做好出版社的图书数据管理路径分析

李琛

中国科技出版传媒股份有限公司, 北京 100717

DOI: 10.61369/TACS.2025010039

摘要 : 为推动出版社在大数据时代下的转型, 探讨出版社图书数据管理的可行路径。通过介绍数据管理工作, 明确图书数据管理的重要性。随之阐述图书数据管理的技术流程, 并提出了定时备份图书数据、按规定申请使用图书数据、保护图书数据安全三点建议, 总结当前图书数据管理价值实现与今后的发展趋势, 希望能为出版社开展图书数据管理工作提供技术方面的参考。

关键词 : 出版社; 图书资源; 数据管理

Analyzing the Path of Book Data Management for Publishing House

Li Chen

China Science and Technology Publishing Media Co., Ltd., Beijing 100717

Abstract : To promote the transformation of publishing houses in the era of big data and explore feasible paths for book data management in publishing houses. By introducing data management work, clarify the importance of book data management. Subsequently, the technical process of book data management is elaborated, and three suggestions are proposed: scheduled backup of book data, application for use of book data according to regulations, and protection of book data security. The current value realization and future development trends of book data management are summarized, hoping to provide technical references for publishing houses to carry out book data management work.

Keywords : publishing house; book resources; data management

数字出版技术在业内飞速发展, 出版社图书数据管理成为当下一项重要的工作。随着该项工作的创新, 除了传统纸质档案储存, 还需要利用现代化技术向智能化方向转型, 探索全新的数据管理模式。出版社的图书数据包含诸多内容, 例如元数据、用户行为信息和图书内容资源, 对其进行管理可采用的管理路径不仅要强调管理技术流程的可行性, 还需要提出满足出版社要求的管理策略。有鉴于此, 本文重点探讨出版社图书数据管理的可行路径, 通过多元化数据分析与处理提升图书资源使用率。

一、出版社图书数据管理

图书数据管理, 即采用现代化技术或指定方法, 以图书信息为对象实施储存、检索、维护, 具体涉及以下流程: (1) 数据采集; (2) 数据录入; (3) 数据分类; (4) 数据存储; (5) 数据检索; (6) 数据更新; (7) 数据维护; (8) 数据分析。

出版社开展数据管理, 一方面是为了提高所有图书资源使用率, 另一方面则是要向广大用户提供需要的图书内容^[1]。站在数据管理员角度, 对出版社图书数据进行运维、安全管理, 能提高图书数据质量, 发挥现代化技术优势, 节约数据维护成本。

二、图书数据管理技术流程

(一) 数据采集与整合

图书数据管理中的数据采集与整合, 应用多源数据聚合技

术, 出版社可以将所有图书数据集成处理, 如纸质书的数字化数据、电子书产出数据和用户形成的交互数据^[2]。采用标准化机制清洗或者转换数据格式等处理方法, 构建数据池。另外, 此环节应用智能采集工具, 例如 OCR (Optical Character Recognition, 光学字符识别) 技术, 可对历史纸质类书籍中的内容进行自动识别, 人工智能工具则负责抓取新媒体平台中的用户反馈信息, 搭建多维数据集。

(二) 数据接收与检查

1. 数据接收

出版社图书排版、印制文件等工作由印制管理人员负责, 包括数据的采集、管理等, 所有形成的图书数据利用数据中心进行管理, 而中心的管理者和出版社的印制管理者对接, 按照出版社拟定的《出版资源清单》, 在约定的周期内接收到图书排版、印制文件。

2. 数据检查

图书数据管理环节采用的排版软件源文件主要有: Adobe

作者简介: 李琛 (1984.12—), 女, 汉, 河北秦皇岛人, 出版中级, 硕士, 研究方向: 计算机科学与技术。

Indesign 产生 INDD 文件、方正书版产生 FBD 文件、Photoshop 产生 PSD 和 TIF 文件等。对于所有文件的排版，必须在每一版最后印刷时形成正式的排版文件，再以正式出版要求装订文件，完成封面、扉页等标识。文件如果有独立的插页、插图和卧表等，需要根据其在图书中的页码命名^[9]。例如表格 1 位于图书的第 7 页，即为“P7BG1”。对于补字情况，必须有相应的字库文件，以及插入发排图片文件，此类文件格式以 JPG 和 TIF 等为主，并需检查文件字体、图片分辨率等。

（三）数据存储与架构设计

对于图书数据的储存，建议采取分布式云存储形式，利用分布式数据库、云计算等技术，提高图书数据存取的效率，并实现弹性扩展，有效规避图书数据储存环节的单点故障。采取混合存储策略，图书元信息为代表的结构化数据可以在关系型数据库中储存，而文本和音频等非结构化内容，则采取对象储存的方式，能同时满足性能、成本要求。

