

# 基于 Altmetrics 的开源软件学术影响力评价研究\*

赵蓉英 魏明坤 汪少震

**摘 要** 开源软件在学术领域的应用不断增多,但其在学术领域的影响力还没有引起足够的重视。本文利用 Altmetrics,对开源项目 Depsy 中 Python 社区的部分软件从下载量、文献被引次数以及软件之间的复用次数三方面对开源软件的学术影响力进行评价和分析。研究发现:①随着 Web2.0 的发展,学术影响力的评价不仅包括论文、专著、研究报告等传统形式的学术研究成果,还包括博客、视频、数据集、软件代码等非传统形式的学术研究成果;②开源软件的发展对学术领域的创新具有催化作用;③Python 社区中各软件的影响力差距较大,大部分软件的学术影响力还很低。研究开源软件学术影响力,有利于丰富学术影响力评价的研究内容,完善学术影响力评价体系。图 5。表 3。参考文献 40。

**关键词** Altmetrics ImpactStory 开源软件 学术影响力

**分类号** G250.252

## Evaluation of Academic Impact of Open Source Software Based on Altmetrics

ZHAO Rongying, WEI Mingkun & WANG Shaozhen

### ABSTRACT

The development of Internet and social media has an increasingly stronger influence on academic communication with the advent of Web 2.0 technology. Nowadays academic achievements include not only papers but also non-paper forms of research results, such as blogs, videos, data sets, software code etc. Academic influence is the impact on scientific research achievements and academic status in academia. These non-paper forms of academic research should be also included in the evaluation system of scholars or academic achievements. Open source software, as a non-paper form of academic achievements, plays an important role in scientific research. The development of Altmetrics not only enriches the development of metrology but also provides a new perspective and method for the evaluation of academic influence. At present, there are lots of studies on non-traditional evaluation indicators, but the researches on the evaluation object are relatively few, especially those on open source software, which plays an important role in academic research. The number of downloads, the number of citations, and the number of reuse between applications for academic research are concrete reflections of its academic influence. At the same time, the open source community and the development of computer technology make the research tools of software more and more

\* 本文系国家社会科学基金项目“中国学者国际学术论文影响力评价研究”(编号:16BTQ055)的研究成果之一。(This article is an outcome of the project “Research on the Influence of Chinese Scholars’ International Academic Paper” (No.16BTQ055) supported by National Social Science Foundation of China.)

通信作者:魏明坤, E-mail: weimingkun24@163.com, ORCID:0000-0001-7689-6294 (Correspondence should be addressed to WEI Mingkun, Email: weimingkun24@163.com, ORCID:0000-0001-7689-6294)

abundant.

Firstly, the index database of the website Depsy is obtained in the paper, and the open source community of Python is researched and analyzed through the data provided by ImpactStory. Then the third-party library Python pandas is used for extraction of the required attributes, i.e. software library downloads, the cited academic literature and the relationship between the reuse of software for further study of the academic influence of open source software, and finally the three indicator data were analyzed. From the results of the software community analysis in the Python community, three indicators-downloads, citation counts, and software reuse-successfully screened out popular software in their communities. From the perspective of users, software reflects the ordinary users to use more academic papers, the authors quoted for many times, and software developers to reuse more times. There is a certain correlation between the three indicators, such as software's dependence that will lead to the increasing on the number of software downloads. The increasing number of citations in academic papers has led to an increase in the use of the software in the academic field. The correlation between download and reuse is relatively high, which indicates that the dependency between software has a direct impact on the download; the correlation between citation and reuse is low, which is attributed to the difference between scholars and developers, and the citation index of the software in the academic literature reflects the academic influence.

Finally, we can reach the following conclusions: 1) the evaluation of academic influence includes not only the academic achievements of the thesis, but also the research results in non-paper form. 2) We found that the development of open source software innovation has promoted the creativity in the academic field by research on the impact evaluation of open source software based on Altmetrics. 3) There is a big gap between the software in the Python community, and most of the academic influence on the software is still low. The limitation of this paper lies in the fact that the standard of "academic" software needs to be further refined. We think that the accuracy of the method of matching the keywords, such as the name and introduction of the software library, needs to be further improved. Besides, we need to strengthen the impact of academic software in-depth study in order to draw more comprehensive and scientific conclusions. 5 figs. 3 tabs. 40 refs.

#### KEY WORDS

Altmetrics. ImpactStory. Open source software. Academic impact.

互联网与社交媒体的发展对学术交流的影响越来越明显,随着 Web2.0 技术的发展,学术成果不仅包括论文、专著、研究报告等内容,还包括其他形式的研究成果,如博客、视频、数据集、软件代码等。学术影响力是科研人员的研究成果及其学术地位对学术发展起到的作用。开源软件作为非论文形式的学术成果,在科学研究中有着重要的地位,这些非论文形式的学术研究成果也应纳入学者或学术成果的评价体系中。Altmetrics 的提出<sup>[1]</sup>,不仅丰富了计量学

的发展,还为学术影响力评价提供了新的视角和方法。目前,国内外对各种非传统评价指标研究的文献日益增多,但对评价对象的研究较少,尤其缺少对学术研究起到重要作用的开源软件代码的研究。用于学术研究的开源软件的下载量、被引次数以及软件之间的复用次数是其学术影响力的具体体现。同时,开源社区和计算机技术的发展使得学术研究软件工具不断丰富,软件工具的学术价值及其影响力越来越受到广大学者的关注。

