



2020年新信息环境下 CASHL 资源与服务拓展设计研究

# 基于使用数据计量和使用动机调查 的 CASHL 文献资源保障研究

课题单位：华东师范大学图书馆

课题负责人：熊泽泉

课题组成员：杨莉（华东师范大学图书馆）、彭霞（华东师范大学图书馆）

结项时间：二零二一年六月

**摘要：**文献资源的价值不仅可以通过学者的引用行为反映，也可以通过读者的下载、阅读、分享等行为反映。本研究按照“使用即有用”的思路，一方面对 Web of Science 中“经济学与商学”学科中的高使用量论文进行了科学计量分析，归纳了高使用量文献的一般性特征，在此基础上分析了华东师范大学及 CASHL 对高使用量论文的保障率；另一方面对不同群读者群体展开问卷调查，了解不同读者在文献使用方面的动机，以期更有针对性地开展文献采购与推荐服务。本研究发现，高使用量论文一般为最新发表的论文以及影响因子较高的论文，且使用量与被引量存在显著正相关；华东师范大学对“经济学与商学”学科中的高使用量论文保障率仅为 83.9%，但 CASHL 对“经济学与商学”学科中的高使用量论文保障率达到 99.1%，极大地弥补了单个高校文献保障率不足的问题；文献使用动机调查进一步验证了用户在文献使用时更倾向于使用最新发表的论文及高影响力论文，并发现不同用户群体的使用倾向具有显著差异，提示在进行电子资源推送时应根据用户使用动机的异质性制定更精准的推送策略以提高文献的使用率。

**关键词：**使用动机 使用计量 科学计量 文献保障

# 目 录

引言	4
1 Web of Science 论文使用计量分析	5
1.1 全球高使用量论文特征分析	5
1.1.1 全球高使用量论文基本统计	5
1.1.2 全球高使用量论文发表年份	6
1.1.3 全球高使用量期刊统计	6
1.2 全球高使用量论文使用量与被引量的相关性分析	8
1.3 全球高使用量论文的保障率分析	9
1.3.1 华东师范大学期刊资源对高使用量论文期刊的保障情况分析	9
1.3.2 CASHL 期刊资源对高使用量论文期刊的保障情况分析	11
2 不同群体的文献使用动机调查	12
2.1 半结构化访谈	12
2.2 调查问卷设计	12
2.2.1 调查问卷维度	12
2.2.2 调查问卷数据获取与信效度检验	12
2.3 调查结果分析	13
2.3.1 受试者基本情况分析	13
2.3.2 不同年龄段用户对文献的使用情况	16
2.3.3 文献使用群体在不同背景下的下载与阅读分析	17
2.3.4 不同用户群体在文献使用上的差异分析	21
3 结论与建议	23
3.1 结论	23
3.1.1 CASHL 对全球高使用量论文保障率较高	23
3.1.2 基于引文的文献保障分析并不全面	23
3.2 建议	24
3.2.1 基于文献使用量进行文献资源保障分析	24
3.2.2 基于用户使用动机的差异性开展精准推送	24
3.2.3 多以中文文字介绍外文文献	24

## 引言

资源是图书馆的立身之本，文献资源建设是图书馆的核心工作之一。随着数字技术的飞速发展，电子文献数量激增，使文献信息资源结构发生了重要变化<sup>[1]</sup>。由于电子文献具有可复制、易传递、易携带等优点，深受读者欢迎，并且对于图书馆物理空间的缓解也有裨益，因此如何高效地进行电子文献资源建设成为图书馆资源建设的重中之重，亦成为众多学者关注的热点问题。

传统的关于文献资源保障研究主要基于引文分析<sup>[2-6]</sup>，但是，基于引文分析的文献资源保障研究存在两方面问题：第一，文献发表和引用的滞后性造成文献资源保障服务的滞后性，特别是人文社科领域成果，因此从引文分析中获得的文献保障情况难以准确反映学者在进行研究时的实际资源保障情况；第二，并不是所有的文献需求都能通过引文进行反映，一些文献在课程教学、大众教育等方面发挥着重要作用，但是使用它们的用户并不发表论文，因而这些用户对文献的需求也无法通过引文数量进行反映。

论文使用数据是指学术论文被用户下载、阅读、分享、保存等过程中所产生的可被记录的量化数据。由于对传统纸质论文的使用难以被统计或记录，文献使用数据被学者所关注是在互联网兴起后，电子文献成为学术信息的主要载体，用户对电子文献的使用能够被服务器准确即时记录，因而为科学计量研究者提供了崭新的研究视角，并从多个角度展开了大量研究，如研究论文的使用模式<sup>[7-9]</sup>、运用使用数据探索科研趋势<sup>[10]</sup>、探索使用数据与引文数据的相关性<sup>[11, 12]</sup>从而为传统的引文评价提供补充等。但是，**尚未见利用文献使用数据来研究文献资源保障情况的报道。**

本研究拟利用文献使用数据来研究文献资源保障情况，可以有效地解决基于引文分析的文献资源保障研究中的问题。首先，文献使用行为发生在研究者的研究过程当中，特别是文献下载行为发生在研究的初期阶段，可实时反映研究者对文献的使用和需求情况；其次，文献使用行为并不关注使用者是否发表论文，而只关注文献是否被使用。本研究认为，**文献被使用即发挥了其价值，被使用次数越多，说明其价值越高，在资源建设的角度应该优先提供保障。**

此外，对文献资源保障率进行研究的目的是为了更好地提供文献资源服务。因此，本研究拟从使用动机的角度，进一步调研不同用户的文献使用动机，掌握用户使用学术文献的外部因素和内部因素，从而更有针对性地制定学科服务策略，更加精准地进行文献推送，提高文献的使用率。

# 1 Web of Science 论文使用计量分析

本部分以 ESI 期刊类型作为分类标准，选取 ESI“经济学与商学”（ECONOMICS & BUSINESS）学科领域中 588 本期刊在 WOS 数据库中近 10 年的发文为数据来源，数据获取时间为 2021 年 3 月 15 日，共获取到 327912 条发文记录。再分别依据近期使用量 U1（近 180 天的使用量）和长期使用量 U2（自 2013 年 2 月 1 日以来的使用量）进行排序，选择前 1% 的论文（分别为 3280 条）作为全球高使用量论文数据集，记为 V1 和 V2 两个数据集，本节后续数据分析围绕上述两个数据集展开。

## 1.1 全球高使用量论文特征分析

### 1.1.1 全球高使用量论文基本统计

表 1-1 为高使用量论文数据集 V1、V2 的基本统计，图 1-1 为近期使用量 U1、长期使用量 U2 的绝对数值分布。长期使用量 U2 的各项检验参数均远大于近期使用量 U1，两者的绝对数值分布均表现为严重的非中心对称的偏态分布，即在高使用量论文数据集 V1、V2 中，80% 高使用量论文的近期使用量 U1 和长期使用量 U2 主要集中在 19-39 和 149-274 之间。经 K-S 检验，结果表明，V1 和 V2 两个样本的 TC<sub>1</sub>、TC<sub>2</sub>、U1 和 U2 的绝对数值分布均不符合正态分布。

表 1-1 两个高使用量数据集基本统计

	指标	最小值	中值	最大值	平均值	标准差	K-S Z值	K-S P值
V1数据集	U1	19	27	902	33.18	24.21	0.28	<0.001
	TC <sub>1</sub>	0	7	3277	74.01	174.65	0.34	<0.001
V2数据集	U2	149	196	4447	235.47	146.25	0.28	<0.001
	TC <sub>2</sub>	0	104	3277	151.34	172.08	0.19	<0.001

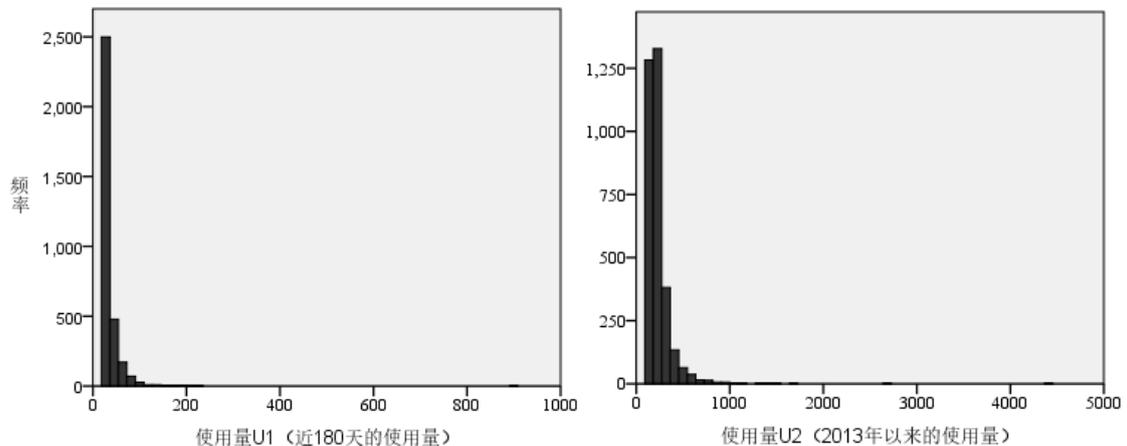


图 1-1 使用量频次分布

### 1.1.2 全球高使用量论文发表年份

图1-2为高使用量论文V1和V2发表年份分布及其占年度论文发表总量的比例。基于近期使用量U1获取的高使用量论文V1多为2019-2021年发表的最新论文，合计有2272篇，占总体的69.27%。而基于长期使用量U2获得的高使用量论文V2多为2016年以前发布的论文，共计2658篇，占总体的81.04%。由于U2代表的是WOS平台上论文的长期使用量，早期发表论文的长期使用量U2会随着时间累积增长。而具有较高近期使用量U1的论文则明显反映出“经济学与商学”学科领域的学者在WOS平台中**更偏好使用最新发表的论文**，尤其是近三年。

