

基于数据中台的文博大数据应用研究

刘韩萍¹, 向诗月¹, 郭曦榕^{2*}

¹四川省文物信息中心, 四川 成都

²成都信息工程大学管理学院, 四川 成都

收稿日期: 2023年11月27日; 录用日期: 2023年12月23日; 发布日期: 2023年12月30日

摘要

随着我国数字文化战略的实施, 数字博物馆建设的不断推进, 文博大数据资源日益丰富, 如何将文博大数据的价值应用在文化公共服务、文物保护、文物和文化以及研学教育是当前文博大数据应用必须要面对的问题。本文旨在利用平台驱动应用的机制, 将数据中台的优势与文博大数据应用的特点相结合, 研究基于数据中台的文博大数据中台的构建及其应用策略, 为文博大数据资源的开发和利用提供技术参考。

关键词

数据中台, 文博资源, 大数据, 数字化

Research on the Application of Culture and Museum Big Data Based on Data Center

Hanping Liu¹, Shiyue Xiang¹, Xirong Guo^{2*}

¹Sichuan Provincial Cultural Heritage Information Centre, Chengdu Sichuan

²School of Management, Chengdu University of Information Technology, Chengdu Sichuan

Received: Nov. 27th, 2023; accepted: Dec. 23rd, 2023; published: Dec. 30th, 2023

Abstract

With the implementation of China's digital culture strategy and the ongoing advancements in digital museum construction, the resources of cultural and museum big data are becoming increasingly abundant. How to apply the value of cultural and big data to cultural public services, cultural relic protection, cultural relics and culture, as well as research and education, is a problem that the current application of culture and museum big data must face. This article aims to utilize the

*通讯作者。

platform driven application mechanism to combine the advantages of data platforms with the characteristics of cultural and museum big data applications, and study the construction and application strategies of cultural and museum big data platforms based on data platforms, providing technical references for the development and utilization of cultural and museum big data resources.

Keywords

Data Center, Culture and Museum Resources, Big Data, Digitization

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着信息时代的不断发展，大数据、云计算、物联网、互联网、人工智能等先进数字科技赋能博物馆，推动了博物馆管理提质增效，加速了文物资源数字化、信息化进程。党的十八大以来，党和国家高度重视中华优秀传统文化的传承、文化遗产的保护、博物馆工作的创新。2021年3月发布的《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》，两次提到“博物馆数字化”。2021年5月，文化和旅游部、国家文物局等九个部门联合印发《关于推进博物馆改革发展的指导意见》提出我国在2035年基本建成世界博物馆强国的战略目标，并对智慧博物馆建设及相关工作做了充分阐述，明确指出要“大力发展智慧博物馆，以业务需求为核心、以现代科学技术为支撑，逐步实现智慧服务、智慧保护、智慧管理”。2021年11月，《“十四五”文物保护和科技创新规划》正式印发，对提升文物科技创新能力进行了“全链条”布局。2022年5月，中共中央办公厅、国务院办公厅印发的《关于推进实施国家文化数字化战略的意见》提出，到“十四五”时期末，基本建成文化数字化基础设施和服务平台，形成线上线下融合互动、立体覆盖的文化服务供给体系。在此背景下如何实现博物馆数字化资源的管理和应用成为现代博物馆数字化建设面临的首要问题。

近年来，国内博物馆先后开展智慧博物馆建设，积累了很多有价值的数据库。如何将博物馆的数据转化为价值，在公共服务、科学研究、文物保护和教育科普中发挥作用是当前文博大数据应用必须要面对的问题。本文旨在利用平台驱动应用的机制，以数据中台为切入点，研究基于数据中台的文博大数据中台的构建及其应用策略，为文博数字化资源的开发和利用提供技术参考。