早在 1983 年,Root 和 Draper 出版了 *Questionnaires as a software evaluation tool* (问卷:作为软件评价的工具),通过问卷法,分别对不同背景和不同类型的用户进行调查,综合分析问卷对软件评价的应用,判断其利弊,更新软件工具的评价方法。1986 年,Macarthur 和 Shneiderman<sup>[2]</sup> 对面向障碍性儿童的读写软件进行评价,主要以读写的便利性和体验的友好度为指标进行评价分析。1995 年,Henderson 等<sup>[3]</sup> 在问卷方法的基础上,结合访谈、口头报告和记录数据,对电子表格、文字处理器和数据库三种不同软件工具进行评价。1999 年 Vlahavas 等<sup>[4]</sup> 建立了一个以顶层软件属性(如质量和成本等)为基础的专家软件工具评价系统。2003 年,Stamelos 和 Tsoukias<sup>[5]</sup> 总结软件评价过程中的各种情境,提出应对不同情境的解决措施,完善了软件评价体系。2006 年,国外出现了专门面向软件评估的专利文献<sup>[6]</sup>,使得软件评估方法形成规范的、专业的体系。可见,国外关于软件工具的评价研究早于国内。

2010 年,Kurilovas 等<sup>[7]</sup> 根据软件的应用范围对不同类型软件进行评价,方便用户选取适合自己研究的软件,实现软件使用的个性化服务。同年,Oppermann 和 Reiterer<sup>[8]</sup> 提出基于 ISO 9241 对软件进行综合评价,并通过实证分析验证了其评价结果与专家评价结果一致。ISO 9241 是关于办公室环境下交互式计算机系统的人类工效学国际标准,由 17 个部分组成,根据人类工效学和可用性原理,分别对各种硬件交互设备属性和软件用户界面设计问题作了详细的规定和建议,可以对一个产品设计符合该标准的程度进行评估和认证。2011 年,Azarian 和 Sidat<sup>[9]</sup> 尝试运用场景矩阵设计一种新的方法对软件质量进行评价,该方法的目的在于识别软件设计和验证阶段的缺陷和不足,从而对软件进行完善。2012 年,Dujmovi 和 Kadaster<sup>[10]</sup> 基于 LSP(偏好逻辑计分)提出一个一体化系统工程环境工具(Integrated Software Engineering Environment,ISEE),并在其定义和分类的基础上对

该工具进行定量评价。

## 1 Altmetrics 与开源软件

### 1.1 Altmetrics

Altmetrics 于 2010 年由 Priem 提出,并定义为“基于网络工具和环境对学术影响力评价的研究”,研究对象可以概括为开放存取平台与学术社交网络中科技论文的各种使用、交流活动。本文认为,Altmetrics 是在网络环境下,基于传统计量学,对其线上平台指标的完善和对评价对象的扩展,是对传统计量学的补充与完善。目前已有大量关于 Altmetrics 的学术文献,这些研究主要概括为以下三个方面。

(1) Altmetrics 的研究价值和意义。Altmetrics 自提出以来,受到很多学者的关注。Priem 作为 Altmetrics 的提出者,肯定 Altmetrics 的价值和意义,推动其相关研究的发展,并认为这是未来的发展趋势。Priem 指出在新的在线学术交流环境下,研究者应创造更多非传统形式的学术成果,在更广阔的平台去分享和讨论,不应继续被“出版”等形式限制,而应更多地关注学术成果的本身<sup>[11]</sup>。Sud 和 Thelwall 认为评价指标的制定是为了筛选出更优秀的学术成果,而网络技术的发展造就了很多新的筛选工具,对 Altmetrics 的相关研究极其必要<sup>[12]</sup>;同时他还认为 Altmetrics 并非是一个全新的学科领域,而是建立在网络计量学基础之上,且早已有不少的相关研究成果。不过也有一些学者认为 Altmetrics 的新型指标不够严谨,也有学者质疑 Twitter 指标的有效性。Colquhoun 认为在学术领域使用 Twitter 的研究者并不多,一些文章被转发往往是因为其标题或者关键词含有非常吸引眼球的词汇(Buzzwords),并不是因为它们具有很高的学术价值<sup>[13]</sup>。

(2) Altmetrics 指标与引文指标的相关性。关于 Altmetrics 这方面的研究较多,而且主要围绕新型指标和传统引文指标的相关性展开。Wee 和 Chia 研究了 18 个学科领域中被引次数

和 Altmetrics Score (Altmetric.com 网站的指数) 前 20 篇的学术论文,发现除了医学领域,其他学科中普遍存在论文的 Altmetrics Score 很高但被引次数较低的现象<sup>[14]</sup>。Bornmann 将 Altmetrics 指标分为三种:微博(Twitter)、线上文献管理平台(Mendeley 和 CiteUlike)、博客<sup>[15]</sup>,并通过荟萃分析法(Meta-analysis)对之前的多数相关性实证研究成果进行收集、总结和分析,发现线上文献管理平台,特别是 Mendeley 上的读者指标,与传统引文指标有着较高的相关性,有极高的附加价值,可作为传统引文指标的补充。事实上关于 Twitter 和线上文献管理平台的指标研究是学者们关注的重中之重,前者是因为相关性的研究结论差异、争论较多,后者则是因为该指标在学术论文中的覆盖率较高,在众多的在线社交媒体指标中研究价值较大。

关于 Twitter 提及量, Priem 选取了三种期刊(PLoS ONE、PLoS Pathogens、PLoS Biology)作为数据源,分别计算了具有社交属性的引用指标与传统引文指标的相关性(包括 Spearman 和 Pearson 两种计算方法),其中关于 Twitter 的有效数据仅为 0.0、0.0、0.1,该数据表明 Twitter 提及量与传统引文指标是不存在相关性或是低相关性的<sup>[16]</sup>。Haustein 等发现生物医学领域的学术论文中有 Twitter 提及量指标的论文特别少,其覆盖率远低于 Mendeley 文献管理平台的统计,而且其分布也因一些领域的研究群体对 Twitter 的青睐程度不同而变得不平衡<sup>[17]</sup>。然而,Shuai 等在研究中发现了 Twitter 提及量与论文早期的引用次数有较高正相关性,并且认为关于两者相关性的计算特别容易受样本选取时间的区间影响<sup>[18]</sup>。Thelwall 等虽然计算出 Twitter 提及量与传统引文指标呈负相关性,但他认为该计算结果并不具有适用性,因为很多学术论文还未积累更多的被引数,而 Twitter 提及量对时间的敏感度较大;同时,Thelwall 通过自己设计的实验发现 Twitter 指标与传统引文指标存在相关性<sup>[19]</sup>。相比 Twitter,线上文献管理平台(Mendeley 和 CiteUlike)中的读者数量指

