

# 大数据科技创新资源建设与共享机制研究 ——以四川革命老区为例

江忠<sup>1</sup>, 江佳玥<sup>2</sup>, 余波<sup>3</sup>

(1. 达州职业技术学院, 四川 达州 635001; 2. 重庆工商大学, 重庆 400067;

3. 达州市科学技术情报研究所, 四川 达州 635099)

**摘要:** 信息获取不仅仅是获取静态的数据, 还要获取具有决策引领、揭示未来规律的情报信息, 使个人和社会对信息类型的需求从单一、静止的数据信息转换为跨专业的智慧型复合信息, 提示潜在信息知识单元的智能化、具体化、动态化、复合化、专业化、快速化等特征, 打造开放、协同、包容、共享的科技创新平台。该共享平台的建设是为了打造大数据产业落地应用的丰富场景, 用数学建模的思想来构造四川革命老区大数据共享生态系统的拓扑结构图。

**关键词:** 大数据; 科技资源建设; 共享机制

中图分类号: TP311.1

文献标识码: A

文章编号: 2096-4706 (2020) 15-0076-06

## Research on the Construction and Sharing Mechanism of Big Data Science and Technology Innovation Resources ——Take the Old Revolutionary Base Areas in Sichuan as an Example

JIANG Zhong<sup>1</sup>, JIANG Jiayue<sup>2</sup>, YU Bo<sup>3</sup>

(1. Dazhou Vocational and Technical College, Dazhou 635001, China; 2. Chongqing Technology and Business University, Chongqing

400067, China; 3. Dazhou Institute of Science and Technology Information, Dazhou 635099, China)

**Abstract:** Information acquisition is not only static data, but also decision-making guidance, revealing the future rules of intelligence information, so as to transform the individual and social demand for information from single and static data information to interdisciplinary intelligent composite information presentation, which indicates the characteristics of potential information knowledge unit, such as intelligence, concreteness, dynamic, compound, specialization, speediness. To build an open, collaborative, inclusive and shared platform for scientific and technological innovation. The construction of this sharing platform is to create a rich scene for the application of big data industry. The topological structure of big data sharing ecosystem in old revolutionary base areas of Sichuan is constructed with the idea of mathematical modeling.

**Keywords:** big data; science and technology resources construction; sharing mechanism

## 0 引言

全球数据量变化遵守摩尔定律, 2020年, 其将达到或超过 35 ZB。随着 5G 技术的广泛普及、云计算中心的广泛使用、物联网信息技术的发展和人工智能技术的成熟, 我们将进入数据的 ZB 时代, 传统科技情报信息的研究模式急需变革, 广大用户对海量数据的服务类型和质量有更高的期望, 科技信息的采集、大数据的存贮、数据格式的转换、信息的查找展现、具有人工智能的挖掘和非常规应用都面临着急剧的变化。

为满足互联网时代非结构化数据存储需求而产生了以列为存贮主体的技术, 而大数据技术的产生是为了解决海量数据集分析的问题。大数据内容丰富, 数据之间的关联需要机器学习、人工智能挖掘、统计分析, 形成有意义的数链, 变“废”为“宝”, 使其为科技研究提供了不竭的源泉, 顺应并提升已有的研究思想、方法、策略、功用和平台。大数据技术是信息产业的一次推动时代前进的巨潮。

## 1 研究目的和意义

由于历史、地理、人文等多种因素的叠加, 四川革命老区在文化、经济、社会发展等多方面处于滞后状态, 想在大数据时代变革中异军突起, 就要借助科技的力量, 通过对大数据资源的建设和对共享机制的理论研究, 为四川革命老区的进一步发展提供理论、实践参考, 对老区的综合发展研究起到理论的支撑作用。达州职业技术学院和达州市科学技术情报研究所位于四川革命老区腹地, 重庆工商大学是四川革命老区脱贫的联系单位, 今年是脱贫攻坚决胜之年, 为了给

收稿日期: 2020-07-16

基金项目: 四川省人文社会科学重点研究基地—四川革命老区发展研究中心 2020 年度项目:  
四川革命老区基于大数据的科技创新资源建设与  
共享机制研究 (SLQ2020SD-024), 项目负责人:  
江忠