2. 文献综述

2.1. 数据中台的研究现状

数据中台是阿里巴巴公司在2015年根据自身业务需要提出来的概念，旨在将阿里巴巴“厚平台，薄应用”变成“大中台，小前台”。2017年5月，阿里集团出版的《企业IT架构转型之道：阿里巴巴中台战略思想和架构实战》详细阐述了业务中台是介于前台与后台之间，采用共享式的方式解决了以往企业“烟囱式”和单体式架构设计产生的多并发、数据分散、试错成本过高等问题。同时将高内聚低耦合、数据完整性、可运营性、渐进性等，列举了建设业务中台的一些原则。随着2018年各大互联网公司纷纷提出中台战略并随之进行组织架构的调整，将“中台”的概念推向了高潮，各类公司纷纷推出了技术中

台、业务中台、AI中台、算法中台、研发中台、组织中台、服务中台、数据中台等理念和技术框架，其核心是抽象、组件化共性的能力，以降低成本、避免重复建设浪费资源。

在知网输入“数据中台”和“中台技术”两个关键字，CNKI指数显示数据中台的学术关注度在2018年均呈现大幅上市的趋势，且对数据中台的关注度要略高于中台技术(如图1所示)。目前国内对于数据中台的研究趋于成熟，数据中台在互联网、金融、电力、媒体等多个行业被广泛应用。数据中台的构建理念是将大数据“用起来”，即通过大数据向管理赋能，通过数据中台将大数据变成实际“可用”的东西，从实用的角度来看，数据中台是大数据运用的平台化、具体化、实践化[1]。数据中台的功能理念是将大数据“统起来”，即通过数据中台打通“数据孤岛”和“数据烟囱”分立的IT(Information Technology, 信息技术)架构，构建数据共享的统一的DT(Data Technology, 数据技术)架构[2]。数据中台的定位理念是将大数据“连起来”，即通过数据中台将前台与后台有机地连接起来，有效解决数据与管理业务的脱节问题。只有将数据“用起来”“管起来”“连起来”才可能让数据发挥其巨大的价值。

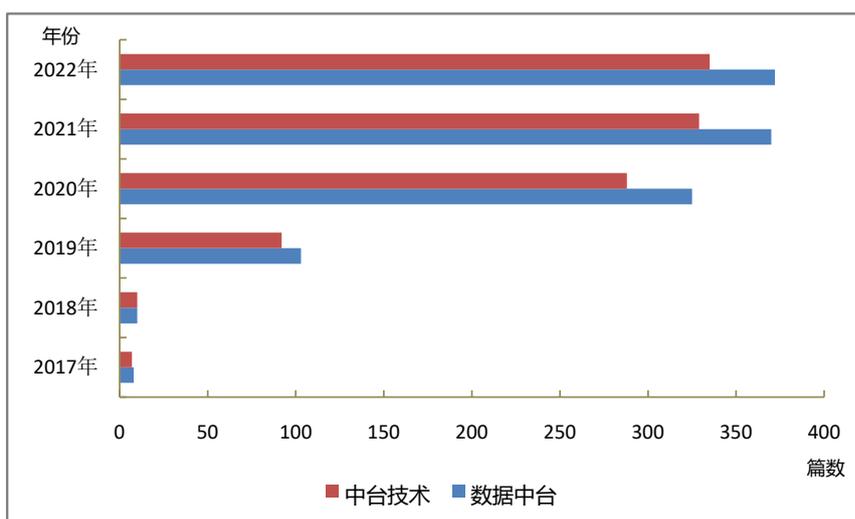


Figure 1. The trend of changes in the amount of document research on big data middle end from 2017 to 2022

图1. 2017~2022年关于数据中台研究的发文量变化趋势

2.2. 文博大数据的研究现状

2020年5月26日，中宣部中央文改领导小组办公室发布《关于做好国家文化大数据体系建设的通知》，指出建设国家文化大数据体系是新时代文化建设的重大基础性工程，也是打通文化事业和文化产业、畅通文化生产和文化消费、融通文化和科技、贯通文化门类和业态，推动文化数字化成果走向网络化、智能化的重要举措[3]。信息采集、存储、处理和可视化技术的不断发展和国家大数据发展战略等政策成为推动文物大数据和博物馆大数据资源建设、开发和利用的两大推力。