标被认为是在学术影响力评价中更权威性的计量指标<sup>[20]</sup>,因为它们与传统引文指标存在共性且数据源获取比较合理<sup>[21]</sup>。Li<sup>[22]</sup>、Torres-salinas<sup>[23]</sup>、Mohammadi<sup>[24]</sup>等学者都对线上文献管理平台与引文指标的相关性进行了分析,得出的结论大体一致,都发现这两者之间存在显著的正相关性。

(3) Altmetrics 在学术影响力评价中的应用。这部分研究内容主要表现在计量指标和计量对象的扩展上,前者指更多种类的数据源对于评价传统文献的适用性<sup>[25-28]</sup>,后者指非传统形式学术成果的分析与评价<sup>[29-30]</sup>。在计量指标方面,Zuccala 尝试利用书评网站 Goodreaders 上读者数量指标,判断学术书籍影响力的大小,通过计算人文学科部分书籍的读者数量和文献中引用次数的 Spearman 指数,Zuccala 发现两者之间存在弱相关性,认为 Goodreader 上的读者数量应该作为评价学术书籍的补充性指标<sup>[31]</sup>;Adie 认为类似于专利、政策文件等“灰色文献”(Grey Literature)是评价学术论文的重要数据源,未来需要识别相关文件,提取元数据以实现对这类文献的追踪、分析,找出专利、政策文件中对学术论文的引用行为<sup>[32]</sup>。在计量对象方面,Kousha 等对 YouTube 上的视频文件在学术文献中的被引情况进行分析,研究了近些年学术文献对视频引用次数的增长趋势、经常被引用的视频类型、不同学科文献对视频引用存在的差异等问题<sup>[33]</sup>;Priem 与 Piwowar 对开源软件进行评价,并发起了一个名为 Depsy 的开源项目,致力于分析开源社区中各软件的影响力,评价指标包括这些软件的下载量、在文献中的引用次数以及软件之间的复用关系等<sup>[34-35]</sup>。

## 1.2 开源软件

开放源码软件(Open-source)即开源软件,主要描述该软件的源代码可以不受许可证的限制,被大众自由获取、修改以及重新发布。从开放源代码的广义角度讲,某些开放的商业软件也属于开放源代码软件范畴。开源软件伴随着

计算机的整个历史,从计算机最早出现开始,最初的软件几乎都是开源的。企业最早的销售产品是计算机的硬件,软件都是捆绑赠送的,所以软件并非企业的盈利载体,它的源代码自然也是开放的。同时,最初软件源代码的开放,还大力促进了计算机行业的快速发展。用户可以自由查看并修改程序,为代码爱好者提供了开放的平台,程序员专注于代码的共享与更新,使得软件质量得到了快速提升。

随着科学技术的不断发展,企业开始以软件为销售主体,众多软件源代码开始了非开放分享模式,该行情促进了著名的 Linux 系统的产生,这也是开源软件发展以来具有历史意义的事件。后期,Bob Yang 成立了红帽公司,专门提供 Linux 系统的开源服务。1998 年,开源软件促进会(OSI)成立,同年 4 月召开了开源峰会,提出关于开源软件的十条定义。从此,开源软件开始流行,朝着规范化的方向发展,并被越来越多使用者接受和认可,并具有以下几个特点。①成本低。开源软件由于其先天的公开性,与商业软件相比,性价比极高。②性能强大。由于其公开性,越来越多的计算机爱好者自发参与到软件的后台建设中,这使得软件在使用过程中拥有更高的性能,并得到不断的完善。③安全性高。开源软件在众多计算机爱好者的自发协助下,不断修护安全问题,减少 bug 的存在,使得软件本身与封闭的商业软件相比,安全性能越来越高。

### 1.3 Altmetrics 与开源软件的关系

Web2.0 时代的到来以及社交媒体的出现加快了学术交流的频率,提高了学术交流的质量,大量开源软件的出现也正是这个时代的产物,越来越多的学者将学术成果在网上公布,如论文、期刊、图书等传统学术载体,而软件、代码、PPT、网站等非传统形式的学术研究成果也越来越受到人们的重视。尤其在 Web2.0 时代,这些非传统形式的学术研究成果在学术研究领域占有重要的地位,对推动学术研究的发展产生巨

大的影响。

ImpactStory 旗下开发的 Depsy 为非论文形式研究成果的学术影响力评价提供了可能,在 Altmetrics 领域处于非常重要的地位。ImpactStory 是 Altmetrics 提出者 Priem 和 Piowar 联合开发的一款免费开源网络工具,旨在帮助研究者分享和评价其所有学术成果的影响力,除了评价传统的期刊文献,还能评价研究者的博客、数据集和软件代码等多种科研成果<sup>[36-37]</sup>。ImpactStory 数据源广泛、种类繁多,包括 Twitter、Mendeley、Delicious、CiteULike、F1000 reviews、SlideShare 以及 GitHub 等,都可以在 ImpactStory 工具上查询。除了数据源的多样化,ImpactStory 工具还能识别多样化的格式并进行处理。ImpactStory 可提供任何类型的学术科研成果评价报告,报告中会显示学术成果在开放获取和社交媒体网络中被提及和讨论的频次,报告中还会提供数据源的链接。ImpactStory 的开发提高了学术领域对开源软件学术影响力的重视,有助于全面评价学术研究成果。Altmetrics 的发展不仅丰富了传统学术研究成果的学术影响力评价指标,也为非传统形式的学术研究成果,如软件、代码、PPT 以及网站等的影响力评价提供了可能,补充了学术研究成果影响力评价的不足。

