

大数据时代文书档案管理及信息化应用

黄金伟

黑龙江省齐齐哈尔市碾子山区跃进街道社会治安综合治理中心

摘要：文书档案管理方式面临许多挑战和限制，如数据隐私与安全、数据质量与准确性、数据存储与备份、数据分析与智能化等问题。为了克服这些挑战，提高文书档案管理的效率和质量，越来越多的机构开始采用信息化技术来管理和应用文书档案。本文通过详细阐述大数据时代文书档案管理的优势和挑战，希望可以为相关研究和实践提供一定的借鉴和参考。

关键词：大数据时代；文书档案管理；信息化；应用

【DOI】10.12252/j.issn.2096-6261.2023.11.096

随着大数据时代的到来，信息化技术的快速发展和广泛应用，对传统的文书档案管理方式提出了新的要求和挑战。在传统的文书档案管理中，纸质档案的存储、整理和检索工作耗费了大量人力和时间，效率较低。同时，纸质档案的保管和共享存在安全隐患和资源浪费的问题。因此，采用信息化技术来管理和应用文书档案，成为提高管理效率和质量的重要途径。本文将重点探讨大数据时代下文书档案管理的优势和挑战，并对信息化应用在文书档案管理中的角色和发展方向进行分析和讨论。

一、大数据时代文书档案管理的优势

首先，大数据时代的文书档案管理可以提高工作效率。传统的文书档案管理需要人工处理大量的纸质档案，工作效率较低。而在大数据时代，文书档案可以通过数字化方式进行管理，实现电子化文档的存储、检索和共享。这样，工作人员可以快速获取所需信息，减少了传统手工检索和整理的时间，大大提高了工作效率。其次，大数据时代的文书档案管理可以实现智能化决策。借助大数据分析技术，可以对海量文书档案数据进行处理和分析，提取出有价值的信息和规律。通过数据挖掘、机器学习等技术，可以发现隐藏在文书档案中的关联性和趋势，为决策提供科学依据。这样，可以帮助管理人员做出更准确、更精确的决策，提高管理水平和工作质量。

二、大数据时代文书档案管理的挑战

1. 数据隐私与安全问题

首先，数据隐私保护是最重要的问题之一。文书档案包含大量用户个人信息，如社保号码、账户信息等。保护用户敏感数据的隐私性是文书档案管理的首要任务。需要建立严格的权限控制机制，确保只有授权人员才能访问相关数据。此外，应采用数据加密和身份验证等技术手段，保护数据在传输和存储过程中的安全。其次，数据安全保护也是一个重要的挑战。黑客攻击、入侵和病毒等威胁对文书档案系统构成了巨大风险。为了应对这些威胁，需要建立完善的安全控制系统，并及时

更新和修复系统的漏洞。定期进行安全测试和审计，持续改进安全性^[1]。

2. 数据质量与准确性问题

在大数据时代的文书档案管理中，数据质量与准确性问题是另一个重要挑战。大量文书档案数据的输入和录入过程中存在人为错误、信息不完整和重复等问题。这些问题对数据质量和准确性造成了严重影响，给后续的数据分析和应用带来了困难。书档案数据中常常存在大量的冗余、错误和不一致的信息。建议采用合适的清洗和整理技术，对文书档案数据进行去重、纠错和规范化处理。通过数据清洗和整理，提高数据的质量和准确性。

3. 数据存储与备份问题

首先，数据存储需求的增长是一个挑战。随着数据量的不断增加，传统的存储设备已经无法满足大数据时代的需求。建议采用分布式存储系统和云存储技术，提高存储容量和性能，确保文书档案数据的安全和可靠。其次，数据备份和恢复是确保数据安全的关键环节。传统的备份方式可能无法满足数据量大、增长快的需求。建议采用增量备份和差异备份等技术手段，减少备份时间和存储空间。同时，定期进行数据备份测试和恢复演练，确保备份的完整性和恢复的可靠性。

