

国家数字图书馆特色资源云平台的建设与实践*

童忠勇

摘要 本文在分析公共图书馆特色数字资源建设、互联网技术发展和相关国家政策文件的基础上,阐述了国家数字图书馆特色资源云平台的建设思路,并基于数字图书馆工程和数字图书馆推广工程的建设成果,利用云计算、大数据、唯一标识符以及关联数据等技术勾勒出平台的总体框架和核心功能,构建了一套基于各省市图书馆优势、面向社会大众、共享开放的特色资源云服务体系,旨在更好地发挥地方特色资源的价值,为全国用户提供精准化、个性化、均等化、公益化的资源服务。图2。参考文献19。

关键词 国家数字图书馆 特色资源 云平台

The Construction and Practice on Cloud Platform of Characteristic Resource in the National Digital Library

Tong Zhongyong

Abstract: After analyzing the current status of characteristic resource construction in public libraries, internet technology development and relevant national policies, this paper expounds the construction idea of the characteristic resource cloud platform for the National Digital Library. The main framework and core functions are built on the basis of the Digital Library Project and the Digital Library Promotion Project. Multiple technologies including cloud computing, big data, unique identifier and linked data are applied in this characteristic resource cloud service system. Characterized services with high quality from various province and city libraries will be able to be open to nationwide readers and will be precisely personalized. Through this platform, the equality and commonweal of public libraries can also be reached. 2 figs. 19 refs.

Keywords: National Digital Library; Characteristic Resource; Cloud Platform

由国家图书馆牵头、各级公共图书馆通力协作的国家数字图书馆已经形成,全国大部分省、市级图书馆都已建立具有地方特色的专题数据库^[1]。截至2017年底,国家图书馆汇聚各省、市图书馆数字资源元数据超过3亿条^[2]。然而,图书馆特色数字资源量的快速增长以及特色资源服务区域化、平台异构化、技术多样化等问题的出现,给读者检索、使用带来了很大的不便^[3]。本文将重点探讨国家数字图书馆特色资源云平台的建设,旨在通过对公共图书馆特色数字资源

的整合和共享,让读者更好地利用图书馆特色信息资源,不断提升公共图书馆馆藏特色数字资源的利用率。

1 国家数字图书馆特色资源云平台建设背景

2011年,数字图书馆推广工程^[4]正式启动,随着服务器等硬件设备的统一配置、应用系统的统一搭建以及专网和虚拟网的同步铺设,各省、市级图书馆资源的加工能力和服务能力逐步加强,

* 本文系国家社会科学基金青年项目“公共图书馆线上线下服务融合研究”(项目编号:16CTQ004)研究成果之一。

并纷纷开展特色资源的数字化工作,极大地丰富了地方图书馆数字化馆藏建设^[5]。2015年,数字图书馆推广工程资源联建项目正式启动,全国各级图书馆充分结合地方特色,依托联建项目建设了一批内容丰富、形式多样的优秀数字资源^[6,7]。

图书馆界高度关注互联网技术的发展和运用。笔者于2018年7月以CNKI为数据源,以“云计算”和“图书馆”为主题词进行检索,重点关注近三年在核心期刊发表的约50篇相关文献,发现图书馆界有关云计算理论的研究成果多集中于云计算在图书馆应用中的模型构建^[8]、数据整合^[9]、资源建设^[10]和服务创新^[11]等方面。与此同时,在云计算技术的实际应用层面,公共文化行业也开展了实质性建设,为国家数字图书馆特色资源云平台的建设提供了参考。如:国家公共文化云于2017年11月29日正式开通,统筹整合文化共享工程、数字图书馆推广工程、公共电子阅览室建设计划三大文化惠民工程,实现了资源共享和云化服务^[12]。另外,随着大数据技术的成熟和应用,图书馆界有关大数据的研究热点集中于大数据环境下图书馆服务体系的创新和服务平台的构建、基于大数据的资源整合和可视化分析以及大数据挖掘视角下的智慧化服务等方面^[13,14],国家数字图书馆特色资源云平台可基于大数据技术,根据读者行为提供定制化、智慧化和个性化服务;近年来图书馆界对关联数据的研究也从理论层面走向应用层面,如上海图书馆利用关联开放数据技术重组图书馆传统资源,构建历史文献数据服务平台^[15],为国家数字图书馆进行特色资源深层次的组织和整合提供了借鉴;移动互联网作为空前广阔的融合发展领域,与广泛的技术和产业相关联,正从技术驱动发展到需求驱动阶段^[16],在此背景下国家数字图书馆应以用户需求为导向,依托移动互联技术促进移动服务泛在化。云计算、大数据、关联数据以及移动互联网等新技术的快速发展和成功应用,为国家数字图书馆特色资源云平台的建设提供了技术保障。

