集生成或从数字档案元数据库中提取而形成,经过数据整理规范审核与补充完善后建立。

保存数字档案元数据是保证数字档案可靠和可用的一项重要措施。元数据库建设按照数字档案元数据采集规范要求建立。元数据采集方式主要是通过通过对电子文件或数字档案的背景、结构和管理过程信息进行自动生成和适当人工添加而形成。

内容数据集是数字档案资源建设的主体,它是通过数据库、数据仓库等技术方法将档案全文按照一定的分类、排序方式排列形成的集合。内容数据一般通过与目录数据挂接方式实施有效管理,随着信息技术,特别是检索技术的发展,将来也会采用其他技术方法对内容数据进行有效管理。对于由电子文件归档形成的电子档案,其内容数据还应与其元数据建立持久有效的联系,防止非法修改,采取技术措施,确保其可靠和可用。

数字档案资源库建设应遵循以下标准规范:

- (1)GB/T20163-2006 中国档案机读目录格式;
- (2)DA/T13-1994 档号编制规则。

5. 保障体系建设

5.1 安全保障体系建设

安全保障体系建设是数字档案馆建设的基础工作,数字档案馆的安全包括数字档案数据的安全和信息系统及其网络平台的安全。数据安全就是要保证数字档案信息的可靠、可用、不泄密、不被非法更改等,系统及其网络平台安全就是要保持系统软硬件的稳定性、可靠性、可控性。

安全保障体系建设主要通过两方面途径实现。一是按照信息安全等级保护的要求,采用相应安全保障技术方法,配备必要的软硬件设施。数字档案馆系统一般要求达到二级(系统审计保护级)以上安全保护标准。数字档案馆系统集成商应具备相应的保密资质,并严格按照有关安全保密规范要求项目进行设计、系统开发和项目施工。建设、监理单位应当加强项目建设过程中的档案信息安全保密工作。二是建立健全数字档案馆安全管理制度,并严格遵照实施。数字档案馆系统安全隐患包括数据窃听、电磁泄漏、电力中断、载体损坏、自然灾害、非法访问、计算机病毒、黑客攻击、系统超负载、假冒身份、权限扩散、数据篡改、操作失误等,应当采取相应的技术措施和管理手段应对这些安全隐患。应当高度重视数字化加工、电子文件接收等过程中的安全保密管理工作。同时应当制订应急预案,完善灾难恢复机制,提高应急处置能力。

应当遵循的有关信息安全规范:

- (1)GB17859-1999 信息安全技术计算机信息系统安全保护等级划分准则;
- (2)GB/T24363-2009 信息安全技术信息安全应急响应计划规范;
- (3)GB/Z24364-2009 信息安全技术信息安全风险管理指南;
- (4)GA/T671-2006 终端计算机系统安全等级技术要求

5.2 标准规范体系建设

严格遵照信息化和档案管理等方面的法规和标准是实现数字档案馆各项功能的必要前提。数字档案馆建设要严格遵守国家及行业相关规章制度和技术标准。

各地在建设数字档案馆过程中,具体技术应用或工作流程如果尚无相关上位标准规范,可以向国家档案行政管理部门提出相关标准规范制修订建议,或根据本地区实际情况制订相关地方标准或企业标准,并积极上报有关部门,上升为行业标准或国家标准。