



# 中文高下载量论文基本特征研究\*

——以图书情报学为例

□文佩丹 杨新涯 涂佳琪

**摘要** 论文的下下载量代表了读者对论文的关注度。为查明论文的下下载量特征,探究高下载量论文的普遍规律,唤醒对论文下载量的重视,文章以图书情报领域的 18 本 CSSCI 期刊为例,以期刊下载量排名前 100 的论文为样本,采用 VOSviewer、COOC13.6、SPSSPRO 等工具软件对论文的年代分布、关键词共现、关键词演进、关键指标及代表论文进行全面分析。结果表明,论文下载量可以及时反映论文的短期关注度,与发表时间、研究内容、期刊影响力相关,高下载量论文在上述方面具有比较明显的特征,为学术研究和论文评价带来新的启示。

**关键词** 高下载量 基本特征 论文分析 图书情报学

**分类号** G350

**DOI** 10.16603/j.issn1002-1027.2023.04.014

论文是学者对外发布阶段性或重大研究成果的主要方式,对论文进行分析和评价是把握学术动态和研究规律的重要手段。目前 ESI 高被引论文和热点论文作为学术成果的评价标准之一,受到学者和科研管理部门的普遍认可。高被引论文指被引频次在指定年份发表的论文中进入同领域前 1% 的论文,热点论文指过去 2 年所发表的、在最近 2 个月中被引频次排在同领域前 0.1% 的论文<sup>[1]</sup>,从定义可以看出二者均是基于被引频次定义的评价指标。

截至 2022 年 6 月 7 日,中国知网数据显示,《元宇宙视域下的智慧图书馆服务模式与技术框架研究》《元宇宙场域下虚拟社区知识共享模式研究》《从 VR/AR 到元宇宙:面向 α 世代的沉浸式儿童绘本交互设计研究》等多篇发表于 2022 年的以元宇宙为研究主题的学术论文,其下载量均已超过 4000 次,但被引频次为 0。这一突出现象引发我们对论文下载量的关注和思考,新技术、新理论、新概念跨界融合的研究可能会在短期内成为关注焦点,为行业注入新的发展动力,论文下载量在快速判别近期研究焦点上具有可靠性。基于此,本文以图书情报学为

例,展开对中文高下载量论文的基本特征研究。

## 1 研究现状

### 1.1 论文下载量的探讨

论文下载量与论文被阅读次数对应,反映了读者对文献的真实需求,应作为评价论文应用价值的一个重要指标<sup>[2]</sup>。以论文下载量为指标进行论文评价可以克服被引频次的时滞性、“马太效应”等局限。已有的关于论文下载量的研究,普遍关注论文下载量与被引频次的关系<sup>[3-4]</sup>,探讨使用下载量预测被引频次的可能性<sup>[5-11]</sup>。部分研究探讨论文下载量随时间的变化规律,如王超等对《中文核心期刊要目总览》中的部分学科期刊论文进行了下载量分布特征的研究<sup>[11]</sup>;方红玲以中国知网学术文献总库和中国知网引文数据库的 6 个学科 30 种期刊 2005 至 2009 年的下载量为数据来源,对各学科下载量的峰值年代进行研究<sup>[12]</sup>;俞立平等对期刊论文的下载量一年内的时间规律进行研究<sup>[13]</sup>。也有学者研究了论文的微信阅读量与知网下载量的关系<sup>[14]</sup>,以及优先数字出版对论文下载量的影响<sup>[15]</sup>等。

\* 重庆市社会科学规划项目“基于科研产出大数据的重庆市支柱产业发展前沿研究”(编号:2020PY26)和国家社会科学基金项目“智慧图书馆的零数据模型及应用研究”(编号:19BTQ011)的研究成果之一。

通讯作者:文佩丹,wenpeidan@cqu.edu.cn。



## 1.2 论文的特征分析

论文特征分析以文献计量为基础,按分析对象进行分类,论文特征分析大体集中在以某期刊刊发的论文为对象、以某研究主题或学科领域文献为对象以及以某指标下产生的高分论文为对象三个方向。如张慧等以《中国图书馆学报》1998—2017年的论文为研究对象,采用引用量、下载量、引用下载比等指标,对其进行计量和品质研究<sup>[16]</sup>;任通<sup>[17]</sup>、黎思敏<sup>[18]</sup>均以我国智慧图书馆相关学术论文为对象,对其进行计量规律分析,袁毅等以人工智能学科文献为对象,基于社会网络分析法对其跨学科特性进行分析<sup>[19]</sup>;宋超、董建军均以ESI热点论文为对象,对其进行特征分析<sup>[20-21]</sup>;加西亚—多拉多(Garcia-Dorado, D)等对期刊高下载量论文的研究内容进行解析,以期刊为研究对象,分析期刊中高下载量论文的类型和领域,以及将高下载量的论文与高被引频次的论文进行对比<sup>[22-24]</sup>。

从文献调研来看,随着网络信息化的发展,论文下载量指标逐渐受到重视,成为网络计量学的重要指标之一。但目前对于论文下载量的研究,多数还是集中在探讨下载量与被引频次的关系,以及下载量随时间的变化特征等内容上。以高下载量论文为对象,对其特征的研究尚有欠缺,少量研究论文仅对论文的类型和领域进行了初步分析,但高下载量论文反映了读者的选择偏好,体现了论文的应用价值,因此有必要系统地开展专题研究。