《国家“十三五”时期文化发展改革规划纲

要》指出要“运用云计算、人工智能、物联网等科技成果,催生新型文化业态”^[17];《文化部“十三五”时期文化产业发展规划》也指出要“大力培育基于大数据、云计算、物联网、人工智能等新技术的新型文化业态,形成文化产业新的增长点”^[18]。基于互联网思维的服务转型已成为国家数字图书馆发展的必然趋势。

在图书馆特色数字资源建设、互联网技术发展以及相关国家政策等多重因素的驱动下,国家数字图书馆于2018年初正式开展特色资源云平台建设,依托各图书馆在地方特色资源收集、整理方面的专业性和优势,利用数字图书馆工程和数字图书馆推广工程累积的技术和经验,统筹规划,采用科学标准的规范体系,加强知识层面的深层次关联和整合,构建出一套共享、开放、各图书馆共同参与的数字图书馆特色资源云服务体系。

2 国家数字图书馆特色资源云平台设计

2.1 总体设计

国家数字图书馆特色资源云平台致力于地方文献的数据挖掘和品牌包装,将分散在各级公共图书馆的资源按一定的知识管理规则和服务目的进行二次组织,并针对各类人群提供多种多样的资源服务。具体来说,可以从数据、用户、系统、服务四个层面对公共图书馆特色资源进行整合:数据层采用统一的元数据描述框架进行加工和制作,构建公共图书馆特色资源数据池,采用混合云模式进行数据存储和组织,通过唯一标识符技术进行资源的统一调度;用户层通过统一用户管理系统对公共图书馆的用户进行实名认证和单点登录,打通各级图书馆之间的用户数据;系统层采用Web Service等框架,标准化集成与资源相关的应用程序和服务;服务层应注重关联化、全媒体化、公益化和互动化,依托技术手段将各种公共数字文化资源、可利用的商业资源融合到公共图书馆特色资源的服务体系中来。依托国家数字图书馆特色资源云平台,各图书馆可通过定制化、组

件化的功能模块搭建,遴选资源,快速构建本地的特色资源服务。

2.2 框架设计

国家数字图书馆特色资源云平台采用“云+端”的宏观框架,即云计算概念加上互联网化的智能终端。云端通过构建连接公共图书馆和各地读者的空间关系网络,实现集图书馆员资源上传、组织及发布,读者在线阅读、视频观看及科研人员学习、研究为一体的统一开放平台。国家图书馆作为云端的实施主体,搭建云资源发布平台,汇聚各省、市图书馆特色资源,为各图书馆的资源发布和线上服务提供技术支撑,实现公共图书馆特色资源的云化和大数据化,缩小地方特色数字资源建设和服务的鸿沟,减少公共图书馆管理和维护的成本,提供安全的运行保障机制。

平台基本架构(见图1)描述如下:底层搭建云计算基础环境,构建 LINUX 集群,提供弹性计算服务,计算资源在集群内的弹性伸缩,并在多个云服务器之间均衡请求分发。在云基础环境之上,搭建云数据中心,各图书馆上传的数据均汇聚在此,可进行资源的二次组织、聚类及分析。在应用层面,以分布式文件系统、任务调度系统为核心,基于弹性计算服务、关系型数据库、开放存储服务、开放结构化数据服务以及开放数据处理服务,为各图书馆提供各类特色资源的储存及发布服务,同时在云上为各图书馆开放统一的资源检索引擎。在服务层面,面向读者提供集 WEB 端、移动端和电视端为一体的全终端、跨平台的终端揭示服务。