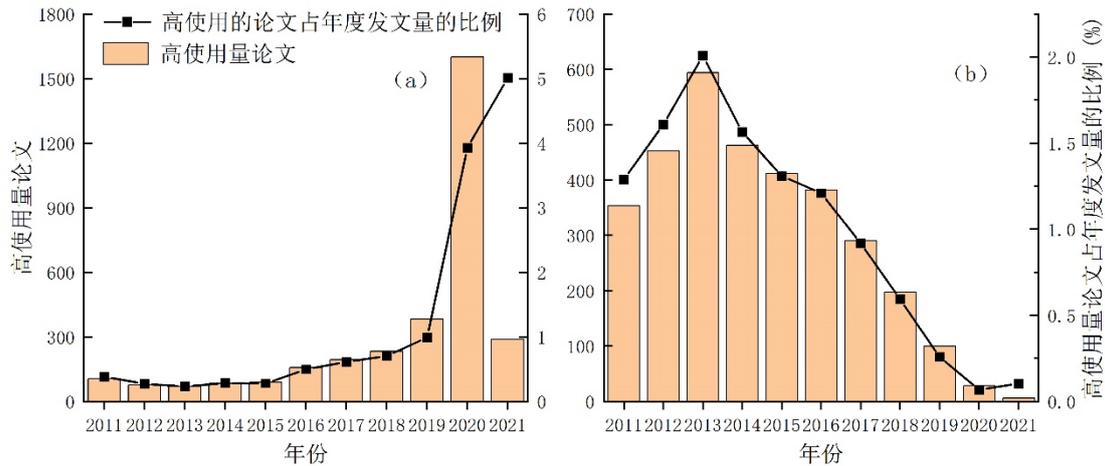


图 1-2 高使用量论文 V1、V2 发表年份分布及占年度发文总量的比例  
(a: 近期使用量 U1, b: 长期使用量 U2)

### 1.1.3 全球高使用量期刊统计

高使用量论文 V1、V2 共分布在 284、265 个期刊。表 1-2 为高使用量论文排行前五种期刊，JOURNAL OF BUSINESS RESEARCH、ACADEMY OF MANAGEMENT JOURNAL、STRATEGIC MANAGEMENT JOURNAL、JOURNAL OF MANAGEMENT 四种期刊均具有较多的高使用量论文，但 MANAGEMENT SCIENCE 针对近期使用量 U1 上具有较多高使用量论文，而 ORGANIZATION SCIENCE 则是针对长期使用量 U2 具有较多高使用量论文。这就表明前四本期刊是“经济学与商学”学科领域近十年来无论是近期使用量还是长期使用量上均表现较高的经典期刊，而 MANAGEMENT SCIENCE 则是近几年新晋热门期刊，使用量持续增长，相对地，ORGANIZATION SCIENCE 期刊在 WOS 平台的使用量则逐渐减少，甚至 2020 年和 2021 年均未有论文列入高使用量论文数据集 V2 中。这其中可能存在两点原因：

(1) 期刊载文量的影响。MANAGEMENT SCIENCE 近十年期间，载文量一直处于上升阶段，由 2011 年的 140 篇增长至 2020 年的 301 篇，而 ORGANIZATION SCIENCE 的载文量总体表现为下降趋势，由 2011 年的 104 篇减少到 2020 年 70 篇。

(2) 期刊影响因子的影响。MANAGEMENT SCIENCE 的影响因子近十年同样表现出较为明显的上升趋势，2011 年的影响因子仅为 1.733，而 2019 年上升至 3.931 (图 1-3)。ORGANIZATION SCIENCE 则由 2011 年历史最高影响因子 4.338，逐渐跌至到 2019 年 2.782

(图 1-4)。

表1-2 高使用量论文排行前五的期刊及其占期刊发文总量占比与排序

高使用量论文 V1-近期使用量 U1		
期刊名	高使用量论文	占期刊发文总量比例/排序
Journal of Business Research	143	3.10%/51
Academy of Management Journal	140	17.48%/2
Strategic Management Journal	131	10.95%/6
Management Science	120	5.48%/24
Journal of Management	111	12.31%/3

高使用量论文 V2-长期使用量 U2		
期刊名	高使用量论文	占期刊发文总量比例/排序
Academy of Management Journal	275	34.33%/1
Strategic Management Journal	207	17.31%/5
Journal of Management	199	22.06%/3
Journal of Business Research	124	2.69%/56
Organization Science	97	11.28%/10

期刊载文量和期刊影响因子会在一定程度上影响科学研究人员的使用量，如谢娟等（2018）曾研究中指出 JCR 分区 Q1 论文的平均使用次数约为 Q4 论文的 5 倍，本研究也侧面反映了研究人员偏向使用影响因子较高的期刊论文。

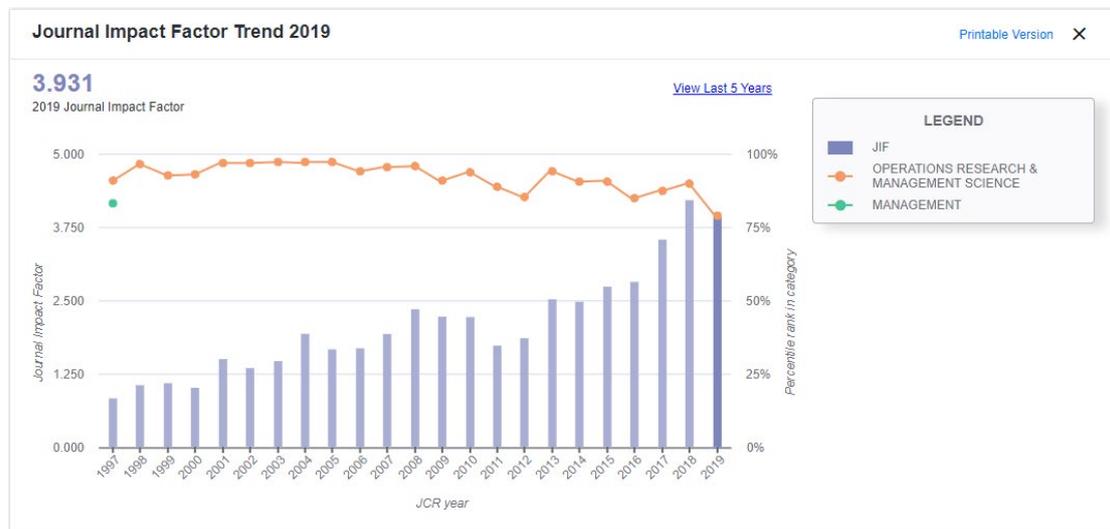


图 1-3 MANAGEMENT SCIENCE 影响因子变化趋势

2.782

2019 Journal Impact Factor

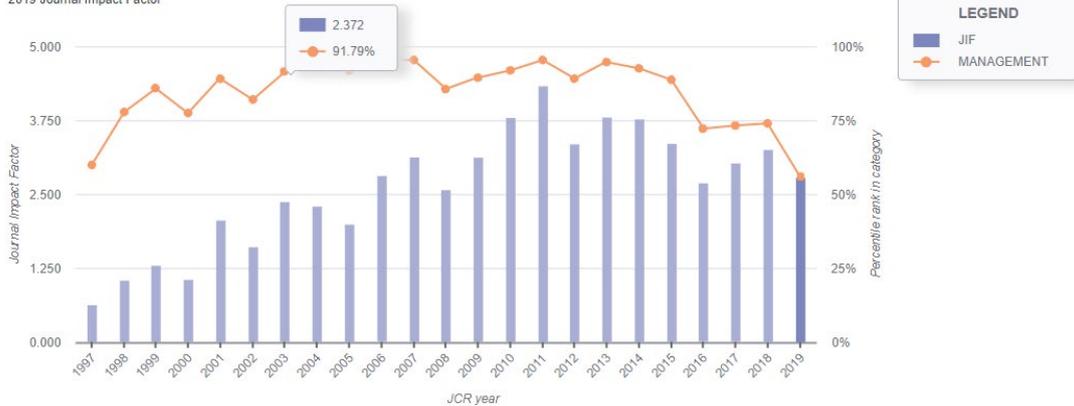
[View Last 5 Years](#)

图 1-4 ORGANIZATION SCIENCE 影响因子变化趋势

## 1.2 全球高使用量论文使用量与被引量的相关性分析

前文中 K-S 检验可知使用量和被引量均不符合正态分布,因此相关性分析采用 Spearman 相关检验。图 1-5 是从论文角度的高使用量论文的使用量与被引量的散点图,长期使用量 U2 与被引量 TC 之间的 Spearman 相关系数 (R 值)要略高于近期使用量 U1,分别为 0.407、0.142,显著性系数均  $P < 0.01$ 。图 1-6 是从高使用量论文发表的期刊角度的使用量与被引量散点图,长期使用量 U2 与被引量 TC 之间的 Spearman 相关系数 (R 值)同样高于近期使用量 U1,分别为 0.816、0.697,显著性系数均  $P < 0.01$ 。从上述结果可知,相比于近期使用量 U1 而言,长期使用量 U2 与被引量 TC 之间存在更强的相关性。这一相关性的差异也说明,被引量能够一定程度上反映读者的长期使用倾向,但是在反映读者近期使用倾向上则略显不足。

