

资源发现系统用户行为特征分析

——以北京大学图书馆为例

□ 吴亚平* 李晓东 艾春艳

摘要 基于元数据预索引的网络级资源发现系统是用户统一检索、发现学术资源的主阵地,多年以来积累了大量的、支持横纵向比对的用户行为数据,综合来看是可用的数据集。笔者通过剖析用户发现与获取电子资源的环境和行为数据集,运用统计分析、热点演化分析、聚类分析等方法,分析得出用户在发现系统上的检索量呈波动上升趋势,检索深度不断增强;发现系统能体现出用户周期性的、突发性的、结合时事热点产生的资源检索需求;来自不同供应商的电子资源发现量分布呈现长尾特征;用户通过不同渠道偏好访问不同电子资源平台,发现渠道的通畅对电子资源利用有明显的促进作用。最后,提出完善用户行为数据追踪、质量控制、关联分析等建议,以增强基于数据优化工作及服务的能力。

关键词 资源发现系统 用户行为 特征演化 聚类分析

分类号 G250

DOI 10.16603/j.issn1002-1027.2023.01.010

用户需求是图书馆一切工作的出发点和归宿,大学图书馆必须要了解用户所在,通过多种方式全面准确地了解用户需求^[1]。图书馆的资源建设和服务存在规律,应努力通过大数据技术等新兴技术手段,发现更多的隐藏在图书馆日常工作中的规律,在规律的指引下,开拓更多的甚而来不及命名的原创型服务^[2]。美国国会图书馆在其《2022—2026 数字馆藏战略评论》(*Digital Collections Strategy Overview 2022—2026*)中也明确强调要通过研究用户行为来拓宽、深化数字馆藏的利用^[3]。深入了解用户需求已成为共识,笔者以北京大学图书馆为例,剖析、分析在电子资源发现环境中积累多年的用户行为数据,并基于用户特征和需求提出用户服务优化建议。

1 相关研究

发现系统的内容集成、结果优化、统一发现等方面的优势在逐渐加强,服务场景也在不断拓宽,如嵌入到学习管理平台^[4],嵌入到数字教参系统^[5],用户使用量稳步上升,用户行为数据规模不断扩大,对其进行分析利用的重要性更为凸显。发现系统用户行

为研究是伴随发现系统出现、发展的核心议题,国内外学者已关注到不同类型资源在发现系统中的用户使用差异^[6]、不同终端界面的用户行为差异^[7]等多方面的问题。其中,用户检索行为分析是一个重要的方面,如搭建搜索日志中间平台^[8]或嵌入 Google Analytics^[9]等网页分析插件收集用户的搜索日志,对用户搜索行为进行各有侧重的分析^[10],基于用户基本信息、访问行为、搜索行为等数据,通过智能计算,挖掘出用户的兴趣度^[11]等。值得关注的是,当前围绕发现系统的定量研究偏少^[12],也有学者提出,发现系统本身积累多年的大量用户行为数据,能够反映电子资源的发现、使用效果,支持对不同的数据资源进行横向比较等,是可用、可对比的比较理想的分析数据源,并呼吁对其加大分析利用^[13]。一项 70 所高校图书馆对“发现系统需要扩展哪些功能”的意见反馈显示,最受欢迎的功能包括可发现统计分析、可视化分析、热点研究,此外,获取全文数、页面点击数、按数据库商统计等方面的统计需求也很高^[14]。这表明,图书馆越来越重视且越来越需要通过发现系统多年积累的用户数据来探究用户行为特

* 通讯作者:吴亚平,邮箱:wuyyp@lib.pku.edu.cn。



征、用户关注的热点、资源的发现情况等,而对于期待发现系统扩展的统计功能,其中部分功能已具备,如获取全文数、分面点击数、按数据库提供商统计等,这体现出图书馆员对发现系统统计功能的了解不够深入,或者尚未对相关统计数据进行了摸排和分析利用。因此,对已有的发现系统统计分析功能、统计数据进行了梳理、分析,为优化服务提供有益参考,具有一定必要性和可行性。

2 概念界定与数据遴选

2.1 概念界定

资源发现系统的用户行为反映用户利用发现系统工具的活动方式和偏好习惯,本文讨论的发现系统工具主要包含北京大学图书馆的“未名学术搜索”和电子期刊/图书导航。文中的电子资源主要指发现系统工具支持发现的电子图书和电子期刊,并将用户通过发现系统工具获取的全文资源量记为电子资源平台的发现量。电子资源平台对应发现系统中的 Provider 概念,即提供支持发现元数据的供应商对应的所有电子资源。对用户行为特征的分析,首先扫描用户发现和利用电子资源全文的环境,梳理

