

图书馆大数据与微服务的技术融合体系研究*

李白杨¹, 白广思²

(1. 武汉大学信息管理学院, 武汉 430072; 2. 郑州大学图书馆, 郑州 450001)

摘要: 现今大数据、微服务研究较多, 大数据与图书馆微服务结合技术体系方面研究有待加强。图书馆大数据与微服务的技术融合体系有大数据、读者需求、数据处理、微服务业务和服务反馈5个层次。技术融合的基础是用户服务, 大数据与微服务具有密切相关性, 两者能够通过技术融合实现双向发展。基于大数据的微服务能为读者和科研用户提供更加高效、优质的服务。

关键词: 技术融合; 图书馆; 大数据; 微服务

中图分类号: G250

DOI: 10.3772/j.issn.1673-2286.2016.1.010

随着互联网技术的不断创新, 网络越来越深刻地改变着人们信息获取和阅读方式。一方面, 由用户产生的数据所构成的大数据持续发展, 形成了蕴含巨大开发价值的信息“富矿”; 另一方面, 微博、微信、微视等微信息服务的出现使信息传播更加快捷, 信息受众群体更加广阔, 信息生态开始出现蝴蝶效应。本文立足于图书馆的大数据^[1]和微服务^[2], 认为图书馆的服务模式正在突破传统形式, 其服务内容更加广泛、服务效果更加优越、服务范围打破了时空界限, 基于微博、微信的图书馆微服务成为服务读者的重要方式。此外, 大数据技术的出现为解决信息不对称现象提供了一套新的技术方法, 其研究与应用已经渗透到图书情报领域的方方面面, 图书馆所拥有和积累的大数据也变得越来越有价值。

1 研究背景

1.1 相关研究

以“标题=大数据”分别在Web of Science核心集、CNKI中文核心和CSSCI核心数据库中进行多次检索, 将检索结果限定在图书情报学科, 去重、删除专题导言等不必要的文献, 然后利用CitespaceII对得到文献进行可视化处理, 结果如图1和图2所示。

对比图1和图2可以明显地看出大数据在国内外



图1 Web of Science核心集中图书馆情报领域大数据的研究主题

图书情报领域都是热门研究主题, 成果较多, 与大数据相关的研究主要有数据挖掘与分析(data mining、analytics)、社交媒体与社会数据(social media、social network、social data)、知识服务与创新(knowledge service、innovation)等。上述研究主题都是与图书馆微服务密不可分的, 其中数据挖掘是方法, 社交媒体是工具, 知识服务是内容。

