

广西科技大数据平台关键技术的研究意义 与建设要点分析

文彬宏

(广西科技信息网络中心,广西 南宁 530012)

摘要:本文针对广西科技大数据平台建设特点,进行合理性的分析,简单介绍合理运用广西科技大数据平台的重要价值,提出广西科技大数据平台关键技术要点和注意事项,通过加强系统构架扩展性与功能定制,并有效解决科技数据管理过程中数据缺乏问题,可以确保广西科技大数据平台安全运行,充分发挥其重要价值,取得的良好效果,旨在为相关人员提供一定帮助。

关键词:广西科技大数据平台;关键技术;建设现状

中图分类号:G311

文献标识码:A

文章编号:1004-7344(2021)08-0243-02

0 引言

大数据技术的良好运用,已经成为推动各行各业快速发展的核心动力,结合大数据特点可以得知,通过运用此项技术建立完善的研究机制,能够确保大数据技术在科技管理领域得到更好运用,为科技行业的全面改革与发展奠定良好基础。因此,本文重点探讨广西科技大数据平台关键技术的研究意义和建设要点,内容如下。

1 研究意义

通过打造广西科技大数据平台,对各项数据进行集成分析,实现数据的全方位、多结构、实时性分析目标,对于顶层设计与统筹协调内容,加大优化力度,确保各项资源得到良好利用,减少“碎片化”等问题的出现。结合国家和自治区的改革方案要求,建设基于业务协同和大数据服务的公开统一的自治区科技大数据信息管理平台。

2 广西科技大数据平台特点分析

(1)科学打造大数据能力平台,实现纵向与横向的共同开放与共享目标。广西科技正在朝着信息化方向快速发展,要全面实现时间跨度、空间范围和实时共享等数据支撑,通过建立大数据能力平台,可以促进不同科技创新主体之间的良好沟通,形成稳定的大数据考核体系。例如,在此数据平台中,通过联合工商、税务、银行、企业客户评价等信息,促进不同领域和产业之间的有效交流与合作,确保大数据信息实现全面共享^[1]。

(2)合理分配各项数据资料,确保数据核心价值得到更好挖掘。通过利用广西科技大数据平台,能够更好地挖掘广西科技数据核心价值,以海量数据为基础,快速生成自动预测体系,由于系统所收到的数据不断增多,工作人员可以在短时间内快速搜索到信息,并将具有良好价值的信息有效整合,经过可视化处理后,做出科学决策。

(3)利用广西科技大数据平台,合理确定出数据指标。通过利用科技数据,能够更加直观地展示科技产业与项目内部的各项指标。当前阶段,广西地区仍然存在统计口径不一等现象,而广西科技大数据平台的建立,采取“互联网+”平台模式,并结合自治区区域发展的特色,进行分项的扩展,实现科技指数全面精准化、直观化科技指标^[2]。

3 广西科技大数据平台关键技术

3.1 系统构架扩展性与功能定制

现有科技政务数据治理和云服务平台在系统构架扩展性、定制性(弹性)方面存在较大问题,使系统难以修改和扩展,难以快速响应用户/业务部门对功能的修改和功能扩展需求。采用柔性治理技术框架,采用微服务和松耦合技术构架,提高系统的定制能力、可扩展性和运行性能^[3]。

当前阶段科技政务数据治理和云服务平台在系统架构扩展性和定制性方面存在一定问题,导致系统无法满足修改与扩展要求,通过采取柔性治理框架,并运用微服务与松耦合技术构架,可以明显提高系统的可扩展性与综合运行性能。大数据平台总体架构见图1。

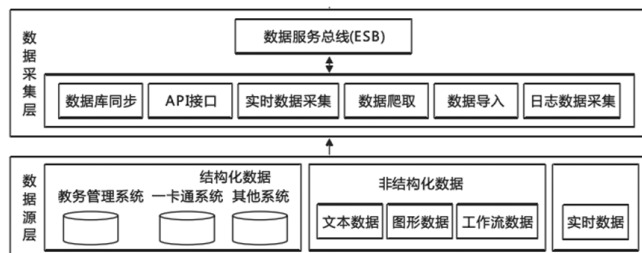


图1 大数据平台总体架构

此项目通过利用云数据技术和云服务技术,确保各项资料得到有效分配,实现动态配置目标,对科研人员来讲,云存储空间

表1 广西科技信息资源目录

序号	信息资源名称	备注
1	广西科技项目数据库	
2	部区会商制度信息数据库	
3	厅市会商制度信息数据库	
4	专业机构管理数据库	
5	联席会议和综评委管理数据库	
6	科技计划管理制度数据库	
7	科技专家库	
8	科技人才库	
9	专利侵权纠纷拆裁决数据库	