从不同渠道获取电子资源全文的路径及相应产生的、可获取的用户行为数据;其次从数据完整性、可靠性、可比性、数据对应行为的典型性等方面对数据集进行评估,选取分析数据集,通过统计分析、聚类分析等方法^[15],得到用户访问、检索、兴趣演化、热门发现电子资源平台、用户访问不同电子资源平台的渠道偏好等行为特征;最后基于结论提出用户服务、资源揭示、数据追踪等方面的优化建议。

2.2 电子资源发现与获取环境及用户行为数据分析与遴选

北京大学图书馆自 2011 年引进 Summon 以来^[16],逐渐形成了以“未名学术搜索”学术资源发现门户、360core 电子期刊/图书列表、360link 全文链接服务以及 Journal Linker 原文查找工具为核心的电子资源发现与获取环境。电子资源全文发现与获取包含 7 种行为路径,可归纳为通过进入电子资源平台查找并获取全文、通过在发现系统检索并点击检索结果获取全文、通过其他基于 360link 链接服务器构造 OpenURL 获取全文、通过电子期刊/图书导航检索查询到达资源详情页获取全文四大类别,相应产生的、可获取的行为数据标注详见图 1。

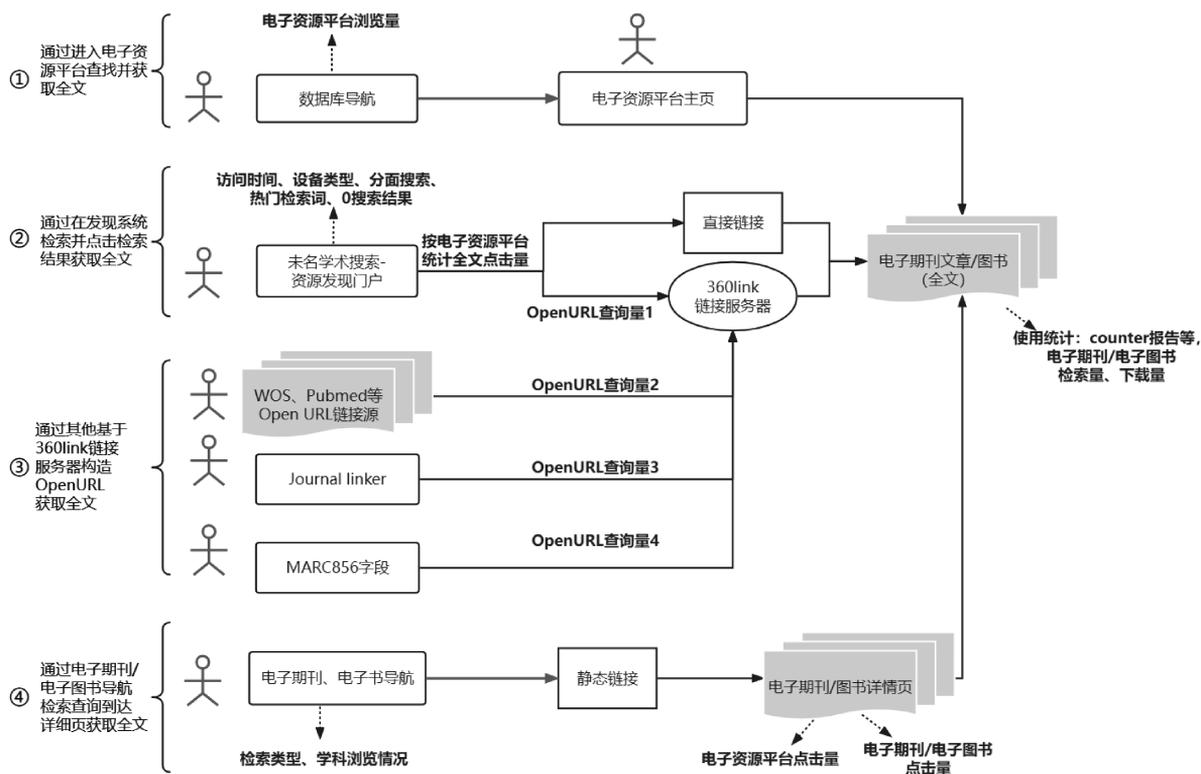


图 1 电子资源全文发现与获取路径与行为数据(目前可获取数据相应字体已加粗)

通过梳理发现,以上数据目前还难以全面反映用户的发现行为,表现为:部分用户获取全文行为数据暂缺,如在“未名学术搜索”的搜索结果页面通过直接链接获取全文的量等;数据粒度未完全统一,如在电子期刊/图书导航中支持按期刊来统计发现量,但“未名学术搜索”却无法按期刊来统计发现量。