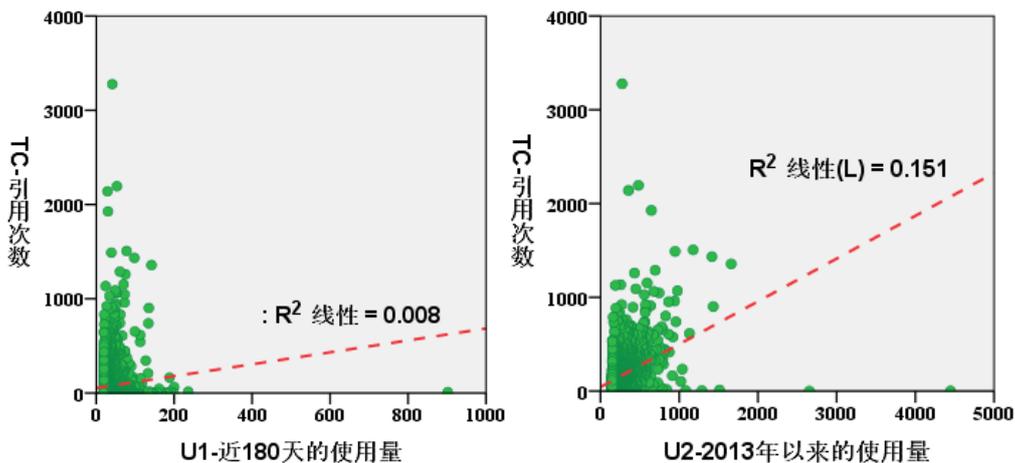


图 1-5 论文角度的近期使用量 U1、长期使用量 U2 与被引量 TC 的散点图

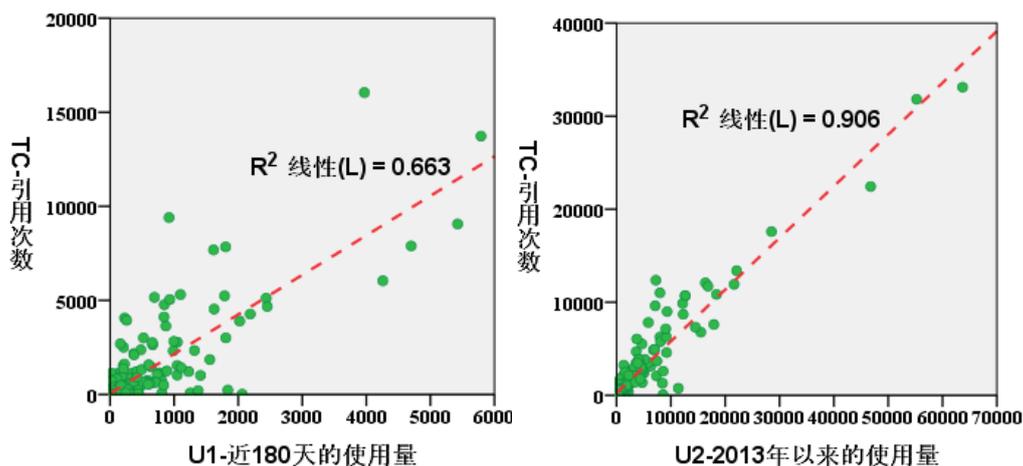


图 1-6 期刊角度的近期使用量 U1、长期使用量 U2 与被引量 TC 的散点图

### 1.3 全球高使用量论文的保障率分析

#### 1.3.1 华东师范大学期刊资源对高使用量论文期刊的保障情况分析

本部分是考察华东师范大学购买期刊对高使用量论文集 V1 和 V2 所在期刊集合的保障情况。“经济学与商学”学科领域的高使用论文集 V1、V2 所在期刊不含重复项共计 329 种，华东师范大学可保障 276 种，资源保障率达 83.9%。在可保障的 276 种期刊中，其中有 11 种 OA 刊。缺藏 53 种期刊详细信息见表 1-3，其中属于 Q1、Q2、Q3 和 Q4 分区期刊分别有 13、17、15 和 8 种。

表1-3 华东师范大学缺藏高使用论文期刊（53种）

期刊名称	ISSN	CASHL 是否保 障	IF	分 区	近期高 使用量 论文数 -U1	长期高 使用量 论文数 -U2
Academy of Management Annals	1941-6520	是	11.75	Q1	49	50
Engineering Construction and Architectural Management	0969-9988	是	2.16	Q2	42	1
Management Decision	0025-1747	是	2.723	Q2	31	14
Accounting Review	0001-4826	是	3.993	Q1	28	17
European Journal of Innovation Management	1460-1060	是	2.613	Q2	28	2
Journal of Service Management	1757-5818	是	4.662	Q1	25	14
International Journal of Operations & Production Management	0144-3577	是	4.619	Q1	18	12
Supply Chain Management-an International Journal	1359-8546	是	4.725	Q1	16	18
International Journal of Logistics Management	0957-4093	是	3.325	Q2	14	1

Journal of Intellectual Capital	1469-1930	是	4.805	Q1	13	0
International Journal of Physical Distribution & Logistics Management	0960-0035	是	4.744	Q1	12	12
China Agricultural Economic Review	1756-137X	是	1.775	Q2	12	0
Business Process Management Journal	1463-7154	是	2.121	Q3	11	2
European Journal of Marketing	0309-0566	是	2.135	Q3	9	4
Asia Pacific Journal of Marketing and Logistics	1355-5855	是	2.511	Q2	9	1
Chinese Management Studies	1750-614X	是	1.036	Q4	9	0
Journal of Product and Brand Management	1061-0421	是	1.832	Q3	9	1
International Journal of Entrepreneurial Behavior & Research	1355-2554	是	3.529	Q2	10	0
Leadership & Organization Development Journal	0143-7739	是	1.977	Q3	8	1
International Marketing Review	0265-1335	是	2.907	Q2	7	2
Journal of Business & Industrial Marketing	0885-8624	是	2.497	Q3	6	2
Journal of Services Marketing	0887-6045	是	3.195	Q2	6	0
Review of International Political Economy	0969-2290	是	2.312	Q1	6	2
International Journal of Technology Management	0267-5730	是	1.348	Q3	6	6
International Journal of Conflict Management	1044-4068	是	1.806	Q2	6	0
International Journal of Manpower	0143-7720	是	0.953	Q3	5	2
Mis Quarterly Executive	1540-1960	是	4.088	Q1	5	4
Journal of Organizational Change Management	0953-4814	是	0.967	Q4	4	1
Marketing Intelligence & Planning	0263-4503	是	2.164	Q3	4	1
International Journal of Managing Projects in Business	1753-8378	是	1.989	Q3	4	0
Accounting Auditing & Accountability Journal	0951-3574	是	3.497	Q1	3	0
Mit Sloan Management Review	1532-9194	是	2.706	Q2	3	29
Jahrbucher Fur Nationalokonomie Und Statistik	0021-4027	是	0.658	Q4	3	0
Journal of Labor Economics	0734-306X	是	3.356	Q1	3	1
Journal of Service Theory and Practice	2055-6225	是	3.512	Q2	3	0
Review of Environmental Economics and Policy	1750-6816	是	6.487	Q1	3	25
Employee Relations	0142-5455	是	1.641	Q3	2	0
Journal of Advertising Research	0021-8499	是	2.169	Q2	2	5
Post-communist Economies	1463-1377	是	0.875	Q3	2	0

Academia-revista Latinoamericana De Administracion	1012-8255	是	0.739	Q4	2	0
Journal of Korea Trade	1229-828X	是	0.628	Q4	2	0
Pharmacoeconomics	1170-7690	是	3.563	Q1	1	1
Journal of Medical Economics	1369-6998	是	1.958	Q2	1	0
Transformations in Business & Economics	1648-4460	是	1.621	Q2	1	1
Journal of Portfolio Management	0095-4918	是	0.709	Q4	1	0
International Journal of Bank Marketing	0265-2323	是	2.8	Q2	1	0
Review of Keynesian Economics	2049-5323	否	0.689	Q4	1	1
Baltic Journal of Management	1746-5265	是	0.958	Q3	1	1
Journal of Law & Economics	0022-2186	是	1.719	Q3	1	1
Journal of Forest Economics	1104-6899	是	0.958	Q3	1	0
Journal of Electronic Commerce Research	1526-6133	是	1.875	Q3	1	1
Entrepreneurship Research Journal	2194-6175	是	1.643	Q4	1	2
American Journal of Health Economics	2332-3493	否	1.902	Q2	0	4

### 1.3.2 CASHL 期刊资源对高使用量论文期刊的保障情况分析

针对“经济学与商学”学科领域 329 种高使用论文期刊，CASHL 已保障 326 种，保障率达 99.1%。<sup>[13]</sup>发现 114 种 CASHL 缺藏的 ESI 经济学与商学学科期刊，他们建议优先补藏的 24 种 ESI 经济学与商学期刊，其中 19 本为本研究中高使用论文期刊，目前皆已被 CASHL 收录。说明 CASHL 对高使用论文期刊的保障率相较于 2018 年已有所提高。

CASHL 仅缺藏 3 种高使用论文期刊（表 1-4），分别为 REVIEW OF KEYNESIAN ECONOMICS、KOREAN ECONOMIC REVIEW、AMERICAN JOURNAL OF HEALTH ECONOMICS。其中 KOREAN ECONOMIC REVIEW 属于 OA 刊，建议在 CASHL 馆藏目录中增加链接。2 种期刊在近期高使用论文集 V1 中有 2 篇论文，3 种期刊在长期高使用论文集 V2 中有 7 篇论文。

表1-4 CASHL缺藏的3种高使用论文期刊

期刊名称	ISSN	OA 刊	IF	分区
Review of Keynesian Economics	2049-5323	否	0.689	Q4
Korean Economic Review	0254-3737	是	0.174	Q4
American Journal of Health Economics	2332-3493	否	1.902	Q2

## 2 不同群体的文献使用动机调查

### 2.1 半结构化访谈

开展正式问卷调查前，研究团队对部分师生进行了半结构化访谈。访谈中主线问题围绕“一般会因为什么原因去下载一篇文献的全文”展开，进而从访谈中抽取与下载动机相关的因素。部分被访谈者的回复内容摘录如下：

教师 Y：好奇；对某一概念或领域的相关知识或动态不清楚时。搜索出多篇文献时，如果让我挑选要下载下来看的原因：篇名极为符合我的需求且文献来源（期刊或学位论文）较为权威；其次如果是一般期刊，我会点击进去看看摘要，然后再决定下载。

教师 G：想详细看论文的研究内容；比如从摘要看跟我想研究的主题非常相关，或者是前期研究。

教师 W：看完题目摘要之后感觉有有用的东西值得深入看一下。

学生 Y：和我要研究的问题相关的；对我的研究问题有启发性的；该论文研究的问题有意思的；可以作为写作模式参考的；现在暂时不用但可以作为以后研究问题的参考论文。

学生 W：一般我会对这个研究方法感兴趣会下载下来，其次我会根据我所写的主题，跟这个主题相近的我会下载下来。

通过对被访谈者回复内容的提炼，可以提炼出如下关键词：

表 2-1 半结构化访谈内容关键词

类型	关键词
文献要素相关	篇名，摘要，文献来源，主题
文献使用主体相关	好奇，需求，启发，有意思，感兴趣

可见，文献的外部要素，以及使用者对于文献价值的感知/期许，都是促使大家进行文献使用（此处主要表现为下载）的因素。因此调查问卷的设计也主要围绕这两方面展开。

### 2.2 调查问卷设计

#### 2.2.1 调查问卷维度

通过前期大量的文献整理和分析研究，并结合半结构化访谈，提炼了用户使用学术文献的可能动机及其对应的尽量全面的潜在变量，并制定了量表。量表涉及三大维度：“文献使用动机-下载”（共 27 项）、“文献使用动机-分享”（共 29 项）、“文献使用动机-管理”（共 17 项），另外量表还调研了用户对“数据库使用偏好”（共 16 项）以及用户的“基本信息”（共 4 项）<sup>1</sup>。