利用引证关系对检索到的文献进行处理, 得到与

* 本研究得到国家社科基金一般项目“大数据时代图书馆微服务应用研究”(编号: 15BTQ024)资助。

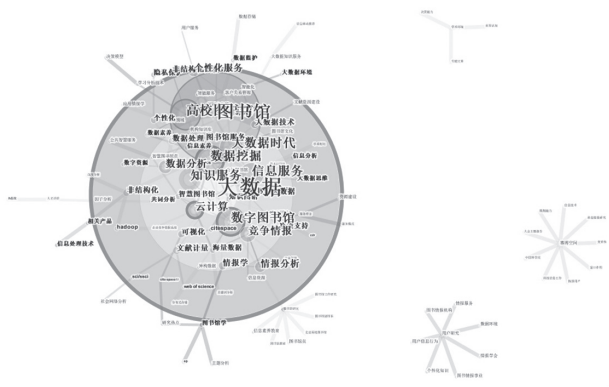


图2 CNKI中文核心和CSSCI核心图中
书情报领域大数据的研究主题

主题相关的主要研究成果有：国外方面，Kwon等^[3]研究了大数据分析对于图书情报机构的业务扩展；Agarwal和Dhar^[4]研究了图书情报机构应用大数据所面临的机遇与挑战；Martinez和Lara^[5]研究指出社会媒体的发展是形成大数据的主要数据源，而大数据的应用离不开社会媒体的支持；Bohloul^[6]等研究了利用社交媒体类微服务工具提供大数据知识服务的SoMABiT模型；O' Leary^[7]认为大数据的整合必须将来自移动设备的传感器、应用数据等纳入其中。国内方面，韩翠峰^[8]提出大数据时代图书馆的一些服务创新策略；李广建等^[9]从情报学视角研究了大数据环境下情报分析与服务的发展；耿晓光和安冬梅^[10]研究了微博在图书馆中的应用，同类的还有李斯^[11]、龚雪竹^[12]等的研究。黄浩波^[13]、孔云^[14]、李白杨^[15]等研究了微信在图书馆中的应用。截至目前，国内外图书情报界还在持续研究大数据与微服务这两个热点话题，可以看出图书馆的微服务发展早于大数据，但大数据对微服务的发展形成有力推动，两者的融合是图书馆扩展业务、强化服务和向着智慧型图书馆转型的重要条件。

1.2 大数据与微服务的关系

从字面上看，大数据与微服务似乎是相反的。实际上，在图书馆的信息服务体系中，大数据与微服务是相辅相成、互相依存的关系。图书馆大数据是其微服务体系的基础条件，微服务是大数据技术成果展现的具体方式和为读者提供优质服务的媒介。微服务不仅是利用微信息技术进行知识传递，也是图书馆的一种技术变革与服务变革，是传统内容向着精深化知识服务的转变。因此，高质量的微服务离不开大数据技术的支撑。具体来

说，图书馆的大数据应用包括对读者借阅习惯、服务轨迹、科研趋势、成果转化等方面的深层挖掘，也包括对图书馆的未来发展、服务模式提供可行的趋势分析，还可以深度挖掘读者在交互型知识服务过程中的潜在需求数据，从而提高图书馆微服务的针对性和有效性。

2 大数据与微服务的技术分层

要实现图书馆大数据与微服务技术融合，首先要理清两者涉及的技术层面，具体如下。

2.1 大数据层

随着数字资源在图书馆馆藏资源中所占比例越来越大，其积累的数据总量是惊人的。大数据层是指由图书馆所拥有的各种类型的数据所共同构成的大型数据块集合，其中包括结构化数据、半结构化数据和非结构化数据。以郑州大学图书馆为例，目前其拥有的大数据资源类型包括海量数字文献资源元数据、读者基本数据、科研成果元数据、借阅信息数据、荐购系统数据、RFID用户服务数据、无线接入服务数据等，具有总量大、数据结构复杂、更新频率快等特点，符合大数据的特征。大数据层是利用大数据技术开展数据挖掘以进行微服务的基础资源。

2.2 读者需求层

读者需求是图书馆服务体系的核心价值所在，只有知道读者需要图书馆提供什么样的服务，才能建设务实、高效、受读者欢迎的图书馆微服务体系。读者需求层具体分为两个方面：一方面是读者的显性需求，即读者对图书馆服务的具体需求，需要图书馆利用问卷调查、访谈等方法进行了解和研究；另一方面，是读者对图书馆的隐性需求，一般是对文献资源的需求和附加服务的需求，需要利用大数据进行分析和研究。

2.3 数据处理层

对大数据的采集、存储、处理是大数据与微服务两者技术融合体系中的实际操作环节之一。目前比较成熟的采集技术一般为分布式数据处理技术，如Hadoop、MapReduce、Spark等。无论采用哪种语言或

技术,通常在大数据处理上具有并发化、分布化、高效化、迭代化等特点,可对图书馆所存在的异构大数据进行统一的挖掘处理。此外,不少高校图书馆部署了云计算服务,其弹性存取技术能够良好地兼容大数据处理。

2.4 微服务业务层

从数据处理层向微服务业务过渡的实质是从信息的处理走向信息的发布,在这个层面,基于大数据的图书馆信息服务范围开始由馆内向馆外。目前,国内高校图书馆积极应用微信息技术,按照工具划分主要包括图书馆官方微博、图书馆微信公众号、独立的移动图书馆等多种形式;按照内容划分主要包括社交工具和图书阅览平台两种。图书馆微服务的核心是运营,包括内容的制作、更新和推送三个方面。良好的运营可以将图书馆大数据资源所提取的信息及时、准确地反馈给用户,对于提升图书馆利用效率和服务质量具有较好的辅助^[16]。总体而言,图书馆的微服务业务离不开两个支撑:一是新媒体支撑,二是移动互联网支撑。