虽然当前数据集存在以上不足,但仍然有积累多年的、高质量的、支持横纵向对比的、不可忽视的数据集,包含了用户依托发现系统工具检索、发现电子资源行为的多个重要方面。可归纳为:(1)检索行为数据,即“未名学术搜索”和电子期刊/图书导航中的访问、检索方式、热门检索词等数据;(2)电子资源平台的发现数据,即按同一供应商统计在“未名学术搜索”和电子期刊/图书导航检索结果中的点击量;(3)数据库导航访问数据,即数据库导航中不同电子资源平台的浏览量;(4)使用量统计数据,包括 Counter 报告中的电子期刊/图书检索量、下载量等电子资源利用量。下文详细展现基于以上数据的分析过程、结论。

3 用户行为特征与演化分析

3.1 用户的资源检索需求在稳步上升,检索技巧的应用能力逐步提升,整体体现出周期性规律

“未名学术搜索”中的用户检索行为包括基于统一搜索框的基本检索和在检索结果基础上进行的分面检索,2017—2021年(以下简称近五年),用户检索资源的频率在波动上升(2020年受疫情及短期停课等影响,数据有所减少),每个基本检索、分面检索会话中用户的检索量在上升,2021年分别达到2.92、4.51,详见图2,这表明用户逐渐熟练使用检索工具,能够频繁切换和应用检索词,使用深度在增强。近五年分面检索量/基本检索量的比例均值为26%,即1次基本检索有0.26的几率继续使用1次分面检索,分面检索类型按使用量降序排列依次为 Subject Terms(主题词)、Discipline(学科)、Content Type(内容类型)、Library(馆址)、Language(语言)、Is Full Text(是否有全文)、Is Open Access(是否开放获取)等,其中 Subject Terms(主题词)分面占到了八成以上(83%),可见主题是用户检索资源的重要维度。

用户在一年中有两个使用高峰期,不同检索类型具有基本一致的走势,详见图3。整体来看,春季学期开始和秋季学期末是用户检索的高峰期,这与相关研究中通过嵌入网站分析工具 Google Analytics 得到的用户访问趋势基本一致^[17]。

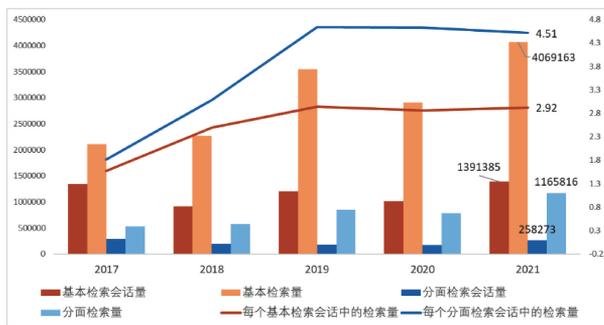


图2 2017—2021年基本检索与分面检索走势

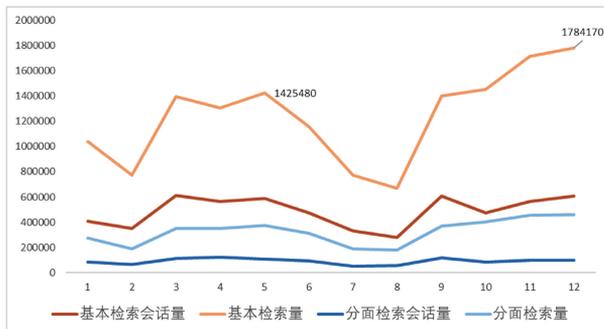


图3 2021年按月度统计基本检索与分面检索平均量走势

用户逐渐将电子期刊/图书导航作为一种相关资源的检索入口。2012—2021年(以下简称近十年)中,自2011年 Summon 系统在图书馆启用后,起初用户使用呈快速增长趋势,2013年达到巅峰,而后有部分回落并逐渐趋于平稳,2019年后整体呈稳步缓慢增长趋势。用户行为类型可细分为浏览行为(按学科浏览,按A—Z浏览)、精确检索行为(标题等于、ISSN等于)、模糊检索行为(标题开头等于、标题包含),其中模糊检索行为占比最多且呈上升趋势,浏览行为次之,且在缓慢下降,精确检索行为最少,趋势较为平稳。可见,用户有模糊需求且进行模糊检索的行为在增多,尤其是“标题包含”的检索;泛化浏览的行为有所下降;用户有明确、细化的需求,且进行精确检索的行为在降低,详见图4。

