智慧图书馆人工智能系统的安全风险防控与优化策略

任永前

上海大学 上海 200444

【摘　要】：智慧图书馆作为信息化社会发展的重要产物，正依托人工智能、大数据、云计算和物联网技术加速建设，其服务模式、资源管理方式和用户体验均发生深刻变革。人工智能系统在智慧图书馆的引入，使自动编目、智能推荐、读者行为分析、人机交互服务等功能大幅提升，但与此同时，智能系统在数据采集、模型运行、算法决策和平台交互中存在的安全风险也逐渐凸显，如数据泄露、算法偏差、系统攻击、虚假身份操控以及用户隐私滥用等问题。这些风险不仅威胁资源系统和用户数据安全，也严重影响图书馆的公共服务职能与社会信任度。本文首先分析智慧图书馆人工智能系统的运行特点，系统梳理在数据管理、系统算法、平台安全与使用行为等方面的主要风险；其次探讨风险形成的技术与管理机制；在此基础上提出数据加密与权限管理、算法审查与优化、平台安全加固、多主体协同治理及人员安全意识提升等综合策略，以构建更加安全、可靠、可持续的智慧图书馆人工智能系统。研究旨在为智慧图书馆的安全体系建设、人工智能系统应用优化以及智能服务的规范发展提供理论基础和实践路径。

【关键词】：智慧图书馆；人工智能；系统安全；数据保护；风险防控

DOI:10.12417/2982-3846.25.03.013

引言

随着人工智能技术的飞速发展，图书馆数字化、网络化和智能化建设进入新阶段。智慧图书馆依托人工智能实现资源自动处理、智能检索、用户行为分析与精准服务等功能，使图书馆的管理效率与服务能力显著提升。然而，大规模数据采集、深度用户建模、算法驱动服务与开放网络环境带来了前所未有的安全威胁。智慧图书馆作为信息服务机构，其数据具有公共性、长期性与敏感性，一旦遭受网络攻击、数据泄露或算法滥用，将严重损害用户隐私与系统运行安全。目前，国内外学者已对智慧图书馆的建设框架与人工智能应用进行研究，但针对安全风险的系统性研究仍然相对不足，在机制分析、风险识别与防控策略方面缺乏体系化研究。为此，本文围绕智能系统在智慧图书馆中的应用特点，深入分析风险来源、运行机理及外部环境影响，并提出多维度防控策略，以期为智慧图书馆人工智能安全体系构建提供学术与实践参考。

1 智慧图书馆人工智能系统的运行特征与发展机遇

1.1 人工智能驱动图书馆服务智能化水平全面提升

人工智能技术在智慧图书馆中的应用主要体现在资源加工、用户服务与管理调度三个方面。其智能编目、自动分类与知识图谱构建能力使海量文献资源的组织效率显著提高，减少了人工操作的复杂度；智能检索与语义分析支持对读者的深度查询需求，实现智能回答、个性化推送、跨资源关联检索等功能；智能机器人与虚拟助手提升了用户交互体验，使图书馆服务更加便捷、友好。这些技术的融合不仅提升管理效率，也推动图书馆由资源中心向知识服务中心转变，为智慧化建设提供重要支撑。

1.2 大数据支撑下的用户行为分析增强服务精准化能力

智慧图书馆在服务过程中会采集并处理大量用户数据，包括借阅记录、检索行为、阅读偏好、访问路径与空间活动轨迹等。人工智能模型经过长期训练可以建立用户行为画像，实现精准推荐、需求预测与资源配置优化。在此过程中，大数据分析增强了图书馆对用户需求的理解能力，使图书馆能够更加科学地制定服务策略，如优化书架排布、提升热门资源供给能力等，从而实现服务供给的需求导向与精准化。

1.3 系统智能化驱动图书馆管理模式发生结构性变革

人工智能技术深度参与图书馆管理流程，实现从传统管理到智能调控的跃升。智能监测系统能够对图书借阅、馆内环境、设备运行状态等进行实时监控，使运维更精细、更高效。智能调度系统可优化人力资源配置，如通过数据分析调整工作岗位、安排值班模式或分配资源处理任务。此外，基于智能技术的知识管理平台将图书馆的资源管理与学术服务进行深度融合，提高了学术研究与知识服务的整体效率。

2 智慧图书馆人工智能系统的主要安全风险分析

2.1 数据泄露风险在多场景、多节点中不断累积

智慧图书馆人工智能系统依赖用户数据进行学习和服务，这些数据在采集、存储、分析和传输中均存在被窃取或滥用的可能。由于智慧图书馆用户涵盖教师、学生、研究人员及社会公众，其身份信息、阅读偏好、检索内容等均具有敏感性。一旦数据在传输中被截获、在存储中遭遇攻击或因权限管理缺失被内部人员泄露，将导致用户隐私受损。此外，图书馆共享的跨系统平台（如校园网、政务平台）也可能成为安全隐患的延

