



大数据时代高校图书馆事实数据库建设的思考

□黄运红*

摘要 教育部高校图书馆事实数据库由教育部高等学校图书情报工作指导委员会组织开发,是一套旨在采集我国高校图书馆事实数据的网上填报与统计系统。为更好地满足大数据时代图书馆的发展需求,在剖析事实数据库的统计指标、电子资源规范与系统管理运维等问题的基础上,本研究就完善指标体系、保障数据填报率、提高数据准确率、增强数据利用率提出改进建议,为完善高校图书馆事实数据库建设提供参考。

关键词 大数据时代 高校图书馆 事实数据库 数据统计 统计指标

分类号 G258.6

DOI 10.16603/j.issn1002-1027.2020.05.005

1 引言

伴随着大数据时代的到来,图书馆在数字化、智能化的发展建设中,积累了大量的馆藏数据、服务数据、管理数据,以及读者数据等。有效分析和利用这些数据,可以从中获取大量价值不菲的信息,辅助图书馆分析用户需求、优化业务流程、调整服务模式、评估管理效能、支撑科学决策。然而这些宝贵的数据,如散落海底的珍珠,尚未被充分挖掘、采集、分析与利用。“美国 McKinsey Global Institute 发布的 *Big Data: The Next Frontier for Innovation, Competition and Productivity* 调查报告,指出尽管全球数据飞速增长,但有将近 87.5% 的数据未得到真正利用。……相较于图书馆所拥有的不断增长的数据量而言,图书馆能够分析的数据比例在不断降低”^[1]。因此,建立一个完善的图书馆数据统计与管理平台,充分采集和利用大数据资源,有助于图书馆总结业务规律、分析读者需求、探索图书馆发展趋势,使大数据在推动图书馆管理创新与服务创新中发挥更大作用。

教育部高校图书馆事实数据库(以下简称事实数据库)是专门为高校图书馆设计的数据统计系统,在我国图书馆数据采集平台中占有极其重要的地位。随着图书馆转型和业务拓展,以及数字资源、自助设备的应用,图书馆又产生和积累了大量新型数

据。然而事实数据库自 2001 年发布以来,只在 2010 年进行过一次重大修订,修订后数据指标已使用近十年,对大数据时代图书馆创新服务和资源利用等情况的统计明显不足,部分指标需要增补、调整,有待进一步更新和完善。

检索中国知网(CNKI)收录的文献发现,关于高校图书馆事实数据库,大多针对第一版数据库指标进行论述^[2-4],对修订后的事实数据库鲜有剖析。本文通过梳理、分析近几年数据库填报使用中发现的问题,提出改善统计指标与管理模式的建议,期望对完善高校图书馆数据统计指标体系、优化图书馆考核评估体系、改进图书馆数据统计与分析应用的管理模式提供有益参考。

2 事实数据库概况

事实数据库是高校图书馆事实数据采集与统计系统,由教育部高等学校图书情报工作指导委员会(以下简称图工委)组织开发,旨在收集全国高校图书馆(以下简称高校馆)事实数据,记录高校馆的发展状况和发展历程,为图书馆评估提供数据基础,助力高校馆的宏观管理与决策。事实数据库于 2001 年正式上线。在图工委的大力推广下,2004 年全面实现网上填报,结束了高校馆纯手工上报数据的时代。

* 黄运红, ORCID: 0000-0003-0827-8691, 邮箱: huangyh@lib.bnu.edu.cn.