利用中国知网知识元检索功能，输入“文物大数据”和“博物馆大数据”两个关键字，CNKI指数显示博物馆大数据的学术关注度始终高于文物大数据，在2015年以后关注度大幅增加，明显高于了文物大数据，与博物馆大数据和文物大数据相关主题的发文章在2020年达到顶峰，近几年来研究热度正在逐步下降，具体如图2所示。文博大数据的研究早期侧重于博物馆大数据应用分析，如沈红军(2015)研究了大数据在文物信息化、古建保护、数据的分析利用、互联网推广等方面的应用[4]；王瑛(2015)研究了大数据在智能导览系统、数字博物馆建设中的应用[5]；张久君(2015)研究了大数据在历史文化遗产管理与利

用中的应用[6]。随着数字博物馆和智慧博物馆建设的不断推进,文博大数据的研究更加具体化,王晓丽(2022)以山西博物馆为例研究了博物馆大数据在创新公共服务方面的应用[7];周虹霞(2022)研究了基于大数据与微服务的博物馆智慧服务[8];岳娜(2020)研究了“大数据”背景下智慧博物馆发展现状及对策;曾熙(2020)研究了文化遗产大数据二维分类框架[9]。

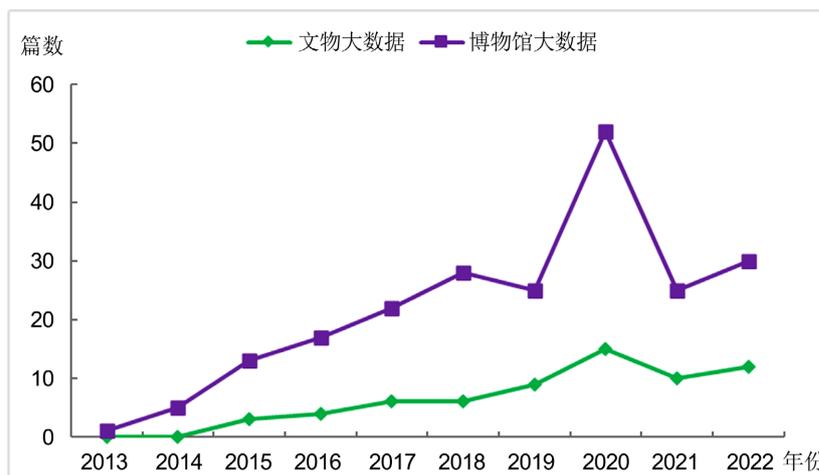


Figure 2. The trend of changes in the amount of document research on museum big data from 2012 to 2022

图 2. 2012~2022 年关于文博大数据研究的发文量变化趋势

3. 文博大数据的应用需求

3.1. 文博大数据的来源

伴随着我国数字文化战略的推进,利用数字技术实现文物“活化”保护、科研以及博物馆数字化建设已成为文博行业进行数字化转型的重要抓手,文物和博物馆的大数据应运而生,文物大数据通常包括可移动文物和不可移动文物的信息、现状、结构、体量、纹理、材质、色彩等数据,另一方面数字博物馆的建设将博物馆进行虚拟化生成了博物馆大数据,包括文物大数据、博物馆数字化版本、博物馆数字化管理与服务也产生了大量的数据。当前文博行业关注的重点是如何将文物数字资源进行汇聚、保护和传承利用,并且服务于文物行业的防、保、研、管、用。

3.2. 文博大数据的应用方向

实体博物馆与数字化博物馆的数实融合是驱动文博大数据应用的根本原因,文博大数据将为行政管理、展览策划、科学研究、保护规划、社教活动、文创产品研发、用户体验提升等提供数据支撑。

对博物馆来说一方面借助各类二维和三维数字化技术将实体文物藏品进行数字化,生成数字化文物藏品进行多维全方位展示,丰富藏品数据展示手段;另一方面需要对数字化普藏品资源进行查询及统计分析、动态管理与共享发布、维护藏品基本信息、附属信息、文物故事、讲解词等内容,同时能够满足展览策划、馆际合作、业务研究等需求。