伸点，使数据泄露风险呈现链式扩散特征。

2.2 系统遭受网络攻击的风险显著提高

智慧图书馆的网络连接程度高，各类智能设备、应用程序与服务平台的接入点多，容易成为黑客攻击对象。攻击方式包括 DDoS 攻击、恶意软件植入、数据库入侵、API 接口攻击等。人工智能系统由于模型结构复杂，亦可能遭受攻击者通过逆向分析模型参数、投毒训练数据、干扰模型输出等方式进行破坏。更为严重的是，一旦系统核心服务被攻击，将造成服务中断、数据损坏以及用户无法访问资源等后果，严重影响图书馆正常运行。

2.3 人工智能算法存在偏差、误判等潜在风险

人工智能模型的运行依赖数据，若训练数据存在偏差、不完整或受污染，可能导致系统在推荐、检索或用户画像方面作出错误判断，产生算法歧视、标签化推荐等问题。这些偏差不仅影响用户体验，也可能对图书馆公共服务公正性造成挑战。此外，深度学习模型的决策过程具有“黑箱化”特点，难以解释其判断依据，使风险难以被及早识别和纠正。算法问题的累积会对智慧图书馆的服务质量造成潜在影响。

3 智慧图书馆人工智能系统安全风险的形成机制

3.1 数据结构复杂化增强了安全防护难度

智慧图书馆的数据不仅数量庞大，而且类型复杂，包括结构化数据、非结构化文献、文本检索记录、行为轨迹等。这些数据在不同平台间频繁流动，使安全边界难以固定。数据接口越多，系统暴露面越大，从而增加潜在攻击点。同时，跨系统的数据交互往往依赖第三方平台或外部接口，而接口标准不统一、风险等级划分不明确也为安全风险埋下隐患。

3.2 智能系统对高频数据依赖形成潜在风险链条

人工智能系统需要大量数据驱动，用户的持续使用为其提供训练基础。这种依赖性使得一旦数据集遭受污染（如恶意投毒、批量输入错误数据），模型将持续产生错误输出。此外，模型运行需要消耗大量资源，并依赖后台服务器和云环境，一旦外部环境发生故障（如网络中断、服务器宕机），将形成综合性安全风险链条。

3.3 安全管理制度与技术更新不同步造成管理漏洞

人工智能技术发展速度快，而图书馆的管理制度往往更新较慢。许多图书馆在制度中未明确界定数据采集边界、算法责任、敏感信息处理方式等内容，出现制度滞后与技术超前的矛盾。同时，图书馆工作人员对人工智能技术的理解有限，缺乏足够的安全意识，使得操作规范与制度执行力不足，导致安全管理漏洞。

4 智慧图书馆人工智能系统安全风险的技术防控策略

4.1 构建多层次数据安全保护体系

为避免数据泄露，应采用全链路加密技术，包括数据采集端加密、传输端 SSL/TLS 加密、存储端的密钥管理与访问控制。同时，加强数据库防护，如采用数据脱敏、分级存储、智能审计等方式管理敏感数据。此外，通过建立数据备份与容灾系统，可提高对突发事件的恢复能力。

4.2 提升人工智能算法的透明度与可控性

应建立算法审查机制，对模型的训练数据、参数调整及输出结果进行定期检测，避免算法偏差持续累积。采用可解释人工智能（XAI）技术，提高算法的透明度，使系统在推荐与检索过程中能提供合理依据。同时，可通过引入对抗训练方法增强模型的鲁棒性，防止其在遭受攻击时发生异常。

4.3 加强智慧图书馆网络平台的安全建设

需对智慧图书馆系统进行全面的网络安全防护，包括部署防火墙、入侵检测系统、安全网关等，并定期进行漏洞扫描与补丁更新。对于接口和 API，需要加入认证机制与调用限制，并监控异常接口访问行为。此外，可引入零信任安全模型，将系统访问从“认证即信任”转变为“持续验证”，增强平台整体安全性。