## 2 中文高下载量论文基本特征分析框架

文献计量学是以文献体系和文献计量特征为研究对象,采用数学、统计学等方法开展文献的分布结构、计量关系、变化规律等研究的一门科学<sup>[25]</sup>。中文高下载量论文基本特征分析属于文献计量学的范畴,文献计量学理论和方法对其有指导意义。因此,本研究基于文献计量学梳理中文高下载量论文基本特征分析框架。

### 2.1 分析维度

文献计量学研究范围广泛,其中与本研究高度相关的内容包括:以书目、文摘和索引为对象,选取刊名(书名)、出版年、出版地、出版类型等研究者认为有价值的项目进行分析;以文献指标为对象,从文献利用角度出发,分析图书馆或情报单位的阅览数、借阅数、文献复制数、使用者的人数及其分布等文献指标;文献情报流规律的研究,如文献的增长规律、

老化规律、引用规律等;以及在知识管理背景下,拓展出的对知识的计量与评价,比较典型的是关键词与主题词的统计和计量分析。

本文基于以上对文献计量学研究内容的理解,结合中文高下载量论文实际情况,拟从年代分布、关键词、关键指标和代表论文四个维度分析其基本特征。年代分布反映了高下载量论文随发表年份变化的累积或削减规律,是高下载量论文基本特征的重要内容。对高下载量论文关键词的分析可以考查论文下载量与研究内容、研究热点的关系,也是高下载量论文基本特征的重要方面。关键指标的统计和分析,有助于快速发现和揭示高下载论文在代表性指标上的特征和规律,帮助厘清论文下载量与其他相关指标的关系。代表论文分析是在年代分布、关键词、关键指标分析的基础上,对异常或特殊论文的分析,可以进一步完善中文高下载量论文基本特征的分析维度,形成由面及点的分析框架和层次。

### 2.2 分析内容

在四个分析维度下,具体的分析内容包括:(1)中文高下载量论文的年度分布情况,以及与之相关的文献增长与文献老化等现象和规律。文献老化可采用共时半衰期来定量描述,指某学科(专业)现时尚在利用的全部文献中较新的一半是在多长时间内发表的<sup>[25]</sup>,从广义上讲半衰期的概念可以推广,因此本研究将共时半衰期概念拓展至高下载量论文的高下载半衰期,即现时高下载量论文中较新的一半论文的年龄。(2)中文高下载量论文的关键词共现和关键词演进规律,关键词共现用于表征关键词的出现频次,帮助识别热点研究内容,关键词演进规律用于呈现研究主题随年代的变化情况,进一步明确高下载量论文与研究内容的关系。(3)下载量、被引频次、量引比等的最大值、最小值、均值、中位数、离散系数等统计学指标的分析和对比,量引比指论文下载量均值与被引频次均值的比值,代表平均产生1次被引所需的下载次数。(4)最高下载量论文、零被引论文、当年高下载量论文三类代表论文的分析,可以探究最高下载量论文和当年高下载量论文的特征以及论文量引背离的原因。通过以上内容的分析,重点关注和探求下载量与论文年龄的关系、下载量与研究内容的关系、下载量与期刊的关系以及下载量与被引频次在反映研究热点和论文评价中的优势和区别。



### 2.3 分析方法

文献统计分析法和数学模型分析法都是文献计量学的重要研究方法,中文高下载量论文的基本特征分析也基于这两类方法展开。采用文献统计分析法完成文献原始数据的采集、整理,以及年代分布图的绘制、关键指标的计算和分析等基本的数理统计工作;采用数学模型分析法建立论文下载量与其他相关因素的数量关系,如回归分析等。除此之外,采用知识图谱的方法对中文高下载量论文的关键词进行可视化描述和分析,更加直观地反映高下载量论文在研究内容方面的基本特征。

根据分析维度、分析内容和方法的梳理,构建起中文高下载量论文基本特征分析的框架,如图1所示。总体来说,中文高下载量论文基本特征分析就是采用统计分析法、数学模型分析法和知识图谱等多种方法,围绕样本文献的年代分布、关键词、关键指标以及代表论文,开展文献的增长与老化规律、关键词共现与演进、关键指标的统计与对比、代表论文典型特征的分析等研究,以探明论文下载量与论文年龄的关系、下载量与论文研究内容的关系、下载量与期刊影响力的关系以及下载量和被引频次在反映研究热点和论文评价中的区别与优势。