#### 2.2.2 调查问卷数据获取与信效度检验

初期本研究进行了小规模问卷发放与回收，并对数据进行了信效度检验，问卷总体信效

<sup>1</sup> 本报告分析部分仅对与本研究密切相关的下载维度进行分析。

度均在 0.9 以上（显著水平为.05）。经过预调研后，在各高校等科研单位进行大规模调查问卷发放与回收，最终获取 525 份调查问卷，删除无效问卷（答题时间过短、答案选项单一）后最终获得 480 份有效问卷，问卷有效率达 91.4%。

问卷总体信度达 0.977，三个维度的信度分别为 0.932、0.994、0.988，信度较好；问卷的总体结构效度达 0.970，结构效度良好。问卷信效度检验具体如表 2-2 至-表 2-4 所示。

表 2-2 问卷总体信度检验

	N	%	Cronbach's Alpha	基于标准化项的 Cronbachs Alpha	项数
有效	480	100.0	.977	.969	93
案例 已排除 <sup>a</sup>	0	.0			
总计	480	100.0			

表 2-3 问卷各维度信度检验

	N	%	维度	Cronbach's Alpha	基于标准化项的 Cronbachs Alpha	项数
有效	480	100.0	下载	.932	.933	27
案例 已排除 <sup>a</sup>	0	.0	分享	.994	.995	29
总计	480	100.0	管理	.988	.989	17

表 2-4 问卷结构效度 KMO 和 Bartlett 的检验

取样足够度的 Kaiser-Meyer-Olkin 度量	.970
近似卡方	53424.303
Bartlett 的球形度检验	df
	2628
	Sig.
	.000

## 2.3 调查结果分析

### 2.3.1 受试者基本情况分析

本次调研的 480 个有效样本中，受试者身份涉及专任教师、本科生、硕士研究生、博士研究生、教辅专技人员、其他（如行政管理人员、辅导员、机关工作人员等）。首先，各类人群受试者的数据分布中，本科生人数最多（图 2-1），这一群体具有使用文献频率高但是发文量少的特点；受试者年龄分布 30 岁以下人群占比最多，50 岁-60 岁占比最少，未获取到 60 岁以上的受试者（图 2-2）；从受试人群的专业分布来看，文理科人数占绝大多数，工农医等专业人数较少（图 2-3）；最后，从论文发表情况看，已发表过论文及准备发表的受试者约占 61%，其中以第一/通讯作者发表过论文的受试者高达 42.7%，未发表过论文的受试者约占 39%（图 2-4）。