四川革命老区扶贫、扶智,响应四川革命老区发展中心的号召,申请课题专门研究四川革命老区科技资源建设和共享的问题,为四川革命老区的发展作出职业学院、研究所应有的担当和作为。

过去三十年,拉动中国经济三驾马车是投资、消费、出口,那么未来三十年,拉动中国经济三驾新马车是大数据、云计算、智能制造。在这种大背景下,本研究的意义如下。

(1) 在加快革命老区脱贫致富奔小康的背景下,及时调整四川革命老区发展战略和信息资源建设分布格局,积极应对大数据时代的新形势和新要求。

(2) 优化整合四川革命老区的科技资源。针对四川省革命老区科技资源分布广泛、难以共享且缺乏对已有资源深度分析的问题,提出建立数字化科技资源平台,通过集成分散的科技信息资源实现资源共享。在目前越来越激烈的国际竞争环境中,科技资源也越来越被人们所重视。而要实现科技资源共享、缓解科技资源的有限性与其高效性之间的矛盾,就要促进地区间的协同发展以及科技资源共享,来推动整个革命地区的科技实力,提高地区在国内、国际竞争中的话语权。基于模型设计,本文构建了四川省革命老区科技信息资源平台,实现了数据的集成、统计分析以及动态可视化数据展示、APP应用开发。

(3) 为四川革命老区的改革和发展提供指导方向,进一步改善四川革命老区的发展质量。

## 2 研究的主要思路

(1) 研究大数据资源、集成管理、大数据资源集成以及大数据基础设施建设的基础理论,作为研究主题的理论支撑系统,构建面向四川革命老区大数据资源集成的大数据基础设施建设结构,要按照一定的逻辑方法,逐步、细化地研究“大数据资源”“集成管理”“大数据资源集成”“大数据基础设施建设”“大数据资源共享机制”等专业名词的基本概念和语境意义,作为后续研究的理论支撑。

(2) 研究大数据资源集成服务模式案例,探索面向四川革命老区大数据资源集成的基础设施建设结构的构建思路,综合运用前人的研究成果,对大数据资源集成服务模式案例进行归纳分析,并阐述大数据资源集成与大数据基础设施建设的有机联系。在探索过程中逐步明晰大数据基础设施建设问题的本质,即是大数据基础设施建设的系统性或结构性问题,进而依据构建思路,系统地创建出四川革命老区大数据基础设施建设的集成结构。

(3) 研究有关四川革命老区大数据基础设施建设的发展现状及问题,提出符合当地大数据基础设施建设的保障机制和对策建议。

(4) 通过文献分析法、数理统计法、对比分析法三种研究方法介绍跨区域科技合作的发展现状、科技合作运行机制、存在的问题、跨区域的经验借鉴、优化运行机制的对策等。通过分析跨区域科技资源分布的特点、科技合作的基础、科技合作共享进展情况来进行研究。

(5) 研究传统结构化数据与非结构化数据的区别联系,研究怎样对非结构化数据进行存储,如何把非结构化数

据转换为结构化数据进行存储。对实现物理存储的软件进行比较、总结。

(6) 通过四川省革命老区科技信息资源平台的建设,可以摆脱大数据杂乱涌现的混乱局面,极大地提高了现有科技资源的使用效率。

## 3 大数据资源建设的有力支撑技术

### 3.1 云计算为大数据提供了存储计算加工应用的平台

云计算可更好地开发、使用互联网,可以利用云计算设置了资源消费支付模式,借助互联网提供动态且按需分配的虚拟化资源,是一种更高效、稳定、弹性的模式;科技工作者也可创新资源需求访问、上传、获取、利用多种资源。在大数据资源处理流程中,数据挖掘分析是很重要的环节,体现了大数据的开发价值。作用于大数据的数据挖掘分析利用数学方法与计算机技术算法,实现快速过滤、归并、整合,深度获取潜在的数据模型。大数据分析需要云计算技术作为重要支撑,云计算技术可以为大数据分析提供动态伸缩的、安全的云存储和分布并行计算资源池,也可提供深度应用开发的APP。