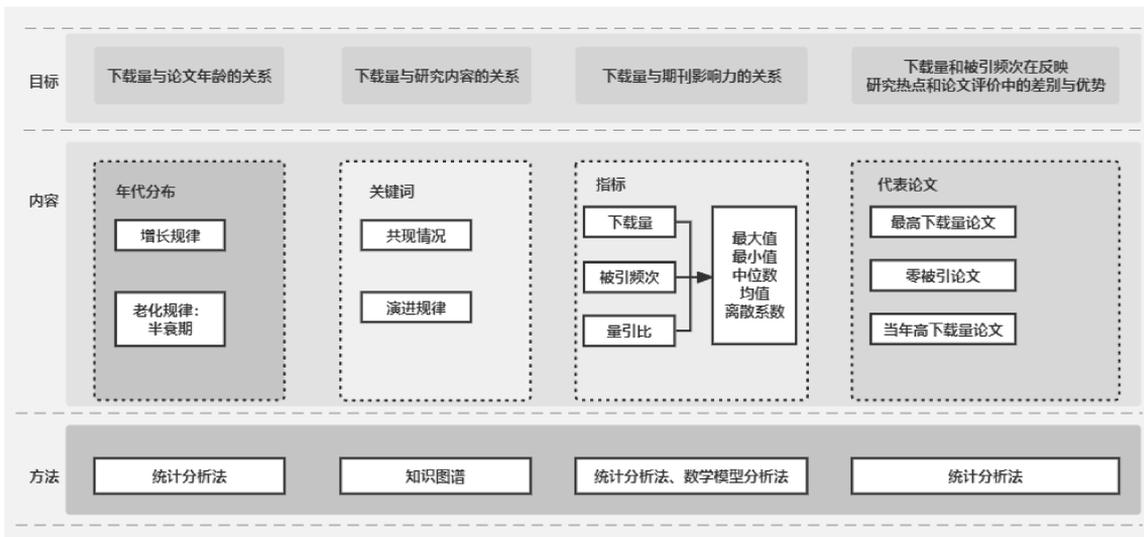


图1 中文高下载量论文基本特征分析框架

## 3 中文高下载量论文基本特征实证分析

实证分析内容包括高下载量论文的年代分布、关键词共现分布、关键词演进以及代表论文分析等,分析指标包括下载量、被引频次、量引比等。

### 3.1 数据与工具

将期刊累计下载量排名前100的论文定义为该刊的高下载量论文,依据文献集中规律<sup>[25]</sup>,即某一学科的大部分文献高度集中在少数期刊中,以中国知网数据库总库收录的图书情报领域的18本CSSCI来源期刊为论文来源,分析中文高下载量论文的基本特征。

数据的采集流程:在中国知网数据库总库以文献来源为检索条件,分别对图书情报领域的18本CSSCI来源期刊进行检索,并以下载量降序排列,将排名前100的论文以Refworks格式文件导出,并将

其题名、文献来源、关键词、摘要、发表时间等关键信息整理汇编入EXCEL表格以备数据分析。数据的采集时间为2022年6月7日,采集数据共计1800条。

数据分析所采用的工具包括EXCEL、VOSviewer和SPSSPRO分析软件,这些分析工具分别在数据汇总与基础统计、关键词密度图绘制、关键词演进规律以及回归分析等数理统计与分析方面具有便捷性和可靠性。EXCEL用于高下载量论文年代分布分析和关键指标的统计对比,VOSviewer用于关键词共现分析,COOC13.6用于关键词演进规律分析,SPSSPRO用于数据的正态检验、相关性分析和线性回归分析。

### 3.2 基本特征概述

#### 3.2.1 年代分布

论文的下载量与发表时间有关,分析高下载量

论文的年代分布,可以查明高下载量论文的时间特征。所采集的 1800 篇文献的年代分布如图 2 所示。

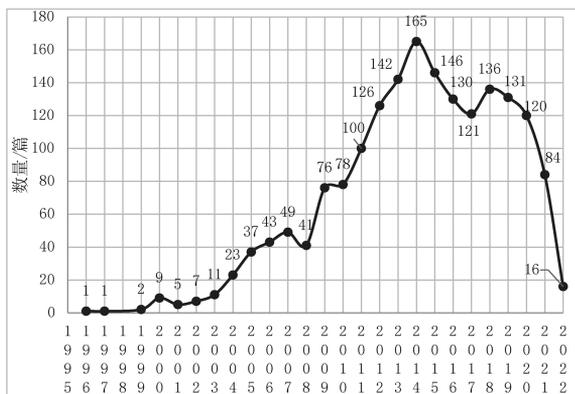


图 2 图书情报学中文高下载量论文年代分布

由图 2 可看出,图书情报学中文高下载量论文的年代分布区间为 1996 年至 2022 年,并且在 2014 年达到峰值,即自发表起大约经过 8 年时间的论文有更大可能成为高下载量论文,发表时间超过 8 年的论文成为高下载量论文的概率反而会降低。新发表的论文更容易受到关注,被读者下载阅读,这与许多学者研究发现的文献老化现象相符<sup>[26-29]</sup>,根据半衰期的定义,计算得出图书情报学中文高下载量论文的高下载半衰期约为 7—8 年。高下载量论文中,2008 年以前的论文明显少于 2009 年及以后发表的

论文,即发表时间在 1—13 年的论文明显多于 13 年以上的论文,其占比达到 86.39%,这是数字阅读的发展与文献老化共同作用的结果。我们可以将发表时间在 1—13 年的论文称为下载量“活跃论文”。2021 年发表的论文数量已达到 84 篇,在年代分布表中的排序位于 11/26,证明论文的下载量可以在 1 年时间达到较可观的水平。

