

大数据环境下广电信息工程的数据共享理论研究

景 婷¹ 杭佳宇²

1. 浙江业视数智科技有限公司 浙江 杭州 310000

2. 杭州闪秀文化科技有限公司 浙江 杭州 310000

摘 要: 大数据技术正蓬勃发展,广电信息工程碰上数据规模猛增、类型繁杂等新状况,挑战与机遇并存。数据共享对提升广电信息工程效能、推动行业创新意义重大。咱们来剖析大数据环境下广电信息工程的数据共享,讲讲内涵、关键技术,探讨共享模式、面临的问题及解决办法,给广电信息工程在大数据时代高效共享数据提供理论支持,助力广电行业数字化转型。

关键词: 大数据;广电信息工程;数据共享;数据安全

引言

大数据时代,广电信息工程产生、积累的数据量呈指数增长,数据价值愈发凸显。数据共享整合广电系统内外各类数据资源,促进信息流通与协同创新,提升广电业务质量和效率。可广电信息工程涵盖内容生产、传输、分发等环节,数据来源广、格式多样,实现高效数据共享难题不少。研究大数据环境下广电信息工程数据共享理论,对推动广电行业可持续发展意义重大。

1 大数据环境下广电信息工程数据共享的内涵与特征

1.1 数据共享的内涵

广电信息工程里,数据共享内涵丰富、意义重大。其涉及范围广,涵盖广电业务流程产生的各类数据,像节目制作时的原始影像、音频素材这类素材数据,用户收视行为里的频道选择、观看时长、互动操作等收视行为数据,还有传输、播放设备的工作状态、性能参数等设备运行数据。这些数据在不同部门、系统、业务主体间传递、交换与使用,为的是实现数据最大价值。这可不是简单的数据传递,而是基于统一标准规范,统一数据格式、编码规则、元数据定义等,保证数据在不同环境都能被精准理解、有效利用^[1]。数据共享核心在于打破数据孤岛,让各环节不再孤立,促进广电行业各环节协同,为内容创作提供丰富素材、创意灵感,为精准营销给出精准用户画像、行为分析,为服务优化提供有力数据支撑,全面提升广电业务质量和效率。

1.2 数据共享的特征

广电信息工程的数据共享呈现多维度特点。数据规模极为庞大,海量视频、音频内容数据,还有用户行为数据等都涵盖其中,数据量随时间持续快速增多。数据类型丰富多样,有结构化的用户信息、收视纪录数据,也有大量非结构化的音视频文件、文本评论之类。不同

类型数据处理、存储方式不一样,让数据共享变得更复杂。数据时效性突出,广电行业内容更新迅速,用户需求变化快,数据价值常和时间紧密相连。实时收视数据可助力广电企业快速调整节目编排、营销策略,而过时数据价值就大大降低了。数据共享安全性要求颇高,广电数据包含用户隐私、商业机密以及意识形态等重要内容,共享时得保证数据保密性、完整性与可用性,防止数据泄露、被篡改。

2 大数据环境下广电信息工程数据共享的关键技术

2.1 数据采集与预处理技术

数据采集是数据共享根基,广电信息工程数据来源繁杂。拿用户收视行为数据来说,能借助机顶盒、智能终端这些设备采集,得到用户观看节目时长、频道切换频率等信息。内容生产环节呢,会采集设备运行数据、节目制作里的素材数据等。采集时,得保证数据准确、完整,用传感器采集、日志文件收集等多种方式。采集来的数据常有格式不统一、含噪声数据等状况,得预处理。数据清洗就是去掉重复、错误、无效数据,像把设备故障造成的异常收视数据剔除。数据转换是把不同格式数据变成便于处理的格式,比如将不同编码格式的视频数据转为通用格式。数据集成是整合不同数据源数据,形成统一数据集,为后续数据共享、分析打基础。

2.2 数据存储与管理技术

海量又多样的数据面前,传统存储方式没法满足需求了。像Hadoop分布式文件系统(HDFS)这种分布式存储技术,把数据分散存到多个节点,提升存储容量与读写性能,能适配广电大数据存储。HDFS容错性强,就算部分节点故障,数据也能用。数据管理上,数据库技术很关键。关系型数据库适合存结构化数据,像用户信息、节目清单,能保障数据一致性、完整性。非结构化

