

大数据在图书馆管理与服务中的应用

宋春花

(上海图书馆(上海科学技术情报研究所) 上海 200031)

[摘要] 随着网络信息技术的发展, 互联网大数据时代已经到来。图书馆管理与服务的未来发展趋势也将会是以此为基础。本文通过分析大数据的时代特征, 提出了其在图书馆管理中的具体应用以及管理方面的实践方案, 并且深入研究了大数据采集、储存、数据管理等多方面内容, 对图书馆管理与服务有了最新的认识。

[关键词] 大数据; 图书馆; 管理与服务

引言

大数据, 顾名思义是指大量的数据, 其中包括许多种形式, 例如结构化、半结构化以及非结构化数据等。由于其无法被普通的软件工具捕捉到, 所以只有新型的处理模式才具有更加强大的洞察力。大数据技术的含义是将某些数据专业化处理, 它与云计算密不可分, 在云计算的分布式构架基础上进行处理才能够满足所需。

1 图书馆大数据

由于大数据系统是随着大量数据应运而生, 图书馆内书籍具有数量庞大、形式多样化的特点, 因此大数据应用于图书馆中十分有必要, 它能够从不同类型数据中获取具有一定价值的信息。图书馆的电子书刊资源、音频、视频资源种类众多, 其中包含了原生数字资源以及虚拟馆藏等其他资源, 且每年的数量均呈上升趋势。另外, 读者的各种信息和外联的数据随着时间的推移不断产生, 重组后再融合进图书馆大数据存储中。图书馆的数据库除现有的书籍、文献之外, 还存在尚未建设的非结构化、半结构化数据。网络时代的到来使人由信息的获取者变为了信息的缔造者和传播者, 微博、微信、QQ等社交软件均属于大数据的一部分, 移动图书馆的读者产生的浏览历史、检索内容等数据量往往大于标准, 由此, 可以体现出图书馆的数据具有多种结构共存的特征。

2 图书馆大数据处理方案

2.1 大数据采集过程

图书馆的大数据采集过程是指实时接收其自动化系统以及外联网中所产生的全部数据, 相比于传统的采集过程, 它能够根据不同类型的资源进行重新整合。它采集的内容包括没有具体统一格式的结构化、半结构化和非结构化数据, 在采集的过程中需要注意数据要高度相关且规范性, 还要保证其复杂性。总而言之, 采集需要具有分布性, 图书馆的采集过程与此过程大致相同。

2.2 大数据存储与组织

现有的大数据库存已经无法满足时代所需, 且存在一定缺陷, 但是储存仓库在拓展、容错性、分析延迟以及成本费用方面均具有一定优势。如果将其与大数据平台想融合, 可能创造出一种共生的混合储存平台来供人们日常所需。图书馆可以利用传统模式的DW和如今最流行的大数据处理平台架构Hadoop作为存储仓库, 图书馆可以按照业务类型选择相对应的储存模式。这种方式虽然符合大数据储存, 但是对图书馆的基础设施具有较高的要求。对于图书馆来说, 分析与挖掘的主要内容是指运用分布式计算集群对保存在其内部的内容进行分析、汇总, 因此, 图书馆大数据符合其挖掘规律。

3 图书馆管理的实践与探索

3.1 增强资源建设

图书馆中的资源收集归纳工作应该受到重视, 这是大数据应用的前提, 其中的重点是增强资源的建设以及数据关联。关联科学作为一种新型科学研究内容, 经常被引入图书馆的普通资源建设之中, 除此之外, 还要整合特色文献、机构知识库的整合, 利用资源关联整合中的索引功能来实现资源的获取、整合与展现, 如此一来, 便能促使资源整合的发展, 彻底实现满足图书

馆需要的资源共享, 从而挖掘出更有价值的资源内容。另外, 图书馆需要留意非结构化数据, 它可以利用其他技术进行改造, 改造过程中需要采集不同时期产生的零散信息, 并且以各种主体、物体、事件为研究对象进行关联。

3.2 培养馆员的专业性

与传统的图书馆管理服务不同, 大数据时代的管理服务工作发生了本质上的区别, 图书馆可以利用此机会对馆员进行专业培训, 开展数据管理与服务内容。图书馆必须为馆员提供数据库, 并聘请专业人员传授给馆员相关的数据意识与分析管理等各项技能, 还应了解馆员的性格特点等, 根据他们的不同进行针对性的分布培养, 从而提高他们的专业素养, 满足图书馆管理的需要, 最后利用他们的高素质以及专业性为所在的图书馆提供参考、指导工作, 还应该提倡馆员勤奋学习专业知识, 并向外传播, 引导更多的人参与进大数据之中。

3.3 服务中的探索与实践

图书馆可以建立起虚拟馆藏结构, 在其中加入资源、物理数据, 利用虚拟现实技术创建模式, 从而为用户及读者提供更好、更方便的服务, 体验之前从未有过的内容。例如, 将书籍目录、书籍摆放位置、借阅系统等与互联网技术进行关联, 而后创建3D场景, 通过引导, 使读者体验到3D导读模式。RFID是国家最先采取的智能架位导航系统, 它被融入到OPAC系统中, 读者可以通过点击架位号的连接进入到查询的起始位置, 按照其位置, 智能分配出距离最短的路线。读者还可以通过此书观看馆内与之相关书籍的楼层, 从而进一步确定所需的书籍。

图书馆可以将书目与流通、电子文献与检索、社交网络与用户等多种数据相关联分析, 最终得出用户最喜欢的书籍类型, 并进行推广阅读, 从而达到阅读推广效应。资源建设还可以作为文献信息资源建设的前提, 收集不同读者的浏览记录, 分析出大多数人的阅读习惯, 以此作为读者需求, 进而安排馆内书籍摆放位置、购买书籍资源等, 为读者提供更加贴心的服务。

结束语

综上所述, 大数据内容在图书馆内的管理与服务应用十分重要, 它不仅能够为图书馆带来方便, 还能够使读者体验到更加贴心的服务, 其中推广效应能够满足不同读者的需求。大数据时代的到来, 使得图书馆的压力只增不减, 图书馆通过培养馆员, 增强其专业能力, 从而更好地引导读者, 为他们提供更好的服务与帮助。大数据是时代发展的必然结果, 它在人才培养、服务探索、管理决策等方面均发挥一定作用, 但是它正处于发展的初级阶段, 还存在一些问题, 图书馆应该认清其本质, 利用其有利的方面切实改善服务质量, 利用网络化、智能化改变传统图书馆的不便之处。

参考文献

- [1] 王娟. 大数据环境下军校图书馆个性化服务及信息资源的共享[D]. 南京邮电大学, 2017.
- [2] 姜淼. 基于大数据的高校图书馆定制化信息服务研究[J]. 无线互联科技, 2018, 15(21): 75-76.
- [3] 万丽媛. 大数据时代图书馆服务创新内容及策略[J]. 黑龙江科学, 2018, 9(21): 140-141.