事实数据库在推广使用中也逐渐暴露出一些问题,2010年图工委召集专家,牵头对事实数据库的统计指标进行重大修订,增加了分馆、勤工助学学生及志愿者人数、联盟费用、信息服务、信息素质教育等指标,修改或细化了读者总人数,以及图书、报刊、论文、手稿、电子资源等文献类型,并删除了一些已不具有统计意义的项目^[5]。

第二版事实数据库分精简版、基本版和扩展版3个版本,沿用至今。精简版专门针对高职高专院校图书馆,基本版针对普通高校馆,扩展版则针对入选“985工程”“211工程”的重点院校图书馆。3个版本的一、二级指标设置相同。基本版为样本,其余

两个版本在基本版基础上删除、简化或补充、细化而成^[6]。其指标项目有A—J共11个大类,累计统计202个数据指标,如图1所示,图中扇形区域的大小反映了对应数据指标的统计项数量,指标涉及图书馆基本情况、经费、馆藏资源及服务四大方面的数据。虽然事实数据库是我国目前应用最广泛、最具权威性的图书馆数据统计系统^[7],记录了高校图书馆珍贵的事实数据,在研究图书馆现状、未来发展及图书馆工作质量评价等方面发挥着重要作用,但随着图书馆的快速发展,已不能完全满足大数据时代的要求,有待进一步改进和完善。

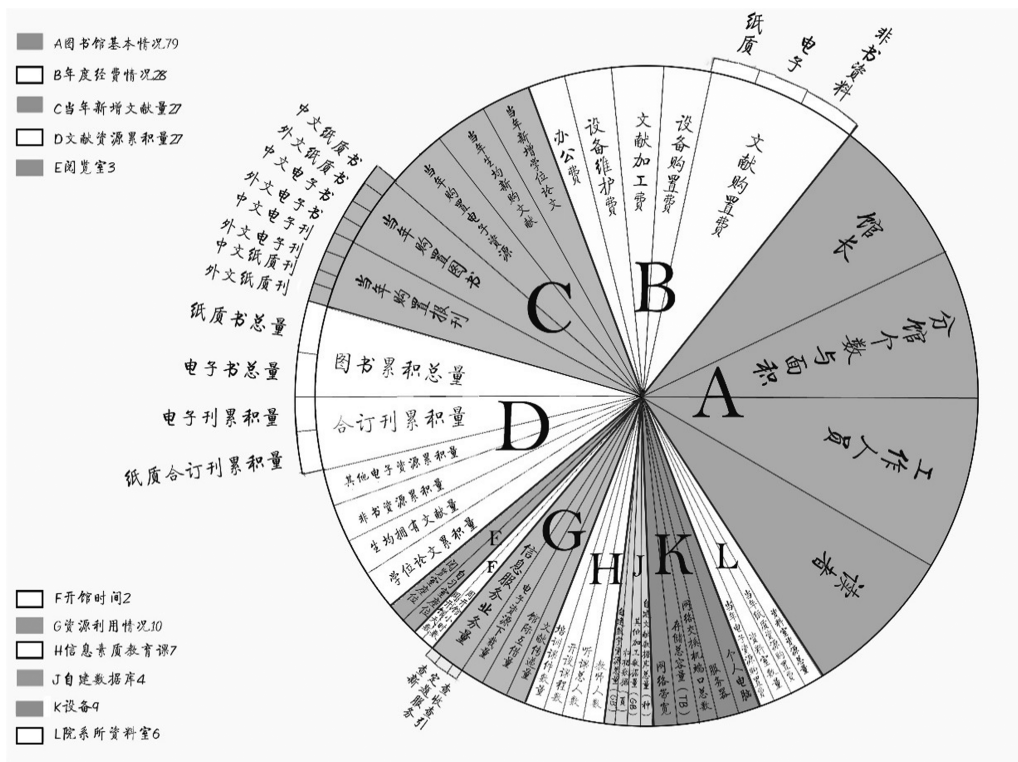


图1 事实数据库包含的数据指标

3 事实数据库统计指标的问题分析

第二版事实数据库经过补充、修改,整体结构比较完整,大部分指标设计合理。不过,笔者在填报使用中也逐渐发现一些问题,比如:个别指标逻辑不清晰,分类不全面,需进一步细化;个别系统自动生成的指标设置不够合理,计算结果有误。有的指标没有含义解释,有的指标解释含混易产生歧义,影响数据的准确性。部分采集的信息不够完整,电子资源计量有待进一步规范。