对文物管理部门而言,一方面需要掌握文物数据和博物馆运营数据的实时变动,通过大数据分析与检索技术,获得各类查询统计、对比分析、报表输出等服务,为管理者实现行业管理、科学决策提供有力的技术支持;另一方面文物管理部门需要整合博物馆、展览、藏品等资源,提供统一的公众服务。

对社会公众而言,需要满足社会公众游览、观赏、研学、交流的需求。一方面需要以便捷的方式获

取博物馆、展览及藏品的多维度动态信息，并且借助大数据分析技术更多的了解藏品背后人物和文物的故事，拓展公众的历史文化知识，提升其对传统文化的认同感；另一方面利用基于 LBS 地理位置服务能够利用大数据分析模型挖掘文物之间的关联关系和时空演化关系，进而利用数字化手段对这些内容进行数字化文创，并实现内容分享、传播和评论等互动。

4. 基于数据中台的文博大数据应用平台构建

4.1. 数据中台的概念

数据中台并不是一个具体的产品或技术架构，而是一种综合解决方案或是一种运营理念，在业界还没有清晰明确的定义。数据中台可以理解为是数据界面的一种架构，通过数据技术将数据从应用界面复制出来，然后通过数据计算、加工，进而为上层数据应用提供统一标准和口径的支撑。数据中台主要是构建规范的、数据全域化的、智能的数据处理架构，进而为前台提供高效的数据分析和应用服务。数据中台主要涵盖了数据资产、数据治理、数据模型、全域数据中心、数据服务等多个层次的体系化建设方法。相比国内的研究，国际上虽然尚未推出数据中台的概念，但针对具体的数据治理、数据服务和数据应用等方面的研究已经比较成熟。

4.2. 数据中台的优势

数据中台定位为数据开发者提供集成的一体化的数据开发平台，总体而言，数据中台能够避免数据的重复计算、提升数据服务力、提高数据共享力和将数据资产化管理。数据中台的优势主要表现在以下几个方面：(1) 数据中台强调数据业务化，让数据用起来，满足企业数据分析和应用的需求。(2) 数据中台增加了以企业的全局视角来梳理数据域的环节，这是数据中台建设中很重要的一环。数据域的梳理正好体现了中台化的能力。(3) 在数据中台的规划阶段就需要去主动迎合业务，需要全面梳理哪些业务场景需要利用数据的赋能才能形成业务闭环，因此，在建设数据中台的同时就必须着眼于业务场景的赋能。(4) 数据中台则将建、治、管、服放到同样的高度，而不仅仅局限于数据建设。

