云计算时代图书馆信息资源的安全防线构建

——评《图书馆信息资源安全:基于云计算环境下》

随着云计算技术在图书馆领域的深度应用,信息资源存储与服务体系正经历结构性变革。云计算依托分布式计算架构,通过 IaaS(基础设施即服务)、PaaS(平台即服务)、SaaS(软件即服务)3 层服务模式,为图书馆提供弹性化的计算资源与存储空间。然而,这种依赖第三方云服务商的架构也引入了新型风险,如互联网基础安全缺陷可能直接威胁云端数据完整性;服务商运营稳定性与安全策略的不足可能导致图书馆核心数据泄露或服务中断;同时,图书馆自身对云端数据的管控能力弱化,进一步放大系统性风险。构建图书馆信息资源安全防线的核心价值在于三重维度的保障:①资源资产维度。云中存储的珍稀数字文献、用户借阅轨迹及版权数据库需防范未授权访问与恶意篡改,避免文化遗产与知识产权受损。②隐私权利维度。读者身份信息、检索行为等敏感数据的泄露可能引发法律追责与信任危机,强化加密与权限管理是履行隐私保护义务的必然要求。③服务延续性维度。抵御 DDoS 攻击、勒索软件第三方云服务商的信誉与稳定性等云端威胁,直接关联图书馆服务的可持续供给,关乎公共知识传播职能的履行,尤其在数智化进程中,跨机构数据共享趋势使安全防线成为突破信息孤岛、释放数据价值的前提条件。

书名:图书馆信息资源安全:基于云计算环境下
作者:黄国彬
出版社:知识产权出版社
ISBN:9787513077095
出版时间:2021 年 10 月
定价:88 元

云计算环境下的图书馆信息资源安全防线,指通过技术与管理双重机制构建的立体防护体系。其核心任务涵盖:① 数据生命周期保护,包括传输加密(如 SSL/TLS 协议)、存储加密(AES/RSA 算法)及敏感信息脱敏处理;② 访问控制体系,基于最小权限原则实施身份认证与操作授权;③ 威胁响应机制,通过入侵检测系统与安全审计实现动态风险监控。笔者在开展“2024 年度河南农业大学本科教育教学改革研究与实践重点项目(2024XJGLX053)研究过程中,认真阅读《图书馆信息资源安全:基于云计算环境下》一书,该书以云计算环境下图书馆信息资源安全为核心研究对象,通过系统性的理论建构与实证分析,构建了完整的学术框架。全书共 10 章,开篇以前言形式明确研究背景、目标、理论基础、方法论及核心概念界定,为后续研究奠定坚实基础。第 2 章文献综述部分采用国内外对比视角,既梳理了国内研究脉络,又剖析了国外前沿动态,最终提炼出国内外研究的共性与差异特征。第 3—第 6 章聚焦实践层面,依次解析了图书馆可利用的云服务产品类型(通用型与专用型)及其技术特征,系统考察了国内外主要图书馆应用云计算的总体概况与典型案例,并深入剖析了通用云服务与专用云服务在图书馆场景中的应用概貌、管理主体及服务对象特征。第 7—第 9 章转向安全维度,首先,界定云计算环境下图书馆信息资源的业务类型与内容体系,进而识别安全需求与风险类型;然后,从个人数据安全、云服务协议不对等、政策法律不足等维度剖析风险成因;最后,基于数据管理生命周期、个人数据安全、虚拟化计算资源安全 3 个维度展开政策法律层面的深度剖析。第 10 章提出具有操作性的应对策略,包括构建适应云计算环境的政策法律体系、签订有利于维护图书馆信息资源安全的云服务协议、制定内部安全管理制度,以及加强国际学术交流与合作。全书贯穿“现状-问题-对策”的逻辑主线,既注重理论深度又强调实践价值,形成了“概念界定-现状分析-风险识别-策略构建”的完整研究闭环,为云计算时代图书馆信息资源安全保障提供了系统的理论指导与实践路径,具有显著的学术创新性与现实指导意义。

作者指出,零信任架构作为云计算环境下图书馆信息资源安全的核心范式,其本质是通过“永不信任,持续验证”原则重构访问控制逻辑。该策略需从身份认证、权限管控、动态评估 3 个维度展开系统性设计。在身份认证层面,应采用多因素认证与生物特征识别技术融合方案,如指纹+动态令牌的双因子认证机制,结合区块链

