

Challenges and optimization paths in the construction of natural resource archive information resources

Hua Li

Natural Resources and Planning Bureau, Changyuan, Henan, 453400, China

Abstract

The construction of natural resource archive information resources is an important way to achieve scientific management and sustainable utilization of natural resources. With the rapid development of information technology, the construction of natural resource archive information resources is facing new opportunities and challenges. Digitization and information technology provide strong support for effectively preserving, managing, and sharing archival information. And due to the complexity of archival data, the decentralization of management models, and the increasing demand for information, higher requirements have been put forward for the construction of archival information resources. Therefore, in-depth research on the challenges and optimization paths faced by the construction of natural resource archive information resources is of great practical significance for improving China's natural resource management level and promoting the rational utilization of natural resources.

Keywords

natural resource archives; Information resource construction; Challenge and Optimization

自然资源档案信息资源建设的挑战与优化路径

李华

河南省长垣市自然资源和规划局, 中国·河南长垣 453400

摘要

自然资源档案信息资源建设是实现自然资源科学管理、可持续利用的重要途径。随着信息技术的迅猛发展,自然资源档案信息资源建设面临着新的机遇和挑战。数字化和信息化技术为有效地保存、管理、共享档案信息提供了强有力的支撑。且由于档案数据的复杂化、管理模式的分散和信息需求的日益增长,对档案信息资源的建设提出了更高的要求。因此,深入研究自然资源档案信息资源建设面临的挑战和优化路径,对提高我国自然资源管理水平,促进自然资源合理利用,具有十分重要的现实意义。

关键词

自然资源档案; 信息资源建设; 挑战与优化

1 引言

当前我国自然资源档案信息资源建设正面临来自各方面的挑战。首先,档案资料分散、不易整合,各部门间档案管理标准不统一,难以实现信息共享;其次,档案资料的真实性与可信性还有待提高,档案资料在录入过程中存在着不完整、人为因素等因素,都会影响档案资料的质量。同时,由于缺乏健全的管理体制与技术支撑,制约着档案信息资源建设进程。所以,只有探究如何从强化数据集成能力、提高数字化建设水平、建立健全管理机制等方面入手,才能有效地解决上述问题,促进自然资源档案信息资源建设向高质量发展。

【作者简介】李华(1980-),女,中国河南长垣人,本科,馆员,从事档案管理研究。

2 自然资源档案信息资源建设的挑战

2.1 档案资料分散难以整合

自然资源包括土地、矿产、森林、水利等不同类型的自然资源,其档案数据分散在国土、林业、水利等多个管理部门和业务领域。各部门分别对相关文件进行独立管理,缺乏一个统一的协调机制,可造成档案信息的有效整合和共享。由于各部门的工作重点、管理方法不同,在格式、标准、保存方法等方面存在差异,这就给集成工作带来困难^[1]。这种分散式的档案管理模式,既制约着档案信息的流通和使用,又造成信息资源的浪费与重复建设。现实生活中,当需要对自然资源进行综合管理与决策的时候,往往很难获得全面、精确的档案信息,从而影响到管理效率与决策的科学性。因此,如何有效地解决我国自然资源档案信息资源建设中存在的问题,成为制约我国自然资源档案信息资源建设的一个重要因素。

2.2 档案资料真实性与可信性

自然资源档案是自然资源的重要载体，其信息的准确可靠性直接影响着自然资源管理工作的科学性与有效性。然而，在现实生活中，档案数据的真实性和可信度却面临着许多挑战。由于自然资源具有很强的动态性，其开发利用状况及自然环境的改变，可能会造成档案数据脱离实际。如土地利用规划调整、矿产资源开发进度改变等，若档案资料不能及时更新，将造成信息滞后，影响档案的真实可信。且在收集整理档案资料时，也会受到人为因素的干扰。如资料输入失误、资料不全或遗失，都将影响档案资料的质量。另外，不同来源的档案数据之间的真实性与可信度也不一致，缺少统一的审核验证机制，导致档案数据在使用过程中的准确性与可靠性难以判定。这不仅影响到自然资源档案信息资源的利用价值，也影响着自然资源的科学管理与可持续利用。

2.3 管理体制与技术支撑不足

从管理体制上看，自然资源档案工作涉及部门众多、层级众多，缺乏统一的协调与管理机制；各部门之间的职责划分不清，沟通和配合不力，使档案信息资源建设工作很难形成合力。同时，我国现行的档案管理体制已不能适应信息时代的发展需要，所以由于缺乏规范化的电子档案管理制度和安全保障制度，导致在管理层面上存在着许多漏洞^[2]。在技术支持上，为有效地存储、管理和共享档案信息，必须要有大数据、云计算、地理信息系统等先进的信息技术支撑。但是，目前很多地区、部门的信息化建设还不够完善，信息化水平参差不齐。部分地区缺少必要的技术设备及专业人才，不能适应档案数字化和信息化的需要；而其他地区虽有一定技术基础，但缺乏对技术更新与应用的持续投入，致使档案信息资源建设缺乏技术支持。由于管理制度、技术支持等方面的缺失，严重影响自然资源管理工作的进程与质量，从而影响自然资源管理工作的整体效能。