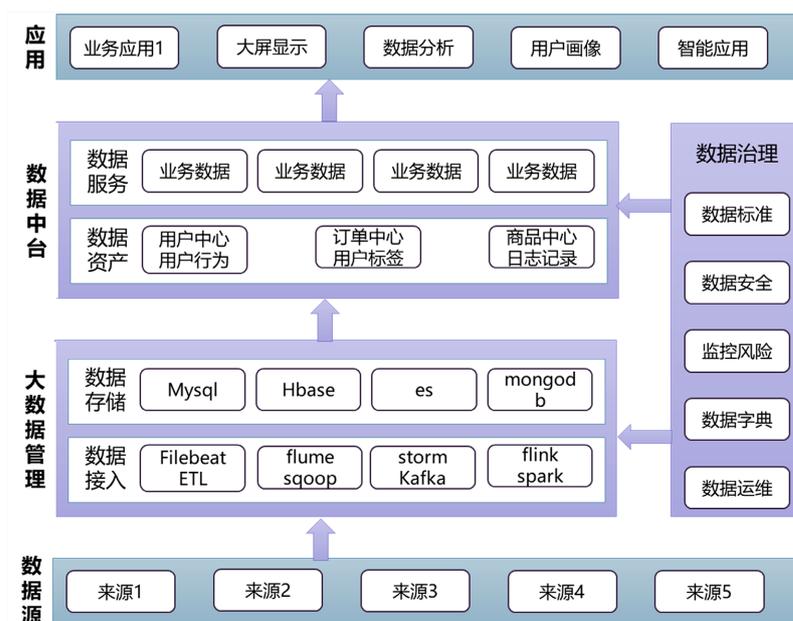


Figure 3. The basic structure of the data center
图 3. 数据中台的基本构架

4.3. 数据中台的架构

数据中台具备“采”、“存”、“管”、“用”四个基本功能。数据接入功能首先需要制定一套数据迁移和采集方案，确定如何将现有数据迁移进来，后续哪些数据实时采集，哪些数据离线采集，以及数据拉取的频率；数据加工存储对接入进来的数据进行整合与完善，打造符合公司业务规划需要的全域数据体系，并传送到大数据平台进行存储；数据治理要保证数据安全同时避免数据污染，因此也需要制定一套数据治理的措施及其规范，包括数据字典、元数据管理；数据安全、权限管理；数据标准、质量管理等；服务应用以成熟的标准体系为基础，以数据 API 的方式为业务提供服务，加速企业从数据到业务价值的转化。为实现上述功能，数据中台的基本构架如图 3 所示。

4.4. 基于数据中台的文博大数据应用平台架构

文博数据中台需要根据文博行业在其业务管理、公共服务、沉浸式体验、科学研究和保护留存等方面的实际需求，结合数据中台的优势，构建以数据中台为核心，业务中台和技术中台为支撑的中台层。文博大数据应用平台采用如图 4 所示的“前台 - 中台 - 后台”的三层架构，其中前台部分主要提供文博大数据典型应用的接口；中台层提供文博大数据应用的数据和技术支撑；后台负责前台接入业务所需的基础性功能处理。

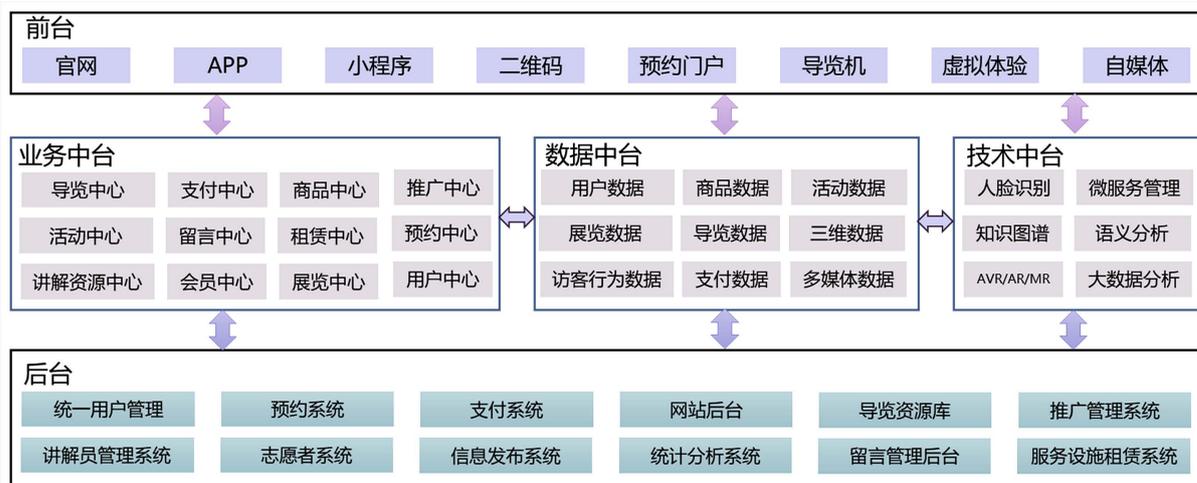


Figure 4. Architecture of culture and museum big data application platform

图 4. 基于数据中台的文博大数据应用平台架构

(1) 前台层

前台层是面向博物馆参观者的接入层，通过业务中台拼接成各个应用前台，为观众提供服务。各终端应遵从公共服务平台的统一技术路线，采用统一的规范接口将各类应用进行接入。根据前面的业务分析，将博物馆面向公众提供的平台划分为网站、APP、微信、小程序、二维码、导览机、预约系统等。

