



■ 本期关注:数据资产入表

出版业数据资产宝藏亟须打开

□本报记者 章红雨

从数据到数据资源,再到数据资产,这些距出版业并不遥远的概念,成为一段时间以来出版人热议的话题。2025新闻出版单位数字出版部门主任联盟会近日在京召开。会上,《中国新闻出版广电报》记者关注到有两个问题引起出版人的热烈讨论。一是数据资产能给出版企业带来哪些经济效益?二是数据资产的商业化运营模式又有哪些?对此,记者采访了中国新闻出版研究院数字出版研究所所长王鹰。王鹰认为,已有迹象表明,敢用数据、会用数据的出版企业,其“双效”与管理水平显著提升。

出版业数据有优势

据2023年新闻出版统计公报显示,2023年,我国图书、期刊、音像制品、电子出版物出版,以及印刷复制、出版物发行营业收入均有所增长。其中,全国出版新版图书21.7万种,较2022年增长5.7%;重印图书32.3万种,增长8.8%。

由此看来,出版业是高价值内容资源富集的行业。但内容资源丰富并不等于数据资源丰富,这需要数字化和数据治理的过程,即把内容资源转化为高质量数据集的过程。

王鹰认为,出版业的数据资源主要分为经营数据和内容数据两大类。

其中,经营数据包括出版物主数据,如图书、期刊、电子音像出版物、网络出版物的数据,生产运营数据则是编、印、库存、采购、物流等环节数据,营销销售数据包括营销推广数据、销售发行数据、客户数据、渠道数据等。另外,还有财务数据,如凭证、台账、报表等,行政管理数据含有员工信息、公文通知、组织机构等。这些数据是出版企业内部使用或作为行业统计监管依据使用,在财务报表中按照数据入表要求一般计入无形资产(除销售数据外)。

内容数据方面,主要来自出版单位拥有版权的出版物的内容,包括图书期刊的图文数据,融入排版文件、印刷文件、电子书文件等,条目数据如XML、JSON等,及富媒体资源数据如音视频、动画、3D模型等,还有以上数据资源的相关元数据,主要用于对外销售交易,在财务报表中按照数据入表要求一般计入无形资产。

“我国出版业经过多年的数字化转型升级,已积累了丰富的内容数据和经营数据。相比于其他行业,出版业数据具有规模庞大、质量较高、类型丰富,有较好加工基础、产权清晰等特点。”王鹰认为,这些特点所具有的优势,使出版业成为一座数据宝藏。但目前这些数据价值远没有释放出来,大部分还处于“沉睡”状态。究其原因,是这些数据资源只是通过出版产品和版权实现其价值,没有转换成数据资产,并通过数据资产的运营最大化地释放数据资源价值。出版数据资源以产品和资产的形式一起汇入国家文化专网进行展示、流通、共享和交易,非常有利于文化大数据体系建设,从而推动并健全现代公共文化服务体系,创新实施文化惠民工程。

出版业数据资源

数据资产入表五大步骤

经营数据

- 出版物主数据
- 生产运营数据
- 营销销售数据
- 财务数据
- 行政管理数据

内容数据

- 图书期刊的图文数据
- 条目数据
- 富媒体资源数据
- 以上数据资源的相关元数据

第一步: 确定需要进行治理的数据资源的范围和边界。

第二步: 明确数据的存储环境、存储方式、数据结构、数据规模、数据冗余程度及元数据整理。

第三步: 将准备好的数据进行集中采集并进行清洗加工,以完成数据归集、数据脱敏、数据对齐、数据格式与结构标准化统一处理等任务。

第四步: 对清洗加工后的数据进行质量检查与管理,保证数据的一致性、准确性和完整性。

第五步: 将清洗质检后的数据入库存储,对数据进行多维度、多主题的分类、标注、关联和挖掘等更深层次的加工。

数据入表需五步骤

对于守着数据宝藏的出版企业而言,哪些是有价值的,哪些又是无价值的,甄别起来还是有一定困难。

对此,王鹰建议数据管理可考虑这样五个步骤。第一步:确定需要进行治理的(即要进行数据资产化的)数据资源的范围和边界。第二步:明确数据的存储环境、存储方式、数据结构、数据规模、数据冗余程度及元数据整理。第三步:将准备好的数据进行集中采集并进行清洗加工,以完成数据归集、数据脱敏、数据对齐、数据格式与结构标准化统一处理等任务,从而实现消除数据杂质和冗余,提升数据质量,增加数据价值的目标。第四步:对清洗加工