三、大数据时代文书档案管理信息化技术的应用

1. 大数据技术在文书档案管理中的应用

首先，大数据技术在文书档案收集与归档方面的应用。传统的文书档案管理通常依赖于纸质档案和手工记录，工作效率低下且易于遗漏和损坏。而借助大数据技术，可以使用OCR（光学字符识别）技术对纸质档案进行扫描和数字化，将文本内容转化为可搜索和管理的电子档案。此外，通过大数据技术，可以自动进行文档分类和归档，便于检索和管理，提高工作效率。其次，大数据技术在文书档案保护与安全方面也发挥重要作用。文书档案中可能包含大量的敏感信息和个人隐私，需要进行有效的保护和安全管理。大数据技术可以通过数据

加密和权限控制，确保文档的安全访问和使用。同时，利用大数据技术的数据监控和审计，可以实时监测文档

的使用情况和权限变更，发现异常行为和安全风险，并及时采取相应的安全措施^[2]。

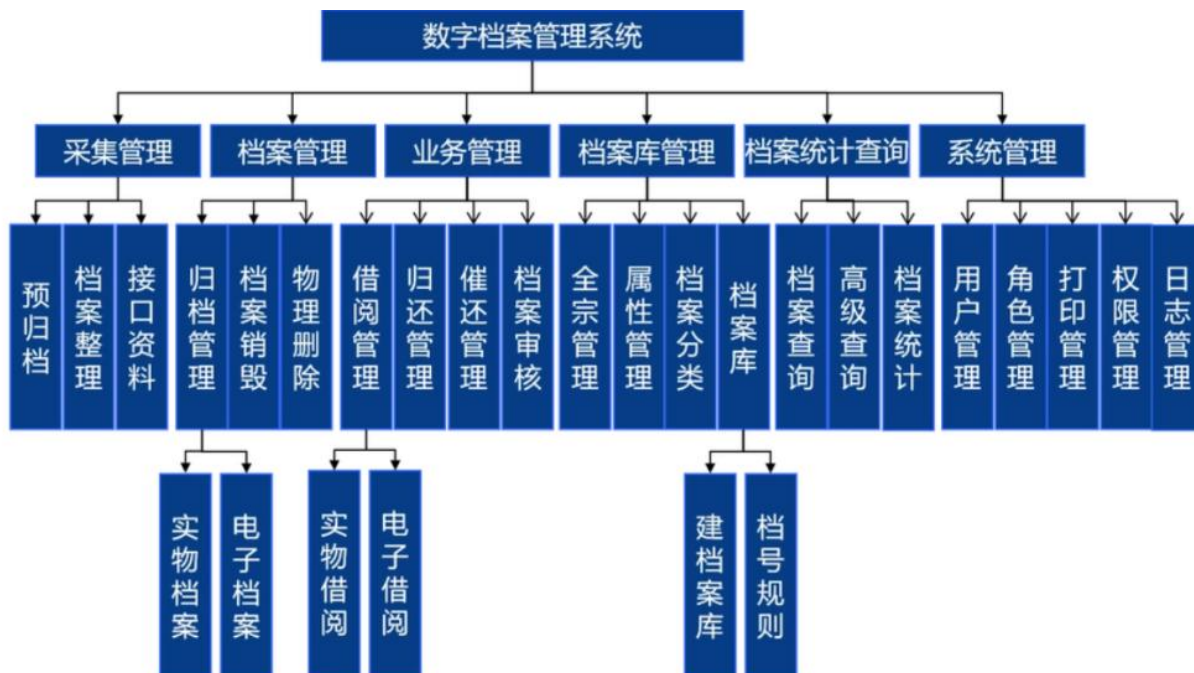


图1 数字化档案管理系统

2. 人工智能技术在文书档案管理中的应用

首先，人工智能技术在文书档案检索与查询方面起到关键作用。传统的文书档案检索通常依赖于关键词的匹配，往往存在检索结果不准确、信息匹配度不高的问题。而人工智能技术可以通过机器学习和深度学习算法，实现文书档案的语义理解和智能检索。例如，通过构建文档内容的语义表示模型，可以实现基于语义的文书档案检索，提供更准确、个性化的检索结果，提高用户的满意度和检索效果。其次，人工智能技术在文书档案利用与价值挖掘中的应用也十分重要。传统的文书档案利用主要依赖于人工筛选和阅读，效率较低且主观性强。而人工智能技术可以通过自动分析和挖掘文书档案中的大量数据，发现隐藏的关联关系和知识模式，提供决策支持和智能化分析。例如，通过自然语言处理和机器学习技术，可以实现对文书档案数据的语义分析和情感识别，对文书档案中的信息进行智能化挖掘和分析，为用户提供更全面、准确的决策支持^[3]。

3. 区块链技术在文书档案管理中的应用

首先，区块链技术在文书档案管理中的去中心化特性。区块链技术通过去中心化的分布式存储和共识机制，实现了文书档案管理过程中的数据可信和防篡改的特性。每个节点都有完整的文书档案记录，所有的文书档案修改和更新都将以分布式共识的方式进行验证，并通过加密算法保障数据的安全性。这种去中心化的特性使得文书档案管理更加透明和可信。其次，区块链技术