### 3.2.2 关键词共现

通过 VOSviewer 分析得出 1800 篇文献关键词共计 6910 个,对参数进行优化调节设置,关键词至少共现 15 次(共现次数的设置不影响高频关键词结果,仅为筛除低频关键词以确保共现图的清晰度),得到 95 个符合阈值要求的关键词。对 95 个关键词进行中英文去重,并去除图书馆、大学图书馆、公共图书馆等与研究实质性内容关系不大的关键词,剩下 57 个有效关键词,其共现情况如图 3 所示。图 3 显示,样本高下载量论文中与大数据相关的研究最多,其次是网络舆情,再次是阅读推广、知识图谱和影响因素方面的研究。这些研究分别代表了图书馆学和情报学不同的研究热点。

### 3.2.3 关键词演进规律

采用 COOC13.6 的趋势变化功能开展关键词的演进规律研究,初步分析发现 2006 年及以前当年关

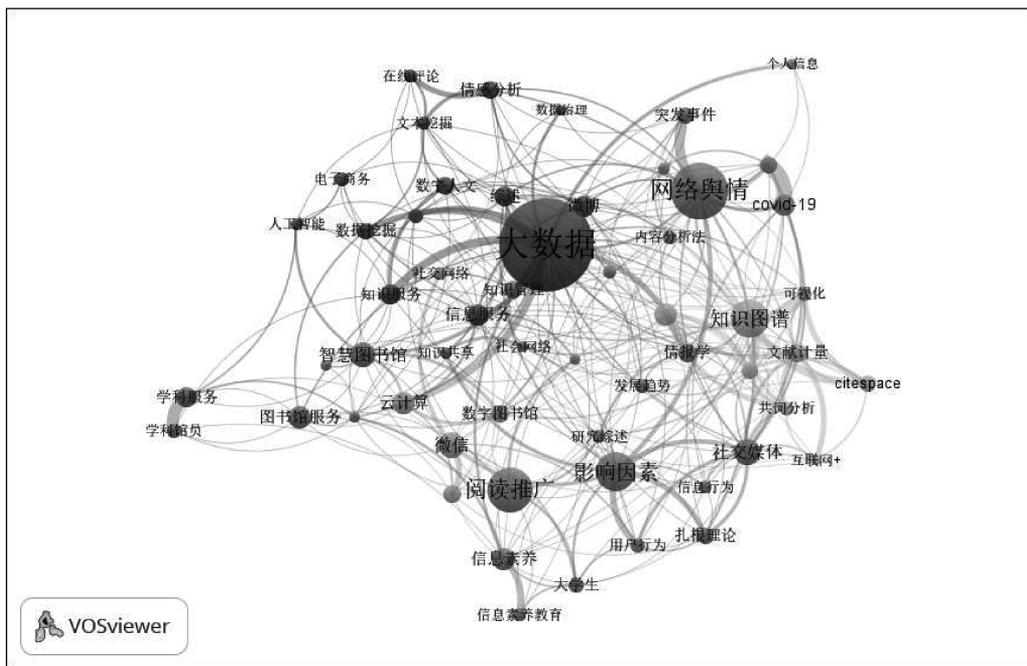


图 3 图书情报学中文高下载量论文关键词共现图



关键词共现均未超过 5 次,这一方面受文献总量的影响,另一方面也表明这段时期的研究未形成明显的热点。因此,仅对 2007 年及以后的样本文献进行关键词演进规律分析,去除图书馆等不能代表研究实质内容的关键词后,每年共现次数排名前 5 的关键词如图 4 所示。从图中可以看出,大数据对图书情报领域产生了深刻的影响,在 2012 年开始成为关注对象,并且在此后的 7 年内都是研究焦点。网络舆

情、阅读推广等的研究也具有比较好的持续性,其受关注的时间跨度均大于 5 年。另外,关于移动图书馆、知识图谱、用户画像、影响因素等的研究也曾在某些年份受到过短暂的关注,但其持续性欠佳,受关注年限均不超过 3 年。同时,我们还特别关注到新冠肺炎和元宇宙的研究,分别在 2020 年和 2022 年出现突发性增长。由此可见,高下载量论文在反映研究热点变化方面具有较好的时效性和客观性。