### 3.2 机器学习是对四川革命老区大数据加工的利器

数据挖掘是在海量数据库中自动地发现有有用信息的过程,数据挖掘技术用来探查大型数据库,发现先前未知的有用模式。机器学习可以提升数据自动挖掘和筛选的速度。随着大数据时代数据容量的迅猛增加和数据结构联系隐蔽性的提升,如何高效地计算和处理数据,成为机器学习关注的焦点。关于大数据的机器学习要同数学模型结合起来用,采用更加科学高效的数学算法将增强机器学习的效率。

### 3.3 用情报学的思想方法算法来分析提炼大数据

回顾过去,科技资源研究主要是在数据、信息的处理层面上,缺乏引领性、智慧性。情报学是研究信息数据采集、信息数据处理、信息数据分类、信息数据结构化、信息数据建模、情报信息生成、情报信息检索、情报信息智能化服务、情报信息互通以及情报科学系统所包括的概念、理论、公式、技术、规律及方法的专门学科。把数据处理升华为情报学研究,提供“耳目、尖兵、参谋”的决策研究<sup>[1]</sup>。

## 4 大数据背景下科技创新资源的建设

### 4.1 大数据资源运用现状

根据实际状况,现在大数据的实际科研工作及数据加工分析中,运用的操作平台如下所示。

(1) 数据平台:在获取数据的基础上,数据平台的主要功能在于对数据信息进行收集、分类、归并及储存等,这种繁琐性操作却是为数据的下一步分析提供食粮和保证。

(2) 分析平台:数据分析应属于大数据分析中最为具有意义和关键的环节,且是大数据分析中的庞大数据呈现其价值的核心环节<sup>[2]</sup>。

(3) 展示平台:通过技术完成数据分析后,需要通过通用的展示平台对数据源进行运用和互动。展示平台的主要功能是完成大数据分析后的科研、推广及利用。

## 4.2 大数据科技创新资源建设面临的问题

重庆工商大学江佳玥对数据格式研究比较深入,现以四川革命老区为例,总结大数据科技创新资源建设遇到的困难与挑战:

(1) 主观上,很多科技情报研究所、信息技术研究所仍然把提供纸质、电子文档文献和科技动态信息当作传统主要工作。

(2) 缺乏广泛的大数据来源,难以进行业务分析。

(3) 数据往往来自不同的操作系统和硬件平台,缺乏事实上统一的接口和协议标准,根本不能相互适配和通信。

(4) 数据信号受到干扰,其质量参差不齐,数据的完整性、正确性、同一性、低延迟性难以保证,从而导致对其分析加工的结果的合理性较低。

(5) 数据结构没有准确建立,导致不能建立数据分析模型,缺乏高效、有的放矢的人工智能、神经网络算法,跨行业、跨作业的数据分析难以进行。

(6) 在数据应用中不能保证信息安全,隐私可能被泄露,有可能与现行的法律法规冲突;数据失真、网络数据、驱动攻击对资源建设不可不防。大数据科技创新资源如果能成为众多行业、单位和个人的核心、关键资产,必须增强其可靠性、安全性。

## 4.3 四川革命老区基于大数据科技创新资源的建设构想

### 4.3.1 大数据科技创新资源建设模型

四川达州职业技术学院江忠在四川革命老区课题研究中,通过反思四川革命老区传统大数据资源的建设和利用过程以及大数据时代的需求,构造一个模型来说明大数据科技创新资源的建设,如图 1 所示。