技术实现身份凭证的不可篡改存储。北京大学图书馆的“总-分馆体系”多源身份认证系统已验证,通过微服务架构实现跨系统身份联邦认证,可有效解决 9 大类 43 种用户身份的统一管理难题。权限管控需遵循最小特权原则,通过属性基加密(Attribute-Based Encryption, ABE)实现细粒度访问控制,如在数字资源访问场景中,可采用基于角色的访问控制与基于属性的访问控制混合模型,并结合零信任网关实现网络隐身,确保只有授权设备才能访问特定应用端口。动态信任评估有必要构建多源数据融合的信任度量模型,全面整合用户行为分析、设备指纹、网络流量特征等数据,且利用机器学习算法实现信任评分的实时计算。在具体操作中可采用随机森林算法对用户登录时间、访问频率、操作路径等特征进行建模,动态调整访问权限。亦可定期通过红蓝对抗演练进行压力测试,确保在遭遇结构化查询语言(Structured Query Language, SQL)注入、跨站脚本等攻击时,系统能自动触发权限降级或会话终止机制,确保系统安全。

笔者认为,在云计算环境中,图书馆需建立覆盖数据采集、存储、传输、使用及销毁的全链条防护机制。第 1 阶段,数据采集阶段。采用差分隐私技术对用户行为轨迹脱敏,如通过拉普拉斯机制向借阅记录注入噪声,确保单条数据无法关联至特定用户身份。进一步利用 k-匿名化技术,将用户群体划分为至少 k 个不可区分的子集,有效防御关联攻击。清华大学图书馆在用户画像系统中实践该方法,显著降低了隐私泄露风险。第 2 阶段,数据存储阶段。实施分层加密体系。对身份证号等敏感身份数据采用国密 SM4 算法全盘加密;行为轨迹数据则应用同态加密技术,支持密文状态下的统计分析,实现“数据可用不可见”。存储架构需融合分布式存储与区块链技术,如通过星际文件系统(Inter Planetary File System,IPFS)实现去中心化存储,并依托区块链的不可篡改性保障数据溯源可靠性。国家图书馆在古籍数字化平台中应用该方案,强化了数字资源的版权保护与防篡改能力。第 3 阶段,数据使用与传输阶段。强制启用 TLS 1.3 协议实现端到端加密,并引入量子加密技术以应对未来量子计算攻击。跨机构数据共享场景中,采用安全多方计算技术实现联合建模,确保协作方无法获取原始数据,如公共图书馆与出版商合作训练推荐算法时,MPC 协议可在保护用户借阅记录隐私的前提下完成模型优化。第 4 阶段,数据销毁阶段。遵循 NIST SP 800-88 标准,结合物理销毁(如硬盘消磁)与逻辑销毁(如密钥销毁)双重机制,确保数据不可恢复。

为保护图书馆信息资源,需推动灾备体系建设的多层次架构建设。灾备系统需遵循“3-2-1”黄金法则:保留 3 份数据副本,存储于 2 类不同介质,且至少 1 份异地保存。业务影响分析阶段,依据 DRI 十大最佳实践量化恢复优先级。如古籍数字化业务要求恢复时间目标(Recovery Time Objective,RTO)≤4 h、恢复点目标(Recovery Point Objective,RPO)≤15 min;普通借阅服务可放宽至 RTO≤24 h、RPO≤1 h。除此之外,可采用“同城双活+异地灾备”混合云架构,如同城数据中心通过同步复制技术实现故障无缝切换;异地中心采用异步复制抵御区域性灾难。中国国家图书馆的“两地三中心”方案中,生产数据实时同步至同城灾备中心,异步复制至异地中心,确保单点故障下业务持续运行。在渐进式灾备演练机制下,分 3 阶段推进:①桌面推演,通过沙盘模拟验证应急预案完整性;②模拟演练,基于仿真系统测试灾备组件可用性;③实战演练,模拟真实攻击场景(如勒索病毒),检验备份数据恢复能力。为了业务连续性组织保障,亦可设立专职灾备指挥小组,明确业务与 IT 部门职责分工,定期开展恢复培训,如哈佛大学图书馆构建“热站(核心业务 4 h 恢复)+温站(12 h)+冷站(24 h)”三级体系,确保灾难分级响应效能。

总之,云计算环境下图书馆安全防线的演进,标志着信息资源保障体系从被动防护向主动防御的战略转型。随着零信任架构与机密计算技术的深度融合,未来防护体系将依托硬件级可信执行环境实现数据全程密态处理,彻底规避传输与存储环节的中间层攻击。在治理层面,需构建跨机构协同的安全治理框架,通过 ISO 27001 与 GDPR 等国际标准的本土化实践,建立覆盖云服务商、图书馆、监管机构的三方责任矩阵。尤为关键的是,安全防护需与数字人文发展深度耦合——古籍文本区块链存证、读者行为联邦学习分析等场景,既要求防护技术适应新型知识生产模式,更需通过《个人信息保护法》《数据安全法》的合规性设计平衡数据利用与隐私保护。安全防线的终极价值不仅在于风险消解,更在于为数字人文创新提供可信基石。当技术防线、制度规范与人文诉求形成有机整体,图书馆方能在保障文化遗产传承的同时,履行数字时代的知识枢纽使命。

(张华/河南农业大学图书馆/副研究馆员)