和半结构化数据,比如音视频、文档数据,MongoDB、Cassandra这类非关系型数据库(NoSQL)优势更大,数据模型灵活、扩展性强^[2]。数据目录管理系统是数据共享重要支撑,它给数据提供统一索引与目录服务,方便用户快速找和访问数据。借助数据目录,用户能了解数据来源、格式、含义等元数据信息,提升数据共享效率。

2.3 数据传输与交换技术

广电信息工程数据共享里,数据传输与交换技术让数据在不同系统、部门间高效可靠传输。网络传输协议是数据传输基础,像TCP/IP协议族,保证数据在网络里稳定传输。为提升传输效率,用光纤网络这类高速网络技术,它带宽大、速度快,能满足大量数据快速传输。数据交换技术实现不同系统数据格式转换和交互。Kafka、RabbitMQ等消息队列技术,在数据交换时起缓冲、异步处理作用,能解除不同系统依赖关系,提升系统稳定性、扩展性。靠消息队列,数据按规则排队、分发,保障数据有序传输、处理。

3 大数据环境下广电信息工程的数据共享模式

3.1 基于平台的集中式共享模式

平台身负重任,像个数据大集合站。把广电企业内的内容制作、分发、市场、技术等部门,外加节目供应商、广告商这些外部伙伴产出的数据,全汇聚到一起集中存管。数据格式上,定下统一的音视频编码规则,像视频编码用某某高清标准,音频按某音质规范,还有各类数据文件格式,好让不同源头的数据在平台能匹配。数据质量有明确要求,得准确、完整、及时。平台会仔细筛选汇聚来的数据,去除重复、错漏、没用的数据,就像挑出设备故障弄出的错收视数据。整合时,把零散数据按特定逻辑关系关联,比如关联同一用户不同业务场景的数据。再依据数据特性和用途分类,拿用户数据来讲,分基本信息,像姓名、年龄、住址,还有收视行为数据,例如看节目时长、换台次数等类别。经过这些操作,平台给用户一站式数据共享服务,大大方便了用户获取、使用数据,助力广电业务顺利开展。

3.2 分布式对等共享模式

分布式对等共享模式,其核心在于各参与主体平等,不存在中心节点掌控。在此模式下,各个参与节点既能存储数据,也可处理与共享数据,节点间直接开展数据交换与共享。它借助分布式哈希表(DHT)这类技术,达成数据的分布式存储与查找。DHT技术会依据特定哈希算法,将数据分散存储于各个节点,当要查找数据时,通过哈希计算能够迅速定位到存储该数据的节点。于广电信息工程领域,多个地区的广电分支机构可

运用分布式对等共享模式。不同地区的分支机构拥有各自的独特节目资源以及本地用户数据。要是某个分支机构想要获取其他地区的优质节目素材,并非如传统模式那样得借助中心平台中转,而是能直接从存有该数据的节点获取。这种直接的数据交换方式,让数据共享更具灵活性与自主性,防止了因中心平台故障或者数据传输拥堵引发的共享延迟状况。

3.3 混合式共享模式

混合式共享模式,把集中式、分布式优点巧妙融合。构建一个核心数据共享平台,同时允许各参与主体有一定直接数据交换。像用户基本信息、版权数据等关键且需统一管理的数据,存于集中式平台。因其重要,要严格管理维护,平台负责保障数据安全、完整、一致^[3]。平台会定期备份这些数据、加密存储,还设严格访问控制机制,防止数据被非法获取、篡改。而本地节目制作素材、区域收视数据等时效性强、具局部性的数据,各参与主体依规则分布式共享。这类数据特性决定更适合本地节点间快速共享,提升数据使用效率。跨地区大型节目制作时,核心平台提供版权管理、数据协调服务,确保节目制作中版权合法使用、数据合理调配。各制作团队能直接共享本地制作素材,减少经中心平台传输数据的时间、成本,提高协作效率。此模式既保证数据集中管控、安全,又发挥分布式共享灵活性,提升广电信息工程数据共享效能。