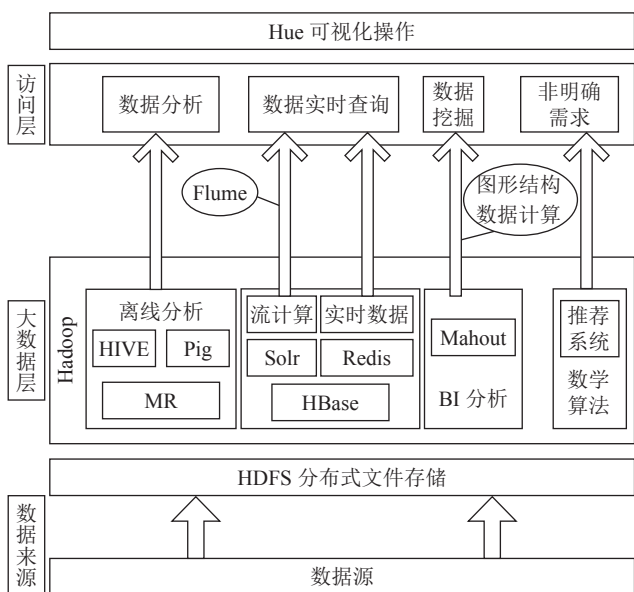


图 1 大数据科技创新资源建设模型

对海量数据放在硬盘中用 MR (MapReduce) 进行编程处理, Pig 是 MR 的一个抽象,它是一个工具/平台,用于分析较大的数据集,并将它们表示为数据流。HIVE 是基于 Hadoop 的一个数据仓库工具,可以将结构化的数据文件

映射为一张数据库表,并提供完整的 SQL 查询功能,可以将 SQL 语句转换为 MR 任务进行运行。Hadoop 分布式文件系统 (HDFS) 被设计成适合运行在通用硬件 (commodity hardware) 上的分布式文件系统。它和现有的分布式文件系统的区别也是明显的。HDFS 是一个高度容错性的系统,适合部署在廉价的机器上。

把大数据分为实时数据和非实时响应数据,实时数据就需要流计算来完成操作,非实时数据可以通过 MR 算法来处理。大量数据需要用 BI 分析、图形结构数据计算来完成数据的挖掘,找出有价值的结论和情报信息,对社会各行各业决策提供科学可靠的依据。BI (Business Intelligence) 是指业务智能或智能业务,其用途是使用现代数据仓库技术、在线分析处理技术以及数据挖掘和数据展示技术进行数据分析以实现业务价值。对未来一些潜在和不明确的需求,就需要用数学算法来构建一个数学模型,按照数学模型对大数据进行分析加工、用户相似度计算:如使用协同过滤算法得出非明确需求,由此可构造一个推荐系统。推荐方法包括基于统计的推荐、基于内容的推荐、专家推荐、协同过滤推荐和混合推荐。推荐系统的应用之一是构建一个全新的商业和经济模式。

### 4.3.2 数据采集技术

收集数据是数据挖掘、实时流处理、建立数据库、数据仓库、MR、Spark、人工编程及智能分析的前奏,高效、准确的数据采集方案对大数据挖掘研究具有重要意义。当前,不同行业有各自的数据采集工具和方法。科技资源对数据处理的安全性要求很高,可以加强有关企业、事业单位和科学研究机构的联系,统一系统接口,规范数据格式和传输协议,用加密软件相互传输数据。大数据采集可通过网络爬虫等方法,也可以通过传感器、天网等硬件系统来采集数据。推荐应用软件来采集,即是要求建立统一、规范、高效、兼容性较强的数据采集系统,使采集到的各种类型数据和数据库软件接口能够相互匹配并自动转换,并能同建立的四川革命老区大数据科技创新资源系统的各种类型的数据库通信。

### 4.3.3 数据存储技术

科技创新资源系统就是用数据库把收集到的数据按要求存储起来,根据需要建立关系型和非关系型的数据库,根据要求来进行后续的分析管理和使用。随着 SQL 数据和 NoSQL 数据量的连续暴增,以及大数据的容量大、类型多样、价值密度低、通用性差等特点,要求本系统有大容量存储设备、快速读写性能以及安全可靠和弹性可伸缩的特点,此前传统关系型数据存储系统的设计显然不能满足大数据 APP 的要求。目前大数据存储主要通过分布式文件系统、关系型数据库技术、NoSQL 数据库技术、云数据库、键值数据库、文件型数据库等实现。

### 4.3.4 大数据技术与人工智能技术相互共生

四川革命老区科技创新资源的利用和建设需要使用大数据技术,大数据技术的发展催生了 Pig、HIVE、Impala、Flume、Mahout、Hama 等技术的诞生与成长。使用了数学



算法的人工智能、机器学习与大数据分析挖掘技术相互依赖、相得益彰。

#### 4.3.5 数据挖掘

沃尔玛著名的“啤酒与尿布”营销案例,就是用了数据挖掘技术。在数据挖掘中建议使用分类法、回归分析法和Web数据挖掘法,学会使用R语言、SPSS、SAS软件对数据进行操作、计算、分析。