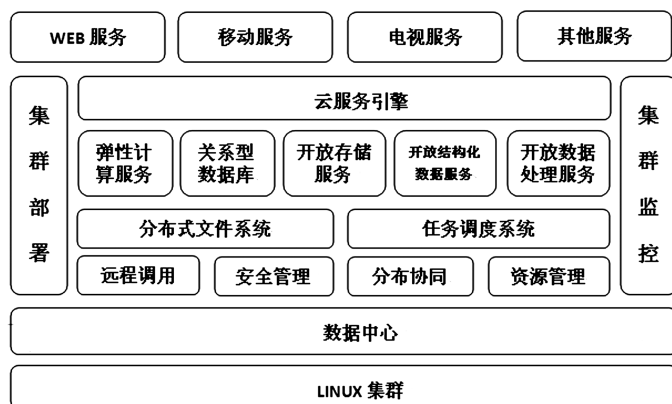


图1 国家数字图书馆特色资源云平台架构

2.3 技术架构

在技术架构上,由底层云计算、云存储、云缓存以及云数据库构建的 IaaS 层(基础设施即服务)为平台的安全稳定运行提供基础保障,各省、市图书馆可在云上动态获取资源。PaaS 层(平台即服务)通过对特色资源的数据采集、组织、发布、调度、搜索、分析、管理以及用户身份认证、工作流处理、短信提示及报警等服务进行统一管理,建立开放合作的机制,面向各图书馆提供标准的接口规范,为应用服务提供支撑。SaaS 层

(软件即服务)面向终端读者,组合 PaaS 层核心服务,搭建公共图书馆特色资源服务门户。各图书馆可在云上开展本地专题库等云应用服务,建设包括地方志、民国资料、家谱、老照片等在内的多种特色文献,采用唯一标识符统一调度资源,将内容经由云平台上传、组织和发布;社会公众可以使用云搜索功能,随时查找所需的资源。与此同时,通过对用户、业务及操作的大数据分析,提供智能化推荐服务,基于云门户实现跨终端、跨平台的应用。具体框架如图2所示。