3 自然资源档案信息资源建设优化路径

3.1 强化数据集成能力

对各类自然资源档案的数据源进行梳理，制定统一的数据格式和编码规范。利用 ETL(Extract, Transform, Load) 工具对不同的数据库和文件系统中的数据进行抽取，并根据规范对其进行转化和清洗，剔除重复和错误的数据。通过接口对接和共享等手段，实现多源数据的有效集成^[3]。在此基础上，建立数据仓库，存储和管理整合后的数据，便于后续的数据挖掘和分析，为自然资源管理部门的决策提供全面和精确的数据支撑。

例如，自然资源厅为对全省土地、矿产、林业等多类档案资料进行集成，在数据整合之前，土地档案资料是分散在各地土地管理部门自行建立的资料库中，格式多样，有传统的 DBF 格式，也有新兴的 SQLite 格式；矿档数据采用 Oracle 数据库，采用行业专用标准编码，采用独立的省矿管

局数据库；森林档案资料由各林业部门保存，主要采用文件系统格式，包括 JPEG（森林资源野外照片）和 CSV（记录树种、数量等）。为实现数据的整合，可组织一支专家队伍，历时 3 个月对全省的数据源进行梳理，建立一个统一的数据格式，将所有的结构化数据都存储在 MySQL 数据库中，而非结构化的数据则统一存储在 Ceph 中。编码格式在国家统一标准的基础上，结合本省的自然资源特点，做了一些微调。使用 ETL 工具 Kettle 对 1000 多条数据提取规则进行配置。从自然资源局数据库中抽取前一天的数据，每天凌晨两点进行更新，每次抽取的数据量在 500MB 左右，涉及土地出让、变更登记等各种信息。对于矿产数据，每周末都会进行全量采集，数据总量约为 2GB，内容包括矿产储量、采矿许可证等。森林数据提取包括森林资源动态监测数据，每季度约 1.5GB。在变换清洗阶段，编写复杂的 SQL 脚本，去除 15% 的重复数据，纠正 5% 左右的错误坐标数据。同时，构建基于微服务框架的数据整合平台，实现各地市及相关部门之间的 API 接口，每月实现 10TB 左右的数据共享。采用星型模型的方法，建立数据仓库，其中事实表存储土地交易量、矿产产量等关键数据，而维度表则包含时间、地理和资源类型等信息。通过数据整合，自然资源厅编制土地利用规划的时间由原来的 3 个月压缩到 1.5 个月，预测精度提高 20%，为合理配置自然资源奠定坚实的基础。

3.2 提高数字化建设水平

制定一份详细的数字化档案计划，按照重要性和使用频度将档案分类，分批进行数字化。选用专业的扫描机，根据档案的材料和纸张大小进行调整，保证扫描后的影像清晰完整^[4]。通过光学识别（OCR）将扫描文本转换成可编辑的电子文件，提高检索效率。利用元数据管理系统，将档案名、形成时间、责任者等描述性信息加入到数字化档案中，完善档案目录系统，方便使用者迅速找到所需要的档案资料。

例如，某大型林业科研院所所有大量的森林资源科研档案，其中包括珍稀树种的研究报告和森林生态监测记录，总计约 5 万余册。为推动数字化建设，应制定详细的计划，并根据档案在科研项目中的重要程度和查阅频率，将其分成第三批。第一批重点数字化近 10 年来濒危树种 10 万份档案，每月平均被检索 200 次，是当前生物多样性保护研究的热点。以富士施乐的 iGen5 数码打印机为扫描设备，针对纸质文件大多是 A4、A3 等纸张，有些年份较长的纸张容易损坏的情况，可设定 300dpi 的扫描分辨率和 RGB 色彩模式，以确保扫描图像清晰，同时不会因为过高的分辨率而对设备和纸张造成过大的负荷。装订成册文件采用平版扫描，散页文件用自动送纸器批量扫描，平均每小时 500 张左右。同时，可利用 ABBYY FineReader OCR 软件对扫描结果进行识别。建立涵盖 5000 余条专业术语的林业专业词汇库，提高识别的准确性。通用文本和林学专业术语的识别正确率分别达到 95% 和 90%。识别完成后，将电子文件导入到自研的元数

据管理系统中,为每个文档增加详细的说明信息。如珙桐的研究报告,增加“珙桐种群生态特性与保护对策研究报告”,撰写时间为“2018年5月”,责任者为“研究小组负责人及课题组成员”等。通过建立一套完善的目录系统,可使科研人员查找文件的平均时间由原来的15分钟缩短到3分钟,极大地提升科研工作的效率。