#### 4.3.6 发挥情报在科学研究中的引领作用

云计算和大数据发展下的情报分析本质还是数据结构的逻辑关系分析,互联网Web 2.0技术的出现,使情报呈现的方式多元化,数据挖掘技术显然使情报展现能力大幅提升,情报价值得到彰显。用这种结构化数据的情报仍然建筑在数据的隐性关系上,缺少人的智慧,分析出情报缺乏对未来的智能预测。随着神经系统、人工智能的进一步研究,按世界最强大脑的思维与理性数据分析沟通联系,就可以排除数据噪音,找到数据临界点,发现环境影响数据的关键因素,这样对情报的分析得更具有科学性、严谨性。

#### 4.3.7 对各种类型资源的加工在APP集成平台上进行

当今数据,如Email、DOC文档、路由交换机信息、医院病人记录、企业管理信息,再延伸到网页、社交互动媒体、分析数据,涵盖音视频、文字、图形、图像等等<sup>[3]</sup>,这些信息缺乏联系,不能建立关联,隐性的规律不易被发现,无法体现其价值,就成了无用的数据;若没有专门的APP和实用的研究系统,就无法实现海量数据的存储、应用、加工、分析等。

### 5 关于四川革命老区科技创新资源的共享研究

目前科学创新手段技术日新月异,大数据需要网络和云计算技术的加工,科技信息量巨大且数据规律紊乱,个人和社会对服务要求更多、更高,5G技术为核心的新一代无线通信网络以及当今移动终端的普及,使得各种软硬件标准不一致,数据格式不区别,这些让资源的传播与共享形成天然的壁垒。

#### 5.1 目前四川革命老区大数据资源共享的壁垒

四川达州市科技情报研究所余波长期工作在四川革命老区科技第一线,对大数据资源共享出现问题总结如下:

(1) 数据不连续,各行各业数据类型繁多。多元关联的大数据起粘合作用,但由于各个部门分享共用数据的技术标准限制形成了“数据孤岛”和“数据烟囱”,降低了大数据产业资源配置效率。

(2) 核心技术缺失。大数据领域整体性、平台级核心技术的创新比较少见。大数据乃人工智能的基础,以大数据和机器学习为基础的深度学习算法等人工智能的核心技术需要突破。

(3) 各地大数据和数字经济重复性发展和布局,一般存在重存储轻应用的状态。由于缺乏统一的大数据发展规划和运行监测系统,各地大数据产业的功能相似,浪费资源。由于部分地区信息化需求不旺,大数据应用单调,仅仅建设1个大数据中心或云中心等基础配置,后期管理和开发跟不上,

势必资源浪费。

(4) 大数据缺乏系统和一致性的标准。基础性标准、数据加工标准、数据安全性标准、产品和平台开放标准、数据收集标准、二次应用开发和科研服务标准等大数据标准体系亟需规范。

(5) 数据安全性意识淡漠,对大数据科技资源的共享发展造成危险,并且没有预防措施。共享使数据、关键信息和大数据创新资源共享平台等面临不可预料的威胁与风险。应加强保护利用大数据技术对海量数据进行挖掘分析所得到的结果,这些结果可能包含涉及国家政治、经济、社会、科技、商业与军事等各方面的敏感信息,需要对研究结果的共享和发布加强安全管理,打造监、管、防三位一体的大数据科技资源安全综合体系。

#### 5.2 对四川革命老区科技创新资源的共享提出一些建议与思考

##### 5.2.1 增强原始创新能力

四川革命老区自主创新能力比较薄弱,特别是关键核心技术受制于人的局面尚未实质性改变。只有加强前瞻性基础研究,提升原始创新能力,才能把握竞争和发展的主动权。加大研发投入,以基础性研究的突破带动引领性原创成果、战略性技术产品的重大突破。形成事实上的行业标准,获得更多的话语权,统一大数据的格式、标准和协议。

##### 5.2.2 使用虚拟化技术创造更高的科技效率

云平台提供基于IaaS、PaaS、SaaS的服务,即不需要自己创建每一个应用。其就像一个大型的网络资源仓库,APP可以在共享资源平台上找到想要的东西,可以用Python等来进行二次开发并发布在APP引擎上,允许物理实体机或终端同时运行多个操作系统,并且应用程序都在独立的内存空间上运行,彼此独立,从而显著降低费用、提升易用性,达到科技资源充分共享的目的。