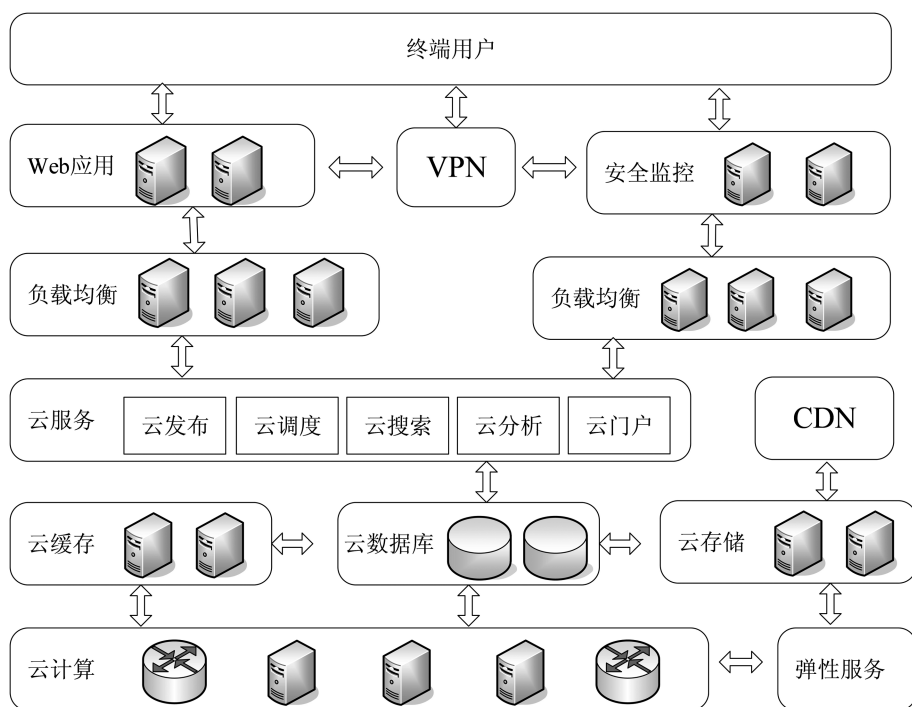


图2 国家数字图书馆特色资源云平台技术架构

3 国家数字图书馆特色资源云平台核心功能

3.1 弹性计算

弹性计算服务作为一种基础云计算服务,提供资源的按需供给和动态管理^[19]。公共图书馆一方面可基于数字图书馆推广工程建设成果,依托推广工程专网构建公共图书馆特色资源服务私有云;另一方面也可以从云服务提供商处采购成熟的公有云服务,如阿里云、百度云、腾讯云等,组建公共图书馆特色资源混合云服务体系,向各成员馆提供安全、稳定、资源可伸缩的弹性计算云服务。计算资源包括云服务器、云数据库、云存储以及负载均衡等。各图书馆可根据自身业务需求,综合考虑自身安全及利益,灵活租用计算资源。基于云计算环境,各级图书馆不再需要搭建服务器及相应的网络等运行环境,在云上可以快速地搭建本地服务,减轻地方图书馆应用系统运维的压力,将工作重心更多地放到读者服务上。

3.2 数据管理

公共图书馆特色资源元数据基本上采用 DC、MARC、ACCESS、XML 等结构化数据来描述,因此在数据管理上需要构建云关系型数据库,按照资源的数据结构来组织、存储和管理数据,实现云上数据库管理,灵活揭示公共图书馆特色资源;同时提供 SQL 窗口、SQL 命令行等便捷操作,支持库表结构的可视化管理,直观地对业务数据进行增、删、改、查的实时监控,减轻公共图书馆对本地数据库维护的压力,使各图书馆更加方便快捷地对地方特色资源进行集中管理和统一服务。

3.3 云存储

随着计算机技术的发展,各图书馆特色资源的数据加工质量越来越高,对大容量存储设备的需求也越来越强烈。各图书馆可根据自身服务需要,灵活申请存储资源,调整服务方式,对安全性、隐私性有较高要求的数据可存放在私有云中;反之,则放入公有云中。国家数字图书馆特色资源

云平台在云存储的设计上,一方面要考虑到资源类型的多样化,包括图书、期刊、报纸、学位论文等文本类资源,甲骨、老照片等图片类资源,政府信息类的网页采集资源,以及音、视频等流媒体资源等;另一方面国家数字图书馆云存储不仅仅是存储,还应该考虑应用层面,云存储应提供资源格式的在线和离线转换或处理,文本类资源支持 TXT、WORD 等转换成 PDF 格式,图片类资源支持 JPG、PNG、TIFF、GIF 等多种格式的转换,流媒体资源支持 FLV、WMV、MPG、AVI 等转换成 MP4 格式。为保证数据安全,格式转换还应提供加水印功能,服务层面应具备灵活的鉴权、授权机制以及防盗链等功能。此外,还应提供标准的 API 接口、SDK 包及客户端工具,可以方便地对 Web 网站或移动应用终端的海量数据进行上传、下载、检索、处理和管理。

3.4 云搜索

云计算环境下,国家数字图书馆提供一种开放的云搜索服务,面向公共图书馆开放搜索引擎,为公共图书馆特色资源服务提供简单、高效、稳定和可扩展的搜索解决方案。国家图书馆充分利用技术优势、人员优势和管理优势,基于云关系型数据库构建开放的搜索服务,在数据层面支持 RDS 数据源的灵活导入以及数据结构和数据处理的自由定制,在服务层面支持在网页中以插件形式灵活嵌入搜索引擎,从而将搜索技术简单化、低门槛化、低成本化和云服务化。各图书馆基于云搜索引擎可指定资源范围进行多条件搜索,并可进行搜索排序、统计分析等操作,使搜索引擎技术不再成为图书馆特色资源服务的业务瓶颈。各图书馆通过在云搜索平台直接配置相应的数据源,即可方便快捷地实现数据的统一搜索。

3.5 云加速

为保证资源内容传输得更快、更稳定,国家数字图书馆特色资源云平台采用内容分发网络(Content Delivery Network, CDN)进行加速,尽可能避开互联网上影响数据传输速度和稳定性的瓶颈和环节。平台租用第三方成熟的 CDN 服务,包括网页加速、大文件加速、视频加速等,解决网

络拥挤的状况,提高用户访问的响应速度,使用户无论身处何地都能快速稳定地使用公共图书馆特色资源服务。

3.6 云分析

云分析服务主要是通过对公共图书馆特色资源云服务平台的业务数据和用户数据进行大数据分析,将结果反馈给用户,进而提升平台的易用性和吸引力。云分析服务通过全程跟踪读者访问数据,如浏览量、访客数、跳出率等,帮助各图书馆及时了解特色资源的服务效果,优化平台服务策略,提升用户体验。