4 大数据环境下广电信息工程数据共享面临的问题与应对策略

4.1 数据安全性与隐私保护问题

广电信息工程数据共享,数据安全性与隐私保护是重大难题。数据共享时,有数据泄露、篡改、滥用风险。恶意攻击者会借网络攻击获取用户收视行为数据,侵犯用户隐私;内部人员也可能因操作不当或违规致数据泄露。为解决这些,利用加密技术加密存储、传输数据,保证数据在存储、传输时的安全^[4]。访问控制技术是保障数据安全的关键,设置不同用户角色、权限,严格限制用户对数据的访问。只有获授权的用户,才能访问特定数据资源,防止未经授权的访问、数据滥用。构建数据安全审计机制,记录、监控数据的访问、操作。一旦察觉异常行为,马上预警、处理。强化员工数据安全培训,提升员工安全意识、规范操作,减少人为因素引发的数据安全问题。

4.2 数据标准与规范不统一问题

广电行业架构复杂,节目策划、制作、传输、播出等环节各有其业务系统,像节目制作、用户管理、广告

投放系统等。数据标准与规范不统一成了数据共享的大难题,不同部门和系统数据格式多样,音视频有MP4、AVI等,编码方式有UTF-8、GB2312,元数据定义差别大,节目名称记录方式各异,这让共享数据难以理解和使用。解决此问题得制定统一标准规范,行业协会和标准化组织要起主导作用,规定音视频用H.264编码的MP4格式、数据文件用CSV格式,统一节目名称记全称、制作时间格式为“YYYY-MM-DD HH:MM:SS”、版权信息涵盖所有者等。企业内部也要建立数据治理机制,梳理规范数据,采集阶段严格按标准采集录入,如收视数据按规定记录,存储管理时转换清洗不合标准的数据,提高数据一致性和可用性。

4.3 数据共享的利益分配与激励机制问题

数据共享时,各参与主体利益诉求有别,怎样合理分配数据共享带来的利益,是影响大家参与数据共享积极性的关键所在。要是没有有效的利益分配机制,数据拥有者可能就不愿共享数据,进而让数据共享难以顺利开展。要建立合理的利益分配机制,就得综合考虑数据的贡献程度、使用价值这些因素。对于提供数据的一方,能够依据其提供数据的质量高低、数量多少以及使用频率等指标,给予对应的经济补偿,或者其他形式的回馈^[5]。拿广告投放业务来说,按照数据提供方所提供的用户精准画像数据的质量好坏、使用成效,给予一定比例的广告收益分成。还得构建激励机制,鼓励各个参与主体踊跃参与数据共享。对于在数据共享中表现出色的部门或者个人,给予荣誉奖励、技术支持之类的激励手段。通过建立良好的利益分配和激励机制,提升各方参

与数据共享的积极性,推动数据共享能够持续发展。

结语

大数据环境给广电信息工程数据共享带来新机遇、新挑战。透彻明白数据共享内涵、特征,运用先进关键技术,选对数据共享模式,应对好数据安全、标准规范、利益分配等问题,就能实现广电信息工程数据高效共享。这不但有利于提升广电行业内容创作水平、优化用户服务,还能助力广电行业在大数据时代创新发展。随着技术不停进步、行业持续发展,广电信息工程数据共享理论会不断完善,给广电行业数字化转型提供有力支撑。持续留意数据共享领域新技术、新趋势,不断探索创新数据共享模式、管理机制,是广电行业在大数据时代保持竞争力的关键。

参考文献

- [1]李运熙,李灵琚.大数据环境下企业会计信息化管理模式研究[J].商场现代化,2022,(17):177-179.DOI:10.14013/j.cnki.scxdh.2022.17.018.
- [2]肖炯恩.跨源多维政务数据共享与服务创新研究[D].华南理工大学,2019.
- [3]吉孙嵘,叶战备.数据审计促进政务数据共享的机理与进路——基于ICA理论的分析[J].商业经济,2025,(02):125-128+143.
- [4]陈怡.企业间强制数据共享:理论、挑战及实现路径[J].中国海商法研究,2024,35(03):83-95.
- [5]李莉莎,谭镇锋.金融数据共享:理论、挑战与机制构建[J/OL].经济与社会发展,1-16[2025-03-13]