5 智慧图书馆人工智能安全防控的管理优化策略

5.1 健全安全管理制度建立标准化运营体系

图书馆在引入人工智能系统后，需要通过制度化建设来确保数据与系统运行的安全性，使各项管理活动具备可遵循的标准。制度内容应覆盖数据管理、访问权限、审核流程与应急处置等关键环节，对人工智能系统涉及的数据采集范围作出明确界定，说明合法合规的要求，使数据使用在源头层面得到控制。算法服务在运行中可能带来偏差、滥用或隐性风险，制度应设定相应的风险评估标准，对模型更新、接口调用与算法解释等内容形成规范，使算法在服务图书馆业务时具备可评估、可控制的特性。人工智能系统的安全治理还需配套日常巡查机制，通过对系统日志、异常访问与处理流程的定期检查，及时发现潜在隐患。问题反馈机制的建立能够让工作人员在发现系统异常或使用障碍时迅速报告，使处理流程形成从发现到整改的闭环管理模式。制度体系在明确性与执行力的共同作用下，能够推动人工智能系统的安全管理真正落地，使智慧图书馆在技术创新与安全保护之间保持良好平衡，为知识服务环境提供稳定可靠的支撑。

5.2 加强人员培训与安全意识建设

人工智能系统在智慧图书馆中的广泛应用，使人员管理成为安全治理的重要组成部分。系统的稳定运行不仅依赖技术本身的先进性，也依赖工作人员在日常管理中的专业素养与规范操作。图书馆可通过定期组织信息安全、隐私保护与应急处置技能等培训，使工作人员熟悉数字资源管理的关键要求，掌握人工智能系统的操作流程与使用规范，在面对潜在风险时形成正确的判断与处置能力。培训过程能够促使员工理解数据泄露、权限滥用与系统误操作的危害，从而在实际工作中更加谨慎审慎。技术人员队伍建设同样具有关键意义，引入数据科学、网络安全等领域专业人才，有助于增强系统维护、故障排查与安全监测能力，使人工智能系统在复杂环境下保持稳定运行。专业技术团队的加入还可以推动图书馆在算法优化、数据治理与风险模型构建方面持续提升，使安全防护从被动应对转向主动预防。

5.3 推进多方协作构建综合安全治理体系

智慧图书馆的安全治理需要在技术、管理与法律层面实现多主体协作，使安全保护形成系统化合力。图书馆可与高校信息中心建立协同防护机制，通过共享安全监测数据、联合开展系统巡检与风险评估，提升整体防护能力。政府部门在数据治理方面具有制度与监管优势，与图书馆在数据合规、访问权限管理、隐私保护规则等方面开展合作，能够使智慧图书馆在处理用户信息与数字资源时更具合法性与规范性。技术企业在安全技术研发与系统构建方面拥有专业能力，通过合作可为图书馆提供更先进的安全架构，例如智能风控系统、攻击检测系统或加密传输方案，使智慧图书馆具备更强的技术安全韧性。用户群体也是安全治理的重要参与者，通过宣传教育、交互式活动或在线提示等方式提升其安全意识，让用户在账号管理、资源使用与信息保护方面养成良好习惯，使安全治理从单向管理转向共同维护。多主体协同模式能够使智慧图书馆在复杂的数字环境中形成稳定可靠的安全生态，为知识服务与文化传承提供坚实保障。

6 结论

智慧图书馆人工智能系统的应用推动了图书馆向智能化、精准化方向转型，但其在数据保护、算法治理、系统安全与管理制度方面仍存在多维度风险。通过对风险特征、形成机制及防控策略的研究可以看出，智慧图书馆的安全管理必须以技术手段为基础，以制度体系为保障，以人员能力为支撑，并通过多方协作构建全面、动态、可持续的安全防护体系。未来，应继续推动人工智能算法的透明化与可解释性研究，加强数据安全立法建设，促进智慧图书馆的安全治理能力持续提升，从而实现技术创新与安全管理的协同发展。

参考文献：

[1] 张春春,孙瑞英.智慧图书馆用户数据合规治理机制研究[J].图书情报工作,2024,68(4):15-26.

[2] 陆康,刘慧,张婧,任贝贝.数智时代智慧图书馆隐私治理的社会理论重构[J].国家图书馆学刊,2024,33(3):84-94.

[3] 吴玉灵,闫东芳.人工智能时代下的图书馆数字安全的风险与应对策略研究[J].图书馆理论与实践,2024(6):123-129.

[4] 王静,王鹏.智慧图书馆生成式 AI 大模型风险治理机制研究[J].情报杂志,2024,43(8):190-197.

[5] 武文秀,徐士贺,马明迪.元宇宙情境下图书馆虚拟数智人:功能、风险与应对策略[J].图书情报工作,2025,69(24):107-116.

[6] 王萍.人工智能环境下智慧图书馆信息安全问题探析[J].图书与情报,2021.

[7] 李宏伟.智慧图书馆建设中的数据安全风险研究[J].现代情报,2020.

[8] 张倩.大数据背景下数字图书馆隐私保护策略研究[J].情报探索,2019.

[9] 陈杰.图书馆人工智能应用的伦理风险与治理研究[J].图书馆论坛,2021.

[10] 刘芳.智慧图书馆平台网络安全管理研究[J].现代图书情报技术,2018.