

大数据背景下大学生使用数字阅读平台习惯偏好

——以豆瓣阅读为例

李鑫 孙嘉茹

(通化师范学院数学学院, 吉林 通化 134000)

摘要:随着互联网时代的到来,大学生数字阅读已成为一种趋势,在我国一系列数字阅读平台中,豆瓣阅读是一家原创文学写作平台,但是相对于京东,掌阅,当当来说,它仍属于较为小众的阅读平台。本文通过阅读查找文献,设计问卷,分析大学生使用数字阅读平台的数据,综合分析其习惯偏好。从而引导大学生修正自己的数字阅读习惯,进一步提升大学生数字阅读的效率。

关键词:大学生;数字阅读行为;豆瓣阅读;大数据背景

随着网络与科技的发展,世界随之进入一个崭新的互联网时代,大众的阅读模式也从传统的纸质媒介转变成新的数字媒介,数字阅读正逐步成为人们阅读的主要途径,大学生群体是数字阅读的主力军,在这种情况下,高校学生的数字阅读行为习惯引起了人们的广泛重视。本文将就数字阅读习惯中一个方面——阅读平台(以豆瓣阅读为例)展开研究。

一、研究现状

在过去的几年里,随着数字阅读的发展,许多领域的学者都已积极地投入到数字阅读的研究中,但是关于数字阅读平台使用习惯偏好数据的文献不多。本文将豆瓣阅读为例从多个方面进行数字阅读平台使用习惯偏好的数据进行公布和分析,为今后相关产业的研究起到抛砖引玉的作用。

二、数据的获取

(一)豆瓣阅读平台

豆瓣阅读在用户授权的情况下,能够记录到各种类型的数字阅读行为数据,经过多年运行,产生了海量的精准数据,所有热门的图书标签共有145个,从数字图书分布看,“科技”“经管”“生活”“文化”“文学”“流行”类的图书分别有9804,11916,19564,32346,21288,25232本,分别占比8.16%,9.92%,16.28%,26.92%,17.72%,21.00%。

(二)数字阅读使用情况

自豆瓣阅读平台上线后,第一天注册人数达728人,两个月后达到3056人,占一年来注册人数的37.3%,截止到2022年7月,东北师范大学有56.2%的在校大学生注册该软件。记录显示注册的学生中有78.3%的人实际进行了阅读,而剩余的学生只是进行了注册,并无实际的阅读行为。

(三)数字阅读数据获取途径

此次调查以东北师范大学在校大一至大四本科生为研究对象,学生在安装完豆瓣阅读APP之后,通过手机号进行注册,在用户

登录时,会被要求授权服务和数据的访问权限,所以在后台,系统会详细地记录授权的用户在该平台上进行的数字阅读行为,其中涵盖了每天注册情况,实际阅读情况,阅读时长,阅读时间区段,阅读类别等一系列大数据。

三、2021年度豆瓣阅读平台“阅”的数据

(一)用户阅读时间

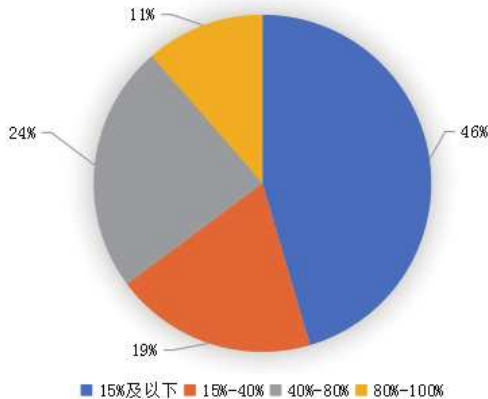
豆瓣阅读APP在用户授权的情况下提供了阅读时长数据,包括“日阅读时长”“周阅读时长”“月阅读时长”“年阅读时长”,可以清晰地知道读者用于阅读的时间。

(二)用户阅读时段

豆瓣阅读APP记录的用户数字阅读时段如图1所示。2021年用户阅读时间段集中在12:00-24:00,其中18:00-21:00为阅读高峰时段,3:00-6:00用户阅读量最低。平台通过掌握这些数据后可以更精准地推送信息,实现用户服务最优化。