3.1 部分统计指标不够精准

3.1.1 图书馆总面积可能不准确

在事实数据库中,统计指标A1(馆舍总面积)的值由系统自动生成,是馆舍1至馆舍10的建筑面积之和,却没有将A1.0.4(分馆总面积)累加进去。事实数据库中关于分馆的定义是:“指设在院系但行政关系隶属于图书馆的馆舍”^[8],许多高校图书馆设有这样的分馆,笔者所在的北京师范大学就在教育学部所在的英东楼设有教育分馆,分馆隶属于图书馆,分馆的工作人员属于图书馆馆员,分馆的文献资料属于图书馆馆藏。分馆面积理应纳入图书馆总面积



中,但系统自动求和时没有累加进去。故而,总面积错了可能填报人员并不自知,确实存在把面积数填在分馆那项,系统中该馆总面积为0的情况。也有的图书馆虽发现不对,但不能修改馆舍总面积的数值,造成系统生成的结果与实际不符。

3.1.2 其他来源新增文献可进一步细化

事实数据库中对文献资源的统计最为详细,充分考虑到各种文献类型,区分了中文与外文、纸质与电子,区分了图书、报刊与学位论文,以及古籍、手稿、拓片、多媒体资料等类型;还考虑到文献来源,区分了当年购置与其他来源新增。然而,其他来源文献统计太过笼统,只有C6(其他来源新增文献)这一条。建议对这一统计项进一步细化:(1)区分中外文。计算文献累积量时是区分中外文的,因此其他来源的文献最好也区分中外文,以方便计算和系统校验;(2)区分捐赠和其他来源。其他来源除捐赠外还有调入、交换等,高校馆每年都会收到不同渠道捐赠的文献,捐赠是其他来源中最主要部分,有时也需要单独统计捐赠图书数量,指标细化区分出来,可以有备无患。

3.2 部分指标解释不够明晰

3.2.1 文献资源累积量指标解释及前后对应不足

在文献资源累积量D的指标解释中强调了“D1—D4各项中,注意纸质文献要扣除剔旧数量”^[9],不过,在实际填报中存在部分院校没有扣除剔旧图书的情况。因为,D(文献资源累积量)、D1(图书累积总量)、D1.1(纸质图书累积量)都是由系统自动求和生成,填报人员不能直接在D中减掉已注销的图书量。D1.1是D1.1.1(中文纸质图书累积量)与D1.1.2(外文纸质图书累积量)的和,因而填表人应在统计D1.1.1和D1.1.2的数量时主动扣减中、外文纸质图书的注销数量,但D1.1.1和D1.1.2的指标解释没有强调要扣除剔旧数量,填报人员容易遗忘,导致数据有误。另外,在统计C7(当年剔除、注销文献)时并没有区分中外文,计算D1.1.1和D1.1.2有难度。建议在D1.1.1和D1.1.2这两项增加扣除剔旧量的提醒说明,或者系统按 $D1.1 = D1.1.1 + D1.1.2 - C7$ 公式自动生成数据,免去填表人手动扣减剔旧文献量的麻烦。

3.2.2 馆际互借和文献传递指标解释不具体

高校馆在记录馆际互借借入量和借出量、文献传递传入量和传出量时,通常分别记录了请求量和

满足量。事实数据库对G2(馆际互借借入量)、G3(馆际互借借出量)、G4(文献传递传入量)、G5(文献传递传出量)没有具体解释,多数图书馆填的是满足量,也有些馆填的是请求量。建议在填报指南里予以明确,使标准统一。