## 2 研究设计

### 2.1 研究方法 with 数据采集

本文首先在 Depsy 网站的可读数据库中获取相关软件的各项指标数据,然后利用 Python 第三方软件 Pandas 提取研究所需要的属性,即软件的下载量、被引用的学术文献以及各软件之间的复用关系,最后分别对这三项数据进行分析。Depsy 网站提供了软件被引用的次数,获取的数据截至 2015 年 10 月(已是最新数据)。在 Depsy 提供的全文数据库的检索式中,通过对单个软件的检索,发现检索结果(截至 2016 年 3 月)比网站的镜像数据(2015 年 10 月)要多。如

numpy,网站给出在 Europe PMC 上的数据是 554 次,而根据网站的检索式得到 643 次,大约半年时间的增幅超过了 15%。为了保证数据的时效性,对引用软件的学术文献进行深入分析,对文中提到的 7 个“高被引”软件在全文数据库 Europe PMC 中进行检索(检索式为" pypi. Python. org/pypi/ \* " OR " github. com/ numpy/ \* " OR " \* " NOT AUTH: " \* ",检索式中“ \* ”代表软件的名称),以获取最新的被引数据(截至 2016 年 3 月)。复用关系是通过 Python 第三方软件 networkx 处理关系数据,生成 gml 格式文件,然后借助 Gephi 软件进行可视化分析。

开源软件通过编程语言应用商店网页上获取,产生的相关下载数据专门由软件的应用平台统计并提供。如 Python 和 R 语言的常用应用平台分别为 PyPI 和 CRAN,这些平台聚集了大量第三方软件,通过简单的命令便可自动寻找资源镜像并下载安装。有些软件的使用依赖于其他软件即软件的复用行为,用户在下载所需软件的时候还必须下载其依赖的软件,有些下载管理工具也会分析软件的依赖关系,自动下载所需要的软件。而这些下载行为都会被应用平台所记录,软件对应的下载数量也可以在网站上获取。同时,软件自身所提供的依赖关系也可被提取,这样通过应用平台的 API 以及下载、安装并使用管理工具,软件的下载和复用指标直接被批量获取。在数据获取过程中对软件的引用指标获取较为复杂,传统学术论文中对软件工具的引用尚未形成规范,参考文献中对软件的引用相对较少,虽然有学者提出 RRID<sup>[38]</sup>(Research Resource Identifier)的标准来规范软件类学术成果的引用,但此标准尚未普及。如果仅仅根据参考文献来判断软件的文献引用显然不够客观,而且目前大部分学术文献数据库并不是开放存取的,要从论文数据库中获取软件的引用指标较为困难。本文通过在学术论文全文数据库中检索与软件名称相关的关键词,来确定软件的引用指标,并通过确定一个固定的检索公式批量获取该项

指标。

本文通过 ImpactStory 旗下开发的 Depsy 提供的对开源社区进行深入的研究和分析,选择 Python 的开源社区作为研究对象。Python 里包含的许多工具库并不局限于学术研究,所以 Depsy 对 Python 所涉及的第三方软件进行分类。具体做法是将这些软件的名字、描述以及标签等元数据与关于“学术”的关键词(例如 chem \* ,scien \* ,bio \* ,doi \* 等)进行匹配,能够匹配的判定为学术型软件,反之则归于非学术型软件。通过这种关键词匹配方法,从数量庞大的 Python 库中快速筛选出大部分学术领域比较常用的软件。最终 Depsy 从 57 243 个 Python 库中筛选出 4 517 个“学术型”工具包,这 4 517 个“学术型”工具包即为本文的主要研究对象。Depsy 本身是开源项目,其原始数据和分析处理后的数据都可以通过 Depsy 所提供的 API 获取,同时 Depsy 还将其完整的数据库公开,因此本文利用 Python 的第三方软件 JayDeBeApi 获取其可读的 Postgre 数据库。

## 2.2 评价工具

本文使用的评价工具为 ImpactStory 旗下开发的 Depsy 网站,该网站致力于评价科学研究,具有较大的影响力。其评价指标的维度具有多样性,评价对象覆盖较多的汇编语言,网站的功能较为全面。Depsy 属于 ImpactStory 范畴的项目,由 Altmetrics 的提出者、同时也是 ImpactStory 的联合创始人 Priem 与 Piwowar<sup>[16]</sup>等发起,于 2015 年 11 月上线,该网站侧重于评价学术开源软件及 Web 应用程序项目。

Depsy 从三个方面评价开源软件的学术影响力。①工具库月平均或最近一个月的下载量。这类数据由 CRAN 和 PyPI 软件提供,CRAN 和 PyPI 分别是 R 语言和 Python 的线上资料库,这些资料库记录用户的下载行为并公开工具包的下载次数。②工具包在学术论文中的引用。此类数据在实际获取过程中较为复杂,如,大多

数软件并没有学术论文的 DOI, 论文中对软件工具的引用并不规范等。Depsy 的解决方案是使用 Europe PMC 和 ADS 全文数据库, 在这两个全文数据库中检索相关工具包的信息, 将一篇文章中对软件的一次或多次提及都算作为对软件的引用。<sup>③</sup>工具包被其他工具包的复用。在开源项目中, 一个项目直接引用其他项目是较为常见的现象, 各软件的复用就形成了一个总体上的依赖网络。通过这种依赖网络可以掌握每个软件被其他软件直接使用的情况, 还可以使用谷歌的 PageRank 算法, 赋予优质项目引用软件的相应权重, 计算出该软件的实际影响力。