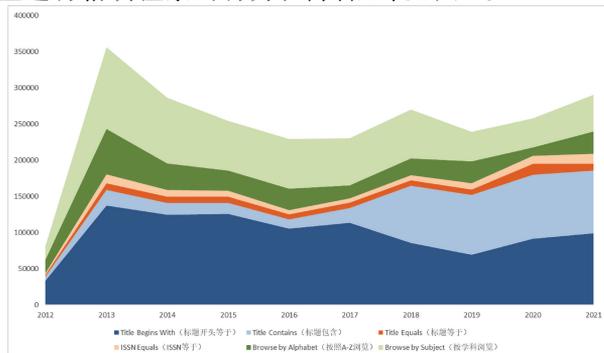


图4 2012—2021年不同浏览类型使用量走势



3.2 “未名学术搜索”能体现出用户周期性、结合突发事件及时事热点的信息检索与资源需求

检索词是用户检索学术资源需求的表达,热门检索词是用户当下资源需求和兴趣的集中体现。“未名学术搜索”记录每月前 500 个热门检索词,2017—2021 年共有 29980 个热门检索词,其中中文检索词占比 77%,高级检索式占比 4%。从热门检索词中挖掘焦点词,即前期不是热门检索词但在当

期出现为热门检索词并在后期保持一定检索热度的词,可以看出用户关注点和信息需求的变化。焦点词需要比较检索词在历史周期、当前及未来周期的热度情况,笔者将焦点词定义为在[前 6 个月—当前月]的历史周期内不是热门检索词,但[当前月—后 2 个月]均为热门检索词,且[当前月—后 2 个月]每月的检索次数均大于平均水平(即 36 次),得到 71 个焦点词,详见表 1。

表 1 2017—2021 年热门检索词中的焦点词

年份	月份	焦点词	年份	月份	焦点词	年份	月份	焦点词
2017	9	无机化学	2019	9	Jane Eyre	2020	9	经济学原理
2018	2	扫描隧道谱	2019	9	菊与刀	2020	9	论语
2018	4	学位论文	2020	1	新型冠状病毒	2020	11	邓小平时代
2018	5	核心期刊要目	2020	2	Covid-19	2021	1	中文核心期刊要目
2018	7	东野圭吾	2020	2	新冠肺炎	2021	1	北大核心期刊
2018	7	中文核心期刊要目	2020	2	海洋科学导论	2021	1	民事诉讼法
2018	7	北大核心目录	2020	3	查重	2021	3	国际组织
2018	7	核心目录	2020	9	习近平谈治国理政	2021	3	尼各马可伦理学
2018	9	汉书	2020	9	刑法学	2021	8	哈利波特
2018	9	爱的艺术	2020	9	庄子	2021	8	普通化学
2019	2	查重	2020	9	社会心理学			

可见,焦点词涉及多个学科的研究主题、经典著作、作者等多个维度,体现出用户周期性的、结合突发事件及时事热点产生的信息需求。如在 2018、2021 年分别发布了 2017、2020 年版《中文核心期刊要目总览》,“中文核心期刊要目”“北大核心期刊”等词成为了相应年度的焦点词,体现出用户周期性的关注度,类似地在集中撰写论文的春季学期,“学位论文”“查重”等焦点词也随即出现。在突发事件及时事热点下,用户突增的资源需求也得以体现,如 2020 年 1—2 月持续出现的“新型冠状病毒”“Covid-19”“新冠肺炎”等焦点词,反映出了抗击新冠肺炎疫情初期,用户对此主题的资源需求大幅提升,以上印证了“未名学术搜索”作为学术资源发现的门户,是用户搜索相关资源的重要入口。

3.3 电子资源发现量分布呈长尾特征,支持发现的资源规模、呈现的结果优先顺序等是关键影响因素

“未名学术搜索”和电子期刊/图书导航是用户发现电子资源全文的两个核心渠道,两个渠道中以不同的电子资源平台为单位统计的 2021 年电子资源发现量分布均呈长尾特征(见图 5,横轴为电子资