4 应用展望

随着数字图书馆推广工程建设的深入,国家数字图书馆特色资源云平台将带动各省、市图书馆特色资源服务模式的创新和服务能力的提升;利用其灵活的部署方式和公开透明的服务资源,各地方图书馆可以更好地把握行业动态,提升竞争力;同时也有利于发挥各地图书馆的群体力量,推进特色资源的共建共享,减少重复加工,实现优势互补,提高图书馆特色资源的利用率。

4.1 创新服务模式

国家数字图书馆特色资源云平台采用“中心—地方”云架构模式,各地方图书馆采用“软件租用”的方式实现数字资源云发布。在发布内容上,不仅支持国家图书馆自建数字资源、征集资源和共建资源等所有裸数据的发布,还支持地方各级图书馆特色资源的发布,按地域、机构、资源类型、接入网络、用户角色实现资源的访问权限控制,并通过汇总地方图书馆提供的元数据来实现中心云平台对外提供资源的统一调度。在发布形式上,不仅支持地方图书馆异地资源发布,将对象数据导入到国家数字图书馆对象资源总库,还支持地方图书馆本地资源发布,以分布式对象数据存储的形式进行资源组织。在专题库建设上,各图书馆可对平台内的数据进行抽取和二次组织,在平台上创建适合本馆服务的资源库,面向读者提供互联网或推广工程专网服务。在读者服务上,地方图书馆向平台云中心进行数字资源与用

户行为数据的汇聚,形成以国家图书馆为中心的数据池和行为池,借助大数据等先进技术面向读者提供智慧化服务。

4.2 提升服务能力

国家数字图书馆特色资源云平台建成后,为推广工程各级图书馆提供 SaaS 级的云服务。在资源整合方面,平台基于唯一标识符技术对公共图书馆的资源统一调度。由于使用统一的系统、统一的标准、统一的数据格式,各图书馆之间的数据传递也将变得简单而及时,各级图书馆可以快速搭建本地专题资源库,提供读者服务。此外,目前国家图书馆已经实现基于唯一标识符技术的电子资源与纸本资源的关联,借助国家数字图书馆特色资源云平台,未来可以将这一成果进一步推广应用到各级图书馆,再与馆际互借平台相结合,可使读者方便、快速地获取相应电子资源的纸本图书。在资源服务方面,平台集成了云搜索引擎技术,将满足大用户量、大并发量的检索、聚类等业务应用需求,可实现资源的篇名、目录、摘要及全文检索,为各图书馆专题库的构建提供系统支撑。在多终端支持方面,平台不仅提供 PC 端的展示,还支持手机、PAD 等移动设备以及电子触摸屏等交互设备的展示。

4.3 灵活的部署方式

在部署模式上,国家数字图书馆特色资源云平台将面向推广工程各级图书馆提供三种部署方式,分别是标准版、增强版和高级版。标准版部署独立的发布模块,提供数字资源发布、权限控制、资源检索等基本功能,作为独立的数据管理后台需借助第三方系统进行数据集成;增强版在标准版的基础上增加部署展示模块,可对外提供独立的资源服务,适用于简单的资源揭示场景;高级版在标准版的基础上部署资源门户模块,提供资源的逻辑库展示及专题库模板定制、用户交互等高级功能,适用于大型数据库的场景应用。

5 总结

国家数字图书馆特色资源云服务体系的构建是一项全方位的系统性工作,不仅是技术发展的

趋势,更是资源管理和读者服务的需要。国家数字图书馆将进一步以优势技术和手段打破不同图书馆、不同类型特色资源之间的界限,树立跨区域资源整合的意识,将分散异构的特色数字资源纳入到云体系中,使公共图书馆中各类特色资源关联起来,为社会公众提供更丰富的数字文化服务。