在文书档案管理中的可溯源性和不可篡改性。区块链基于密码学原理，每个区块都包含了前一个区块的哈希值，当有新的文书档案被添加时，将生成一个新的区块，并链接到之前的区块。这种链式结构使得整个文书档案记录具备了不可篡改的特性，即一旦记录被写入区块链，就无法被篡改或删除。在文书档案管理中，这将有助于确保数据的完整性和准确性。第三，区块链技术在文书档案管理中的智能合约应用。智能合约是一种自动执行合约条款的计算机程序，它可以自动验证和执行文书档案管理中的一系列操作，如审批、访问控制、权限管理等。在区块链技术的支持下，智能合约可以实现透明、可靠的文书档案管理流程，消除了中间环节的信任问题，提高了文书档案管理的效率和准确性。

4. 云计算技术在文书档案管理中的应用

首先，云计算技术提供了文书档案管理的高效、可靠和灵活的存储解决方案。传统的文书档案存储通常依赖于本地服务器或硬盘等物理设备，存在容量限制和数据冗余的问题。而云计算技术通过将文书档案存储在云端服务器上，可以实现无限扩展的存储容量，并保证数据的备份和冗余。同时，云计算平台提供了高可用性和可靠性的服务，确保文书档案的长期保存和安全性。其次，云计算技术为文书档案的共享和协同提供了便利。传统的文书档案共享往往需要通过物理传递或复制文件的方式，效率低下且易丢失。而借助云计算技术，文书档案可以通过共享文件夹或云存储空间进行实时共享和

表1 基于区块链技术的文书档案管理

应用领域	应用场景	应用优势
权限管理与控制	基于区块链技术建立文书档案的权限控制机制，确保只有授权人员才能访问和管理文档。	1. 安全性：通过区块链的去中心化特性，文档的访问权限由网络中的各个节点进行验证和控制，提高了文档的安全性和防篡改能力。 2. 自动化：基于智能合约的权限管理可实现自动化的文档访问控制和身份验证，减少人工管理的成本和风险。
文档溯源与防伪	通过区块链技术记录文书档案的产生和传输过程，实现文档的溯源和防伪功能。	1. 可信度：区块链技术提供了不可篡改的数据记录和溯源功能，确保文档的真实性和可信度。 2. 防伪能力：通过区块链的分布式特性，可以防止文档的伪造和篡改，对于重要的文档具有重要保障作用。
共享与合作	利用区块链技术建立文书档案的共享和合作平台，实现多方之间的信息共享和协同工作。	1. 去中心化：区块链的去中心化特性使得多方可以直接共享和交换文档，无需依赖第三方的中介。 2. 透明性：区块链上记录了共享和合作过程中的各项操作，提高了合作的透明度，减少合作纠纷的发生。