源平台,纵轴为电子资源发现量)。

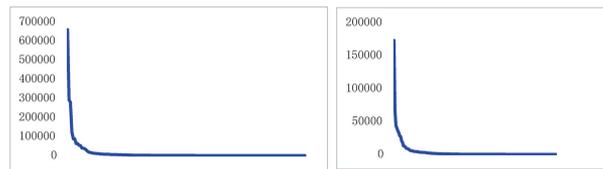


图 5 “未名学术搜索”(左)、电子期刊/图书导航(右)中电子资源平台发现量的长尾分布

以下分别列出 2021 年“未名学术搜索”、电子期刊/图书导航中的热门发现电子资源平台(即发现量总和占总量 80%的头部热门电子资源平台)并分析其主要影响因素。

用户通过“未名学术搜索”可发现馆藏数据、电子图书全文、电子期刊文章全文等多种资源类型。2021 年,“未名学术搜索”中的馆藏目录、机构知识库或一些特有资源如燕京大学学位论文等资源占总发现量的 28%,对剩余的电子资源平台结合二八定律测算出热门电子资源平台,根据热度由高到低排序为:维普、万方、Elsevier、ProQuest、方正、Wiley、Springer Nature、JSTOR、EBSCOhost、知网、Taylor & Francis、SAGE、DOAJ。电子资源平台在“未名



学术搜索”的发现热度与三方面的因素直接相关:一是在 Summon 知识库中所包含的该电子资源平台的电子期刊文章/电子图书全文的数量,即 Summon 的中央索引中心(Central Discovery Index, CDI)中拥有该平台电子资源元数据的数量;二是图书馆的电子资源在发现系统后台知识库的配置情况,是否将可访问的电子资源全部配置完成;三是该电子资源平台在同一检索结果下的全文链接类型和呈现优先顺序设置,如同一资源的链接类型选择顺序为直接链接、OpenURL 链接,同一链接类型下则选择排序优先级更高的电子资源平台。

用户通过电子期刊/图书导航主要来访问电子图书全文和期刊主页进而获取文章全文,2021 年电子期刊/图书导航中的热门发现电子资源平台根据热度由高到低排序为:万方、EBSCOhost、知网、方正、Elsevier、Springer Nature、超星、ProQuest、Wiley、JSTOR、DOAJ、CUP、Ovid,与 2021 年“未名学术搜索”中的热门发现电子资源平台有部分一致性。类似地,电子资源平台在导航中的发现热度差异与以下三个因素直接相关:一是电子资源平台支持发现的电子期刊/图书数量;二是图书馆在发现系统后台知识库的电子资源配置情况,是否将可访问的电子资源全部配置完成;三是同一检索结果下不同电子资源平台的链接呈现优先顺序设置,这一定

程度上引导用户以更大的概率使用排序优先级更高的电子资源平台。此外,学科热度、研究主题热度、用户习惯等差异也很关键。

3.4 用户通过不同渠道偏好访问不同电子资源平台,发现渠道的通畅对电子资源利用有明显的促进作用

用户访问电子资源有多个渠道,通过发现系统工具(“未名学术搜索”、电子期刊/图书导航)、通过导航工具(数据库导航)、直接进入电子资源平台是三个主要的全文获取渠道。为探究不同类型电子资源平台的用户访问渠道偏好,综合分析评估电子资源使用情况,以“未名学术搜索”发现量、电子期刊/图书导航发现量、数据库导航浏览量、电子资源平台检索量、电子资源平台下载量为指标,通过 lg 变换、归一化后,基于 k-means 方法^[18]对 2021 年数据较为完整的 31 个电子资源平台进行聚类分析,得到聚类结果如图 6 所示。对每项指标数值的高低水平采用四分位数(Quartile)来衡量,即把所有数值由小到大排列并分成四等份,处于三个分割点位置的数值即四分位数,将数值 $\geq Q_3$ 的记为高等水平,数值 $[Q_2, Q_3)$ 的记为中等水平,数值 $< Q_2$ 的记为低等水平,同一类别中的电子资源平台按照“未名学术搜索”发现量降序排列,最终得到各类别特征如表 2 所示。