##### 5.2.3 网络端口镜像分析提供科学决策共享大数据资源

在互联网主干网上的Router、FireWall上做多个端口的Ghost镜像,可以抓取经过该区段的主要Packet Stream,然后由人工智能的分析软件对需要的重要网站、路由器接口、防火墙或者IPS数据包进行收集、分类、提取、分析,找到网络上的有价值的流处理信息,为科技创新资源建设的决策起辅助作用。

##### 5.2.4 对传统科技情报信息业务进行深加工

运用人工智能、云技术的大数据科技创新共享平台有如下优势:

(1) 借助物联网、科技创新资源等技术和数据源可以掌握更多的实时数据和静止数据。

(2) 获得的信息比其他方式更快捷、更客观、更全面、更具可信度。

(3) Flume能够快速反应和处理实时数据,对服务对象及时干预。

(4) 人工智能、机器学习、神经算法判断和决策的正确率较高。

### 5.2.5 构建 NoSQL 数据存储和关系型数据库的联合共享系统平台

初级的海量的数据用面向列存储的 NoSQL 数据库存储, 然后用形如 PowerDrill 技术, 把海量数据进行查询分析并转化为结构化数据, 再用数据挖掘技术、机器学习工具把这些信息转换为对科学有引领性的情报, 再把这些情报知识进行共享。

### 5.2.6 打造开放、协同、包容、共享的科技创新平台

由于不同行业部门的数据格式千差万别, 可以用 XML 技术作为存储交换统一类型的语言; XML 不容易显示多种格式, 其可视性差, 可以用 ASP.NET、JSP 等技术开发个性化的桌面界面。

用 ASP.NET、JSP 等技术可把各种关系型的数据库 SQL Server、ORACLE 等的数据库提取出来, 用 XML 语言转换。其不同的驱动引擎屏蔽在连接程序中, 对用户是透明的。这样的 MVC 模型可以打造一个开发、共享的科技创新平台。

### 5.2.7 用开源云技术 OpenStack 和开源大数据技术 Hadoop 构建共享平台

开源技术为四川革命老区提供了经济、方便的选择, OpenStack 既是一个社区, 也是一个项目、一个开源软件、一个提供了部署云的操作平台或工具集。使用 OpenStack 易于构建虚拟计算或存储服务的云, 既可以为公有云、私有云,

也可以为大云、小云提供可扩展的、灵活的云计算。

这两项技术都是开源的、免费的, 全世界通用, 并且生命力旺盛, 用这样的技术来构建四川革命老区科技创新资源平台通用性强, 与最先进的 Google 云有相似之处, 能够少走弯路, 适合人数少的科研团体进行开发和二次开发。

### 5.2.8 加入云平台增强数据的安全性

在大数据创新资源被广泛应用共享的过程中, 部分数据不完整、脏数据、欺骗攻击、数据传输超时长等信息安全问题也随之而来, 因此为了使大数据科技创新资源能够为社会所用, 真正意义上推动共享型数据和情报技术的健康发展, 就必须加强数据、软硬件在信息安全方面的研究, 从而最终推动大数据分析科技创新平台的实际应用。

## 6 构建四川革命老区基于大数据科技创新资源的共享生态系统

### 6.1 四川革命老区基于大数据科技创新资源的共享生态系统的构成

如图 2 所示, 四川革命老区基于大数据科技创新资源的共享生态系统大致有以下组成部分: 主服务器 (Master Server)、名称服务器 (NameServer)、区域服务器 (Chunk Server)、数据库、数据仓库、前端应用工具和访问接口、大数据采集系统、代理访问服务器 (Proxy Server)、用户等。