2.5 服务反馈层

为了避免图书馆跟风运营新媒体和微服务却反而出现与用户实际需求相悖的情况,微服务必须按照一定的科学方法定期进行服务反馈调研。常用的调研有

机器统计和问卷调查两种方式。调查的内容主要是信息的筛选、分析、服务定位是否符合用户需求,调研需要基于科学方法,最终形成分析结果,动态记录图书馆微服务的发展历程,并针对遇到的问题提出相应的解决方案。由此,图书馆内基于大数据技术的微服务流程就实现了从资源向服务、从馆内向馆外、从单向服务向双向交流的新型服务目标。

3 大数据与微服务的技术融合

基于上述5个层面,研究其技术融合的具体路径。

3.1 用户数据互操作

服务用户,是大数据与微服务进行技术融合的基础,其涉及的技术环节是用户数据的交换和互操作。图书馆以用户ID(公共馆为借阅卡、高校馆为学号和教工号)作为通用数据,为互操作带来了便利。根据目前图书馆的业务特点,如图3所示,数据源可分为借阅数据、行为数据、移动数据和社交数据,这些数据都可以采用用户ID进行互操作。一般来讲,用户数据使用SQL Server存储,其他几个数据源的数据库和接口格式分别有Oracle、.NET和JAVA等,使用数据库链接器能够将异构的数据库进行桥接,用户数据是研究大数据环境中数据关联的“钥匙”,是进行大数据分析的基础条件。

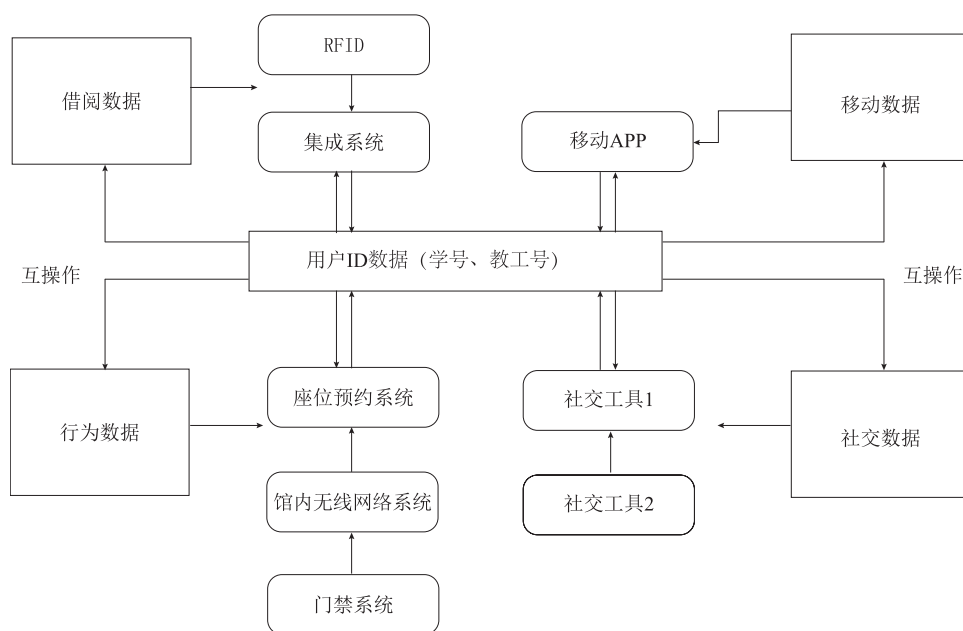


图3 图书馆内基于用户数据的异构数据库互操作

3.2 技术融合体系

如图4所示, 技术融合体系主要由大数据处理和微服务两个主体步骤构成, 并涉及大数据处理、用户需求分析、知识库构建、微服务内容编辑和推送四个关键阶段。阶段一, 利用用户数据将异构的图书馆大数据进行整合处理, 其中包括图书馆的借阅信息、行为记录、馆藏信息、网络资源、移动信息、社交媒体等^[17]。阶段二, 利用调查问卷、荐购系统、社交媒体等对用户的显性服务需求进行调研; 同时利用基本的数据挖掘规则调查用户的隐性需求, 例如用户借阅图书的种类倾向、学科倾向、座位预约的时间分布等。阶段三是对大数据进行时间、空间和语义等多维尺度的分析, 根据图书馆自身开闭馆时间、图书空间分布等情况设计分析参数和阈值。阶段四是微服务的内容编辑和推送, 目前图书馆微信公共号是使用便利、内容扩展强的工具, 通过开源API可以较好地进行基于社交媒体的图书馆大数据服务推送。