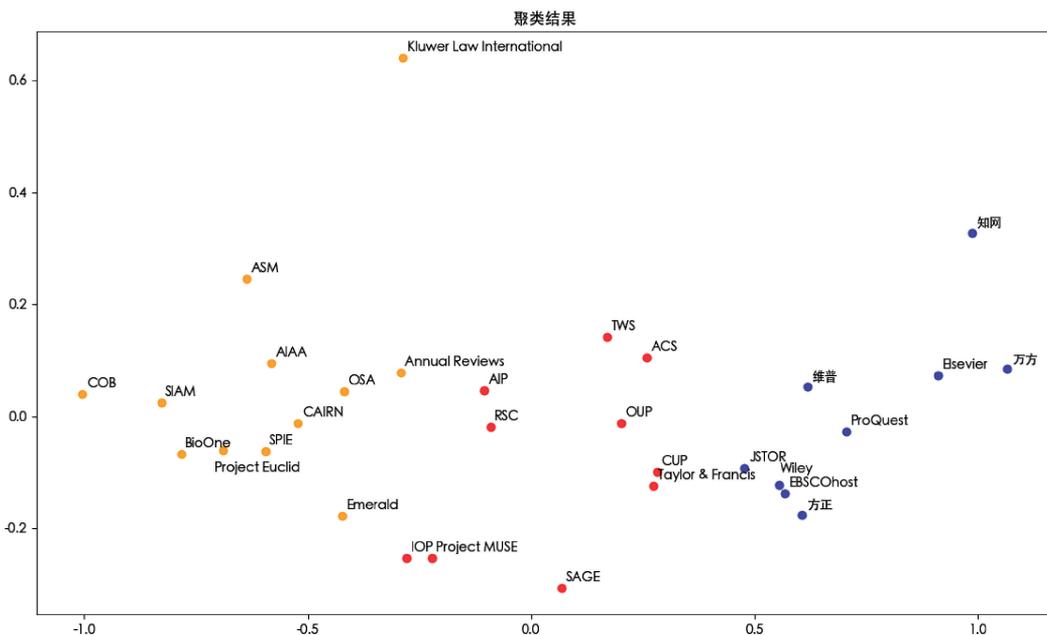


图 6 电子资源平台聚类结果



表 2 各类别特征

类别	电子资源平台	“未名学术搜索” 发现量	电子期刊/图书 导航发现量	数据库导航 浏览量	平台检索量	平台下载量
类别一	维普	高	中	高	高	高
	万方	高	高	高	高	高
	Elsevier	高	高	高	高	高
	ProQuest	高	高	高	高	中
	方正	高	高	中	中	高
	Wiley	高	高	中	中	高
	JSTOR	高	高	高	高	中
	EBSCOhost	高	高	高	中	低
类别二	知网	中	高	高	高	高
	Taylor & Francis	中	中	低	中	高
	SAGE	中	中	低	低	低
	CUP	中	中	高	中	中
	OUP	中	中	中	中	中
	Project MUSE	中	中	低	低	低
	ACS	中	中	中	中	高
	TWS	中	中	中	高	中
	RSC	低	低	低	中	中
	IOP	低	低	低	低	中
	AIP	低	低	中	低	低
类别三	Emerald	低	低	低	低	低
	Annual Reviews	低	低	中	低	低
	OSA	低	低	低	低	低
	SPIE	低	低	低	低	低
	CAIRN	低	低	低	低	低
	AIAA	低	低	中	低	低
	Project Euclid	低	低	低	低	低
	BioOne	低	低	低	低	低
	SIAM	低	低	低	低	低
	ASM	低	低	低	低	低
	COB	低	低	低	低	低
Kluwer Law International	低	低	低	高	中	

表 2 显示了用户通过不同渠道访问不同电子资源平台的差异及电子资源下载量,类别一中的电子资源平台各访问渠道整体均较为通畅,下载量较高,说明读者访问和利用电子资源的情况较好,其中万方、Elsevier 的各个访问渠道都比较通畅,下载量高。类别二中的电子资源平台各类访问渠道通畅程度和全文下载量整体处于中等水平,其中用户更多依赖数据库导航访问 CUP 平台资源,而对于 TWS 则更多选择直接进入平台查找资源。类别三中的电子资源平台各类访问渠道通畅程度和全文下载量整体处于较低水平,其中用户更多选择直接进入 Kluwer Law International 平台查找资源,而对于 Annual Reviews、AIAA 等更多依赖数据库导航来访问资源。值得关注的是,类别一中的大多为综合数据库,资源整体规模较大,类别二、类别三中的专业数据库占比较多,用户相对来讲从发现系统获取全文的比例较低,后续需结合资源规模和不同学科用户的使用习惯差异进行更有针对性的案例分析。

4 研究结论与建议

用户需求是图书馆优化服务的指南针,结合用户在以发现系统工具为主要依托的电子资源利用环境中的行为特征和已有学者的研究结论,提出以下用户服务优化方案。

4.1 把握用户使用高峰期,瞄准、扩大核心用户群体,优化信息素养教育方案、提升信息素养教育效率

从时间和对象上看,春季学期开始和秋季学期末是用户检索、发现、利用资源的高峰期,同时研究也发现,发现工具最好介绍给大学一年级的学生使用^[19],即对于低年级的大学生,发现系统可减少他们的认知负担,减缓其查找资源的学习曲线,以一种广泛搜索的方式作为他们的研究起点,相较而言,高年级学生或者图书馆员则更偏向于将发现系统作为次要选择的研究工具,因此每年的新生入学季无疑是开展培训讲座的又一个最佳时期,能够最大概率匹配用户需求,获得更多关注度。也有研究发现图书馆员和学生的使用率较高,而教师的使用率则低得多^[20],这也启发图书馆需关注使用率低的用户群