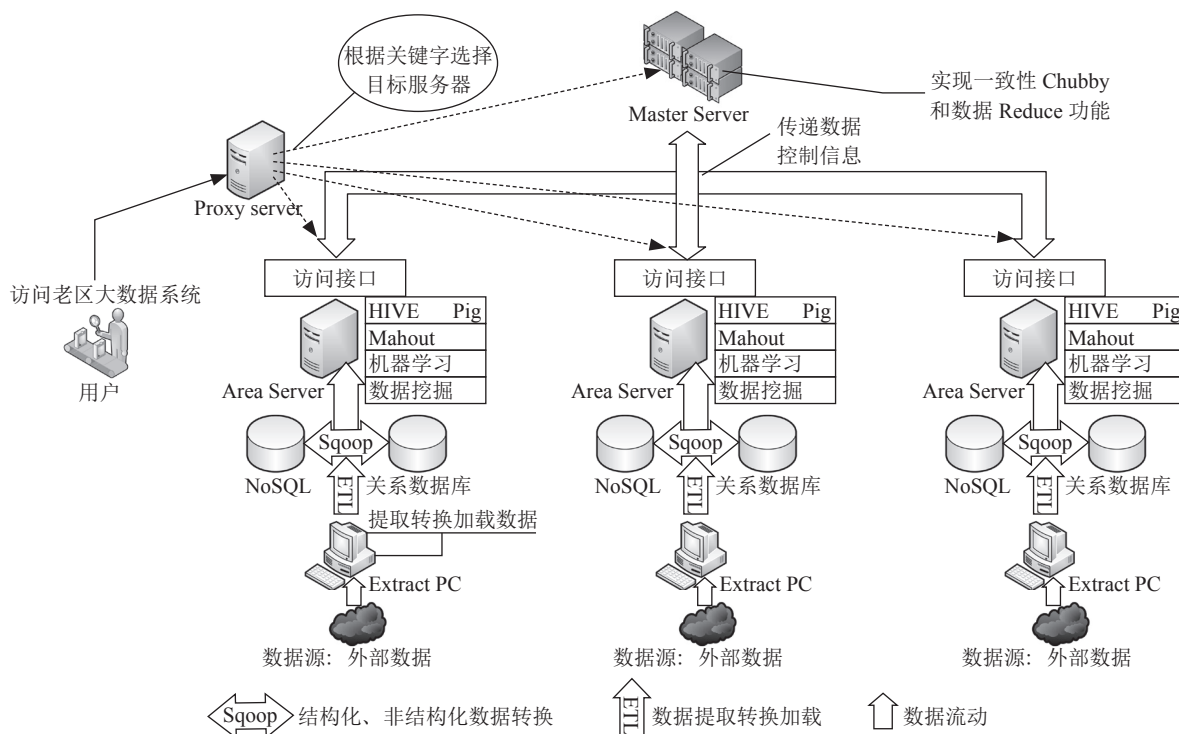


图 2 四川革命老区基于大数据科技创新资源的共享生态系统拓扑结构图

### 6.2 四川革命老区基于大数据科技创新资源的共享生态系统运行机制原理

如果把需要采集外部数据来源定义为  $X$ , 那么四川革命老区的各个市县的数据为  $X_i$ , 由于这些数据是随机的, 相当于执行了函数  $Shuffle(X) = \{X_1, X_2, X_3, \dots\}$ ; 主服务器

把一项任务用函数  $f(x)$  来分发给每一个区域服务器,  $map(f(x), [X_1, X_2, X_3, \dots])$ ; 经过每个区域服务器的运算, 把每个关键字  $key_i$  进行归并求总运算, 把这些结果归并到主服务器, 相当于执行了函数  $Reduce(key_i, [Value_1, Value_2, Value_3, \dots])$ ; 外部数据源  $X_i$  经过大 (下转 85 页)

学生可以通过微信公众平台推送的选课信息了解实时的课程开设信息,也可以在平台的“课程库”中查阅按照不同性质的学分分类的待选课程信息,结合本人已修过的学分的情况选择报名兴趣课程,并可在“我的”“已报名课程信息”中浏览已报名的课程信息。

### 3.4 教务处员工功能界面

教务处员工可通过该系统查看待审核的开课申请,按照学校的相关审核要求进行审核,若审核未通过,则注明未通过原因并退送至开课教师系统模块;若审核通过,后台系统则自动将该课程的相关选课信息在选课前一周推送给学生,供学生浏览、选择、报名。

## 4 结 论

基于微信公众平台的高校学生线上选修课程选课信息系统从现阶段高校学生选课的实际出发,在分析归纳三类不同用户需求的基础上,根据其不同的需求设计系统的三大核心功能模块,并对后台数据库做了详细规划设计。本系统原型的设计和实现,有效地缓解了高校有限的教务处员工处