图2 阅读进度分布



(三) 阅读进度

除了下载之外,数字阅读平台还能将读者的翻页、笔记、注释等具体的阅读行为记录下来,它不仅能知道“读了没有”,还能知道“读了多少”,这是数字阅读一种特有的功能,它会将每个人的阅读进度都记录下来。

如图2数据显示,基本完成阅读的人数占比为23.91%,完全阅读的人数占比为11.40%,从调查结果来看,还是会有很多同学下载后不进行阅读的,从而大大降低了数字化阅读的价值。

三、豆瓣阅读平台资源种类数据

(一) 资源种类利用

如表1为评分≥8.0且评论人数10w+以上的图书(10本)

表1

标签	图书	作者	评分	评论人数
小说	追风筝的人	卡勒德胡赛尼	8.8	269893
外国文学	小王子	圣埃克苏佩里	9.0	219512
小说	白夜行	东野圭吾	9.1	184912
小说	挪威的森林	村上春树	8.0	182859
小说	三体	刘慈欣	8.8	135871
小说	达芬奇密码	丹布朗	8.2	125447
小说	红楼梦	曹雪芹	9.5	117547
小说	活着	余华	9.1	126869
随笔	看见	柴静	8.8	114862
小说	简爱	夏洛蒂勃朗特	8.5	104665

通过图表可以发现,名单中大部分都是开卷有益的好书。

表2为资源种类实际阅读数据

资源类别	实际阅读册	实际阅读率
职场进阶	2811	33.9%
家庭教育	3675	45.2%
互联网+	1195	36.1%
人物传记	652	26.9%
中外名著	739	66.3%
成功励志	792	60.7%
少儿频道	289	68.2%
社会科学	256	17.3%
教材教辅	382	24.5%
影视原著	3673	81.2%

2021年中外名著的实际阅读率为66.3%,而社会科学的实际阅读率只有17.3%,这表明不能仅依照点击率就直接判断资源的利用状态。

(二) 图书阅读偏好

在学生的数字阅读中,其行为呈现出多元化的特征。对于平台来说,要保障读者读到“好书”,数字阅读可以为“好书”这个概念提供一定的依据。豆瓣阅读平台的数据可以反映读者数字

阅读行为。图5统计了2021年阅读进度在70%以上的图书可为平台精准推送提供数据支持。

表3

排行	图书	下载次数	阅读进度在70%以上的次数	占比
1	小王子	687	542	79%
2	白夜行	917	689	75%
3	最好的我们	1046	758	73%
4	盗墓笔记	2173	1565	72%
5	活着	1965	1395	71%
6	后宫甄嬛传	892	553	62%
7	斗罗大陆	1704	903	53%
8	人生海海	516	170	33%
9	月亮与六便士	673	195	29%
10	摆渡人	539	151	28%

五、概要

大学生作为高水平人才的摇篮,理应具备良好的数字信息资源获取、分析、筛选和内化的能力。本文从大学生使用数字阅读平台习惯偏好展开了研究,了解到当代大学生的数字阅读行为习惯,为今后开展数字阅读推广工作打下了坚实的基础。由于大学生的阅读行为、阅读习惯等都比一般群体更加活跃,更加年轻。因此对大学生数字化阅读行为进行深入的分析与研究,不但可以有效地促进文献资源的利用,同时能让我们更好地了解读者偏好的资源种类,也能了解到读者使用数字阅读的效果,从而为读者提供更好的阅读体验。

数字资源对科学技术的发展以及人才的培养有着积极的影响,所以研究数字阅读行为是非常有必要的,随着数字阅读的普及,数字阅读平台会掌握大量精准的阅读数据,并且挖掘数据的能力有所增强,这些数据经过整合之后形成一定的模型,提高资源共享的能力,以精确的服务来达到“悦”读的目的,从而有效地促进阅读,引导大学生形成良好的数字阅读习惯,进一步提高大学生进行数字阅读的效率。

参考文献:

- [1] 葛亮. 互联网背景下豆瓣阅读自出版模式研究[J]. 哈尔滨师范大学社会科学学报, 2019, 10(5): 4.
- [2] 杨新涯, 王莹, 刘义勇, 等. 精细化数字阅读行为数据研究——以重庆大学京东阅读平台为例[J]. 图书馆论坛, 2019, 39(6): 9.

通化师范学院2022年大学生创新创业训练计划项目(202210202066)成果。