体,通过用户访谈等调研方式,深入了解低频用户背后的原因和信息需求,改善其使用情况。

从内容和形式上看,一方面在相关讲座内容的设置上,要避免偏重电子资源平台的界面和使用方法介绍,应认识到教授的内容不应只局限于如何使用界面,更多的精力应放在培养用户的批判性思维和选择优质资源的能力方面。另一方面,用户在电子期刊/图书导航中的模糊检索行为占比最多且呈上升趋势,逐渐将其作为一种相关资源发现的入口,因此还需加强检索策略的培训,帮助用户就某一检索主题优化检索词和检索策略等,以便更好地匹配到资源。

4.2 明确用户访问资源的渠道特征,优化资源配置方案,提升资源的可发现性、可获取性

用户访问电子资源平台的渠道包括通过发现系统工具、通过导航工具、直接登录平台三种,明确用户访问各电子资源平台的渠道偏好差异,可帮助图书馆及时发现和明确原因,以畅通用户访问渠道,提升电子资源的利用率。如对发现系统工具访问渠道不通畅的电子资源平台及时排查电子资源配置的情况,对热门发现电子资源平台增加电子资源配置的更新频率等。同时还需注意结合电子资源规模、学科特征等进行更有针对性的案例分析,在改善现有访问渠道的基础上,针对相应的目标用户开展重点调研,明确其电子资源利用习惯,增加定向邮件推送等多渠道访问方式,提升资源的显示度与利用率。

4.3 及时把握用户信息资源需求,聚合相关资源,开展专题资源推荐服务

发现系统能体现用户周期性的、突发性的、结合时事热点产生的资源需求。相比传统的用户调研等方法,发现系统提供了一种更经济、有效的捕捉用户需求的方式,即通过焦点词透视、追踪用户当前的信息需求,基于用户需求,提供多样化的主动服务,如围绕焦点词聚合相关资源,开展专题资源揭示与推荐服务,提升资源推荐与用户需求的匹配度和有效性;围绕焦点词遴选优质资源,开展专家领读、精读等阅读推广服务等。

5 总结

发现系统工具中多年积累的用户行为数据,能够体现出用户发现、利用电子资源的行为特征的重要方面,对用户服务的优化有实际参考、指导意义。

但受限于当前数据仍无法全面呈现用户使用发现系统的所有行为,以及电子资源平台的发现行为数据与其他数据整合力度不够,本研究仍存在不足,如在对电子资源平台进行聚类分析时,只考虑了电子资源平台总体资源的数据指标,以致于资源规模较大的电子资源平台整体表现较好,未来还需结合电子资源平台的资源量,基于更合理的篇均指标、不同学科的用户使用偏好差异等进行分析,以增强研究结论对业务工作的指导意义。

未来,首先需要认识到完善用户行为数据的追踪、记录方案是用户行为研究的基石,图书馆需协同发现系统技术支持共同制定可行的数据追踪方案、数据质量控制规范、同一指标下不同统计标准数据的整合利用等,还可集成 Google Analytics 等网页分析工具记录用户网页访问数据,与发现平台的数据互为补充。其次要加强发现数据与用户身份数据、资源数据、利用数据等的关联,加强运用调研法、访谈法、行为追踪法等多种用户行为研究方法,针对重点用户群体展开深入研究。最后对分析结论的探讨和应用也需进一步深入,如结合资源规模、学科特征等开展针对性案例分析等,以产生更大的数据效益、增强基于数据优化工作及服务的能力。

参考文献

- 1 陈建龙,邵燕,张慧丽,等.大学图书馆现代化的前沿课题和时代命题——《大学图书馆现代化指南针报告》解读[J].中国图书馆学报,2022,48(1):17-28.
- 2 王波,支娟,陈建龙.图书馆现代化新征程上的创新发展——浅析北京大学图书馆东楼修缮重启的新意[J].大学图书馆学报,2021,39(1):33-43.
- 3 Library of Congress. Digital Collections Strategy Overview 2022-2026[EB/OL].[2022-03-22]. https://www.loc.gov/acq/devpol/Digital%20Collections%20Strategy%20Overview_final.pdf.
- 4 王俊,杨丽萍,毕新.发现工具嵌入学习管理系统的实践与探索[J].新世纪图书馆,2017(1):20-24.
- 5 张元俊,王昊贤,李国俊,等.“信息协同”视角下嵌入数字校园的高校图书馆数字教参服务——以北京大学数字教参资料系统为例[J].大学图书馆学报,2022,40(2):27-35.
- 6 Way D. The impact of web-scale discovery on the use of a library collection[J]. Serials Review, 2010, 36(4):214-220.
- 7 Olsen A. SUMMON user experience: a study of the SUMMON laptop and mobile interface[J]. College & Undergraduate Libraries, 2020, 27(2-4): 420-434.
- 8 朱玲,聂华.通过日志挖掘研究图书馆资源发现服务用户的搜索行为[J].现代图书情报技术,2011(12):74-78.
- 9 杨丽萍,王俊,韦晓莹.资源发现系统的用户信息行为分析——以西交利物浦大学图书馆为例[J].图书馆研究,2019,49(2):92-99.