3.3 建立健全管理机制

建立自然资源档案信息资源管理的组织结构,明确各部门的责任主体,明确各部门的责任^[5]。建立档案管理的全生命周期管理体系,包括文件的生成、存档、储存、转移、销毁等各个环节的标准化。建立质量监控机制,定期对档案管理工作进行检查和评价,对不达标的环节进行整改。健全档案使用制度,明确不同使用者的存取权限,规范借阅查询程序,确保档案的安全性,提高档案的使用效率。

例如,某县自然资源局负责全县土地矿产资源的档案管理工作,在传统档案管理中,可偶然出现遗失和不规范的情况。为建立完善的管理机制,可重组组织机构,成立“档案收集部”、“整理组”、“保管组”和“用组”。档案收集科负责收集各业务单位的档案,按月5日前完成上月档案的收件工作,其收件准确率要在98%以上。整理组将收集到的文件进行分类编目,要求收档后10个工作日内完成整理工作,正确率不超过2%。保管组负责档案的保管工作,保证库房温度 $20\pm 2^{\circ}\text{C}$ 、湿度 $50\pm 5\%$,并配备24小时对档案进行实时监测。建立档案管理的全生命周期管理体系,从建立档案开始,确定各个环节的时间节点和操作规范。土地出让合同文件应在签订后15个工作日内完成存档,永久保存。每5年一次档案转移评估,以确保储存媒体的安全性。建立质量监管机制,成立由局领导和业务骨干组成的督导组,每季检查一次。如在一次季度检查中,发现档案编号有误,约1000份,要求整理组限期一个月内进行整改。完善档案利用体系,把读者分成三类,即内部工作者、科研工作者和公众。所有非机密文件均可供内部人员阅览,研究人员需提前三个工作日提出申请,经审核同意后方可借阅,借阅期限以15天为限,可大大提高档案管理工作的效率,为自然资源管理工作的顺利开展提供强有力的保障。

3.4 加强技术创新与应用

应着眼于大数据、人工智能、区块链等前沿技术,探索其在自然资源档案管理中的应用场景。引入大数据分析技术,深入挖掘海量文档数据,挖掘数据之间的关联性和规律,为资源管理决策提供支持。利用人工智能中的图像识别、语义分析等技术,实现档案内容的自动分类和标注,提升档案管理的智能化水平。为保证档案数据的真实性、完整性和不

可篡改性,可采用区块链技术来保证档案数据的安全性。

例如,省自然资源研究承担着全省自然资源监测和科研的重任,在面对海量的监测数据时,积极进行技术创新和应用研究。以森林资源监测数据为例,通过卫星遥感和无人机等方式,每月生成5TB以上的高分辨率图像、树高、郁闭度等信息。在Hadoop集群中引入大数据分析技术,构建50个节点,配置8核CPU,64GB内存,同时采用MapReduce框架,对5年内森林资源数据进行深度分析。如结果表明,在某一山区,每年七月至九月林火发生率与月平均温度、降雨量和植被覆盖度有很强的相关性。年平均温度在 30°C 以上,降水在50毫米以下,植被覆盖率在70%以上的地区,林火发生的概率比其它月份高30%。在此基础上,可采用基于深度学习的林火识别框架TensorFlow模型。采集10余万幅森林图像数据,包括2万幅林火图像,通过对视频中火灾区域、火蔓延方向等信息进行标注,训练模型精度达到92%以上。该模型能够对无人机拍摄的图像进行实时分析,从而实现火灾隐患的自动识别。还可引入区块链技术,构建联盟链,参与节点为研究院、省林业局及相关研究机构。实现了森林资源监测的原始数据和分析结果的在线存储,确保数据不被篡改。通过该技术的创新应用,全省森林火灾预警时间提前2小时,灭火效率提高25%,有效保障全省森林资源的安全。

4 结语

综上所述,加强自然资源档案信息资源建设,对提高自然资源管理水平,促进资源合理利用,具有重要的现实意义。在此背景下,对档案信息资源进行优化,可以有效地提高档案信息资源的质量与使用效率。未来,随着信息技术的不断进步和管理理念的创新,自然资源档案信息资源的建设必将向智能化、系统化的方向发展,为自然资源的科学管理与可持续利用提供了可靠的保证。

参考文献

- [1] 林剑.大数据驱动的自然资源数字档案建设路径与实践探索[J].办公自动化,2024,29(24):47-49.
- [2] 李治霖.基于SWOT分析的自然资源档案信息化管理策略[J].兰台世界,2024,(12):111-113.
- [3] 李海玲.自然资源档案及其数字化建设策略研究[J].办公室业务,2024,(22):25-27.
- [4] 李佩瑜.数字化背景下的自然资源档案管理路径分析[J].办公室业务,2024,(22):79-81.
- [5] 张敏.自然资源档案信息资源建设存在的问题及解决路径探究[J].兰台内外,2024,(24):51-53.