更大,海量数据得到更好利用,打造一种全新的云服务模式,促进学术交流,提高科研成果知识成果管理水平。从科技管理层面来讲,通过建立大数据平台,可以为管理者和决策者提供精确数据,为后续的管理决策提供重要支持。在大数据时代背景下,数据表现出密集和驱动性特点,此项目主要利用“互联网+”大数据技术,不仅可以提升各类科技信息的准确性,而且能够实现数据的快速分离,为各大科研室提供良好的数据支撑^[4]。

3.2 解决科技数据管理过程中数据烟囱与数据孤岛问题

科技信息管理体系在运行期间,因为不同部门之间的数据无法实现共享,业务的互通性比较差,主要是因为系统开发者对科技业务缺乏足够了解,使不同部门之间的业务数据不统一,同时,不同系统的数据格式与标准存在较大差异,使数据互通与共享方面存在较大难度。为了确保上述问题得到有效解决,工作人员通过建立大数据能力平台,可以实现各项数据的有效融合,加强数据共享。

3.3 研究出适用于科技领域的知识图谱

信息抽取作为创建科技知识图谱的核心步骤之一,通过加强信息抽取,可以确保海量数据得到良好利用,因为大规模数据集人工标注成本特别昂贵,在一定程度上增加了科技知识图谱构建难度。通过采取无监督、半监督的抽取方法替代原有的监督方法,能够显著降低成本。采取先聚类后标注的方法,并建立相应的训练模型,可以有效减少数据标注成本、降低信息抽取对标注的依赖^[5]。

3.4 加强科技大数据分析

通过加强科技大数据分析,可以为各科技部门的评价和决策提供良好支撑。和传统的专家系统评价体系相比较来讲,通过加强科技数据分析,以数据驱动为核心,采取量化评价方式,可以取得良好新效果。如果数据数量过多,而且质量比较好,通过采取数据驱动学习量化评价方法,可以做出正确决策。

在此项目中,通过运用“互联网+”模式和云计算、大数据等最新技术,并对数据库内部各类信息进行大力完善,根据科技体制与科技管理改革要求,优化项目管理流程,实现信息的深度改造,确保科研业务处理效率得到明显提高。同时,通过不断提高科技投入,并对各项数据进行分析预测,进行修正与调整,确保各项数据得到高效利用,制定出完善的解决方案,提高科技管理模式的合理性与规范性。

3.5 优化技术路线

通过采取柔性治理模式,对既有的技术路线进行全面优化,可以明显提高数据的精确性,具体内容如下:

(1)加强柔性流程管控,定制柔性功能,加大柔性数据管控力度,在此项目中,以流程模板为核心,优化流程审批内容,确保各项数据治理流程服务水平得到显著提高。

(2)以参数配置为核心,完善自定义技术框架,对既有的功能模块与技术模块进行优化,确保各类数据得到有效利用^[6]。

(3)加强柔性治理与知识图谱的科技政务数据治理云平台的开发和实现。主要包含三个子系统的开发,分别是基于流程模板化(及可视化)定制与自动化生成,建设科技业务融合管理系统;运用过程数据和知识数据分离的数据存储和管理构架搭建科技政务数据资源管理系统;基于数据可视化、数据挖掘技术和知识图谱技术建设科技决策辅助系统。系统建成之后,在广西科技管理部门进行应用示范,通过利用平台功能,全面建设科技项目管理流程,建设数据查询和分析模块,构建科技创新主体、科研项目、科技人才等主体的知识图谱;实现科技大数据的联合分析,为科技管理部门提供决策参考,广西科技信息资源目录见表1。

4 结语

综上所述,通过对广西科技大数据平台关键技术进行合理分析,例如加强系统构架扩展性与功能定制、解决科技数据管理过程中数据烟囱与数据孤岛问题、研究出适用于科技领域的知识图谱、加强科技大数据分析、优化技术路线等,能够确保广西科技大数据平台稳定运行。

参考文献

- [1] 赵少东,王程斯.基于异构计算与实时可视化技术的综合能源大数据平台研究与应用[J].微型电脑应用,2019,35(11):96-99.
- [2] 罗云芳,唐运乐,闵金花.高校教育大数据平台的构建与关键技术研究[J].广西职业技术学院学报,2020,13(3):110-114,124.
- [3] 杜毅博,赵国瑞,巩师鑫.智能化煤矿大数据平台架构及数据处理关键技术研究[J].煤炭科学技术,2020,48(7):177-185.
- [4] 辛继胜,刘泽华,李昊.课堂教学大数据采集与分析平台架构及关键技术研究[J].广东轻工职业技术学院学报,2020,19(1):11-17.
- [5] 白云,初庆东,倪方映.高等教育质量监测与评估大数据平台建设的 key 问题研究[J].教育教学论坛,2020(2):118-119.
- [6] 李晓明,张海军,许多红,等.基于大数据平台的电网线损与窃电预警分析关键技术分析[J].通信电源技术,2020,37(1):196-197.

收稿日期:2021-01-19

作者简介:文彬宏(1976—),男,汉族,广西桂林人,工程师,本科,主要从事政务信息系统建设及网络信息安全管理。