

自然资源部数字档案室建设总体设计与实践

文/李芳芳 吴玉龙 宋元 米捷 杨玲

随着我国各行业各领域信息化建设的蓬勃发展,数字档案室建设成了各机关档案管理部门的工作热点。近年来,自然资源部高度重视数字档案室建设工作,遵循“资源为先、标准规范、整体推进、确保安全”的原则,构建了自然资源部数字档案室管理信息系统(以下简称“数字档案室系统”),建立了自然资源部档案资源数据库,编制了一系列标准规范。在国家档案局档案馆(室)司的具体指导下,自然资源部数字档案室于2018年被国家档案局批准为“全国示范数字档案室”。本文以自然资源部数字档案室建设为例,介绍数字档案室建设的总体构架、实践经验和未来设想。

总体构架

自然资源部数字档案室采用“云框架”的理念进行建设,依托于已有的网络、计算、存储与安全保障等基础设施,基于数字档案室系统实现了档案管理和服务的统一部署与应用、统一分发与监控。自然资源部数字档案室总体构架包括设施层、资源层、平台层、应用层、接入层等5个层次以及标准规范体系、安全保障体系、实施保障体系等3个体系。

1. 五个层次

(1) 设施层

设施层是数字档案室系统建设的基

础支撑平台,包括机房、网络、储存和备份设备、系统软件等信息系统运行的物理场所和信息资源建设的工作环境。自然资源部数字档案室依托已有的基础设施,加强了身份认证、漏洞扫描、病毒防护、审计跟踪等安全策略,保障了数字档案室系统部署环境的安全稳定。

(2) 资源层

资源层主要实现了对文书、照片、音像等各类档案资源的统一管理。自然资源部数字档案室采用大数据的管理理念,构建了电子档案资源数据库,包括元(目录)数据库、电子文件数据库、多媒体数据库、管理过程库、标准规范库以及系统维护库等,实现了对自然资源部、原国土资源部、原国家测绘地理信息局、原国家海洋局的各类电子档案资源的规范化管理。其中,文书类电子档案包括原国土资源部1998—2018年共计8.7万件档案(全部实现OCR识别)、原国家测绘地理信息局1956—2018年共计4.8万件档案、原国家海洋局1958—2012年共计1.1万卷和1.5万件档案;专业类电子档案包括矿业权管理、建设用地等18类共计4.8万卷档案;照片类档案共计5000余张;音像类档案共计8000余项。电子档案数据总量达10TB。

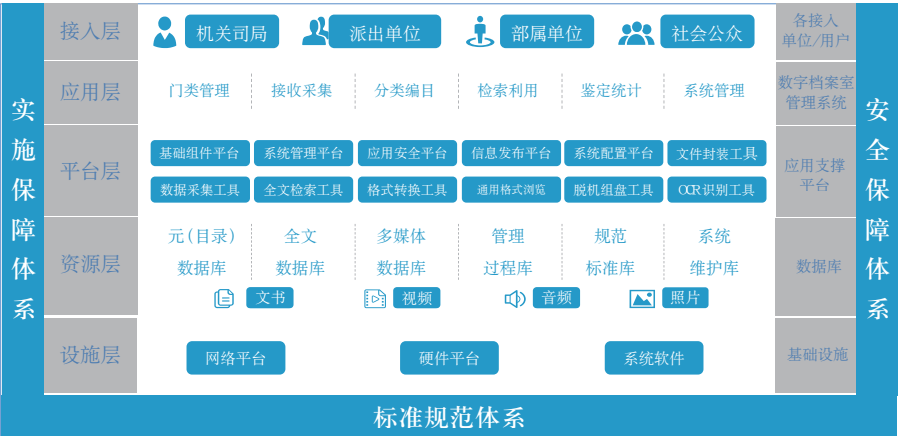
(3) 平台层

平台层包括数据采集工具、全文检索工具、格式转换工具、脱机组盘工具、

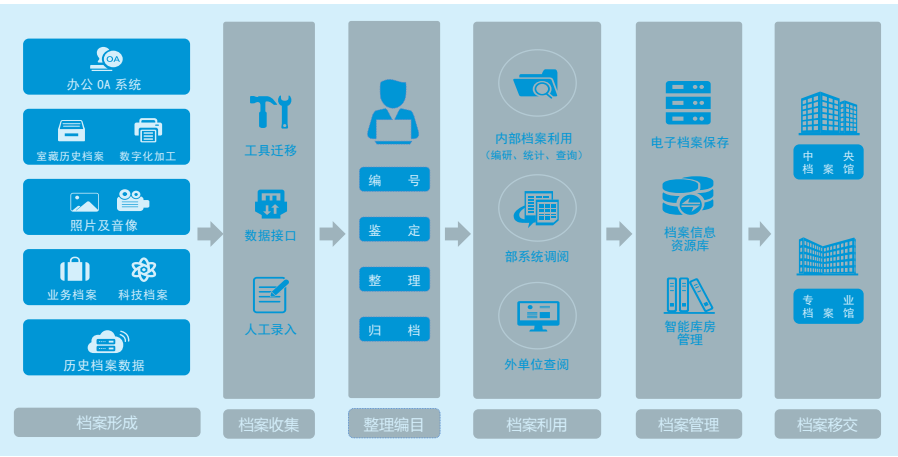
OCR识别工具、文件封装工具、通用格式浏览工具、系统管理平台、基础组件平台、应用安全平台、信息发布平台和系统配置平台等。平台层为应用层提供了支撑,架构在平台层上的应用层一方面可以缩短开发周期,降低系统建设风险,根据需求灵活配置功能;另一方面可以提升系统的性能。