3.3 微服务策略

微服务的基础是泛在化阅读的实现, 也就是说微

服务是图书馆泛在化服务的一种表现形式。在图书馆实体中, 我们通过开设多媒体阅读区、信息共享空间等策略来实现服务的改革。同样, 图书馆可以通过微服务实现泛在化服务的升级。自从2013年图书馆开始使用微信作为主要的微服务平台后, 微信公共服务平台的服务功能也大为扩展, 目前已开始支持各种API接口, 能够良好地支持前端与后台的结合, 其服务不再是单纯的回复数字得到结果, 而是可以将图书馆的资源通过集成系统的API接口大量接入微信。与此同时, 独立的移动图书馆客户端、其他社交网络、移动借阅机器、图书馆二维码互动等技术都成为微服务的重要组成部分, 整体上图书馆的泛在化微服务已经成熟。接下来, 在强大的技术支撑下, 图书馆的微服务策略应是充分利用大数据技术和微平台, 面向用户的需求升级服务质量和资源内容。

4 前景与展望

一方面, 图书馆界对于大数据的研究很多, 但对于大数据究竟有何用、如何扩展图书馆业务等问题还未有确切的答案。另一方面, 国内图书馆的微服务业已成

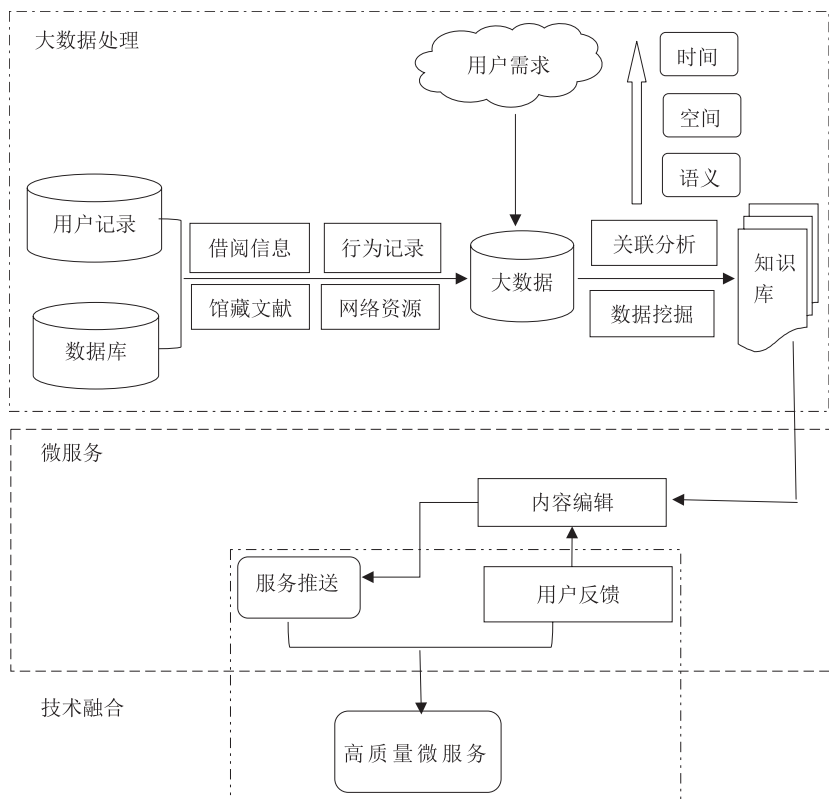


图4 图书馆大数据与微服务的技术融合体系图

为新的业务增长点。从2013年开始,图书馆的微服务已经从移动书目查询发展到移动图书馆阅读,再到如今的移动借阅、个性化参考咨询微服务等。笔者长期扎根图书馆的大数据和微服务建设,认为两者可以成为互相补充的业务增长点。这其中,用户需求成为连接大数据与微服务的关键点,上述的技术融合体系在于以用户数据和用户需求为核心,实现大数据挖掘和微服务推送的有机联结。随着图书馆大数据理论和技术的持续研究,可以预见未来我国图书馆界将更好地实现大数据与微服务的双重增长,缩小与发达国家图书馆服务的差距,实现弯道超车。