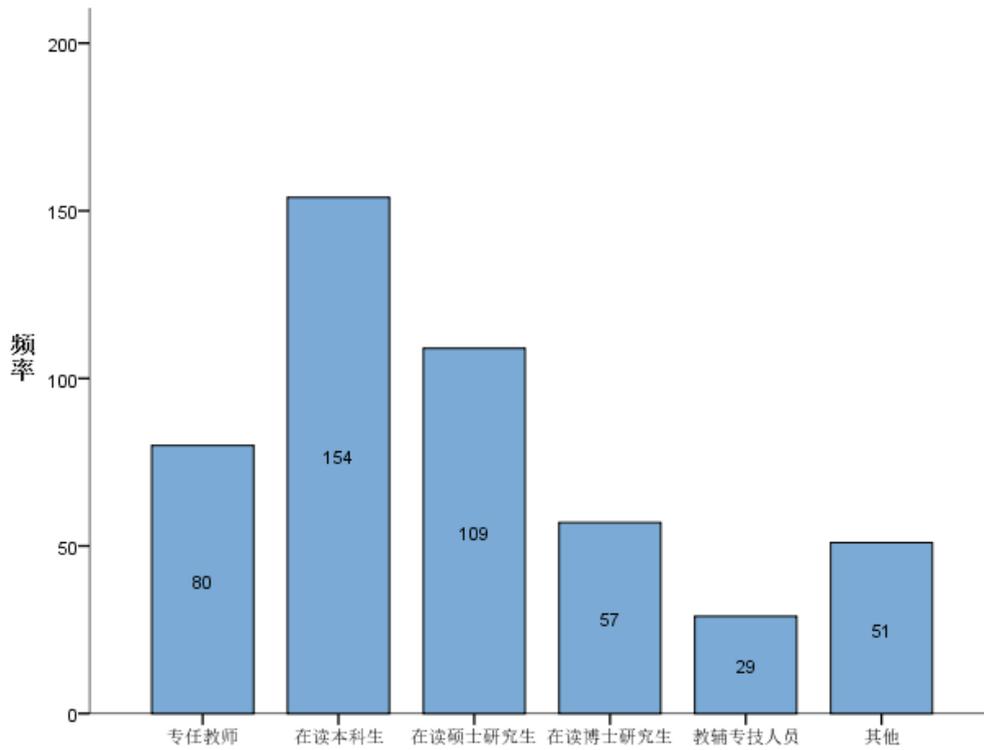


图 2-1 文献使用群体职业/身份特征

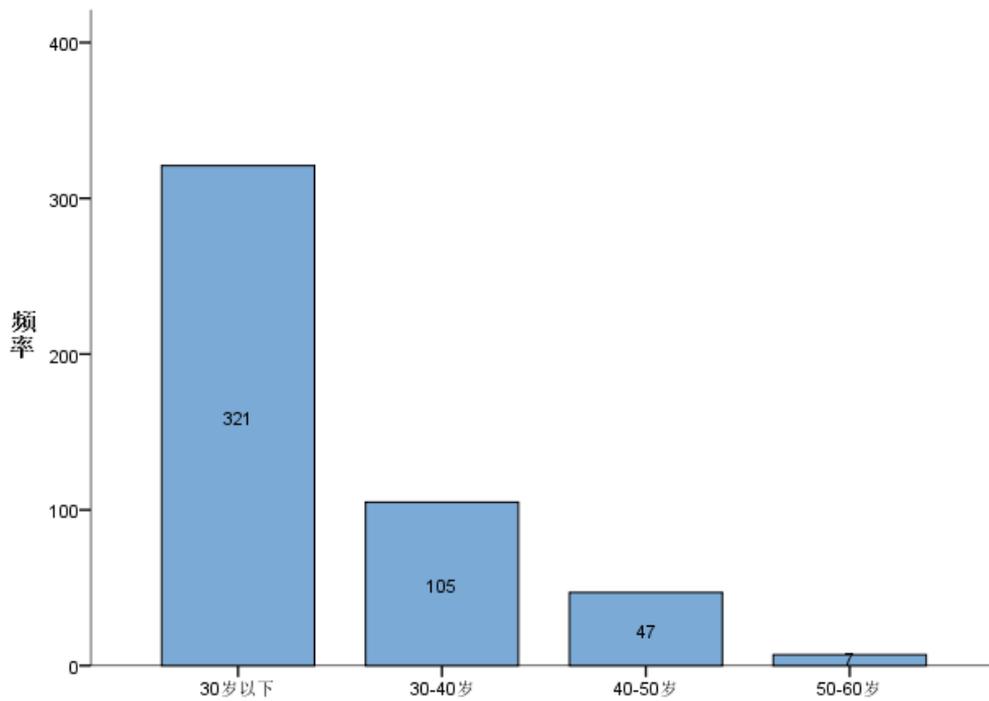


图 2-2 文献使用群体年龄分布

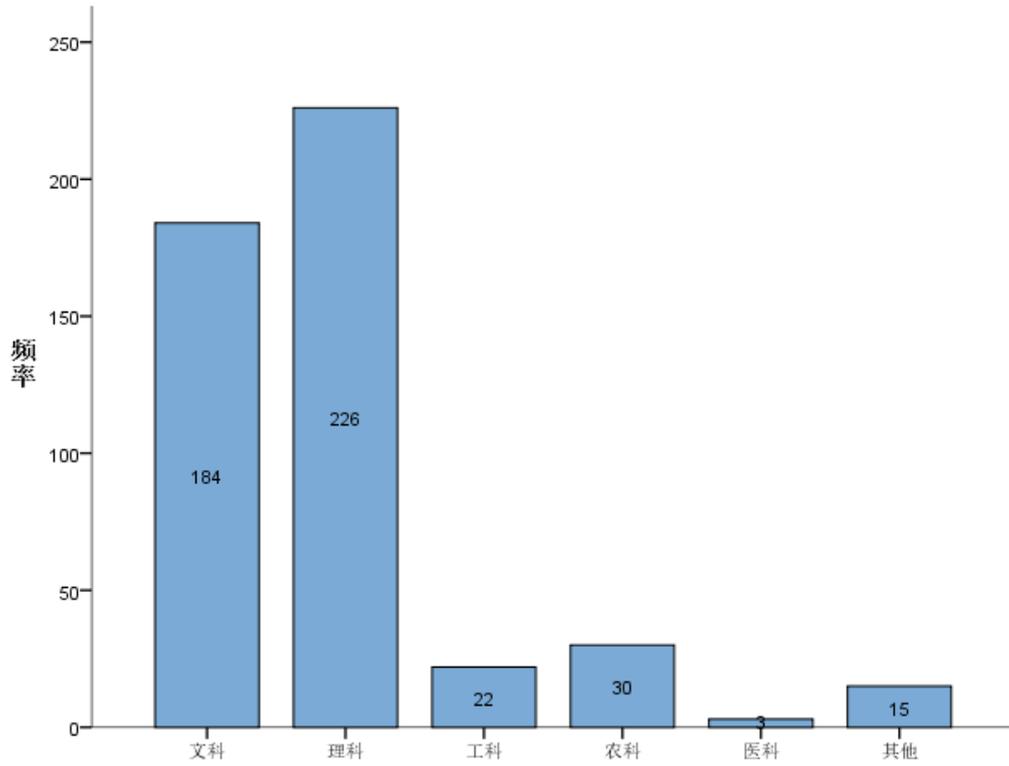


图 2-3 文献使用群体专业情况分布

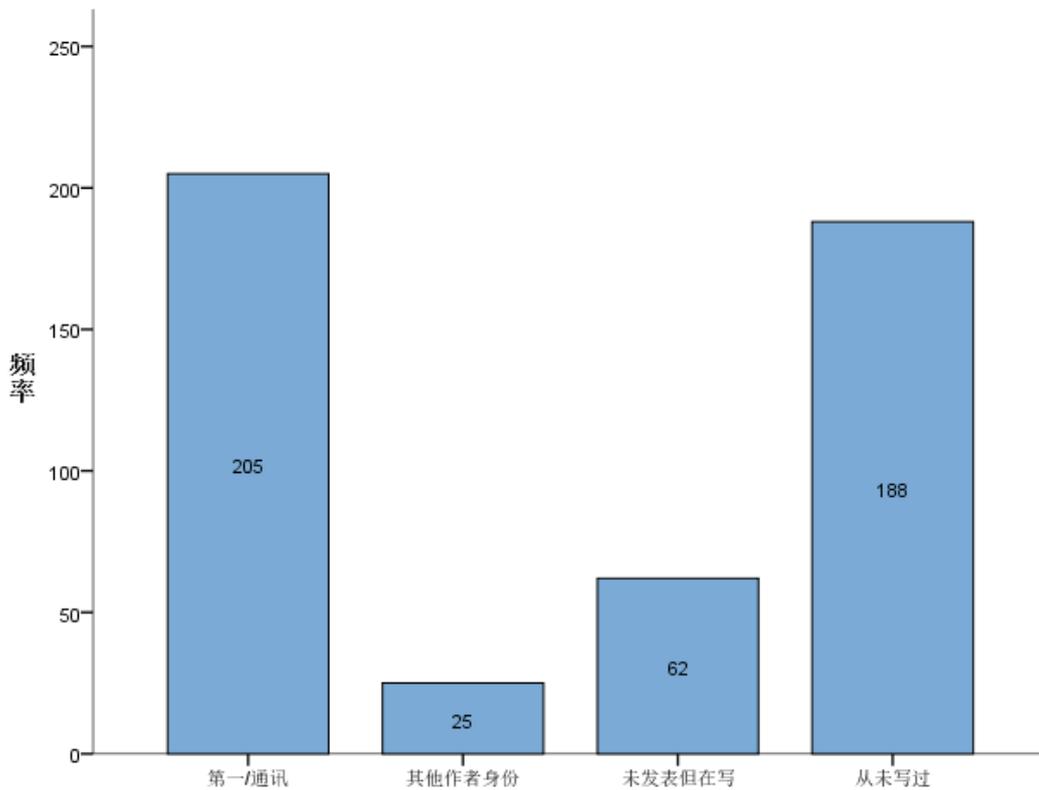


图 2-4 文献使用群体论文发表情况

从上述调研数据的基本情况分析可知，各类数据分布均呈现出非正态、方差非齐性，且各类数据的样本量也存在极大差异，因此，基于上述的数据分析基础，在后续调查结果中，采用非参 Kruskal-Wallis 方法进行差异检验。

### 2.3.2 不同年龄段用户对文献的使用情况

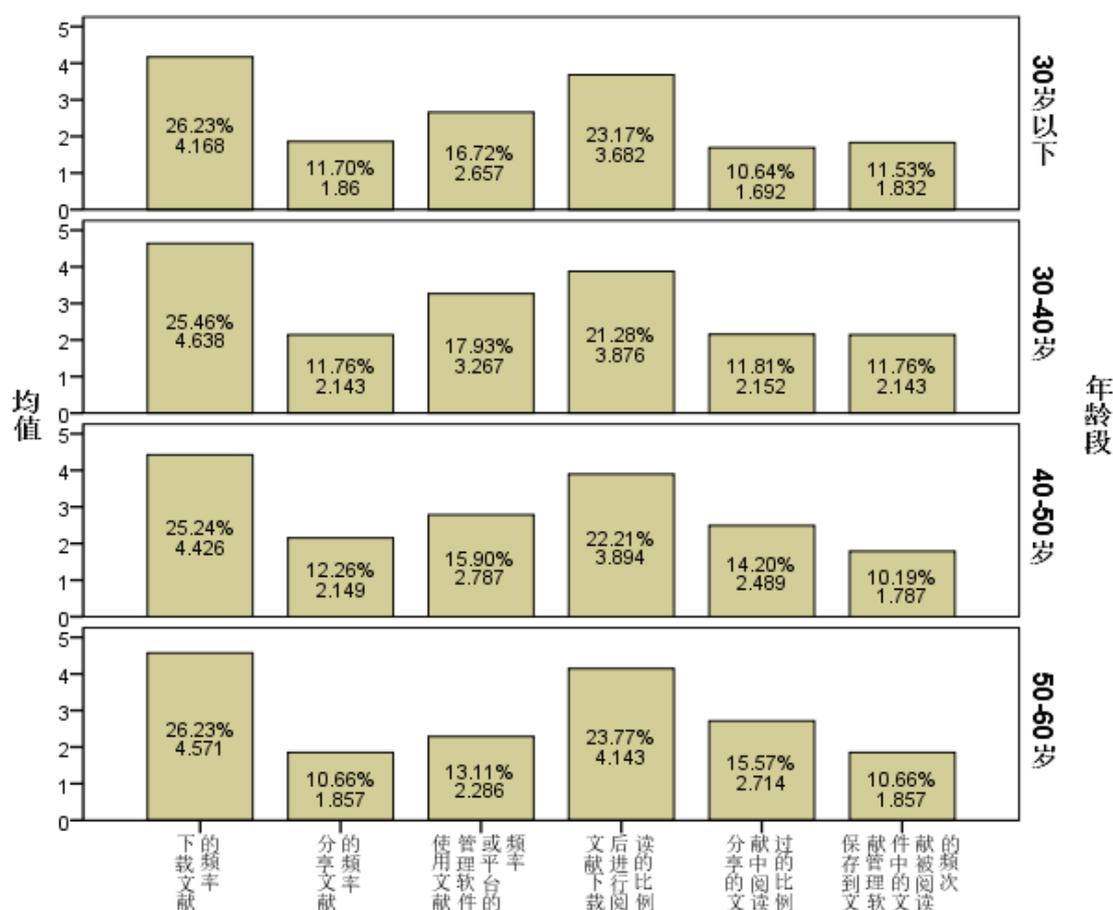


图 2-5 不同年龄段用户对文献的使用情况

在对不同职业/身份、年龄、专业以及发文情况的文献使用者的下载、分享及保存情况进行基本分析后，可以看出，下载作为后续分享、保存的前提，是频率最高的文献使用行为；使用文献管理软件的频率则较低，文献分享行为则最为罕见。由于读者浏览行为数据较难获取，文献下载数据无疑是研究读者文献使用行为的最佳数据来源。

此外，文献使用者获取文献全文的主要方式符合预期：排在前三位的分别是“在全文数据库检索主题等非题名字段后选择性下载”、“在全文数据库中直接检索题名下载”、“通过搜索引擎检索后进入相应数据库下载”，通过文献传递或社交网站或求助他人等其他方式获取全文的情况较少。因此，可以认为通过非数据库平台进行全文获取只是少数现象，基于数据库平台的下载数据可以反映绝大多数用户的文献下载行为，具体数据参见图 2-6。

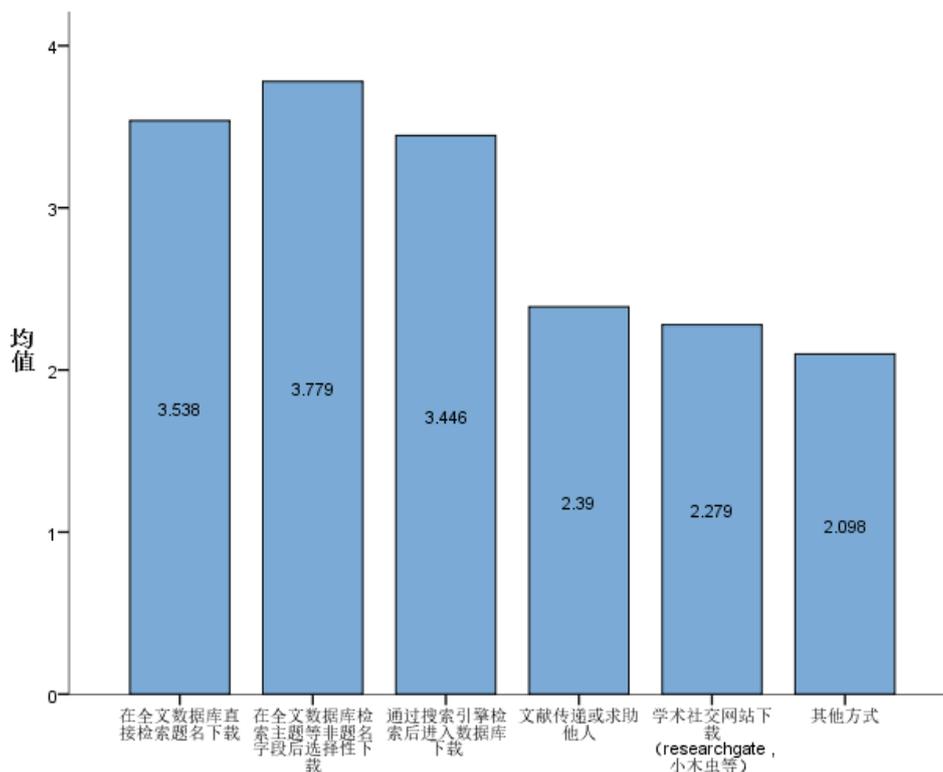


图 2-6 文献使用者获取全文的主要方式

### 2.3.3 文献使用群体在不同背景下的下载与阅读分析

#### 1) 不同职业/身份文献使用群体下载文献的用途和目的

下载文献的用途和目的主要从开展科研的文献调研、提供教学素材、满足个人的兴趣、后续引用以及其他非常见用途这五个方面进行调研，发现不论是何种身份背景文献使用者，其下载文献的用途集中在“文献调研”和“后续引用”这两个以研究为主的用途上(图 2-7)，其次，作为教职员工（如专任教师：3.763、教辅专技人员：3.414）会因准备教学素材而进行文献下载；此外，教辅专技人员和本科生较其他职业/身份背景文献使用者还会“因满足个人的兴趣”而进行文献下载。可见，虽然文献调研和引用是用户下载行为发生的主要目的，但是除此之外，不同用户群体在使用文献过程中具有不同的倾向性，如教师为了准备教学材料、教辅人员及本科生为了满足非学术的个人兴趣等，这些都是非引用倾向的知识获取过程。

