

大数据与图书馆建设

尚 超

(中国医科大学图书馆, 辽宁 沈阳 110001)

摘 要: 大数据时代的到来是信息现代化科学技术发展的必然结果, 也是图书馆建设的发展趋势, 大数据时代的到来更加提醒了我们要重视“数据”这个一直以来存在于我们身边却不受重视的因素。现阶段, 现代化图书馆的建设也离不开大数据的辅助, 信息技术快速发展的时代背景下, 图书馆数据建设也渐渐被提上了日程。分析现阶段大数据与图书情报之间的关系, 掌握大数据在图书馆建设中的发展趋势, 对现代化图书馆的建设工作有着不可小觑的重要作用。

关键词: 大数据; 图书馆; 建设; 信息化

中图分类号: G252

文献标识码: A

文章编号: 1002-1248 (2015) 02-0116-03

Big Data And Library Construction

SHANG Chao

(China Medical University library, Shenyang 110001, China)

Abstract: Big data era is the inevitable result of the development of science and Technology Information Modernization, and is also the trend of the development of Library Construction. Big data era especially reminds us to pay attention to "data" which had been a neglected factor around us. At this stage, the construction of modern library cannot succeed without the assistant of Big data. Under the era ground of rapid development of information technology, the library data construction has gradually been put on the agenda. Analysis of the relationship between the data and information, and mastering the development trend of Big data in library construction, have played an important role in the construction of modern library.

Keywords: Big data; Library; Construction; Information

大数据的概念是由云计算发展而来的, 是伴随云计算科学技术日渐成熟而日趋发展为可控的, 能够从云计算复杂多样且大量的信息中快速找到需求信息的, 主要以互联网为平台的一种数据形态的描述方式。在这里, 海量数据的概念早在互联网诞生之初就已出现, 只是在很长一段时间内, 由于科学技术发展的限制和高新科技的应用效应, 使得人们没有对海量数据的概念进行发掘和深究, 造成数据的重要性在很长一段时间被忽略。而如今, 随着计算机云计算数据技术的深入发展和广泛应用, 在科学技术越来越发达的条件下, 互联网和计算机、移动上网设备的不断普及, 数据产生源和设备条件已然成熟。现阶段, 在国外一些科学技术先进的国家, 大数据已经被广泛应用到社会发展

的相关领域中, 并取得了可喜的成效。在国内, 许多大型企业, 包括互联网企业、电信金融企业、医疗卫生企业、智能交通等企业, 也已经抢先一步进入了大数据使用的时代。文章就大数据时代的到来对图书馆建设的影响做出相关探索。

1 大数据相关介绍

对于大数据的定义, 官方是这样阐述的: 大数据 (big data), 或称巨量资料, 指的是所涉及的资料量规模巨大到无法通过目前主流软件工具, 在合理时间内达到撷取、管理、处理、并整理成为帮助企业经营决策更积极目的资讯。在维克托·迈尔·舍恩伯格及肯尼思·库克耶编写的《大数据时代》中, 大数据指的是不

收稿日期: 2014-09-02

作者简介: 尚超 (1979-), 男, 硕士研究生, 馆员, 中级职称, 发表论文 10 余篇。

用随机分析法(抽样调查)这样的捷径,而采用所有数据的方法。大数据的4V特点包括:Volume(大量)、Velocity(高速)、Variety(多样)、veracity(真实性)。Volume是指数据的规模和容量远远超过了传统时代的数据,例如, Twitter和Facebook等社交网络的数据产量达到每天7TB到10TB,传统的数据处理技术根本没有能力应付这么大的数据量;Velocity是指数据周转的一般速度。IBM一般认为大数据是一种流动的数据而不是静止不动的一些数字信息,我们必须在数据流动的过程中同时进行数据的提取分析和及时处理,根据数据流动的速度来看,根本不可能在一批数据的结束后才开始进行数据的分析;Variety是指数据的来源和类型的多种多样,在进行大数据的分析时,必须能够进行结构化和半结构化的处理,甚至原始数据化的处理也要涉及。大数据的特征是由IDC定义的,即海量的数据规模、高速的数据流转和动态运输、多样复杂的数据类型和数据的极高价值^[1]。