参考文献

- [1] 韩翠峰.大数据带给图书馆的影响与挑战[J].图书与情报,2012(5):37-40.
- [2] 张英.微服务:开创图书馆服务的“蓝海”[J].图书馆建设,2011(7):51-53.
- [3] Kwon O, Lee N, Shin B. Data quality management, data usage experience and acquisition intention of big data analytics [J]. International Journal of Information Management, 2014, 34(3): 387-394.
- [4] Agarwal R, Dhar V. Editorial-big data, data science, and analytics: The opportunity and challenge for IS research [J]. Information Systems Research, 2014, 25(3): 443-448.
- [5] Martinez-Martinez S, Lara-Navarra P. Big data transforms the interpretation of the social media [J]. Profesional De La Informacion, 2014, 23(6): 575-581.
- [6] Bohloul M, Daltier J, Fer M D, et al. Knowledge discovery from social media using big data-provided sentiment analysis (SoMABiT) [J]. Journal of Information Science, 2015, 41(6): 779-798.
- [7] O'Leary D E. Exploiting Big Data from Mobile Device Sensor-Based Apps: Challenges and Benefits [J]. Mis Quarterly Executive, 2013.
- [8] 韩翠峰.大数据时代图书馆的服务创新与发展[J].图书馆,2013(1):121-122.
- [9] 李广建,杨林.大数据视角下的情报研究与情报研究技术[J].图书与情报,2012(6):1-8.
- [10] 耿晓光,安东梅.微博及其在图书馆的应用[J].图书馆学研究,2010(11):38-39.
- [11] 李斯,唐琼.我国高校图书馆微博形象研究:以微博风云榜高校图书馆为例[J].大学图书馆学报,2015,33(1):66-73.
- [12] 龚雪竹.公共图书馆和高校图书馆微博的比较分析[J].大学图书馆学报,2014,32(6):78-82.
- [13] 黄浩波,何卫华,叶青.微信及其在图书馆信息服务中的应用[J].图书馆学刊,2013,35(1):62-64.
- [14] 孔云,廖寅,资芸,等.基于微信公众平台的图书馆移动信息服务研究[J].情报杂志,2013(9):167-170.
- [15] 李白杨,白广思.面向知识咨询的图书馆移动服务设计与实现:以微信公众平台为例[J].图书馆学研究,2013(19):69-72.
- [16] 陈凤娟.大数据时代图书馆的新媒体营销策略[J].图书馆学刊,2014(11):24-26.
- [17] 郭春霞.大数据环境下高校图书馆非结构化数据融合分析[J].图书馆学研究,2015(5):30-34.

作者简介

李白杨,男,1991年生,武汉大学信息管理学院博士研究生,研究方向:信息检索与服务。

白广思,男,1964年生,郑州大学图书馆,研究馆员,硕士生导师,研究方向:数字图书馆、图书馆微服务,通讯作者,E-mail: bgsi@zzu.edu.cn。

Research on the System of Integration of Big data and Micro Service in Library

LI BaiYang¹, BAI GuangSi²

(1. School of Information Management, Wuhan University, Wuhan 430072, China;

2. Library of Zhengzhou University, Zhengzhou 450001, China)

Abstract: There are lots of researches related to big data and micro service, and the technology system of integration of big data and micro service should be strengthened. This system consists of 5 levels which are big data, users' requirements, data processing, micro service operation and service feedback. The technology integration is based on users' service in library, while in fact big data is frequently related to micro service, so they could realize the two-way development by the technology integration. The big data-based micro service could provide the efficient and high quality service for readers and scientific users.

Keywords: Technology Integration; Library ; Big Data; Micro-Service

(收稿日期: 2015-12-28)