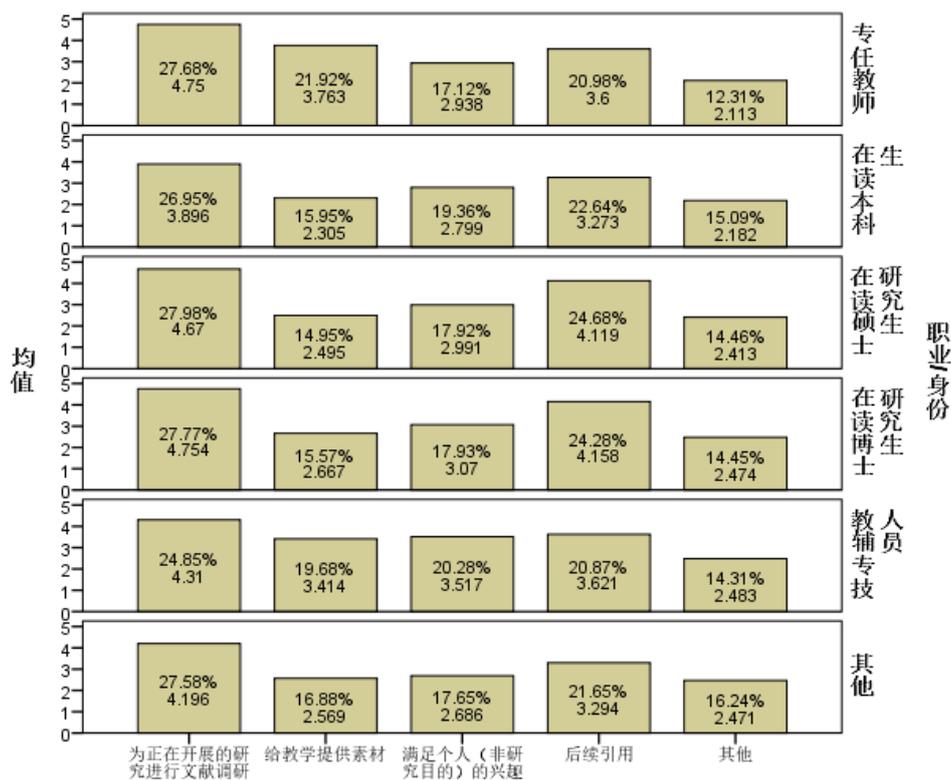


图 2-7 不同职业/身份文献使用者下载文献的用途和目的

## 2) 不同背景下“文献下载后阅读的比例”分析

那么，不同背景文献使用者在文献下载后，其阅读情况如何，进一步数据分析发现：有 70% 及以上的其他人员（如行政管理岗、机关党工委等）、在读硕士研究生其文献下载后阅读比例高于 60%；有 60% 左右的专任教师、在读博士研究生文献下载后阅读比例接近 60% 的；教辅专技人员与本科生的阅读比例较之其他人员相对低一些。具体数据分布如图 2-8 所示。

而针对不同年龄的文献使用者进行分析，发现 50 岁及以下的文献使用者其文献下载后阅读比例高于 60% 的人员占比仅在 50%-60%；而 50-60 岁的文献使用者其文献下载后阅读比例高于 60% 的人员占比高达 71.43%。由此看，年轻人下载文献后具有更多样化的动机或使用形式，具体数据分布参见图 2-9。

不同专业的文献使用者，其下载文献后的阅读比例也有所不同，阅读比例高于 60% 的不同专业的文献使用者占比情况由高到低依次是医科、农科、文科、工科、理科、其他。如图 2-10 所示。

此外，从论文发表情况分析文献使用者的下载后阅读情况也可看出：发表了论文的或是正在写论文文献使用者其文献下载后阅读比例高于 60% 的人员占比也均超过了 60%，尤其是以其他作者身份发表论文的文献使用者竟高达 76.66%，而未写过论文的文献使用者相对阅读比例较低。如图 2-11 所示。