(4) 应用层

应用层是指数字档案室系统的各功能模块,按照档案管理工作流程,设计开发包括门类管理、接收采集、分类编目、检索利用、鉴定统计、系统管理等功能模块,实现各类电子文件从接收、整理、登记、统计、保管、利用到鉴定处置整个流程的管理。数字档案室系统包括收集、管理、保存、利用4个重要环节。其中,收集环节主要实现各类档案资源的接收、整理和归档登记,包括电子文件的收集归档、纸质档案的数字化加工、音像档案的格式转换、业务系统档案的收集归档、历史档案的迁移等;管理环节主要实现对已接收登记的各类档案资源进行规范化、流程化管理,如鉴定、处置、移交等;保存环节主要通过电子文件管理、虚拟库房管理等功能,对电子档案和实体档案进行长期有效管理;利用环节主要实现对档案资源的信息发布和检索利用,通过目录共享、在线申请和浏览权限设置等多种方式,为用户提供



数字档案室系统总体架构图



数字档案室系统业务流程图

档案资源的在线利用服务。

(5)接入层

接入层是指数字档案室系统的终端用户,包括自然资源部各司局、各派出单位、各直属单位等。自然资源部机关各司局是数字档案室系统的主要用户,可实现对文件材料的收集、整理、归档和查阅利用;各派出单位、各直属单位也可通过数字档案室系统实现对本单位文件材料的整理、归档和利用。此外,社会公众可以办理相关申请手续,通过档案管理部门借阅端口实现对档案的查阅利用。

2.三大体系

(1)标准规范体系

标准规范体系既是数字档案室系统建设的指导方针,也是数据交换和共享的基础。通过对档案管理工作的现状进行分

析,结合业务发展和信息化目标,采用规范档案业务管理流程等标准化手段,规范各类档案资源从形成到收集、整理、保管、统计、利用、编研等业务流程,以最大程度实现档案管理的信息化、规范化和标准化。为保障电子文件归档和电子档案管理科学规范,自然资源部在数字档案室建设过程中编制了电子档案管理办法、档案分类方案及档号编制规则、电子文件归档管理指南、电子文件元数据规范、电子文件归档接口规范、数据备份管理办法、应急响应预案等一系列标准规范。

(2)安全保障体系

完善的安全保障体系是数字档案室系统建设和顺利运行的保障,安全保障体系包括网络安全、系统安全、数据库安全、信息安全、设备安全、信息介质安全

和计算机病毒防治等各个方面,可通过制度规范、人员管理、技术控制、权限设置等多种方式确保数字档案室系统与数据资源的安全可控。

(3)实施保障体系

实施保障体系是指在数字档案室系统建设的过程中,从人才、资金、项目管理和制度规范等方面保证项目的顺利实施。

技术创新

1.电子文件自动归档及接收登记

在全面梳理归档流程的基础上,通过在办公自动化系统中增加预归档功能,以及办公自动化系统与数字档案室系统的对接,实现归档流程的规范化。通过系统完成从档案整理人员进行电子文件收集、整理到档案审核人员对归档目录和文件进行审核,再到档案管理人员进行最终检查和接收登记的归档流程,并通过系统实现了归档各环节电子文件和元数据的规范性检查与自动化接收登记等功能。

2.元数据自动捕获

数字档案室系统实现了电子文件元数据的自动捕获,在元数据库建设中,把元数据分为通用元数据和专用元数据。其中,通用元数据包括立档单位名称、档号、年度、题名、日期、责任者、文号、保管期限、密级、件号等。而专用元数据则各不相同,如数字化副本包括扫描分辨率、色彩模式、扫描设备型号等专用元数据;照片类档案包括人物、地点、摄影者、摄影时间等专用元数据;音像类档案包括摄像者、摄像时间、视频比特率、音频比特率等专用元数据;专业类档案由于专业类别不同,元数据项也各有不同,如建设用地档案包括面积、省份、坐标、项目编号、项目名称等专用元数据。上述元数据的捕获,有的是通过办公自动化系统或业务审批系统继承过来的,有的是数字化过程中产生的,有的需要在数据整理和格式转换过程中进行自动捕获,有的需要在发生电子档案接收、登记、迁移、处置等业务行为的同时自动生成。