### 3 开源软件学术影响力评价体系

#### 3.1 开源软件学术影响力评价可行性分析

##### (1) 评价指标的选择

传统的学术影响力评价体系是基于引文指标建立的, 已在学术界得到认可。随着 Web2.0 的出现以及 Usage Metrics 相关研究的发展, 计量学领域的研究者尝试获取其他类型指标来弥补引文指标的不足。引文量是一个规范性很强的指标, 学者对文献的引用是建立在下载或阅读文献基础之上的, 并不是每个学者都将阅读过的文献加以引用。对开源软件而言, 同样也是学术成果, 与学术论文一样被下载、阅读甚至引用。因此, 下载和引用都可作为开源软件的评价指标, 而引用可分为两种类型: 一种是软件被学术论文引用, 另一种是软件之间的引用。软件的下载量一定程度上反映出该软件的知名度以及受欢迎程度, 可帮助用户识别开源社区中使用较多、评价较好的开源软件库; 软件在学术论文中的被引数量体现相关研究领域软件的使用情况, 然而文献中软件引用行为并不规范。一些学者还没有引用软件的习惯, 同时软件引用标准的缺失也给学者引用行为造成一定的障碍, 但这并不影响这些软件在论文中被提及; 软件对软件的引用能够反映出被引用软件在社区中的影响力。在开源软件中这种形式的引用实

际上是一种使用, 一种开源软件在编写的过程中如果需要引用前人写好的软件, 会直接导入并使用这个软件, 与软件在学术论文中的引用相比, 软件对软件的引用行为更加客观、明显。

本文参照传统学术影响力评价体系, 构建衡量开源软件学术影响力的三项指标, 即软件的下载量、软件在文献中的引用、软件之间的引用, 即下载、引用、复用。这三个指标分别从不同方面反映软件的学术影响力, 从整体角度看, 三个指标之间具有相关性, 它们有机地构成了开源软件学术影响力的评价指标体系。

##### (2) 评价方法分析

文献计量学的主要研究对象是文献, 但随着计量学的深入发展, 出现了信息计量学和网络计量学。随着 Web2.0 的发展, 研究对象拓展到 Web 页面等形式多样的网络内容, 同时研究对象也深化到本体、元数据、语义等粒度更细的信息单元。然而, 即便是计量学从纸质文献的主战场转移到互联网, 计量学领域常用的方法依然是适用的, 例如数理统计分析、社会网络分析等。软件工具作为特殊形式的学术成果, 它与学术论文同样具有发布时间、作者、所属学科领域等要素, 同样也存在着引用的关系, 采用计量学的方法来评价其影响力仍然是可行的。

本文评价开源软件的学术影响力, 其载体大部分是计算机汇编语言, 而学术论文的载体是期刊。目前, 学术论文的产出数量远远超过用于学术研究的开源软件的数量。期刊的学科性质比较明显, 不同期刊上的论文可按研究主题划分为相应学术领域文献的集合, 而编程语言大多都不具有学科性质, 虽然开发人员编写的软件可能是针对某一领域, 但没有明显的学科划分。因此, 与学术论文在学科领域的影响力评价不同, 开源软件需要选择相应的汇编语言环境以及适用于学科领域的软件, 在其所属汇编语言的开源社区范围内进行评价

#### 3.2 开源软件学术影响力指标评价体系构建

通过对开源软件学术影响力评价的可行性

分析,得到评价开源软件学术影响力三个层面的一级评价指标,分别为软件的下载、引用以及复用。通过分析一级指标的制定和获取途径及

方法,制定相应的二级指标,最终确定软件学术影响力评价指标体系,如图 1 所示。

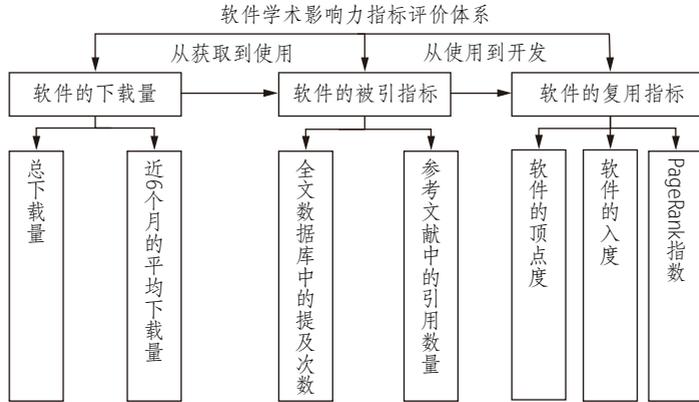


图 1 软件学术影响力指标评价体系

下载是网上获取资源的主要方式,它以利用资源、传播资源等行为为前提,象征着软件的热度、知名度。而下载指标容易受到时间因素的影响,由于该行为的触发不需要很深的决策,这个指标的积累较快。同时,每个软件的诞生时间具有明显差异,这也是影响下载指标的重要因素。鉴于此,本文选取了总下载量以及近6个月的平均下载量作为软件的具体数据指标,前者表示软件总的获取次数,后者则是动态的、实时更新的统计指标。

软件的引用指标是最能体现软件学术影响力的指标,其获取难度较大。目前,学术论文中对软件工具的引用还不够重视和不够规范,使得我们很难获取比较客观的数据。理论上讲,如果学术论文对软件工具引用比较规范,我们只需要获取其引用的参考文献数据即可,这些数据通过第三方软件即可获取。然而在这些参考文献中很难找到对软件工具的引用信息,即便是论文的著者在其研究中使用并且在文中提到了该工具。为了尽可能的客观,通过在论文中检索软件名的相关关键词来判断论文是否对软件有引用行为。由于大部分商业数据库无法提供全文数据,我们只有选取开放存取的全文数据库作为数据源。