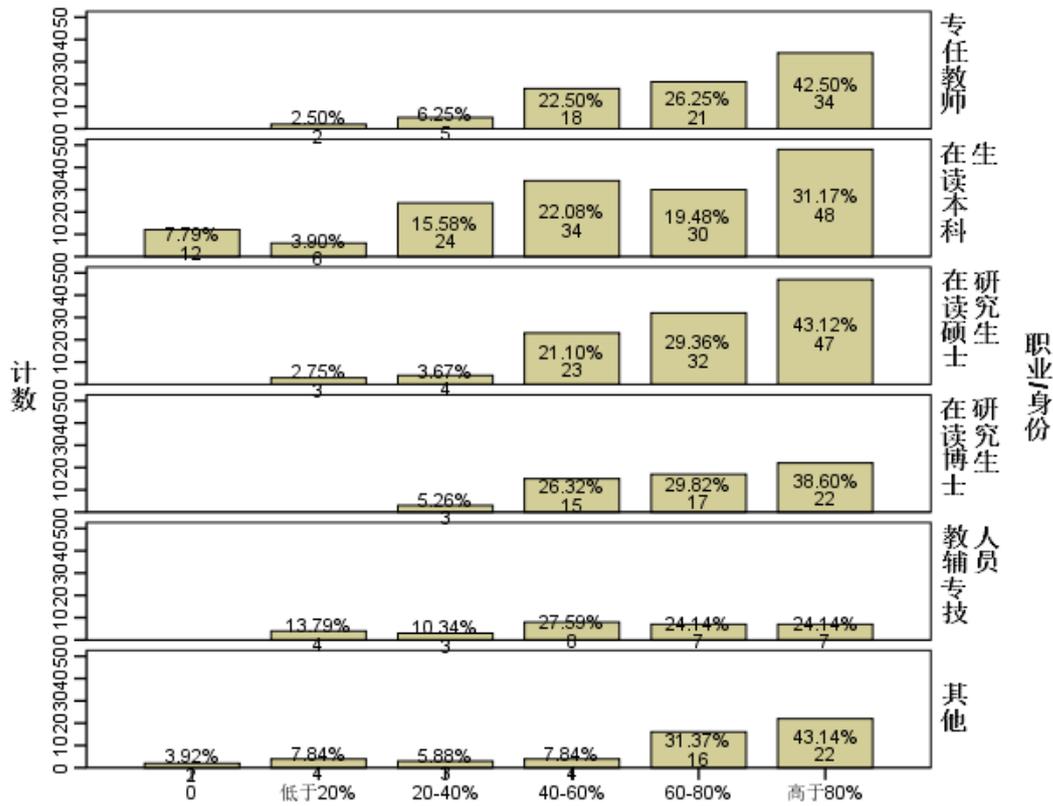


图 2-8 文献下载后阅读的比例

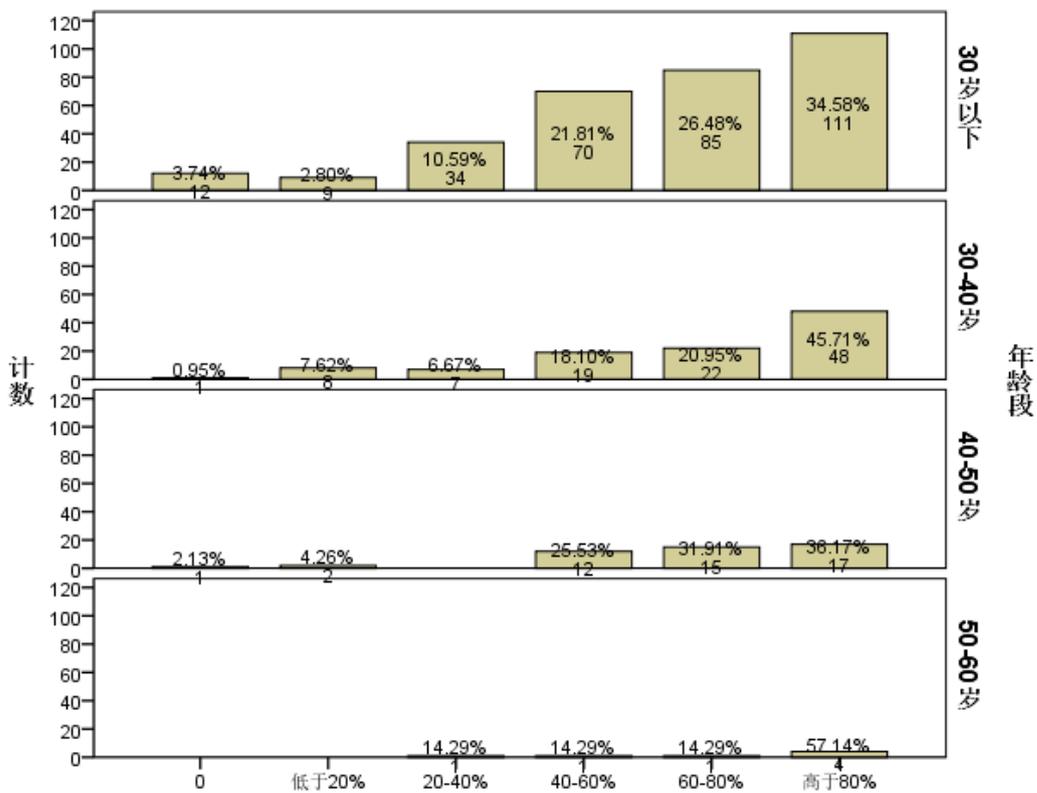


图 2-9 文献下载后阅读的比例

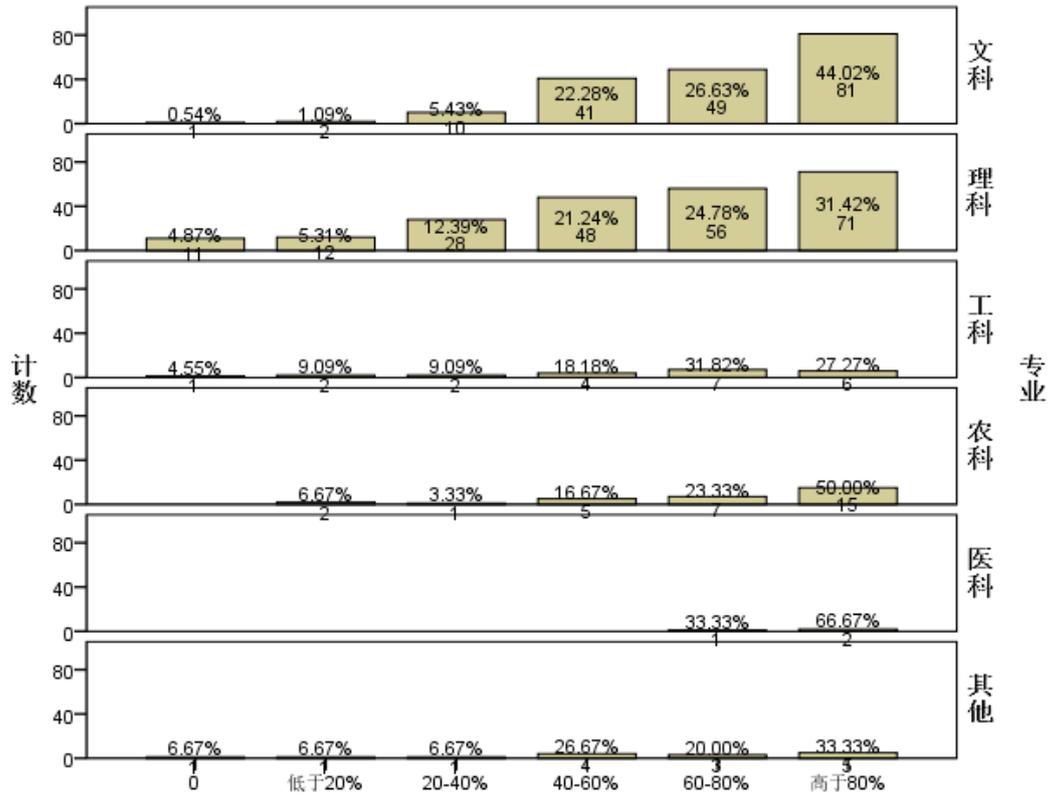


图 2-10 文献下载后阅读的比例

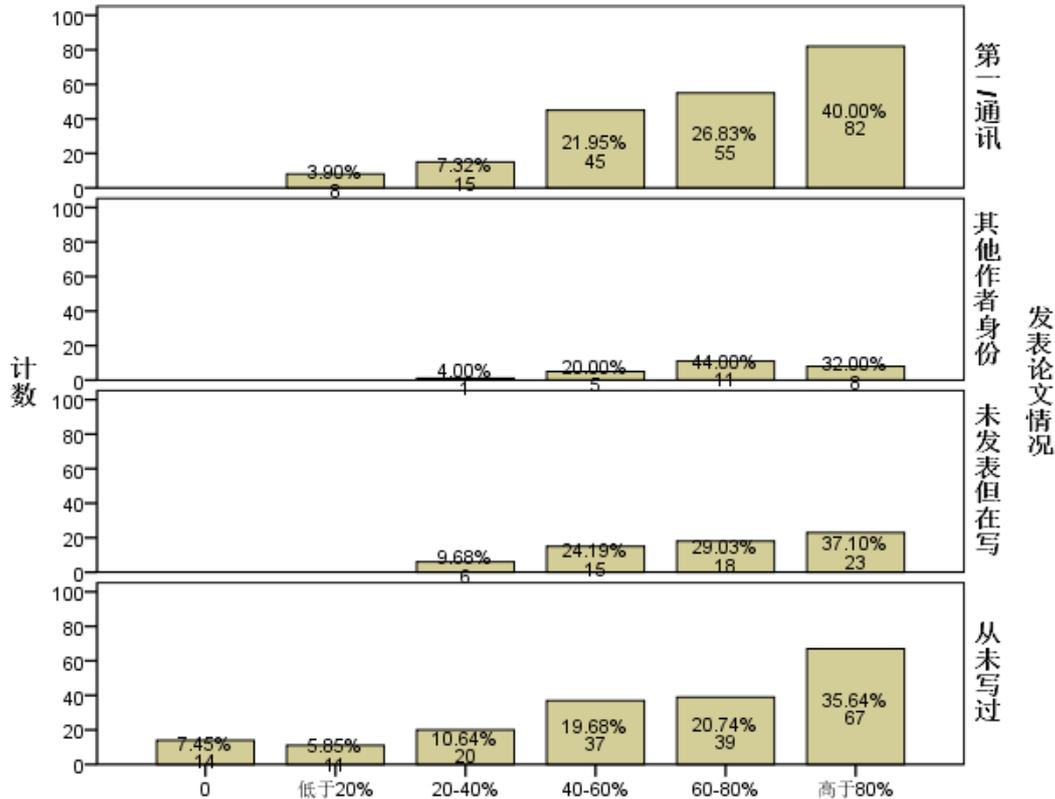


图 2-11 文献下载后阅读的比例

### 2.3.4 不同用户群体在文献使用上的差异分析

#### 1) 不同学科在下载、分享、管理文献上的差异性

通过对数据进行差异检验，发现：下载和管理文献两方面呈现出总体的学科差异性，如文献使用者在“下载文献的频率”、“文献下载后进行阅读的频率”以及“使用文献管理软件或平台的频率”这三方面分别具有显著差异。总体差异如表 2-5 所示。

表2-5 文献使用者在文献下载、管理上的差异检验

变量	样本	检验统计量	自由度	渐进显著性 (2-sided检验)
下载文献的频率	480	45.622	5	.000
使用文献管理软件或平台的频率	480	14.665	5	.012
文献下载后被阅读的频次	480	19.401	5	0.002

显示渐进显著性 (2-sided 检验)。显著性水平是 0.05。

进一步对三组具有显著差异的变量分别进行延后多重比较，发现在“下载文献的频率上”文、农、医科分别与其他学科（如艺、体等）具有显著的学科差异性，理科也分别与文科、农科具有显著的学科差异，如表 2-6 所示；在“使用文献管理软件或平台的频率”上验后多重比较并未呈现出显著的组间差异，学科与学科之间仅有一定程度的差异，没有显著差别；在“文献下载后被阅读的频次”上，仅有理科与文科之间具有显著的学科差异，如表 2-7 所示。

表2-6 不同学科在文献下载频率上的多重比较检验

样本1-样本2	检验统计量	标准误	标准检验统计量	Sig.	调整显著性
其他-文科	131.611	36.081	3.648	.000	.004
其他-农科	153.367	42.492	3.609	.000	.005
其他-医科	251.067	84.984	2.954	.003	.047
理科-文科	69.670	13.342	5.222	.000	.000
理科-农科	-91.425	26.110	-3.501	.000	.007

显示渐进显著性 (2-sided检验)；显著性水平是0.05；仅显示具有显著差异项。

表2-7 不同学科在文献下载后被阅读的频次上的多重比较检验

样本1-样本2	检验统计量	标准误	标准检验统计量	Sig.	调整显著性
理科-文科	49.843	13.211	3.773	.000	.002

显示渐进显著性 (2-sided检验)；显著性水平是0.05；仅显示具有显著差异项。

#### 2) “研究主题与本人研究方向一致”在不同学科、年龄、职业/身份上的差异性

在众多文献特征中，哪些会促使文献使用者下载全文阅读？通过分析发现，“研究主题与本人研究方向一致”评分最高，也即**文献使用者在下载文献时最为看重所下载的文献的研究主题**，若主题与自我研究保持一致，那么就会决定下载；其次，文献发表在高质量期刊上，也是一个重要的促使下载文献的特征；第三个便是文献的被引频次较高，这与下载频次较高这一文献特征相比，文献使用者更看重的是被引频次，而被引频次以及发表在高质量期刊上，都可以认为是对论文质量的一个预判，因此可以认为**文献使用者下载文献时第二看**