3.3 部分指标采集信息不够完整

3.3.1 统计工作人员总数未考虑业务外包情况

从20世纪90年代起,我国图书馆开始将部分业务外包,以强化核心业务,缓解编制紧张、人员不足的压力。外包业务涉及采访、编目、流通等,甚至有的图书馆整体外包^[10]。“全国约一百所高校的调研情况表明,全国曾经或是正在实行外包的图书馆的比例高达80%,相对而言,较发达的地区实行外包的比例也比较高,北京、上海、广州等地都高达95%以上”^[11]。事实数据库统计A3(工作人员总数)时,计算了在编职工、合同制职工、临时聘用职工,2011版还补充了图书馆聘用勤工助学学生和志愿者的情况,把他们每年工作的总小时数,按“8小时×22天×12个月折算一人/年”的公式进行折算,转换成工作人员数量累加到A3中。但是,统计指标却没有考虑服务外包情况。外包人员承担了比学生工和志愿者更多、更深入的图书馆业务,不应忽略,如果进行图书馆人力资源分析时,把存在服务外包和未实施外包的图书馆工作人员总数直接进行对比分析(因为现有系统无法区分),所得结论可能不足以反映真实情况,说服力有限。

3.3.2 计算阅览室和座位未包含多人研究间

E(阅览室)只收集了E1.1(阅览室座位)和E1.2(自习室座位)两个指标,只这两个数据不能完全体现图书馆为读者提供的阅览空间。近年,为更好地满足读者多元化与个性化需求,许多高校开始图书馆空间再造的研究与实践,划分安静学习区与交流研讨区,进行动静空间分区,增设相对封闭的个人或多人研究间与小组交流的研讨间;结合阅读推广活动,提供文化体验空间;具有朗读、录音、交互学习等功能的多媒体视听空间;为培养学生创新创业能力,开辟创客空间等。建议将这些多层次的空间服务纳入事实数据库的统计之中。

3.3.3 反映图书馆核心竞争力的指标欠缺

事实数据库现有指标“仍沿袭着20世纪90年代的设计思路,注重文献典藏功能,重点统计的是馆舍面积、工作人员情况、馆藏数量与经费数量”^[12],



没能完全客观地反映图书馆转型和创新服务带来的变化。伴随着大数据、云计算和人工智能的迅猛发展,当前,以特色资源、数字资源和人力资源为主的资源建设是提升图书馆核心竞争力的基础,将信息技术应用于图书馆设备设施管理的技术驱动是提升图书馆核心竞争力的保障,围绕读者服务、学科知识服务、阅读推广和空间再造的服务创新是提升图书馆核心竞争力的根本动力。许多高校馆尝试在资源、技术、服务等方面转型升级,深挖本馆特色资源,建立学校文库、机构知识库,收藏古籍珍本和善本,以及具有学科特色、区域特色、专题特色的特藏,开发数字人文项目。大数据背景下,传统的图书馆管理系统难免捉襟见肘,一些图书馆已着手调研、实施新一代图书馆服务平台。清华大学图书馆和北京师范大学图书馆已上线 Ex Libris 公司的 Alma 管理平台取代原有的管理系统,华中科技大学图书馆和电子科技大学图书馆应用了 Innovative 公司的 Sierra 系统。在服务方面,以清华大学图书馆为首,高校馆相继组建学科馆员队伍,深入院系或课题组了解学科与科研信息资源的需求,开展学科资源建设,进行学术动态分析,提供学科知识服务;部分高校馆提供专利查新、专利技术分析、专利素质教育等涉及知识产权的信息服务。同时,全媒体阅读推广、官方微信公众号推送和信息素养教育等多维立体的

读者服务如火如荼;基于“第三空间”概念的空间再造及创客空间建设方兴未艾。

建议酌情补充以下指标:(1)A3.1 在编职工增加学科馆员和阅读推广馆员人数的统计;(2)D1.1.1.2 统计善本古籍量;(3)补充特色资源建设情况,统计具有学科、区域、专题特色及特定时期的特藏数据库数量;(4)K6.1 图书馆管理系统名称中增加“Alma”和“Sierra”两个选项;(5)学科服务中可量化统计的数据,如学科服务专定题项目数、学科分析报告份数;(6)知识产权服务中可量化统计的数据,如知识产权培训、知识产权咨询、专利分析报告等;(7)阅读推广活动统计,如实体书展举办期数和参展图书册数、举办阅读交流活动(论坛讲座、读书会、阅读沙龙、真人图书馆、文化体验等)的场次数和参加人数、全年通过线上和线下荐读的书目量;(8)图书馆微信公众号的关注用户量、发布推文篇次阅读量、点赞量、分享人数等;(9)在读者行为方面,增加全年入馆人次的统计。