- 10 王俊,杨丽萍,韦晓莹.发现系统用户检索式行为研究——以西安交通大学为例[J].新世纪图书馆,2020(4):45-50.
- 11 蔡颖,才小川.资源发现系统服务能力提升初探——以天津搜索系统为例[J].图书情报导刊,2020,5(3):31-37.
- 12 李慧芳,孟祥保.近十年国内外图书馆资源发现系统研究与实践进展述评[J].图书情报工作,2020,64(6):120-129.
- 13 孙博阳,金丽萍.开放获取期刊用户利用数据分析[J].图书馆杂志,2011,30(2):41-45.
- 14 程蓓,吴帼帼.高校图书馆电子资源整合揭示及管理机制调查研究[J].图书馆工作与研究,2021(11):76-83.
- 15 黄文彬,吴家辉,徐山川,等.数据驱动的移动用户行为研究框架与方法分析[J].情报科学,2016,34(7):14-20,40.
- 16 周义刚,聂华,廖三三.北京大学资源发现服务推广探析——以“未名学术搜索”为例[J].图书情报工作,2014,58(20):67-72.
- 17 杨丽萍,王俊,韦晓莹.资源发现系统的用户信息行为分析——以西安交通大学图书馆为例[J].图书馆研究,2019,49(2):92-99.
- 18 张建萍,刘希玉.基于聚类分析的 K-means 算法研究及应用[J].计算机应用研究,2007(5):166-168.
- 19 Tonyan J, Piper C. Discovery tools in the classroom: a usability study and implications for information literacy instruction[J]. Journal of Web Librarianship, 2019,13(1):1-19.
- 20 Nichols A F, Crist E, Sherriff G, et al. What does it take to make discovery a success?: a survey of discovery tool adoption, instruction, and evaluation among academic libraries [J]. Journal of Web Librarianship,2017,11(2):85-104.

作者单位:北京大学图书馆,北京,100871
收稿日期:2022年5月30日
修回日期:2022年11月10日

(责任编辑:关志英)

Research on Users' Behavior of Resource Discovery Service

—A Case Study of Peking University Library

Wu Yaping Li Xiaodong Ai Chunyan

Abstract: The Web-Scale Resource Discovery Service based on pre-index metadata is the main entrance for users to search and discover academic resources uniformly. Over the years, a large amount of users' behavior data that supports horizontal and vertical comparisons has been accumulated. This paper analyzes the environmental and behavioral data sets of users' discovery and utilization of electronic resources and takes statistical analysis, cluster analysis and other methods to draw following conclusions. First, the user's retrieval quantity shows an upward trend, and the retrieval depth continues to increase. Second, the discovery service reflects the user's periodic and sudden resource retrieval needs combined with current events and hotspots. Third, the database discovery volume presents a long-tailed distribution. Fourth, users prefer to access different electronic resource platforms through different channels, and the Discovery Service plays a significant role in promoting the utilization of electronic resources. Finally, suggestions are put forward to improve user behavior data tracking, quality control, and correlation analysis to enhance the ability to optimize services based on data.

Keywords: Discovery Service; User Behavior; Characteristics Evaluation; Cluster Analysis



(接第 20 页)

Construction and Exploration of Three-Dimensional Librarian Training Mode from the Perspective of Activity Theory

—Taking Library and Information Work Committee of Anhui Higher Education Institutions as an Example

Chu Jiewang Zhu Aiyu Wang Jingwen Si Guangyun Meng Fansheng

Abstract: In the context of the national high-quality development, high-quality development becomes the goal and requirement of libraries in the new development stage. As one of the most active and dynamic elements, professional development of librarian is directly related to the development of the library cause. This paper uses the framework of activity theory to analyze some activities carried out by Library and Information Work Committee of Anhui High Education Institutions in the aspect of librarian training, summarizes its successful experience, and puts forward suggestions on its possible optimization.

Keywords: Library and Information Work Committee of Anhui High Education Institutions; University Library; Librarian Training; Activity Theory