(2) 业务中台

业务中台主要提供业务应用支撑，将博物馆将各个系统、业务的前端应用与后端服务解耦，将各个功能中的服务进行梳理、并沉淀。其次将不同前端的重复、类似的服务进行整合，在单个服务的完善和增强的过程中注意服务的通用性。

(3) 数据中台

数据中台主要提供数据资源管理的能力。数据中台定位于为各专业、各单位提供数据共享和分析应

用服务。以数据分析域、管理域为基础，沉淀共性数据服务能力，通过数据服务满足横向跨专业间、纵向跨不同层级间的数据共享、价值挖掘、分析应用和融通需求。博物馆面向观众推出了网站、APP、小程序、导览、预约、购物等多个系统，随着时间的积累沉淀了大量的数据资源，通过建设数据中台实现数据互通共享，进行用户信息资源共享和板块间交叉营销。

(4) 技术中台

技术中台是为博物馆提供技术开发支撑能力。主要是把一些公共、底层的技术能力抽象出来，与业务逻辑分离，并形成各种接入式基础服务，同时可以为多个业务线提供服务。

(5) 后台

后台是指系统的后端平台，通过后台的各系统对前端业务进行业务处理。博物馆公共服务后台主要由网站系统、活动管理系统、导览管理系统、展览管理系统、统一用户管理系统、公共服务系统多个系统进行支撑。

4.5. 文博大数据平台的功能

(1) 提供灵活的个性化检索

个性化检索本身就具有对应的灵活性，但是传统的数据资源检索模式，对于数据内部结构和语义的读取和理解能力有限。在此基础上，开发了基于自然语言处理和推荐算法的数据中心技术和数据检索中心，可以很好地了解用户的兴趣，向客户推送的信息更加符合客户的个性化需求，从而进一步保存构建关键词，满足用户的个性化信息需求，提高用户体验。

(2) 创建博物馆藏品的三维模型

利用数据中台技术能够为所有博物馆藏品建立三维模型，创建精确的结构化信息，包括年龄、材料和形式，分析和挖掘以往研究记录中的数据，并添加许多半结构化和非结构化数据。通过社交媒体与互联网的互动，可以广泛聆听不同领域的声音，为博物馆研究的学术研究提供新的视角，从而增加博物馆研究的深度和广度，促进各类博物馆的发展。

(3) 定制数字交互界面

访问者可以在其中输入/输出相应的文本信息，以获得所需的文本、图像或视频内容。参观者通过点击屏幕上的各种功能，能够选择合适的参观展览路线。现代博物馆利用超声波技术来确定游客在某一时刻的位置，然后通过声音或图像展示向普通观众解释物品和收藏品的细节。因此，数字智能系统的交互页面具有高度的个性化和个性化。

(4) 实现双向互动

数据中台可以让博物馆访客在虚拟环境下对博物馆藏品进行移动，360度观察藏品，同时数字化藏品可以根据访客的指令进行回应，如语言解说。同时也可以根据人说话的语义进行分析判断，满足访客的求知欲和好奇心。

5. 基于数据中台的文博大数据应用

5.1. 利用数据中台的数据管理能力，实现文博数据资产管理

数据中台具有强大的数据管理能力，博物馆里最关键的是藏品数据，数字藏品数据是文博大数据的重要组成部分。在智慧博物馆的未来发展中，博物馆要建立起“采存管用”全链条数据管理流程，构建起跨部门、跨层级的一体化数据资源管理体系，不断丰富自身的数字资源积累，并逐步扩大对外开放力度，真正让文博资源进入百姓生活，促进中华优秀传统文化全民共享。必要的情况下，可引入区块链技术，解决在数据资源开放过程中可能遇到的安全、版权、真伪、留存、利用等问题，实现藏品数据的永久性