3. 四性检测功能

四性检测是指在档案管理的不同环节配置不同的检测方案以检测数据的真实性、完整性、可用性、安全性。例如,在归档环节,真实性检测要检测文件来源和内容的真实性,检测题名、日期、密级、责任者、保管期限等元数据是否准确等;完整性检测要检测归档总量、元数据完整性、电子文件内容完整性等;可用性检测要检测电子文件内容、元数据、软硬件环境等是否可用等;安全性检测要检测病毒、归档过程是否可控、归档载体是否安全等。通过一系列检测条件的设置,保障档案数据的真实、完整、可用、安全。

4. 系统间数据传输接口开发

自然资源部各类业务信息系统与档案管理系统对接方式有2种,分别为中间表方式和XML方式。其中,中间表方式是业务信息系统将已经办结的、待归档的条目及电子文件信息写入中间表中,档案管理系统定时或者实时从中间表中取得数据;XML方式是业务信息系统将已经办结的、待归档的条目及电子文件表单信息拼装成既定XML文件,档案管理系统定时或者实时从指定位置取得XML数据进行相关处理。

实践经验

自然资源部在开展数字档案室建设期间,不断攻坚克难、敢于突破创新,结合自然资源档案数据体量大、种类多的特点,打造“档案云”服务体系,在实现存量档案规范收集、整理和利用的同时,积极探索和推进各类电子文件在线归档。目前,实现了文书类电子文件在线收集、整理、检查和归档全流程电子化;对于采矿权登记审批、建设用地预审等网上业务审批系统,开发了相应的电子归档接口,实现了从资料申报、业务审批、公告

数字档案室系统元数据方案设置和四性检测界面图

发布等各个环节电子文件的自动收集和归档。此外,还探索了网页类电子文件这种新档案类型的分类方案、归档范围、保管期限以及利用方式,并通过网页电子文件收集管理系统和数字档案室系统的对接实现了网页类电子文件定期在线归档。总而言之,自然资源部通过数字档案室建设提升了档案管理工作水平,对机关行政职能的发挥起到了支撑作用。在推进数字档案室建设过程中,工作重点和难点主要有以下3个方面:

1. 档案知识与信息技术的融合

档案管理工作是一项专业性很强的工作,各类档案的整理、编目、装订、修复都需要专业的档案知识作支撑。与此同时,电子档案有其特殊性和先进性,除了要遵循传统载体档案收集、整理、利用的通用要求外,还要符合电子文件元数据、系统间数据交换、自动检测及数据安全等方面的要求。这就要求档案工作者既要掌握档案管理的知识和技能,又要具备利用信息技术实现海量档案数据规范收集、管理的技术和能力。

2. 业务工作流程的梳理和实现

日常档案管理工作中的接收、编目、整理工作主要是针对传统载体档案而言的,实现电子文件在线归档需要在了解传统载体档案工作流程的基础上,结合信息技术的特性,通过系统在线实现和优化归档流程,如电子文件整理、电子档案登记确认等。运用信息技术可以提高归档人员和档案管理工作者的工作效

率,同时还可以提高档案管理工作的质量和规范化水平。

3. 加强在线归档前端监督指导

按照档案管理相关要求,档案整理工作是由档案形成人员和部门来完成的,文书、照片、音像等各类电子档案都要由其形成人员进行收集、整理和归档。同样,实现电子文件在线归档首先需要办公自动化系

统及相关业务系统增加档案预归档功能,实现电子文件的收集、整理和保管期限确定。这对档案管理工作者的前端监督指导能力提出了更高的要求,需要档案工作者与信息化建设人员深入研究每一类档案的特点,指导办公自动化系统及相关业务系统开发预归档功能,以确保各类档案数据的齐全、完整、规范。

未来设想

随着数字档案室系统的应用与各类档案数据的增加,自然资源部数字档案室将从以下3个方面进一步完善:一是智能安全高效的海量档案数据管理。需要针对各类档案设置不同的长期安全管理和备份策略,并针对文书、照片、音像等各种形式的电子文件实现自动备份、自动检测、应急恢复等功能。二是智慧档案管理体系的建设与升级。应用大数据、语义分析等技术,实现档案管理的智能化,例如采用更加科学的方式实现档案保管期限确定智能化、档案整理和归档智能化、档案鉴定处置智能化等。三是通过海量信息提取与分析技术,实现档案结构化与非结构化数据编研分析及产品定制,为自然资源部机关和其他用户提供精准的、高质量的档案信息服务,真正体现档案数据的价值和作用。

作者单位:自然资源部信息中心

责任编辑:李俊豪