后的数据进行质量检查与管理,保证数据的一致性、准确性和完整性。第五步,将清洗质检后的数据入库存储,对数据进行多维度、多主题的分类、标注、关联和挖掘等更深层次的加工,进一步提升数据的可用性、知识性和内在价值。

上述步骤完成后,出版企业需要明确数据的三种权益。其中,持有权是通过数据资源持有权进行分级确权登记,给予受限持有、公开持有等确权等级,并进行必要的信息披露。加工使用权包括数据加工权和数据使用权,提供一种灵活的有限产权,为市场主体免除必须持有数据才能加工使用数据的顾虑。经营权则基于特定

场景进行数据版权合规高效流通,形成数据准入标准(合规标准、质量标准),促进数据产品接入、信息披露、资本化。

同时,王鹰提醒出版企业还应建设主数据和元数据管理的组织和体系。比如,制定管理的制度和标准;明确管理的职责和分工,对数据进行精确采集、校验、清洗、去重、整合、对齐及补齐等优化操作(例如对加工完成的电子书的内容进行校对、修正);对主数据和元数据进行多元化分析,保障管理系统的稳定运行和数据的妥善存储,加强数据的安全防护等。还有,要对数据资产进行确权固定、分级分类,对重要的数据在存储、传输等环节进行加密,对数据的敏感部分进行删除、遮蔽或替换,对数据采用私有化部署等。

酷、腾讯视频等)、网络社交平台(微信)等,网络出版企业的数据资产主要业务是基于用户的数据资源。目前已有一些评估模型可供参考,例如DEVA模型、曾李青模型、国泰君安模型等。

数据能给出版业带来哪些经济效益?王鹰认为,数据收益除带来科学管理外,还可以作为资产来资本化,资产资本化有这样几种途径:数据资产质押融资、数据资产证券化、数据资产作价入股、数据资产保险、数据资产增信融资、数据资产授权使用、数据资产信托等。

“总之,在推动数据要素市场化大背景下,推动出版业数据经济发展和数据价值最大化已是当务之急。”王鹰说。

数据资产化:

不是选择题而是必答题

□本报记者 李美霖 见习记者 蒲添

在数字经济与文化强国战略的交汇点上,出版业正经历着从“内容提供商”向“数据服务商”的转型。财政部《企业数据资源相关会计处理暂行规定》落地后,全国出版单位加速数据资源入表进程。

在近日举行的2025新闻出版单位数字出版部门主任联盟会暨第十五届中国国际数字出版博览会筹备研讨会上,来自全国200余家出版单位的负责人围绕“数据资产化”这一核心议题展开深度讨论,共同描绘出版业数字化转型的新图景。

推进试点尤为重要

数据资产入表为出版业打开了价值重估的新维度。出版业需要发挥行业数据优势,助力数字中国建设。

中国新闻出版研究院数字出版研究所所长王鹰认为,出版业的数据资源相对于其他行业具有多方面的显著优势,是一座尚未充分开发的数据宝藏。通过行业数据资产化和生产要素化的转变,出版业也将在数字中国建设过程中扮演更重要的角色。

以数传集团与四川人民出版社合作的“元宇宙图书”项目为例,该项目经北方国家版权交易中心下发国内出版行业首个数字版权登记凭证,实现确权,并在北方国家版权交易中心收益权交易平台实现元宇宙图书收益权挂牌转让成交。这一案例不仅验证了“科技+版权+金融”的协同效应,更将图书内容转化为可交易流通的标准化数字资产。

选择数据资源基础好、资产实力强的出版集团和大型出版机构作为试点单位,推进试点工作并建库试点体系尤为重要。人民交通出版社财务总监缪之湘提出了数据资产管理的优化路径。

她建议,坚持以数据资源持有权为前提,以数据加工使用权为重点、以数据产品经营权为关键,实施数据资产全生命周期管理。同时,通过建立健全出版社内部数据资产各项规章制度,搭建数据治理框架,并明确责任与核查机制,严格数据资产质量管理;通过提升从业者专业素养、设立新岗位、发展数据产业链来培育市场生态体系。

面临确权难与计量难挑战

尽管前景广阔,出版业数据资产化仍面临确权难、计量难两大核心挑战。

出版社积累了一定数量或规模的数据,甚至建成了数据库,并为数据做了标识,但为何会出现“数据孤岛”?与会专家认为,其根本原因在于标识标准在底层技术上不具备关联功能。

成本归集与计量,是数据资产入表的难点。根据财政部的规定,数据采集、脱敏、清洗、标注、整合、分析、可视化等加工过程中的支出,是成本的主要构成要素。数字化时代,出版行业的数字资产来源于版权作品,数字资产首先进行传统内容数字化处理,包括基于内容生产的数字版权产品开发、运营和管理。同步提取有价值可变现的数据元素,再形成基于数字内容产生的数据要素产品。