参考文献

- 1 魏大威,董晓莉.利用区块链技术驱动国家数字图书馆创新升级[J].图书馆理论与实践,2018(5):98-103.
- 2 张福俊,等.国内外数字资源发现平台比较研究——以 Trove、DPLA、Europeana、WDL 和文津搜索为例[J].国家图书馆学刊,2018(1):86-96.
- 3 吴高.地方公共数字文化特色资源建设现状调查与思考——以全国文化信息资源共享工程省级分中心特色数字资源建设为例[J].图书馆建设,2016(1):39-46.
- 4 周和平.加快实施推广工程 建设覆盖全国的数字图书馆服务体系——在数字图书馆推广工程馆长培训班上的讲话[J].国家图书馆学刊,2012(5):5-13.
- 5 张丽玲.省级公共图书馆网站地方文献信息资源建设调查与分析[J].图书馆学研究,2012(18):45-50.
- 6 林增兴.近二十年我国图书馆特色数据库研究综述[J].图书馆工作与研究,2018(1):61-66.
- 7 郑云霞.图书馆数字资源的共建共享——以数字图书馆推广工程为例[J].新世纪图书馆,2018(1):48-51.
- 8 董雪迪,等.基于云计算的国家数字学术资源多元保存模型构建研究[J].情报科学,2018(6):35-39.
- 9 徐延华.云计算技术环境下数字图书馆资源的整合[J].山西档案,2018(4):100-102.
- 10 段杨军.云计算环境下的图书馆资源库建设[J].图书馆,2016(8):108-111.
- 11 王琛,等.云环境下的图书馆移动阅读服务创新[J].山西档案,2018(3):108-110.

《青少年蓝皮书——中国未成年人互联网运用和阅读实践报告(2017—2018)》发布

2018年9月10日,中国社会科学院新闻与传播研究所、社会科学文献出版社和中国少年儿童发展服务中心、中国青少年宫协会联合发布《青少年蓝皮书——中国未成年人互联网运用和阅读实践报告(2017—2018)》。报告显示,在2017年:

看课外书是未成年人最喜欢的课外活动之一,高于上网、看电视和体育运动。未成人课外阅读以纸质图书为主,年平均阅读纸质图书13.8本,是成年人的3倍,其中26.2%的未成年人看课外书1—5本,24.8%看课外书6—10本,19.4%看课外书11—20本,1.9%为0本。未成年人获取课外书的信息渠道主要以学校和同学朋友为主。

未成年人主要阅读的三类图书为:文学名著、历史知识/历史人物类和科普类。

城乡未成年人在阅读数量、阅读质量等方面存在不均衡状况,年阅读量在10本以上的城市学生比例为48.1%,农村学生为30.0%。

资料来源

- 1 蓝皮书:未成年人年平均阅读纸质书13.8本是成年人3倍[EB/OL].[2018-09-14].http://news.china.com.cn/txt/2018-09/10/content_63368651.htm.
- 2 青少年蓝皮书:中国未成年人互联网运用和阅读实践报告(2017—2018)[EB/OL].[2018-09-14].<https://www.ssap.com.cn/c/2018-09-10/1071885.shtml>.
- 3 《青少年蓝皮书》:与父母感情不好 孩子上网时间更长[N/OL].[2018-09-14].http://zqb.cyol.com/html/2018-09/11/nw.D110000zgqnb_20180911_3-09.htm.

(国家图书馆研究院 提供)

-
- | | |
|--|--|
| 12 国家公共文化云正式开通[J].国家图书馆学刊,2018(1):47. | 17 文化部“十三五”时期文化发展改革规划[EB/OL].[2018-06-16]. http://zwgk.mcprc.gov.cn/auto255/201702/t20170223_491392.html . |
| 13 陈军营,等.中外图情领域大数据近十年(2007—2016)研究现状与发展趋势分析[J].情报科学,2018(7):104-110. | 18 文化部关于印发《文化部“十三五”时期文化产业发展规划》的通知[EB/OL].[2018-06-16]. http://zwgk.mct.gov.cn/auto255/201704/t20170420_493300.html?keywords= . |
| 14 曹霞.国内大数据知识服务研究热点与进展[J].西南民族大学学报(人文社科版),2017(9):235-240. | 19 郭毅可,韩锐.云计算中的弹性算法:概要和展望[J].上海大学学报(自然科学版),2013(1):1-4. |
| 15 孙翌,等.上海图书馆学情报学研究综述:2014—2016年[J].图书馆杂志,2017(12):27-36. | |
| 16 傅耀威,孟宪佳.移动互联网技术发展现状与趋势[J].科技中国,2017(12):60-62. | |

(童忠勇 高级工程师 国家图书馆)

投稿日期:2018-08-06