3.4 电子资源计量不够规范

3.4.1 电子资源计量标准不统一

电子资源数据是事实数据库填写中最令人头疼的部分,一方面电子资源数据库种类繁多、数量庞大,另一方面电子资源计量还处于探索阶段,缺乏统一规范的标准。国内外电子资源计量标准存在许多版本,比较有代表性的如表 1 所示。

表 1 国内外电子资源统计标准

	名称(简称)	发布机构	发布时间	简要说明
国内	高等学校图书馆数字资源计量指南(高校指南)	高校图工委和 CALIS 管理中心	2004 年、2007 年	规定了高校图书馆数字资源计量的指标
	图书馆数字资源统计标准和应用指南(国图指南)	国家图书馆	2010 年	定义了数字资源统计利用的核心指标
	中华人民共和国文化行业标准——图书馆数字资源统计规范(WH/T 47—2012)	文化部	2012 年	以国图指南为蓝本
国外	国际图书馆统计标准(ISO 2789)	国际标准化组织(ISO)、国际图书馆协会联合会(IFLA)	始于 1974 年,最新版 2013 年	适用于各类图书馆的国际标准,规定了电子图书馆绩效服务指标
	基于网络的信息资源利用统计评价指南(ICOLC Guideline)	国际图书馆联盟联合组织(ICOLC)	1998 年	国际标准,规定电子资源统计必须包括登录数、检索数、浏览的次数等
	信息服务和使用:图书馆和信息服务机构统计指标——数据字典(NISO Z39.7)	美国国家信息标准协会(NISO)	始于 1968 年,最新版 2013 年	美国图书馆的国家标准,规定了电子馆藏与电子服务的统计指标
	电子资源绩效评价项目(E-Metric)	美国研究图书馆学会(ARL)	2001 年	针对研究性图书馆的电子资源使用及服务评估指标体系
	网络化电子资源在线使用统计(COUNTER)	英国 COUNTER 项目组	始于 2002 年,最新版 2019 年	规范定义了电子资源需要统计的数据量、统计数据计算的策略、输出报告的内容、格式等 ^[13] 。被国外数据库商广泛采用



事实数据库依据的“高校指南”是针对高校馆制订的电子资源计量标准,但统计规范不够细致,对数据库商缺乏约束。目前,高校馆在统计电子资源数据时,很大程度上依赖数据库商提供数据,而不同数据库厂商或数据库平台提供的数据库格式各异,标准不一,差别很大,导致图书馆无法对所购电子资源进行有效的横向统计分析^[14]。计量标准的不统一、不规范,增加了统计工作的难度,导致事实数据库采集的不同院校图书馆的电子资源数据缺乏可比性,数据的准确性也难以判断。

3.4.2 电子资源利用统计不全面

关于图书馆电子资源利用情况,事实数据库只统计了G6(电子资源下载量)一项。下载量固然是反映电子资源使用情况的重要指标,但只统计这一项并不够,浏览量、检索量、被拒访问量、媒体资源点播量、成功请求量等也是较重要的指标。而且,G6没有明确的指标解释,统计者往往根据各自的理解提供数据,有的认为下载保存到本地为下载量,有的把全文浏览即视为下载量,有的把文献传递量记入下载量。对同一用户点击下载同一文献的情况,有的重复累计,有的则去重了。还有的统计者对指标自行定义,将“没有下载数据的则选取与下载量含义最为贴近的指标替代,并认定为等同下载”^[15]。这种理解的偏差,必然对统计结果造成不同程度的影响。