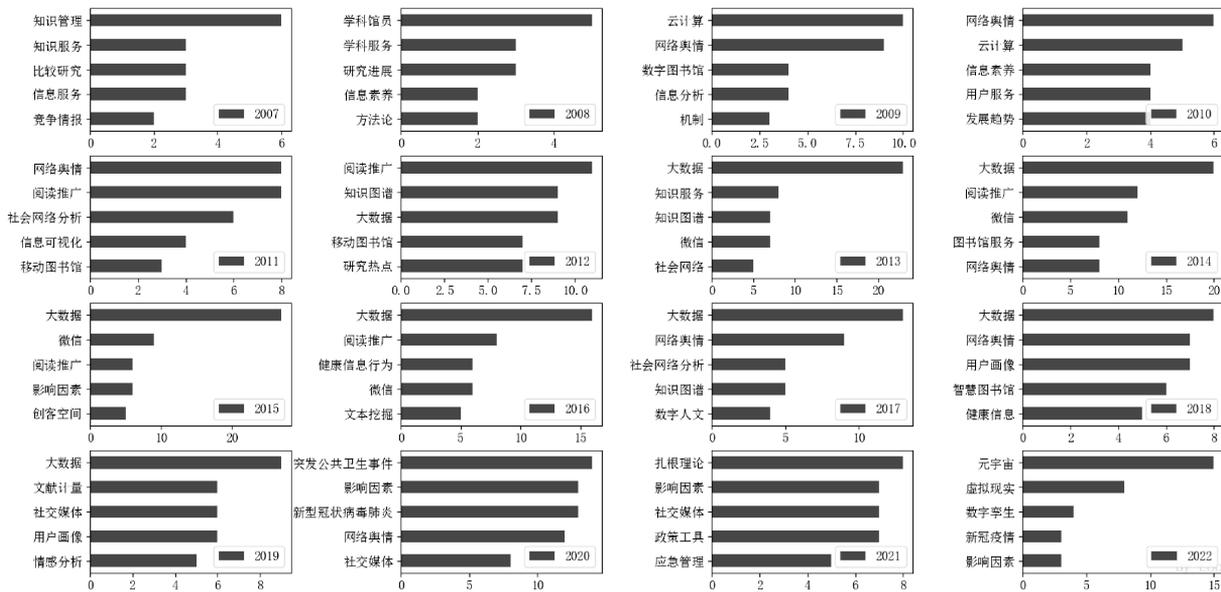


图 4 图书情报学中文高下载量论文关键词演进图

### 3.2.4 关键指标对比

对各期刊高下载量论文的下下载量、被引频次、量引比等关键指标及其离散系数进行统计分析,按期刊下载量最大值排序,结果如表 1 所示。部分文章被引值为空,在统计分析时统一计为 0。

表 1 显示,《中国图书馆学报》的高下载量论文在下载量和被引频次上都表现突出,除被引频次最小值和量引比外,其他指标都位居 18 本期刊榜首。为此,我们采用 SPSSPRO 软件对各期刊高下载量论文的下下载量均值和期刊复合影响因子的相关性进行分析,以判断文献下载量与期刊影响力的关系,正态性检验发现下载量均值符合正态分布,但期刊复合影响因子不符合正态分布,选择 Spearman 相关系数进行分析<sup>[30]</sup>,得到二者的相关性在 10% 水平上显著,相关系数为 0.465,表明各期刊高下载量论文的下下载量均值与期刊复合影响因子存在中等强度的相关性。进一步进行线性回归分析(最小二乘法),结果如表 2 所示,拟合线性关系为:下载量均值 =

1398.996 + 707.799 × 期刊复合影响因子。

我们还注意到《图书与情报》下载量最小值仅为 224,与其他 17 本期刊的下下载量最小值相差甚远,值得特别关注。《图书与情报》下载量最小值的论文为《公共图书馆存在的理由:来自图书馆使命的注解》,文章调研了世界范围内主要政策性、纲领性文献对公共图书馆使命的阐释,在此基础上分析教育、信息服务、文化传播、促进社会和谐、培育信息素养、培养阅读兴趣、扫盲等使命对公共图书馆存在的理由的效度,以引发对公共图书馆存在理由的思考。论文下载量为 224,被引频次为 141,计算得出量引比仅为 1.59,表明该论文没有引发普遍关注,但其下载与被引的转化率较高。其受众主要为从事相关研究、对该领域有调研需求的群体,原因在于该篇论文为典型的理论研究论文,选题小众、立意高、理论性强。

表 1 的指标中,离散系数最大的是被引频次最小值,区间为 0—11,表明各期刊在被引频次最小值上存在统计学意义上的最大离散程度,并且最小值

2023 年 第 5 期  
大学图书馆学报



为0的期刊占比达到66.67%,说明大部分期刊的高下载量论文中有零被引文章的存在。离散系数最小的是量引比,说明各期刊产生一次引用所需的下载量差别相对较小。

表1 各期刊高下载量论文关键指标对比

期刊	下载量				被引频次				量引比
	最大值	最小值	中位数	均值	最大值	最小值	中位数	均值	
中国图书馆学报	38266	3345	5300	7085.43	1662	3	112.5	168.49	42.05
情报理论与实践	23944	3027	3711	4816	489	0	72	92.14	52.27
图书情报工作	19123	3295	4333	5217.69	740	1	109.5	136.11	38.33
图书情报知识	18073	2229	2893.5	3817.04	441	1	69	83.52	46.17
现代情报	16866	2996	3960	4778.13	529	0	59	79.5	60.1
情报杂志	15972	3651	4780.5	5800.86	569	0	107	132.63	43.74
图书与情报	15293	224	2847.5	3601.94	393	10	80	94.82	37.99
信息资源管理学报	11610	1058	1589	1900.79	678	0	15	26.27	72.36
图书馆学研究	11407	2144	2832.5	3544.12	298	0	57.5	68.8	51.51
大学图书馆学报	11094	2227	3031	3719.66	674	11	89.5	124.59	29.86
图书馆杂志	11034	1723	2369.5	2900.48	262	0	64.5	70.49	41.15
图书馆建设	10987	1606	2074	2481.67	526	0	50	68.25	36.36
情报科学	10771	2801	3512.5	4104.39	308	2	84	106.43	38.56
数据分析与知识发现	10440	1735	2348	2802.19	580	0	57.5	84.04	33.34
情报资料工作	8476	1617	2107	2465.77	271	0	49.5	53.41	46.17
图书馆论坛	8067	2042	2599	3037.05	372	0	47.5	57.94	52.42
国家图书馆学刊	6814	1170	1625	2081.95	416	0	40	55.05	37.82
情报学报	5586	1508	2033	2321.19	312	0	29	43.51	53.89
离散系数	0.54	0.43	0.36	0.38	0.60	2.16	0.41	0.42	0.23