数据资产的确权问题涉及“确有无”“确多少”“确长短”“如何确”四个基本层面。缪之湘强调,在这一过程中,一方面,传统评估方法在数据资产评估中的应用存在局限性;另一方面,出版单位内部尚未建立完善的数据治理体系,数据质量、数据安全等仍是治理重点。

北方国家版权交易中心副总经理、国家新闻出版署出版融合发展(辽宁)重点实验室副主任李昊千认为,目前出版业数字资产管理能力薄弱,版权保护与法律风险高发,商业模式单一与产业链协同不足,另外,出版社数字化转型投入不足,传统出版企业受资金和风险意

识制约,对新技术应用保守,难以支撑长期创新。

用友文化传媒行业数智化咨询专家杨峰提出了实现出版行业数据资产入表的三个关键路径,即数据资源化、数据产品化和产品电子化。出版社应借助自有或第三方数据平台,实现资产成本的快速归集与分摊、数据资产的自动确权 and 计量,最终形成绩效报表披露。

作为国内领先的数字化转型服务商,用友为出版行业数字化转型与一体化服务提供解决方案。通过指数融合、一体化平台构建、产业协同、精准数据支撑和数据服务,实现全产业链的一体化管理和数字化运营。

计量体系的创新探索同样紧迫。山西春秋电子音像出版社开发的《五台山文化遗产精品壁画与塑像》数据库借助现代数字图像技术,对五台山不同历史时期的佛教塑像、壁画,进行较为系统和细致的展示。该数据库版权经由北方国家版权交易中心版权交易平台完成挂牌,这一案例为文化数据资产化提供了可复制的估值范式。

中国标准化研究院副研究员高昂认为,出版业需要根据由ISO组织正式发布的国际标准ISO 55013《资产管理 数据资产管理指南》来建立、实施、保持与持续改进数据资产管理体系。通过定义数据采集、存储、处理和销毁的流程,以确保数据的一致性和安全性。通过数据质量管理,确保数据的准确性、完整性和及时性,以支持组织的决策过程。

驱动出版业数字化转型

数据资产化和数据要素概念的出現,第一次建立了出版业与内容产业之外的行业进行价值交换的通路——以“数据”为生产要素的形式开始在社会进行流通交易。

在数据资产化浪潮下,出版业正通过技术创新与模式重构开辟增长新空间。例如在全球快速崛起的卡牌市场——不少出版社基于优质IP,打造艺术卡牌产品与“藏书票”,通过数字化“藏书票”、数字卡牌的发行,同步绑定实体纸质图书兑换、实体卡牌兑换、集卡权益兑换等的销售和互动。

这种“数实融合”模式,既保留了传统出版的文化价值,又开辟了增量市场空间。同时也印证了数字资产与消费场景结合的巨大潜力。AI技术的深度应用正在重塑产业格局。出版业实现数据资产化后,行业优质的科技文献内容数据资源和文化内容数据资源可以封装成高质量的语料库和数据集产品,作为数据要素进入市场交易流通,以供大模型训练使用。同方知网数字出版中心项目经理何朝辉认为,AI Agent(人工智能代理)作为智能化工具,正在深度重构出版业的创作、生产、分发及服务全链条,其核心价值在于降本增效、创新体验、精准触达、版权保护,助力出版业实现内容形态多元化、服务智能化、运营数据化。

“数字资产是出版业的‘数字田地’,版权经纪人则是智慧农夫,二者协同催生‘内容种植—价值收割—生态再生’的良性循环。”李昊千说,版权经纪人通过专业化运作,将出版单位版权资源开发形成大量数字资产,数字资产的市场反馈为版权经纪人提供数据资源支撑与技术赋能。版权经纪人通过市场化运营实现数字资产的价值转化。二者共同构成出版业数字化转型的核心驱动力。

随着数据资产入表政策的深化,出版业将加速从“内容提供商”向“数据服务商”转型。会上,不少专家认为,“数据资产化不是选择题,而是必答题。出版业唯有拥抱变革,才能在重构中把握先机。”未来,出版行业或将呈现数据要素深度融入生产流通、技术标准与评估体系逐步完善、数实融合场景持续创新的三大发展趋势。