3.4.3 电子学位论文归类欠妥

事实数据库依据“高校指南”(2004年),将学位论文纳入电子书统计中。2007年修订的“高校指南”,“电子图书(包括与图书类似的出版物)”已不包括电子学位论文^[16]。学位论文没有出版号,不属于正式发行的出版物,建议把电子学位论文从电子图书中剥离出来,单独统计。

4 事实数据库管理运维的问题分析

4.1 指标修订更新不及时

伴随着大数据时代来临,网络技术、量子计算、人工智能快速发展,图书馆的资源、技术和服务实力显著提升,在空间文化建设、电子资源建设、信息服务能力、阅读推广能力等方面的业务不断拓展,馆内自助设备、电子资源利用、网站及新媒体访问每天产生大量数据。然而,事实数据库的统计指标已使用近十年没有进行过重大修订,部分指标已不能完全

客观地反映当前图书馆的服务能力与核心竞争力,急需根据图书馆的业务发展和变化进行修订,使其更好地满足大数据时代的数据采集与统计需要,更准确地揭示图书馆资源、技术、服务的发展现状。

4.2 填报率较低、准确性不高

事实数据库自2001年实现网上填报,初期只有456所院校在系统中录入了数据,填报率非常低。在图工委的大力宣传和推广下,尤其是修订升级以后提交数据的院校逐渐增加,如表2所示。

表2 2013—2018年事实数据库填报情况

时间(年)	2013	2014	2015	2016	2017	2018
院校数量(所)	755	707	821	932	1106	1215

表2的统计原则是:所有数据都没有填写的视为没有填报,只要有数据,哪怕这所院校只有一个指标项填写了数据,都计入填报数据的院校数量中。2019年6月,教育部公布的全国高等学校共计2956所,事实数据库中填报数据的院校只占三分之一左右,整体填报率不理想。而且,不少院校只填写了部分数据,没有数值的空项较多,部分指标的填写率低于填报数据院校数量的20%。又由于指标含义不够清晰、培训不到位、各馆理解不同、基层采集数据困难等因素,已报数据准确性也不够高,有些数据能看出明显错误。