表2 线性回归分析结果

线性回归分析结果 n=18									
	非标准化系数		标准化系数	t	P	VIF	R <sup>2</sup>	调整 R <sup>2</sup>	F
	B	B/t	Beta						
常数	1398.996	782.782	—	1.787	0.093 *	—	0.378	0.339	F=9.736
复合影响因子	707.799	226.835	0.615	3.12	0.007 * * *	1.000			P=0.007 * * *

注: \* \* \*、\* 分别代表1%、10%的显著性水平。

### 3.3 代表论文分析

以中文高下载量论文基本特征概述为基础,选取具有代表性的论文进行个性化的分析和比较,以挖掘高下载量论文在某些方面的典型特征。

#### 3.3.1 最高下载量论文比较

18种期刊最高下载量论文对比信息如表3所示。数据显示,18种期刊的最高下载量论文有1/3是关于已有研究现状的综述类文章,这些文章不仅下载量高,被引频次也高,这6篇综述类文章的被引频次区间为278—424,下载量区间为10440—

23944,属于名副其实的下下载量和被引频次双高论文。由此可见,综述类文章较易成为双高论文。表中发表年份最近的论文发表于2021年,研究主题均为元宇宙,同时关注到2011年和2012年关于大数据、智慧图书馆的文章也是期刊最高下载量文章。2021年和2011年正是元宇宙和智慧图书馆在图书馆界的研究元年,这说明高下载量论文确实具有鲜明的时代特征,新技术、新概念、新理论的提出往往会在图书情报领域引起强烈反响,论文下载量也的确可以快速反映研究热点。



表 3 各期刊最高下载量论文比较

期刊	论文题名	下载量	被引频次	发表年份
中国图书馆学报	论信息安全、网络安全、网络空间安全	38266	324	2015
情报理论与实践	国内网络舆情研究的回顾与展望	23944	424	2009
图书情报工作	国内知识图谱应用研究综述	19123	398	2013
图书情报知识	突发公共事件网络舆情研究综述	18073	278	2014
现代情报	浅谈高校图书馆与大学生信息素质教育	16866	38	2007
情报杂志	绩效管理理论研究综述	15972	347	2010
图书与情报	基于 5T 理论视角下的企业微博营销策略及应用分析——以欧莱雅的微博营销为个案研究	15293	124	2012
信息资源管理学报	元宇宙视域下的用户信息行为: 框架与展望	11610	17	2021
图书馆学研究	我国网络舆情研究与发展现状分析	11407	298	2009
大学图书馆学报	当代大学生阅读问卷调查分析	11094	104	2011
图书馆杂志	图书馆需要怎样的“大数据”	11034	262	2012
图书馆建设	未来图书馆的新模式——智慧图书馆	10978	526	2011
情报科学	B2C 电子商务网站用户体验评价研究	10771	198	2013
数据分析与知识发现	K-means 算法研究综述	10440	395	2011
情报资料工作	突发事件驱动的应急决策知识库结构研究	8476	27	2015
图书馆论坛	元宇宙是图书馆的未来吗?	8067	15	2021
国家图书馆学刊	中国互联网络信息中心发布第 46 次《中国互联网络发展状况统计报告》	6814	118	2020
情报学报	多维视角下应急情报管理体系的知识库构建研究	5586	29	2017

### 3.3.2 零被引论文分析

在高下载量论文中,有部分零被引论文存在,这部分论文属于典型的量引背离论文,对其进行分析,可以探究量引背离论文的规律。样本高下载量论文中的零被引论文共计 36 篇,发表时间不足 1 年的占比 83.33%,发表时间超过 2 年的占比 16.67%,发表时间超过 5 年的仅 3 篇,占比 8.33%,分别是《访谈提纲》《中文网络报刊资料的检索技巧》《如何在网上进行课题检索》,这 3 篇文章均为研究性不强但与实际工作高度相关的文献。由此可见,论文产生被引需要时间的累积,在 1 年内产生被引还比较困难,但 2 年以上的文章一般都可突破零被引。通过分析高下载量论文中的零被引现象,可以看出下载量受发表时间和研究价值的影响更小,能够反映出研究内容的真实关注度,但被引频次更能体现出研究的质量和水平,反映论文的中长期学术影响力。

### 3.3.3 2022 年高下载量论文分析

1800 篇高下载量论文中,有 16 篇论文发表于 2022 年,如表 4 所示。发表时间最晚的论文是 2022 年 4 月 25 日发表于《图书馆论坛》的《元宇宙赋能虚拟图书馆: 理念、技术、场景与发展策略》,该论文不

足 2 个月便成为该期刊的高下载量论文,下载量为 2781 次,被引频次为 0。论文中有 12 篇被引频次为 0,占比达到 75%,其余论文被引频次也不超过 3 次,可见与下载量相比,用被引频次衡量论文关注度的确存在滞后性。有 10 篇论文以元宇宙为研究主题,这与目前的研究热点相符,也表明下载量可以在短期内准确提示研究热点。

### 3.4 基本结论

通过前述分析和研究,可得出如下基本结论:

- (1) 论文的下下载量与发表时间有关,会随着时间累积增长,但事实上新发表的论文比年代久远的老旧论文有更大可能成为高下载量论文。
- (2) 论文下载量与论文类型和研究内容有关,研究综述具有累积下载量和被引频次双高的优势,而热点研究具有短期下载量高的优势。
- (3) 期刊间论文下载量和被引频次等关键指标存在差异,下载量均值与期刊复合影响因子存在相关关系,相关系数为 0.465,二者的拟合线性关系为: 下载量均值 = 1398.996 + 707.799 × 期刊复合影响因子。

(4) 在反映研究热点上,论文下载量比被引频次更及时,样本分析显示有文章在 2 个月内成为了高



下载量论文,而总体来看论文在1年内产生被引还比较困难。

影响力方面比下载量具有更强的客观性,样本分析显示研究性不强的论文发表时间超过5年也难以突破零被引。

(5) 论文被引频次在反映论文质量和中长期影

表4 2022年发表的图书情报学中文高下载量论文

期刊	论文题名	下载量	被引频次	发表日期
情报科学	沉浸理论视角下元宇宙图书馆“人、场、物”重构研究	3320	2	1月1日
情报理论与实践	元宇宙视域下的智慧图书馆服务模式与技术框架研究	4060	0	1月11日
	元宇宙场域下虚拟社区知识共享模式研究	4612	0	3月18日
情报资料工作	在线健康社区中基于SOR模型的用户参与行为影响因素研究	2187	0	3月17日
数据分析与知识发现	在线社区用户画像及自我呈现主题挖掘——以网易云音乐社区为例	3675	0	3月2日
	我国智慧图书馆的发展现状与发展趋势研究	2373	0	1月6日
图书馆建设	元宇宙时代图书馆、档案馆与博物馆(LAM)的技术采纳及其责任创新:以NFT为中心的思考	3026	2	1月15日
	从VR/AR到元宇宙:面向α世代的沉浸式儿童绘本交互设计研究	4095	0	2月16日
	元宇宙赋能虚拟图书馆:理念、技术、场景与发展策略	2781	0	4月25日
图书馆论坛	农村公共文化服务体系与乡村旅游的融合发展——以浙江德清总分馆改革为例	2609	1	1月25日
	元宇宙赋能的图书馆社会教育:场景、审视与应对	2606	0	2月10日
	元宇宙中的数字记忆:“虚拟数字人”的数字记忆概念模型及其应用场景	3880	0	3月16日
	图书馆即教育:元宇宙视域下的公共图书馆社会教育	2258	0	1月24日
信息资源管理学报	基于协同创新理论的政府大数据治理框架的构建——基于G省的案例研究	1715	0	2月17日
	突发公共卫生事件国际媒体涉华报道特征及相关因素分析——基于COVID-19数据	1089	0	2月18日
中国图书馆学报	元宇宙之治:未来数智世界的敏捷治理前瞻	4311	3	1月11日

## 4 启示与展望

### 4.1 研究启示

通过前述分析,可以在学术研究和论文评价两方面获得启发,对未来的研究和评价具有指导意义。

在学术研究方面,可以重点从以下方面着手开展研究:(1)关注社会普遍关注的问题,将研究融入大社会,不局限于小学科,如新冠肺炎与网络舆情;(2)结合国家战略,做国家需要的研究,如全民阅读与阅读推广;(3)关注技术变迁给行业带来的影响和变革,如大数据与智慧图书馆、元宇宙与图书馆服务等。

在论文评价方面,多维度的分析均表明,论文下载量比被引频次更能及时反映论文的关注度。特别

是对于发表时间不超过1年的论文,下载量将更能体现其关注度和影响力;而论文被引频次则在评价论文质量和中长期影响力方面具有比较优势。因此,常规的仅以被引频次评判论文价值的方法具有一定的局限性,应尽可能注意短期关注与长期价值、学术水平与社会效应各方面的平衡,体现评价的动态客观性,起到以评价促进研究的作用。

### 4.2 值得进一步研究的问题

中文高下载量论文基本特征的分析为论文评价带来了新的启示,限于篇幅和研究内容,本研究没有进一步探讨单篇论文评价指标的选取与构建,具体包括评价指标的考查与筛选、指标的组合构建以及指标的广泛试验等内容。但这对于改变目前论文评