理庞杂的学生选课信息的压力,能够为学生精准选课、教师自助申请授课和教务处员工轻松管理提供一个简单、方便的平台,提高了高校日常教学工作的效率。

### 参考文献:

- [1] 张晓芬, 马楠, 王世华. 基于 Android 平台高校移动选课系统的开发与设计 [C]// 中国计算机用户协会网络应用分会 2014 年第十八届网络新技术与应用年会. 2014: 98-101.
- [2] 王书森. 大学生选课管理系统的设计与实现 [D]. 天津: 天津大学, 2017.
- [3] 贺维彬. 基于 MySQL 数据库的在线选课系统 [J]. 电脑知识与技术, 2014, 10 (5): 883-885.
- [4] 沈涛, 杨露. 基于 WeChat 公众平台的大学选课系统设计与实现 [J]. 信息与电脑 (理论版), 2019 (8): 91-93.
- [5] 赵波. 选课管理系统的设计与开发 [J]. 云南民族学院学报 (自然科学版), 2002 (4): 245-248.
- [6] 邓万友, 孟雅杰, 贾焕军, 等. 学分制下网络选课系统的构想与实现 [J]. 黑龙江高教研究, 2001 (1): 92-94.

作者简介: 张松 (1980—), 男, 汉族, 安徽宿州人, 讲师, 博士, 研究方向: 信息资源管理、管理信息系统。

(上接 80 页) 数据系统采集、加工、转换, 相当于执行了函数  $ETL(X_i)$ , 把结果在存储为 NoSQL 数据表, 可以通过 Sqoop 工具将其转换为关系型的数据表; 用户提交查找的关键字信息  $key$ , 通过路径选择服务器来实现路由, 相当于通过  $Proxy(key)$  函数得到访问的 IP 地址, 可以均衡负载, 加快访问速度。

### 6.3 四川革命老区基于大数据科技创新资源的共享生态系统技术支持

根据四川革命老区的区情, 建议选择 Hadoop 开源大数据技术, 该技术成熟、使用面广; 同时选择两种数据库, 非关系型的数据库为科学研究带来快捷高效, 关系型数据库满足传统经典的数据服务; 对大数据加工一般采用 MR 技术, 这种技术是对磁盘的数据进行操作, 速度较慢, 可对海量数据进行加工; 如果采用基于内存的大数据处理技术 Spark 会对系统的速度有很大提升, 最好同时采用两种技术。

## 7 结 论

进入大数据时代的科学研究已经数据模型化, 向数据分析、挖掘转移, 对于仅擅长文献、信息、数据的收集、分类、汇总、分析是不够的, 要提升服务于决策支持的情报科学水

平, 就必须建设科技创新共享资源平台, 刻不容缓。

四川革命老区各地在使用信息和加工数据的过程中, 对软硬件资源、技术分享、市场的需求可能存在共同点。市县科学技术信息研究机构的责任划分不明确, 缺乏有效的责任分担机制, 也缺乏有效的沟通、协同、分享和统筹机制。四川革命老区各市州科技信息资源协同创新发展路径还不明确, 没有形成整体规划, 需促进形成统一标准的四川革命老区大数据资源开发共享协议。

### 参考文献:

- [1] 刘彤, 蒋继娅, 吴素研. 科技情报与信息技术 [M]. 北京: 北京科学技术出版社, 2010.
- [2] 赖茂生, 赵丹群, 韩圣龙, 等. 计算机情报检索: 第 2 版 [M]. 北京: 北京大学出版社, 2006.
- [3] 唐国纯. 云计算及应用 [M]. 北京: 清华大学出版社, 2015.

作者简介: 江忠 (1966—), 男, 汉族, 四川渠县人, 副教授, 本科, 理学学士, 研究方向: 高等数学、初等数学教育、信息安全及云计算等; 通讯作者: 江佳玥 (1998—), 女, 汉族, 四川达州人, 初级会计师, 管理学在读研究生, 研究方向: 财务管理; 余波 (1972—), 女, 汉族, 四川渠县人, 副研究员, 本科, 研究方向: 科技情报信息研究。