4.3 统计和分析功能有限

事实数据库目前提供了数据填报、信息查询、统计汇总和数据分析等功能,具体如图2所示。

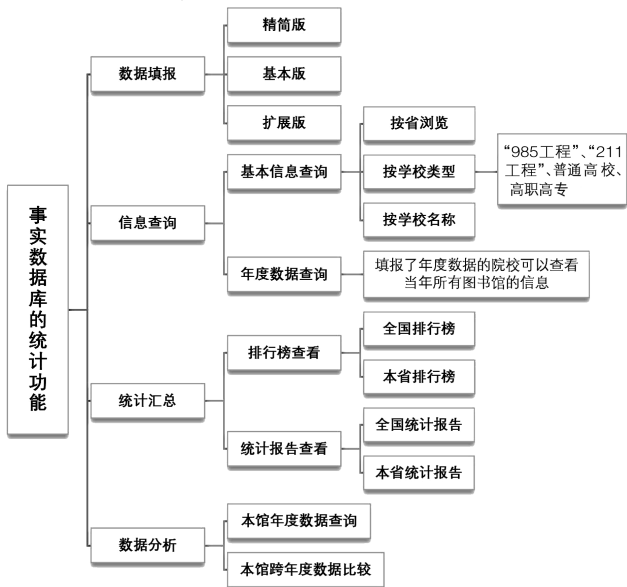


图2 事实数据库的统计功能



这些功能满足了各高校馆填写本馆数据、查询全国图书馆基本信息和年度数据的基本要求,但使用中仍存在不便之处,提供的统计与分析功能有限。(1)各高校馆填报数据时,建议最好一次性填写完成,因为提交之后再修改,不仅操作不方便,修改后的数据有可能存不上,在生成数据时可能出现1个指标项同时有多个数值,需要人为判断哪个数值更合理、更有效。在导出EXCEL表时,由于数据个数增加了,导致数据项串行,几个图书馆的数据导出后不能直接进行比较,须先进行人工纠错,删掉多余数据;(2)进行跨年度数据比较时,图书馆用户只能看到本馆的,省图工委可以查看本省跨年度数据比较和馆际数据分析,只有教育部高校图工委才有权查看所有图书馆的跨年度数据,如此设置限制了绝大多数图书馆用户对事实数据库数据的深度分析研究;(3)统计汇总中提供的排行榜不多,更新较慢。截至2019年12月底,系统中公布的2018年度全国排行榜只有3个,分别是“2018年馆舍总面积排行榜”“2018年电子资源购置费”“2018年在编职工”,2017年度没有公布全国排行榜。笔者能查看的本省排行榜只更新到2016年度,2017年度和2018年度都没有。点击“统计报告查看”,显示“暂时没有数据”。

5 完善高校事实数据库的建议

5.1 完善指标体系,及时修订数据指标

科学、规范、完善的图书馆统计指标体系是全面展现图书馆发展现状,对图书馆进行科学管理和精准评估的重要依据,对推动我国高校图书馆发展意义重大。应在事实数据库的基础上,整合图书馆其他统计渠道的数据,及时修订事实数据库中设置不得当、逻辑不清晰、分类不全面、表述欠准确的统计指标,尤其要尽快制订电子资源统计规范,完善相关指标解释,并增补反映图书馆新增技术和业务的统计项目,不断完善指标体系。

5.2 强化管理机制,保障数据填报率

一直以来,图工委对推动图书馆数据统计工作发挥了重要作用,但管理机制不完备,缺少行政法规的强制力,执行力度比较低。美国国家教育统计法赋予国家教育统计中心(NCES)执行全国图书馆统计工作的权利,法律的强制力使得被调查的各个图书馆必须配合,保证了图书馆统计数据的有效回收

(回收率基本在90%以上),并且捏造数据是一种违法行为,在一定程度上避免了数据造假情况的发生^[17]。我国可以一方面推动图书馆统计立法,另一方面赋予图工委更大的权力,健全图工委的监管机制,使其具有更强的推动力和执行力,为开展统计工作提供根本性保障。第三,建立激励机制,激发全国各院校统计和填报数据的热情与动力,提高事实数据库的填报率,全面、准确地记录我国高校图书馆发展状况。

5.3 加强培训推广,提高数据准确率

统计数据是事实数据库极其重要的组成部分,数据质量决定了它的价值。高校馆落实数据统计工作遇到的问题主要来自三个方面:(1)有些馆领导不够重视,没有组织足够的人力从事统计工作,馆里缺乏数据收集、保存机制;(2)基层统计人员流动性大、不够专业,对事实数据库统计指标产生误读或误解。现在图书馆很少设专职人员从事统计工作,多由办公室人员兼任,日常有大量繁杂的行政事务干扰,且图书馆多采用岗位聘任制,人员流动性较大,可能造成统计工作缺乏连续性,基础数据不完整;(3)统计数据涉及图书馆多部门的交叉业务,来源多、统计口径不一,导致数据重复统计或遗漏。另外,虽已进入大数据时代,很多高校仍采用Excel表格对日益庞大复杂的数据进行管理和统计,不仅工作量大、效率低下,还容易出错,影响了数据的准确性。

因此,不仅要着力完善统计指标,还应加大培训和宣讲力度,使领导更加重视,并给负责统计的馆员提供专业、系统的指导,加强各院校统计人员的业务沟通与经验分享。同时,电子资源数据多由数据库商提供,还应加强对数据库商的规范性指导,使其提供数据时采用统一、规范的标准。另外,还可以考虑开发一个供高校馆使用的、方便统计人员进行数据采集的数据管理系统,辅助图书馆将采编、流通、信息服务、学科服务、综合管理等业务整合起来,自动生成统计报表,真正帮助统计人员提高工作效率,提高数据的填报率和准确率。