软件的复用指标可以看作是软件之间的引用指标,它体现软件在开源社区中的影响力。鉴于属于相同类别之间的引用,这类指标具体表现为各软件的依赖关系。这种依赖关系与学术论文之间的引用关系一致,一个软件可以引用其他软件,也能够被其他软件引用。本文根据社会网络分析原理,对软件之间的依赖关系构造有向的社会网络图,软件的顶点度表示与之有关系的软件数量(包括引用与被引用),而软件的入度和出度分别表示该软件被引用和引用的软件数量。该指标体系中主要选择了软件的顶点度和入度,特别是入度,它与计量学中传统引文指标相同,对衡量开源软件的影响力尤为重要。而出度代表该软件的原创性,如果一个软件的入度较高、出度较低,说明该软件的质量较高,在该指标中采用入度和出度的比值代表软件的质量。另外,由于度指标只能反映出特定的关系,并不能表示各软件之间还存在其他间接的引用关系。本文为了进一步完善该指标,以谷歌的 PageRank 算法弥补入度指标的不足。

上述三个一级指标不仅从软件广度反映其影响力,还从软件的深度反映彼此间的差异。从软件的下载到引用再到复用的过程,可以看

作是从获取到使用再到开发逐渐深入的过程。可见,本文提出的指标评价体系是从广度和深度全面反映软件的学术影响力。

#### 4 开源软件学术影响力评价结果

Depsy 提供的原始数据包括 Python 软件库中软件的下载量、学术文献引用次数以及软件之间的复用关系,据此,本文主要从这三个方面来评价“学术型”软件的影响力。

#### 4.1 Python 社区内软件的下载量分析

软件的下载量代表软件的使用频度,软件的下载量反映以下三个方面内容:软件应用的研究领域活跃度,软件的知名度,软件本身质量。学术领域主要借助 Python 进行数据分析、文本挖掘、深度学习等研究。如 numpy、scipy、nltk 等分析工具的下载量极大,位于前列。在 4 517 个软件中,最近一个月(2016 年 3 月)下载量排名前 15 的软件如表 1 所示。

表 1 Python“学术型”软件最近一月下载量 TOP15

软件	下载量	描述
iPython	557 570	功能丰富的交互式 shell
numpy	536 880	用来存储和处理大型矩阵的科学计算包
pandas	265 381	基于 numpy 的高级数据结构和工具的数据分析包
djangorestframework	221 140	基于 django 的 rest 框架
iPython_genutils	207 745	iPython 的补充工具包
scikit-learn	188 694	强大的机器学习资源包
scipy	171 521	易于使用、为科学和工程设计的工具包
networkx	171 294	用于复杂网络创建、社会网络分析的工具包
matplotlib	163 769	用于绘图、功能强大的可视化软件包
posix_ipc	155 778	解决进程间通信的工具包
ndg-httpsclient	142 438	为 httplib 和 urllib2 使用 OpenSSL 提供支持的工具包
PyVISA	118 393	用于端口控制的软件包
Nltk	116 777	自然语言处理及文本挖掘的工具包
h5py	86 205	将数组数据转换为 HDF5 格式的工具包
Ipdb	79 347	基于 iPython 的 Python 调试工具

注:①该数据是 2016 年 3 月份的镜像数据,截至 2015 年 10 月 Depsy 并没有实时更新这部分数据,网站中的其他数据同样存在这个问题,而且随机一个月的数据不具有很强的代表性,获取该类数据过去几个月的平均值会较好;②排在前 15 名的工具包中,有些软件严格意义上不能算作“学术型”软件包,例如 iPython 是用户使用 Python 进行下载的交互式工具包。另外像 djangorestframework 这类负责网页开发的软件,并非直接用于学术研究。

通过分析,很多学术领域常用的软件下载量排名没有达到前 1%,如用于科学计算的 numpy、scipy 以及 pandas,用于机器学习和社会网络分析的 scikit-learn 和 networkx,用于数据可视化的 matplotlib 等。这些软件都是学术领域中

处理数据、分析数据的常用软件。为了进一步了解这些软件的下载情况,绘制了下载量累积函数曲线,如图 2 所示。通过对数据的分析,4 517 个软件中一半的下载量小于 63 次,90% 的软件下载量小于 2 293 次,而 1% 的软件下载量

大于 18 995 次。可见,虽然 Python 的“学术型”软件数量很多,但使用次数较多的软件并不多。造成这种现象的原因有三个方面:软件质量参差不齐,Python 社区强大的软件生态环境致使 Python 软件库增长迅速,但是优质软件与庞大的 Python 软件资源库相比显得较为单薄;许多软件的开发和使用以其他软件为基础,用户在下载软件时往往还需要下载其依赖的软件,这会提高被依赖软件下载的数量;产生较早的优质软件优势累积,最终产生马太效应,优质软件在社区中有较高的评价,其影响力不断提升,这样会有更多的用户去下载和使用。

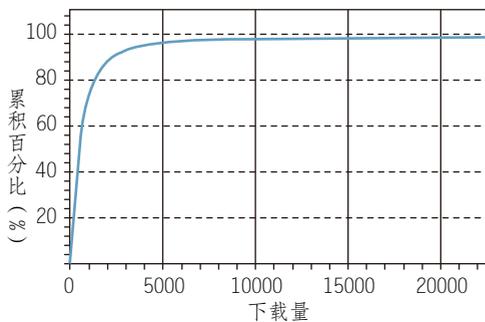


图 2 Python“学术型”软件 2016 年 3 月下载量累积函数曲线

#### 4.2 Python 社区内软件在学术论文中被引次数分析

与软件的下载量不同,软件在学术文献中的被引次数直接与学术有关,表现出软件的学术影响力。但由于学术论文中引用软件不规范现象普遍存在,以及商业数据库的非开放存取,本文在 Depsy 上收集的被引数据远远少于客观引用的次数。理论上讲,软件的被引次数应该与下载量存在一定的相关性,鉴于此绘制了关于下载量和被引数的气泡图,如图 3 所示。该图由 Python 的第三方可视化软件 bokeh 绘制,图的纵横坐标分别代表下载量和被引数,气泡的大小与被引次数有关,气泡越大被引用的次数越多。图中标注了被引次数大于 200 的软件名称,它们从高到低依次是 scipy、numpy、matplotlib、bioPython、networkx、openbabel 和 scikit-learn 7 个