保存、知识产权保护、文物信息追溯、藏品信息鉴伪、文物数据传播共享、数据资产利用等,为博物馆数字资源开放增加保护屏障,最终实现观众获取文博资源的极大丰富和博物馆藏品利用率的显著提高。

5.2. 利用数据中台的多源数据整合能力,实现多维度文博数据的聚合服务

随着智慧博物馆服务的多元化发展,数据中台需要能够对数据进行整合和完善,提供适用、适配、成熟、完善的一站式大数据平台工具,在简便有效的基础上,实现数据采集、交换等任务配置以及监控管理。博物馆数据中台必须具备在数据集成与运营方面的能力,能够接入、转换、写入或缓存来自访客、博物馆管理部门以及文物藏品等多种来源的数据,实现数据的深层次的挖掘和应用。

5.3. 利用数据中台的跨平台融通能力,打造文博数据驱动文博服务的数据生态系统

文博大数据中台架构能够实现文物和博物馆的数据化、数据资产化、资产服务化、服务应用化、应用智能化,持续赋能文博数据应用闭环,提高文博数据的服务效果与质量。文博数据服务的过程是由访客、导游、管理者、资源与设施等要素组成的复杂系统活动,文博数据的产生不仅来源于单个要素的作用,更多的是多个要素之间的全面协同调配、融通循环。因而,在文博大数据中台架构的深入应用过程中,要促进文博数据服务过程中各要素之间的数据融通,推动文博大数据与文博大数据应用之间的深度耦合,打造文博数据驱动博物馆智能服务的数据生态系统。

6. 总结

数字经济的发展态势下,数字化转型成为各行各业必须面对的问题,对于博物馆来说,如何改变传统的数据收集、分析、服务方式,拓宽数据利用渠道,为博物馆业务更新升级提供扎实的基础对博物馆的转型和智慧博物馆的建设至关重要。数据中台在传统前台和后台之间的纽带作用为文博行业的数字化转型提供了有效的技术路径。虽然目前我国大数据技术在文博领域的应用还处于起步阶段,但其在博物馆旅游发展、管理模式以及我国文博行业中发挥的效用已初具规模。大数据分析挖掘方法在文博行业的应用,将激发文博行业的潜能,文博大数据的潜在价值必将被挖掘和释放,不但能够满足当代公众对于文博行业的个性化需求,还能够参与到我国现代化文化事业的建设中来,成为将博物馆数字资源最大效益化呈现的重要手段。

参考文献

- [1] 张弛. 数据中台在应对突发公共事件中的核心价值研究[J]. 华中科技大学报(社会科学版), 2021, 35(1): 77-84.
- [2] 石进, 等. 基于总线结构的政务数据共享应用模型研究[J]. 现代情报, 2019, 39(6): 111-119.
- [3] 郭全中. 国家文化大数据体系建设的相关要点分析[N]. 中国新闻出版广电报, 2020-06-08(004).
- [4] 沈红军. 大数据在博物馆中的应用[J]. 信息技术与信息化, 2015(7): 63-64.
- [5] 王瑛. 大数据时代下的博物馆信息化建设思考[J]. 信息与电脑(理论版), 2015(15): 72+74.
- [6] 张久君, 李公立. 大数据背景下历史文化遗产的智慧管理与利用[J]. 信息与电脑(理论版), 2015(22): 96-99+118.
- [7] 王晓丽. 博物馆大数据在创新公共服务方面的研究——以山西博物院为例[J]. 文物鉴定与鉴赏, 2022(17): 52-55. <https://doi.org/10.20005/j.cnki.issn.1674-8697.2022.17.013>
- [8] 周虹霞, 李华飙, 周宇阳, 余景景. 基于大数据与微服务的博物馆智慧服务研究[J]. 博物馆管理, 2022(3): 47-54.
- [9] 岳娜. “大数据”背景下智慧博物馆发展现状及对策[J]. 中北大学学报(社会科学版), 2020, 36(2): 128-132.