价的“刊本位”和“高被引”现象非常重要,值得深入研究。

### 参考文献

- 1 Clarivate. Glossary[EB/OL]. [2023-05-04]. <http://esi.help.clarivate.com/Content/glossary.htm>.
- 2 杨春华,王桂枝,马红月.换个角度看服务:从下载量看图书馆论文的价值与需求[J].图书与情报,2010(4):100-103,111.
- 3 O'Leary D. On the relationship between citations and appearances on "top 25" download lists in the *International Journal of Accounting Information Systems* [J]. *International Journal of Accounting Information Systems*, 2008, 9(1): 61-75.
- 4 O'Leary D E. The relationship between citations and number of downloads in *Decision Support Systems* [J]. *Decision Support Systems*, 2008, 45(4): 972-980.
- 5 熊泽泉,段宇峰.论文早期下载量可否预测后期被引量——以图书情报期刊为例[J].图书情报知识,2018,35(4):32-42.
- 6 王超.期刊论文被引量与下载量关系研究[J].情报探索,2020(6):33-39.
- 7 谢娟,龚凯乐,成颖,等.论文下载量与被引量相关关系的元分析[J].情报学报,2017,36(12):1255-1269.
- 8 刘雪立,方红玲,苗媛,等.五种综合性眼科学期刊论文下载量与被引量的关系及部分论文的引量背离现象[J].中国科技期刊研究,2010,21(5):629-632.
- 9 Jahandideh S, Abdolmaleki P, Asadabadi E B. Prediction of future citations of a research paper from number of its internet downloads[J]. *Medical Hypotheses*, 2007, 69(2): 458-459.
- 10 Zavos C, Kountouras J, Zavos N, et al. Predicting future citations of a research paper from number of its internet downloads: the *Medical Hypotheses* case[J]. *Medical Hypotheses*, 2008, 70(2): 460-461.
- 11 王超,李书宁,李晓娟.期刊论文下载分布特征及其机制研究[J].情报科学,2016,34(12):59-63.
- 12 方红玲.我国科技期刊论文被引量和下载量峰值年代——多学科比较研究[J].中国科技期刊研究,2011,22(5):708-710.
- 13 俞立平,江菲.学术期刊论文下载量的年内时间规律研究[J].情报杂志,2020,39(4):140-144,126.
- 14 赵文青,宗明刚.学术论文微信阅读量与知网下载量的关系研究[J].中国科技期刊研究,2019,30(9):967-975.
- 15 徐云峰,徐红星,陆海燕,等.优先数字出版影响发表时滞和下载量的定量分析[J].编辑学报,2013,25(6):607-608.

- 16 张慧,叶鹰.1998-2017年《中国图书馆学报》论文的计量研究与品质分析[J].中国图书馆学报,2018,44(5):34-46.
- 17 任通.2010年至今我国智慧图书馆研究领域的计量规律研究[J].高校图书馆工作,2022,42(3):37-44.
- 18 黎思敏.我国智慧图书馆研究领域的知识图谱可视化分析[J].图书馆研究与工作,2021(1):23-28.
- 19 袁毅,季泽豪.人工智能跨学科特性研究——文献计量分析的视角[J].图书馆杂志,2022,41(6):46-52.
- 20 宋超,陈悦,汪玲,等.热点论文分布特征与影响因素分析——兼评时间窗口与学科间差异[J].图书情报工作,2019,63(16):84-94.
- 21 董建军.中国作者发表的ESI热点论文分析与启示[J].情报探索,2021(10):119-127.
- 22 Garcia-Dorado D, Schluter K D, Martinson E A, et al. Which papers are most interesting to the readers of *Cardiovascular Research*? —Information from download monitoring [J]. *Cardiovascular Research Oxford Univ Press*, 2005, 65(1): 1-5.
- 23 Coats-Andrew J S. The top papers by download and citations from the *International Journal of Cardiology* in 2007 [J]. *International Journal of Cardiology*, 2008, 131(1): E1-E3.
- 24 Coats-Andrew J S. Top of the charts: download versus citations in the *International Journal of Cardiology* [J]. *International Journal of Cardiology*, 2005, 105(2): 123-125.
- 25 邱均平.文献计量学(第二版)[M].北京:科学出版社,2019:9.
- 26 邱均平.我国科学文献的时序分布与老化研究[J].图书情报工作,1989(4):5-12.
- 27 李霄.材料科学文献半衰期研究[J].武汉理工大学学报,2009,31(1):153-156.
- 28 刘雪立.科技期刊下载量半衰期的建立及其文献计量学意义[J].中国科技期刊研究,2012,23(4):561-564.
- 29 Burton R E, Kebler R W. The "half-life" of some scientific and technical literatures [J]. *Journal of the Association for Information Science and Technology*, 1960, 11(1): 18-22.
- 30 张厚粲,徐建平.现代心理与教育统计学[M].北京:北京师范大学出版社,1988:112-115.

作者单位:重庆大学人文社会科学高等研究院,重庆大学图书馆,重庆,400044

收稿日期:2022年12月23日

修回日期:2023年5月16日

(责任编辑:支娟)

## A Study on the Basic Characteristics of Highly Downloaded Chinese Papers

—Taking Library and Information Science as an Example

Wen Peidan Yang Xinya Tu Jiaqi

**Abstract:** The number of papers downloaded represents the readers' attention to the papers. In order to find out the characteristics of the number of papers downloaded, explore the general laws of high number of papers downloaded, and arouse the attention to the paper download, this paper takes 18 CSSCI journals in the field of library and information as an example, selects the top 100 papers in the number of downloaded journals as a sample, uses VOSviewer, COOC13.6 and SPSSPRO to analyze the age distribution, keyword co-occurrence, keyword evolution, key indicators and representative papers of papers, and furthermore discusses the selection and construction of evaluation indicators for single paper. The results show that the number of downloaded papers can timely reflect the short-term attention of papers, and the number of papers downloaded is related to publication time, research content and journal influence. The papers with high downloads have obvious characteristics in the above aspects, which brings new enlightenment for the scientific research and paper evaluation.

**Keywords:** High Download; Basic Characteristics; Paper Analysis; Library and Information Science