软件包,除 bioPython 和 openbabel 外,其他 5 个软件在前文下载量分析中都有提到。bioPython 和 openbabel 分别是开发计算分子生物学的软件和转换各种化学数据格式的软件接口,这两个软件分别属于生物学和化学领域的专业软件。上述 7 个软件属于“学术型”软件,可见通过文献被引数的筛选,对“学术型”软件的甄别更加准确。与其他 5 个软件相比,bioPython 和 openbabel 的被引数量较少(分别是 406 次和 210 次)。主要原因是两者的学科属性较为鲜明,属于学科领域内使用的软件,而其他 5 个软件都属于万能型的数据分析类软件,在各学科中都有应用。其中,scipy 以 1 164 次被引数量位居首位,比排名第二的 numpy(788 次)高出很多。而 scipy 之所以有如此明显的优势归因于其丰富专业的数学公式及函数模块,包括统计、优化、整合、线性代数、傅里叶变换、信号和图像处理、常微分方程求解器等。

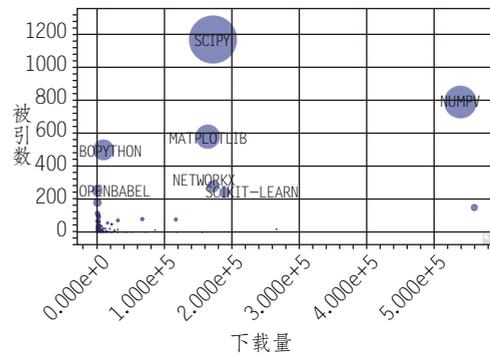


图 3 Python“学术型”软件下载量与被引用次数

软件在文献中的提及或引用不仅反映学者对“学术型”软件的重视程度,还反映出软件在学术领域的影响力不断增强。为了对文中 7 个“高被引”软件的被提及数量有更清晰的认识,本文通过发展曲线图展示其近几年的被提及情况,如图 4 所示。图 4 为 scipy 等 7 个软件从 2008 年起在学术文献中被提及次数的变化情况,在图中可以明显观察到软件被提及次数在增长,这表明近年来学者对软件的重视程度越来越高,这种重视程度表现在对软件的应用、对

软件所在领域的研究、对软件的引用等。如 bioPython 和 openbabel 分别应用于生物和化学这两个传统学术领域,它们在文献中的被提及数量持续增长,表明该领域学者对软件处理数据、分析数据的重视;又如 scikit-learn, 诞生于 2010 年,是服务于机器学习领域的专业软件,在 2012 年到 2015 年迅速增长,一定程度上反映了机器学习领域近些年的研究热度。然而,学者

们对软件在文献中被引用的重视程度无法在数据中得到明显的反馈,仅凭数据也很难看出学者们对软件引用规范的重视。实际上,随着科技的不断发展,无论是对软件应用的重视还是对软件所在领域的关注,都会使软件在学术界的地位越来越高,“学术型”软件将在大量学术文献中被提及。

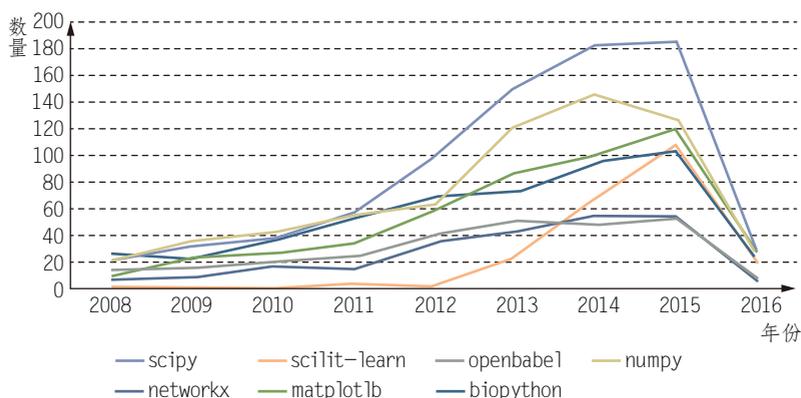


图 4 2008 年至 2016 年 (3 月) 7 个软件在 Europe PMC 的被提及数量曲线

### 4.3 Python 社区中软件之间的复用分析

软件除了被学术论文引用之外,还有可能被其他软件所复用。软件之间复用的现象在开源社区中尤为明显,对于应用广泛的编程语言 Python 也不例外。通过有向社会网络可反映 Python 社区中影响力较大的群体。Depsy 对 Python 社区中各软件的复用关系进行相关分析,通过利用 Python 的软件管理工具 wheel 下载软件的源代码,分析源代码中 setup.py 或 requirements.txt 与导入软件包相关的文件,找出软件所需要的其他软件包,这些软件包即为该软件复用或依赖的软件。为了使研究对象的范围聚集于 4 517 个“学术型”软件中,本研究对 Depsy 所提供的数据进行提炼,并利用 Python 的第三方软件 networkx 处理并生成可用于社会网络分析的 gml 格式文件,最终将其导入到可视化软件 Gephi 中,进行分析并绘制可视化图谱。

通过分析发现,在 4 517 个节点中只有

1 985 个节点的度不为 0,也就是说剩余 56% 的软件与其他软件是没有任何联系的,这并不代表这些软件与其他软件不存在复用关系,而是因为这些依赖或复用的软件没有归为“学术型”软件,所以对于软件性质的判定对结果有较大的影响。为了更加清晰地观察软件之间的关系,要对节点做进一步的筛选,隐藏度小于 5 的节点及其相关联的边,最后留下了 293 个节点和 993 条有向边,同时显示节点的标签,调整节点和标签的大小与节点的入度 (Indegree) 成正相关,如图 5 所示。图中 numpy、scipy、matplotlib 以及 pandas 很明显占主导地位,事实上,这 4 个处于中心位置的软件也是各行业做数据分析必备的软件,它们经常被各种数据分析同时使用并相互协助。numpy 负责高效处理和存储数组格式的数据,scipy 提供丰富的数理计算公式及函数,pandas 中的数据结构 DataFrame 对数组数据的抽象极大简化了数据分析的繁琐操作,