5.4 优化统计平台,增强数据利用率

经过20多年的数据采集和积累,事实数据库已拥有比较翔实的数据资料,但因多种原因,这些数据没有公开,尚未充分挖掘和分析利用,目前依据事实数据进行分析、研究的文献还不多。可以通过以下措施尽可能多地采集有效数据,挖掘数据价值,提高



数据利用率:(1)优化统计平台,使数据录入、修改及导出更加便捷;(2)进一步开发统计功能,增加数据分析维度,生成可视化图表;(3)组织专家整理分析数据,及时发布统计报告等。

6 结语

通过不断改进和完善,事实数据库的统计指标会更加科学合理,采集的数据更加准确有效,提供的功能更加便捷丰富,为辅助管理决策、支持教学科研、加强过程监控、优化考核评估等提供强有力的数据支持,从而可以在图书馆事业发展中发挥更大作用。

参考文献

- 1 樊伟红,李晨晖,张兴旺,等.图书馆需要怎样的“大数据”[J].图书馆杂志,2012,31(11):63-68,77.
- 2 谈鹤玲,魏本力.教育部高校图书馆事实数据库统计功能及其指标研究[J].科技情报开发与经济,2015,25(16):105-107.
- 3 胡秀梅,何雪梅,李睦,等.美国研究图书馆协会统计与“高校图书馆事实数据库”统计的对比研究[J].大学图书馆学报,2011,29(6):95-98.
- 4 何雪梅,李睦,胡秀梅,等.“高校图书馆事实数据库”统计指标修订情况概述[J].大学图书馆学报,2011,29(6):86-91.
- 5 黄运红,李书宁,弓建华.高校图书馆数据统计指标体系探析[J].高校图书馆工作,2019,39(6):38-42.

- 6 姚晓霞.高校图书馆事实数据库的设计理念及其实现[J].大学图书馆学报,2013,31(5):114-117.
- 7 同5.
- 8 教育部高校图书馆事实数据库系统[EB/OL].[2019-08-02].
<http://libdata.scal.edu.cn/Index.action>.
- 9 同8.
- 10 李红.实行整体外包服务的三个公共图书馆的建设[J].图书馆杂志,2017,36(1):33-37,32.
- 11 敖雪蕾.研究型高校图书馆的外包服务管理研究[D].上海交通大学,2015.
- 12 同5.
- 13 杨巍,叶仁杰,吴元业,等.COUNTER Release 5 的新特征及其应用研究[J].大学图书馆学报,2020,38(1):18-25,41.
- 14 张计龙,殷沈琴,汪东伟.基于 COUNTER 的电子资源使用统计中的标准问题探讨与研究[J].图书馆理论与实践,2016(5):95-100.
- 15 祝红艺,罗红彬.数字资源计量规范研究——基于事实数据库专项数据填报的思考[J].现代情报,2016,36(1):114-117.
- 16 高等学校图书馆数字资源计量指南(2007年)[EB/OL].[2019-08-02].
<http://www.scal.edu.cn/tjpg/201311191006>.
- 17 陈琳.美国图书馆统计工作研究[J].新世纪图书馆,2010(5):77-79,93.

作者单位:北京师范大学图书馆,北京,100875

收稿日期:2020年2月8日

修回日期:2020年8月2日

(责任编辑:支娟)

Study on the University Library Fact Database in the Era of Big Data

Huang Yunhong

Abstract: The University Library Fact Database is designed and developed by the steering committee of the university library of the Ministry of Education. It is a set of online filling and statistics system designed to collect fact data of the university libraries in China. In order to better meet the development needs of libraries in the era of big data and through analyzing the statistical indicators, electronic resource specifications, and system management of the fact database, this paper puts forward suggestions on how to improve the statistical indicator system, how to guarantee the data filling rate, how to improve the data accuracy rate and how to enhance the data utilization rate, with hope to provide reference for improving the construction of the fact database.

Keywords: Big Data Era; Academic Library; Fact Database; Data Statistics; Statistical Indicators