（四）数据分析与智能应用

图书数据的分析是获取用户需求的重要手段，采用用户行为洞察方式，以大数据技术分析所有用户的阅读偏好、购买图书的趋势，分析结果为出版社选题策划、制定营销推广策略提供参考^[4]。应用人工智能算法实现个性化的书籍推荐，面向广大读者，为其推送定制化书单，并能实时监测当下市场的热点，根据监测所得信息对数字资源发布策略做出调整。

（五）生成图书目录

出版社图书的元数据是“提供关于信息资源或数据的一种结构化的数据，是对信息资源的结构化的描述”，图书数据的使用者、出版社数据系统均可利用元数据管理工具检索、理解、应用数据。出版社构建图书数据管理系统，以 Office 办公软件作为辅助，可自动生成图书目录，将目录导入到系统，为图书数据管理提供依据。

三、出版社的图书数据管理策略

（一）定时备份图书数据

根据出版社要求将图书数据分类，设立“总文件夹”，再以排版、图片和字体等文件，依次设立子文件夹，所有文件存放对应文件夹，设置定时数据备份，例如每日、每周进行一次数据备份，保证图书数据完整^[5]。出版社的所有图书数据可设置移动硬盘备份、独立服务器备份，以免数据丢失造成的严重后果。

移动硬盘为外部存储器，比较方便携带，支持随时随地备份数据，也是日常办公比较常见的数据备份工具^[6]。相比于其他的数据储存设备，移动硬盘容量大，且读写速度快。若移动硬盘中数据损坏或遗失，由专业技术人员恢复数据，为图书数据安全性提供保障。独立服务器本质上属于物理服务器，为用户个人使用。对比共享服务器、虚拟主机，出版社使用独立服务器备份数据，能保证硬件资源完整性，而且也获得控制操作系统的权限。独立服务器无需和其他用户共享图书数据，所以数据更安全^[7]。

综上，出版社备份图书数据采用移动硬盘和独立服务器，前

者能满足随身携带的需求，后者能满足数据安全性需求，具体可结合图书数据备份要求选择。

（二）按规定申请使用图书数据

出版社中的工作人员如果需要使用图书数据，需按规定流程提出申请，即在办公自动化系统中进行相关操作^[8]。申请者提交了申请之后，数据中心的数据管理人员、数据所属部门管理者等在系统后台依次操作，审核数据使用申请，完成图书数据交接。提交的数据使用申请需清晰标注图书数据的名称、数量以及规范文件格式等内容，申请者、图书数据管理人员等签字后，完成申请操作。提交电子版申请以及数据后续使用情况均需储存，以便后续核对与检查。

（三）保护图书数据安全

保护图书数据安全，是出版社图书数据管理的重要策略。对于已经储存的图书数据，必须定期检查、更新，出版社中运行的信息系统、安全网关等也需定期检查维护，技术管理人员查看图书数据及管理系统的访问权限、操作权限^[9]。系统用户访问行为、数据日志信息等，由技术人员负责检查与存档，定期安排管理系统运维、安全防护软件升级、高危漏洞修复，保证系统始终在安全的运行状态，规避图书数据的安全风险。关于出版社图书数据安全问题，主要分为内部业务安全问题和外部业务安全问题两类，分别提出如下建议。

1. 图书数据内部业务安全问题的解决方法

保证图书数据内部业务安全，需要建设数据中心内部局域网、后台权限管理系统、内容修改记录日志。对于数据中心内部局域网，要求局域网访问权限与使用权限只对出版社内部人员开放，工作中形成的图书数据第一时间上传到局域网中，以备工作人员检索和调取，实现图书数据在出版社内的资源共享。构建后台权限管理系统，所有管理人员都要设置用户名和密码登录，并按照每个人职责权限的区别设置登录权限，开通权限前直接在系统中提交图书数据的使用申请，统一由管理员处理，通过了审核之后为其开放权限。系统中储存所有申请记录，如果未通过审核，也要注明原因。

因为出版社的图书数据量比较大，需要频繁更新数据，图书数据管理后台应具备内容修改记录日志这一项功能，包括新增图书的内容、图书内容进行的删减和修改、修改前后的内容对比、修改时间等。管理人员可以直接在后台中调取记录，如果在图书数据管理中遇到问题可调取数据进行对比。