由以上分析了解到大数据实际上就是一个在云计算的基础上,具备信息收集优化功能和精准信息定位功能,并最终收集通过网络产生的各种信息数据的重要工具。通过网络收集的数据既包括传统结构化的数据,也包括占主流地位的非结构化和半结构化数据(就是原始数据)。在使用网络进行信息的收集工作时,配合以各种先进的数据分析、计算、建模软件,将收集到的数据更加精准、直观地用图表的方式展现在眼前,为进一步决策打下基础、做好准备。大规模收集网络数据是大数据的基本功能,除此之外,大数据也更加注重数据收集、科学技术和设备三者的有机统一,即数据的增值。

大数据时代是现代化信息技术发展的必然趋势,大数据时代的到来,更加提醒了我们“数据”这个一直以来存在于我们身边却不受重视的因素。这不仅仅是因为大数据本身带来的巨大便捷和信息价值,更加因为其所牵扯到的每一个信息机构都需要朝此方向发展,否则将无法生存。大数据时代的到来,更加受到图书情报界的关注和重视,现阶段,图书馆建设中对大数据的应用还处在探索发展的新时期^[2]。

2 图书馆建设推进大数据应用的进程

2.1 图书情报界开展学术研讨会进行的相关研究。

我国国内第六届图书情报界大数据应用学术研讨会自2012年在上海召开以来,对全国各地的现代化图书馆建设和大数据应用产生了深远的影响。其中包括2013年7月万方数据股份有限公司在四川省成都市

召开的以“立足科技创新,繁荣文化发展”为主题的2013年夏季论坛会议。在此次论坛会议中,来自全国各地的图书馆相关负责人士、业内学者150余人,与万方数据企业的专家们共同就大数据时代图书馆的服务建设问题进行了详细的探讨。包括大数据时代下图书馆服务平台的建设、服务模式的创新发展等问题。对现代化图书馆的发展方向进行了深入的分析和专业探析。除此之外,北京正阿帕比有限公司曾经在北京师范大学举办过大数据时代图书馆创新服务研讨会。在此次研讨会议中,京津地区几乎所有高校的图书馆负责人员以及市区专业图书馆管理人员将近80人出席了会议,共同以“在大数据的压力下图书馆如何应对从而提供更高品质的服务”为主题进行了激烈的探讨。同年7月25日,在江西省高校图工委举办的以“大数据与图书馆大未来”为主题的学术研讨会中,来自江苏省、山东省和浙江省等高校的知名专家就这一严肃的问题进行了专业的分析和探讨。此外,还有各种各样的关于大数据时代图书馆服务建设的学术研讨会议在全国各地纷纷召开,昭示着我们已经迈进大数据与图书馆相结合的时代。例如以“大数据时代的科技信息资源共享”为主题的第八届科技信息资源共享促进国际会议、主题为“大数据时代医学信息学创新发展”的中华医学第十九次全国医学信息学术会议等。这些关于大数据时代图书馆建设相关学会研讨会的召开,都对大数据在图书馆建设的发展中起到了推动作用,对及时认清当前情况,做出相应的图书馆发展对策和计划起着至关重要的作用^[3]。

2.2 图书情报界相关学术论文发表研究情况。

作者从中国知网和万方数据服务平台等资源网站了解到,最早的相关方面论文是《图书与情报》2012年第4期的《大数据时代的图书馆服务浅析》一文,之后的一年多以来,大约共有40余篇图书情报与大数据时代相关方面的研究论文发表在相关刊物上。在这些刊登出的论文中,《图书与情报》2012年第6期、《情报资料工作》2013年第2期、《现代图书情报技术》2013年第4期等,连续刊登出一整组的数据专稿,这突出体现了对大数据与图书馆建设的研究频率在短期内的急速上升,间接反映出了该学术学科的蓬勃发展^[4]。这些论文的研究成果可以概括成四大方面:

(1) 大数据时代对图书馆建设的影响以及图书馆发展的相应对策。大多数文章都包含了对大数据的概念内涵与意义、大数据时代到来为图书馆发展提供的契机、图书馆建设应做出的对策、今后图书馆与大数据结合的发展趋势等相关热点内容的探讨,只是不同

的文章、不同的作者对这一问题有着不同的研究视角。其中,张兴旺等学者针对图书馆所属的学术学科、研究和管理特殊背景方面对大数据时代对图书馆建设的影响进行了探讨;刘明等学者则是从生物医学角度对图书馆在大数据时代到来如何进行建设进行了具体分析;朱静薇等学者是从数据管理、数据技术、队伍建设、服务模式等几大方面分别提出了大数据时代图书馆发展的具体对策。