重所下载文献的质量。此外，此次调查发现，文献使用者并不偏好发表时间较早的文献，这与文献引用动机存在一定的差异。主要文献特征分布如图 2-12 所示。

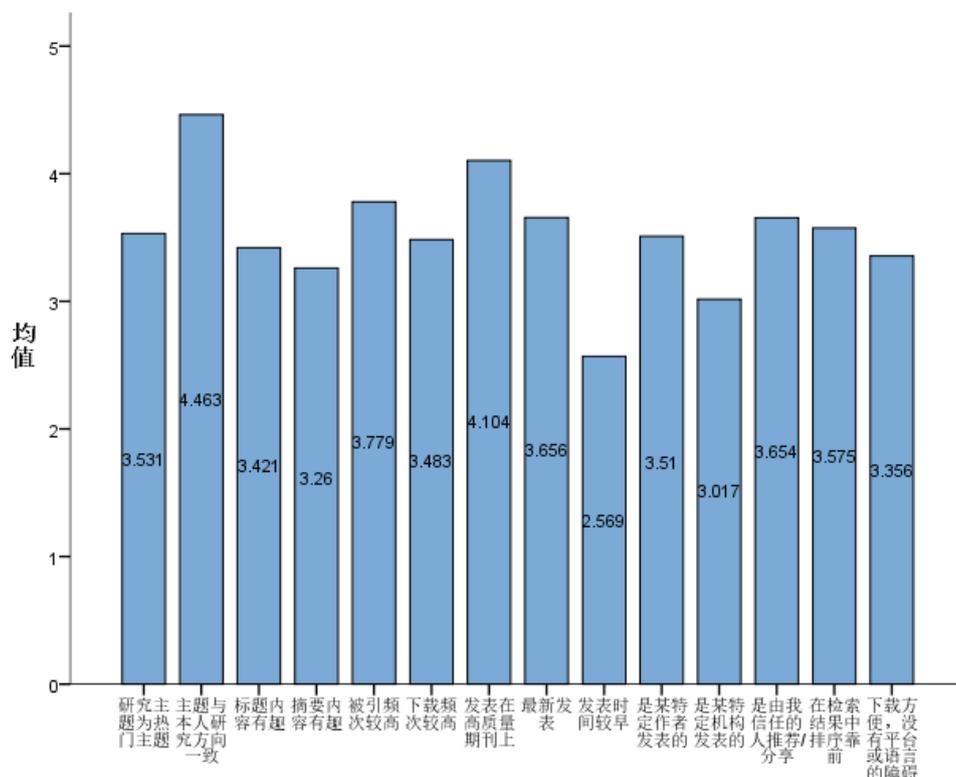


图 2-12 促使文献使用者下载文献的主要文献特征

基于上述分析，“研究主题与本人研究方向一致”是促使众多文献使用者下载文献的首选特征，那么，在不同学科或不同年龄或不同职业/身份特征的文献使用者在进行文献下载时，在该特征的表现上是否具有某些差异？对此，通过非参 Kruskal-Wallis 差异检验发现：在不同学科、不同职业/身份人群上总体存在显著的差异。如表 2-8 所示。

表2-8 “研究主题与本人研究方向一致”在学科、年龄、职业/身份上的差异检验

变量	样本	检验统计量	自由度	渐进显著性 (2-sided检验)
学科	480	20.934	5	.001
年龄	480	5.948	3	.114
职业/身份	480	41.554	5	.000

显示渐进显著性 (2-sided 检验)。显著性水平是 0.05。

进一步验后多重比较分析，学科方面：文科与其他学科存在显著差异、文科与理科存在显著差异，即文科生较其他群体更看重研究主题与本人研究方向的一致性，这也侧面说明理科生的阅读兴趣更加广泛，所下载阅读的文献并不局限于自身的研究主题；职业/身份方面：在读本科生分别与在读硕士、在读博士、专任教师存在显著差异，其余职业/身份人群未表现出显著差异，这也符合预期，即本科生的研究主题主要还是由导师拟定，相对于其他用户群体，并未形成固定的研究方向，因此其阅读兴趣也更加广泛。具体如表 2-9、表 2-10 所示。

表 2-9 不同学科人群在“研究主题与本人研究方向一致”上的多重比较检验

样本 1-样本 2	检验统计量	标准误	标准检验统计量	Sig.	调整显著性
其他-文科	90.008	29.507	3.050	.002	.034
理科-文科	36.902	10.911	3.382	.001	.011

显示渐进显著性 (2-sided 检验)；显著性水平是 0.05；仅显示具有显著差异项。

表 2-10 不同职业/身份人群在“研究主题与本人研究方向一致”上的多重比较检验

样本 1-样本 2	检验统计量	标准误	标准检验统计量	Sig.	调整显著性
本科生-博士生	-63.713	17.037	-3.740	.000	.003
本科生-硕士生	-72.368	13.755	-5.261	.000	.000
本科生-专任教师	77.118	15.144	5.092	.000	.000

显示渐进显著性 (2-sided 检验)；显著性水平是 0.05；仅显示具有显著差异项。

## 3 结论与建议

### 3.1 结论

#### 3.1.1 CASHL 对全球高使用量论文保障率较高

虽然华东师范大学对“经济与商学”学科全球高使用量论文的保障率仅有 83.9%，但是 CASHL 对该学科全球高使用量论文保障率达到了 99.1%，极大地弥补了单个高校文献保障率不足的问题。

对于单个高校，受经费所限，在资源采购时往往侧重本校重点学科，对部分弱势学科的资源保障率可能相对较低，而 CASHL 通过成员高校“优势学科重点部署，联盟全局统筹规划”的保障方式，促进人文社会科学资源的共建共享，为成员高校的教学和科研提供了有利的支撑，是高校实现学术资源保障的重要途径之一。

#### 3.1.2 基于引文的文献保障分析并不全面

从本研究调查问卷来看，许多文献使用者并不发表论文，因此一些文献的价值也无法通过引文进行反映，基于引文的文献保障分析可能并不全面。如本研究中所调查的本科生群体，以及教辅群体，在下载论文时，“满足个人非学术的兴趣”这一动机明显高于其他群体；此外，专任教师下载论文的目的中，“为教学提供素材”也是极为重要的一个动机，而这些都是无法使用动机都无法通过引文进行反映。

被引量和下载量的相关性分析也验证了问卷调查的结论，即存在高使用量低引用量的论文，也存在高引用量低使用量的论文，两者并不完全一致。

因此，通过对文献使用数据（特别是全文下载数据）的分析，能够弥补引文分析的不足，为更有效地保障“有用”文献提供数据支撑。

## 3.2 建议

### 3.2.1 基于文献使用量进行文献资源保障分析

被引量能够一定程度上反映读者在较长时间窗口内的文献使用倾向,但是在反映读者近期使用倾向上则略显不足。因此建议图书馆和 CASHL 在进行文献资源保障分析时,考虑分析近期的全球高使用量论文及图书,这些论文及图书反映了全球科研工作者对于知识的即时需求,理应优先保障。考虑到许多图书馆进行资源采购时是按年采购,因此可参考数据建议为前一年的文献使用数据,以反映文献的最近使用情况。

同时也可以对低使用文献进行特征分析(这也是本研究后续需要扩展的内容),从用户的使用动机上看,低使用文献并不一定表示文献使用价值低,有可能是因为对其揭示度差等原因导致用户无法获知,加强对低使用文献的揭示也是需要进一步需要开展的工作。

### 3.2.2 基于用户使用动机的差异性开展精准推送

虽然用户在使用文献时总体上倾向于使用于与自身研究主题一致的文献以及高影响力文献,但是不同的用户群体也表现出了差异化的文献使用动机。如理科生的阅读兴趣更加广泛,所下载阅读的文献并不局限于自身的研究主题,而本科生也并未形成固定的研究方向,使用主题范围更加广泛,因此基于非学术兴趣的阅读也更为常见,传统的根据用户所属专业来进行文献推送可能并不精准。因此可以根据不同用户的身份差异开展更加精准的文献推送。比如目前华东师范大学个人借阅系统根据个人的借阅历史进行图书推荐, ScienceDirect 也可以根据用户的下载历史进行论文推荐。CASHL 也可以利用类似的算法,通过个人账户的文献请求历史,开展期刊/论文的推荐,让读者在不需检索的情况下获得所需文献。

### 3.2.3 多以中文文字介绍外文文献

CASHL 依靠成员馆之间的合作共享方式,能够保障绝大多数国外权威的人文社会科学电子资源,但是宣传力度不足,甚至各成员馆在对本校特有资源的揭示度上也欠佳,一些重要的外文资源由于得不到有效的揭示而未被使用。因此 CASHL 和各成员馆应加大利用新媒体进行资源的揭示,对于外文文献,最好是进行简单翻译后进行宣传,以降低读者的阅读门槛,提高读者的阅读兴趣。比如目前 CASHL 微信公众号中设有“馆·舍”栏目,进行成员馆空间资源的介绍,也可以设置“馆·藏”栏目,介绍成员馆特有电子资源。酒香也怕巷子深,只有被读者所了解,才能被读者所使用。

## 参考文献

- [1] 汪涛,肖希明. 新信息环境下的文献资源保障系统建设[J]. 图书与情报, 1999(01):33-36.
- [2] 曲岩岩,童旭. 高校特色重点学科文献资源保障实证研究——以哈尔滨工程大学“核学科”为例[J]. 图书馆研究, 2015,45(02):31-35.

- [3] 李海霞. 基于引文分析的图书馆期刊资源建设策略研究——以哈尔滨工程大学图书馆为例[J]. 图书馆建设, 2011(08):43-45.
- [4] 马建华. 引文分析在图书馆文献采集中的作用——北京大学有机化学专业博士论文的文献计量研究[J]. 大学图书馆学报, 2003(03):70-73.
- [5] 徐志玮, 郑建瑜. 高校化学学科用户对纸本/电子期刊需求研究——以中山大学化学学科用户为例[J]. 图书情报知识, 2010(04):44-50.
- [6] 侯利娟, 郝群, 张立彬. 高校图书馆外文文献资源保障研究——以复旦大学数学学科为例[J]. 图书馆, 2018(01):92-100.
- [7] Wang X, Peng L, Zhang C, et al. Exploring scientists' working timetable: A global survey[J]. Journal of Informetrics, 2013,7(3):665-675.
- [8] Wang X, Xu S, Peng L, et al. Exploring scientists' working timetable: Do scientists often work overtime?[J]. Journal of Informetrics, 2012,6(4):655-660.
- [9] Duan Y, Xiong Z. Download patterns of journal papers and their influencing factors[J]. Scientometrics, 2017,112(3):1761-1775.
- [10] Bollen J, de Sompel H V. Mapping the structure of science through usage[J]. Scientometrics, 2006,69(2):227-258.
- [11] Schloegl C, Gorraiz J. Comparison of citation and usage indicators: the case of oncology journals[J]. Scientometrics, 2010,82(3):567-580.
- [12] Schloegl C, Gorraiz J. Global Usage Versus Global Citation Metrics: The Case of Pharmacology Journals[J]. Journal of the Association for Information Science and Technology, 2011,62(1):161-170.
- [13] 宋姬芳, 单向群, 于淼, 等. 基于 WoS 和用户调查的 CASHL 文献资源保障研究——以 ESI 经济学与商学学科为例[R].2018 年新信息环境下 CASHL 资源与服务拓展设计研究, 2018.

## 特别感谢

感谢 CASHL 管理中心以及中国人民大学图书馆宋姬芳、单向群等老师提供的支持和帮助!