2. 保障数据外部业务安全的问题

出版社图书数据的外部业务安全保护，其一应保护内、外部网络安全，其二需保护图书数据传输安全。

第一，保障内、外部网络安全。保护内部网络安全应重点防范计算机病毒，考虑到计算机病毒有比较强的复制能力和感染性，尤其是在网络中快速传播，导致系统遭受重大损失。所以出版社需安装性能更高的防火墙与网关，保证系统内部安全。针对外部网络的安全防护，则需制定应急响应方案，例如安装预警系统，由该系统负责实时监测网络，如果发现异常行为或问题，可以及时报告给技术管理员^[10]。该系统还能对受控访问区、安全

内网进行监测，检查区域范围内是否有可疑的攻击行为，并对系统访问、连接的安全性及合规性作出判断，防止黑客恶意入侵网络，或者出版社图书数据泄露等现象。

第二，图书数据在传输过程中的安全防护，建议随时查杀出版社工作人员即时通信工具、外部设备等。例如日常工作中使用的聊天软件和电子邮件，储存备份图书数据使用的移动硬盘等。以上工具能随时向外界传输图书数据，是安全风险的重点管理对象。出版社工作人员应尽可能的不使用即时通信工具传输关键图书数据，而是应用有安全保障的外部设备，例如移动硬盘和 U 盘，但要注意此类设备在传输图书数据前，必须全面检查安全性，若发现携带病毒必须马上查杀，保证无安全风险方可传输图书数据。检查过程中若移动硬盘、U 盘提示驱动故障，或者扇区存在问题，工作人员可以使用防护软件予以修复，保护存储数据的安全性、完整性。利用电子邮件传输数据，必须在传输的同时下载备份，工作人员还要形成定期修改邮箱密码、清理无效邮件的习惯，以免储存在邮箱中的数据遭到不法分子盗用。

四、出版社图书数据管理价值实现与发展趋势

基于当前出版社图书数据管理取得的成果，其价值实现主要集中在三个方面：第一，建设数字出版平台，例如在线数据库能

实现纸质资源和数字出版的同步，支持按需印刷、电子书订阅等诸多线上服务，数据库进行数据分析后，增强了学术用户需求定位的精度。第二，拓宽智能服务范围，例如构建个性化学习平台，将所有出版资源整合，并在图书数据的驱动作用下提供交互式体验，提高用户群体的黏性。第三，活化图书资源，利用数据化技术转化出版社所有图书数据资源，为其赋予可检索、可分析的性质，为学术研究提供支持。

现阶段图书数据在数字化时代下呈爆炸式增长，出版社不仅面临图书数据管理方面的压力，还需要解决数据安全保护、跨平台整合等问题。今后建议出版社在图书数据管理工作的实践中，加深对人工智能的应用，例如应用人工智能进行内容创作、评价图书数据管理质量；应用区块链技术全程追溯版权源头；加大跨出版社图书数据共享力度，在图书出版行业真正实现资源协同。

五、结束语

综上所述，出版社图书数据管理是大数据时代下的一项重要工作，利用大数据、云计算技术，可以提高图书数据管理效能，简化数据管理流程。加强对图书数据安全管理的重视，并采集广大读者用户的阅读偏好信息，能为出版社选题策划和营销提供参考，从而切实提高图书数据管理水平。

参考文献

- [1] 薛文静, 王彤. 基于数据分析的馆藏图书质量提升研究——以 L 大学旅游管理专业为例 [J]. 甘肃科技, 2024, 40(10): 89-99.
- [2] 桑昀. 数据分析视域下的图书出版高质量发展 [J]. 出版广角, 2024, (19): 54-59.
- [3] 张小航. 大数据技术在信息管理领域的应用价值与前景分析 [J]. 中国新通信, 2024, 26(19): 78-80.
- [4] 张可. 数字化时代图书档案管理与信息安全研究 [J]. 新传奇, 2024, (18): 126-128.
- [5] 陈江标. 基于大数据的智能图书管理系统的研究与设计 [J]. 电子元器件与信息技术, 2024, 8(04): 116-119.
- [6] 田蕊. 基于大数据技术的图书数字信息资源管理系统设计和实现 [J]. 信息记录材料, 2024, 25(04): 134-136+139.
- [7] 郭彩萍. 大数据时代公共图书馆图书管理的创新与守旧 [J]. 中国民族博览, 2024, (05): 250-252.
- [8] 付静. 大数据助力公共图书馆图书资料管理推陈出新 [J]. 文化产业, 2023, (28): 50-52.
- [9] 李娜. 大数据视域下公共图书馆图书资料管理研究 [J]. 华章, 2023, (09): 174-176.
- [10] 陆爱群, 金水萍. 大数据技术助力图书资料档案管理的信息化建设 [J]. 文化产业, 2023, (21): 148-150.