(2) 大数据时代的图书情报工作。大数据不仅应用在网络信息的收集工作中,还在情报基础、理论竞争情报、情报检索系统等领域应用广泛。对于大数据时代的情报工作,有些学者根据多年来大数据在这方面应用表现出的不足和缺陷,在论文中进行了严肃认真的探究。例如,缪其浩等总结了自己及其所带领的情报工作团队在大数据研究上取得阶段性成功的原因;贺德方结合相关案例给出了情报工作在发展方向上的思考;黄晓斌等对关于大数据情报中企业情报获得的相关模型作出了研究;吴广印介绍了大数据时代下分布式搜索引擎的典型应用;刘炜等的专题论文梳理了目前采用大数据技术发布关联数据的各种方法和路径,将关联数据深入研究到图书馆具体应用的实施、标准规范的研究、工具的使用和与前沿技术结合等相关方面^[5]。

3 大数据时代下的图书馆

在信息网络高速发展背景下,大数据在现代化图书馆建设中的应用不仅仅只是网络数据的工具,更是围绕其进行相关信息发掘、利用的一种资源。这种资源对于图书馆核心竞争力的开发具有重大意义。

首先,多年来的图书馆信息化建设使图书馆内的资源数据大量囤积,许多数据因为信息成本的爆降,引起半结构化、非结构化的原始数据信息资源的浪费。例如声像视频、流媒体数字资源、学术文献的各种科学原始数据图表等,这种现象会造成对原有数据采集处理模式的危害,影响系统重新优化资源配置。大数据是用来克服现有信息中心不能适应复杂多样的信息数据处理以及非结构化的原始信息数据获取效率低下的现状的。也就是说,大数据时代的图书馆是基于云计算等大数据处理方式的图书馆,这种新的基于大数据的图书馆建设方式成本十分低廉、见效是十分快速。

其次,云计算的数据处理方式不仅颠覆了传统的图书馆资源组织建设,更加促进了图书馆的人性化服

务。由于如今各种移动设备、网络终端的不断普及以及 web2.0 技术在图书馆服务中的应用越显重要,只要是网络使用者在人人网、QQ、微信、微博、BBS、博客等社交网络进行过的网络行为和数据浏览都会被大数据时代中的图书馆作为人性化服务的根据,这极大地丰富了图书馆人性化服务信息数据的来源。能够利用大量数据信息去充分挖掘信息背后读者的平时喜好和浏览习惯,促进了图书馆个性化服务的进一步完善,加快了大数据在图书馆服务建设中的应用进程^[6]。

再次,大数据时代的到来不仅促进了现代化图书馆对数据资料分析理解能力的增强,更加促进了图书馆服务质量的大幅度提升,从而促使以信息服务见长的图书馆服务范围和领域得到更好的扩展。例如,在图书馆中开展一些高附加值的科学数据监管服务,当前科学的发展已经进入第四范式,即数据密集型科学模式阶段。在国外一些科学技术先进的国家,大数据图书馆内的科学数据监管服务已经在进行,如今国内也相应的建立了一些科学数据监管服务平台。

综上所述,大数据时代的到来是现代化信息化科学技术发展的必然结果,也是世界化的发展趋势,大数据时代的到来更加提醒了我们“数据”这个一直以来存在于我们身边却不受重视的因素。现阶段,大数据已经成为现代化图书馆建设的重要手段,在信息技术快速发展的时代背景下,图书馆数据建设也渐渐被提上了日程,分析现阶段大数据与图书情报之间的关系,掌握大数据在图书馆建设中的发展趋势,不仅对现代化图书馆的建设工作有积极的指导作用,更对大数据在中国其他方面的应用积累了重要的经验。

参考文献:

- [1] 王丽敏,王淑阁.云计算环境下图书馆管理与服务方式探析[J].兰台世界,2011,(6):70-71.
- [2] 王立东,张辉程.网站在线客服系统应用于图书馆的可行性分析[J].图书馆学研究,2010,(10):82-84.
- [3] 刘丽梅.浅议图书馆的优质信息服务[J].黑龙江科技信息,2010,(13):104.
- [4] 吴晓骏,刘翔.论数字图书馆环境下长尾服务的实现[J].图书馆,2009,(4):96-97.
- [5] 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.
- [6] 宋方.大数据机遇[J].企业管理,2012,(7):99.