协同编辑，不受地域限制。同时，云计算平台提供了权限管理和版本控制，确保文书档案的安全和一致性^[4]。

四、大数据时代文书档案管理的建议

1. 建立良好的数据质量控制机制

首先，建立数据采集和录入的规范化流程。文书档案管理的数据质量始于数据的采集和录入过程。建议建立明确的数据采集标准和规范，确保数据的准确性、完整性和一致性。例如，制定统一的数据录入格式，明确要求采集员在录入数据时进行必要的验证和核对。同时，应提供培训和指导，确保采集人员对数据采集流程和标准有清晰的理解和掌握。其次，建立数据存储和管理的安全机制。数据存储和管理是保障数据质量的关键环节。建议采用安全的存储设备和系统，采用灾备和备份机制，确保数据的可靠性和持久性。同时，采用访问控制机制和权限管理，限制数据的访问和修改权限，以保护数据的安全性和完整性。此外，建议定期进行数据清理和整理，清除无效和过时的数据，保持数据存储的清爽和精简。

2. 增强数据安全与隐私保护能力

首先，建议制定和执行严格的数据安全政策和控制措施。文书档案管理机构应建立详细的数据安全政策，明确管理用户权限、数据访问控制、数据存储和传输的规则和要求。将敏感数据和非敏感数据进行分离存储，确保只有授权人员才能访问敏感数据。同时，加强对工作人员的安全教育和培训，提高他们对数据安全的意识和防范能力。其次，建议采用最新的数据加密和身份验证技术。通过对文书档案中的数据进行加密，可以在数据传输和存储过程中提高数据的保密性和完整性。采用强密码策略和多因素身份验证措施，确保只有授权人员才能访问和操作文书档案数据。此外，可以使用数字签名和数据摘要等技术来验证数据的真实性和一致性，防

止数据被篡改或伪造。第三，建议建立严格的访问控制机制和审计制度。通过设置不同级别的访问权限，将文书档案访问和操作权限授予合适的人员。仅有必要的人员才能够访问和操作特定的文书档案数据，以降低数据泄漏和滥用的风险。同时，建立审计系统和日志记录机制，对文书档案管理人员的操作进行监控和记录，便于追溯和查找异常行为，确保数据安全性和后续跟踪。

五、结语

在大数据时代，信息化应用对于提高文书档案管理的效率和质量起到了至关重要的作用。通过数字化记录和管理文书档案，可以减少人工的工作量，提高整理和检索的效率。然而，在信息化应用中也面临着一些问题和挑战，如数据隐私与安全、数据质量与准确性、数据存储与备份、数据分析与智能化等方面需要进一步加强和完善。因此，为了更好地应对大数据时代的文书档案管理需求，需要持续关注和研究信息化技术的最新发展，并结合实际情况进行合理应用和创新。

参考文献

- [1] 赵婵. 大数据时代文书档案信息化管理与应用研究[J]. 兰台内外, 2021, (30): 25-26.
- [2] 孙海霞. 大数据时代文书档案管理及信息化应用途径研究[J]. 信息记录材料, 2021, 22(05): 190-191.
- [3] 潘晓梅. 大数据时代文书档案管理及信息化应用途径研究[J]. 办公室业务, 2020, (10): 81-82.
- [4] 张丽梅. 大数据时代文书档案管理及信息化应用分析[J]. 办公室业务, 2020, (04): 89+95.
- [5] 彭海燕. 大数据时代文书档案管理及信息化应用分析[C]// 中国智慧工程研究会智能学习与创新研究工作委员会. 2020年社会发展论坛(西安)论文集. 山东高速路桥股份有限公司, 2020: 6.