表2 Python“学术型”软件顶点度(TOP10)

软件名称	入度(Indegree)	出度(Outdegree)	度(Degree)
numpy	849	0	849
scipy	381	0	381
matplotlib	271	0	271
pandas	216	0	216
iPython	107	0	107
networkx	72	0	72
scikit-learn	67	0	67
bioPython	61	0	61
bob	47	8	55
h5py	43	0	43

## 5 结果分析

本研究对 Python 社区的部分软件从下载量、文献被引次数以及软件之间的复用次数三个方面进行评价,提取出影响力较大的软件。在研究过程使用的大部分数据来源于 Depsy 平台,在分析过程中发现在 Depsy 平台获取数据、处理数据存在一些缺陷,通过不断改进数据收集、处理方案,完善评价指标及评价体系,对软件学术影响力评价进行深入研究。

从 Python 社区中软件分析的结果而言,下载量、文献被引次数以及软件之间的复用次数三个指标能够成功筛选出其社区内的热门软件,这些热门软件从使用者的角度体现出普通用户使用较多、学术论文著者引用次数较多以及软件开发者复用次数较多的特点。这三者本身也存在一定的相关性,如某些软件对其他软件的依赖会导致被依赖软件下载次数增多,在学术论文中引用次数的增多会导致学术领域内的研究者对该软件使用量的增加。本文利用 Python 的科学计算包 scipy<sup>[39]</sup>,以 4 517 个“学术型”软件的下载量、文献引用次数以及被其他软件或项目复用次数为数据样本,计算三者下载、引用、复用之间的皮尔逊相关系数<sup>[40]</sup>,如表 3 所

示。表中的数据( $p$  值小于 0.01)证实了三者间存在相关性,其中下载与复用的相关性较高,说明软件之间的依赖关系对下载量有直接的影响;引用和复用的相关性较低,这归因于学者和开发者所青睐的软件存在差异,而且软件在学术文献中的引用指标更多是反映学术方面的影响力。

表3 皮尔逊相关系数矩阵

	下载	引用	复用
下载	1	0.535	0.594
引用		1	0.419
复用			1

## 6 结论与不足

Altmetrics 的发展为学术成果的影响力评价带来了新的活力,学术领域的评价标准已不再局限于引文指标,开始扩展到社交媒体、文献存储、软件代码以及专利政策等数据源指标,非传统的学术成果也将成为评价的主要对象。特别是对于软件代码工具等非传统形式的学术成果,它们经常在学术论文中被使用,但很多情况下并没有被著者规范地引用,而这种引用行为在科学评价中不能被忽略。本文主要从

Altmetrics 的角度,分析开源软件的学术影响力,并构建学术软件影响力指标评价体系,最后具体评价 Python 社区的部分软件的影响力。在评价中容易混淆的是,有些数据源既可以作为计量指标,也可作为评价对象。如维基百科,它既有对学术论文的引用,也有被学术论文引用,在学术影响力评价的过程中需要分清评价对象与评价标准。

学术论文是学术领域中最有价值也是影响力较大的学术成果,但不能代表研究人员的全部研究成果。如对一位研究人员的学术影响力进行评价,其论文作为评价的一部分,还应包括其他形式的学术成果,其中编写的软件代码占比越来越大,在学术影响力评价体系中占有重要地位。现代计算机技术的提高不仅为学术领域研究提供丰富的技术和方法,还孵化出许多使用方便、可靠性高的数据源,这为本文的数据获取提供了极大的便利。Python 社区中各软件

的影响力差距悬殊,大部分软件的学术影响力较低。无论是软件的下载,引用及复用,它们相对应的指数分布不均衡。可见,开源社区中常用的软件在全部软件中占比有限,而且优质软件编写的投入也远远大于一般软件。此外,很多普通软件在编写时都会对优质软件进行复用,这导致了优质软件的影响力不断累积,其优势在不断扩大。在数量众多的开源软件中,由于开发人员的投入有限,缺少合理的激励机制,出现软件质量参差不齐的现象,导致各软件的影响力差距比较悬殊。本文在分析过程中的局限性在于,“学术型”软件的判定标准有待进一步精确化,笔者认为单凭对软件的名字、简介与关键词进行匹配的方法进行分类,其准确度有待进一步提高。今后需要加强对开源软件学术影响力的深入研究,以便得出更为全面、科学的研究结论。

## 参考文献

- [ 1 ] Priem J, Hemminger B H. Scientometrics 2.0: new metrics of scholarly impact on the social Web[J]. First Monday, 2010, 15(7) :5.
- [ 2 ] Macarthur C A, Shneiderman B. Learning disabled students' difficulties in learning to use a word processor: implications for instruction and software evaluation[J]. Journal of Learning Disabilities, 1986, 19(4) :248-253.
- [ 3 ] Henderson R, Podd J, Smith M, et al. An examination of four user-based software evaluation methods[J]. Interacting with Computers, 1995, 7(4) :412-432.
- [ 4 ] Vlahavas I, Stamelos I, Refanidis I, et al. ESSE: an expert system for software evaluation[J]. Knowledge-Based Systems, 1999, 12(4) :183-197.
- [ 5 ] Stamelos I, Tsoukiás A. Software evaluation problem situations[J]. European Journal of Operational Research, 2003, 145(2) :273-286.
- [ 6 ] Iizuka T. Software evaluation method and software evaluation system: US, US 20060253824 A1[P]. 2006.
- [ 7 ] Kurilovas E, Serikoviene S. Learning content and software evaluation and personalisation problems[J]. Informatics in Education, 2010, 9(1) :24.
- [ 8 ] Oppermann R, Reiterer H. Software evaluation using the 9241 evaluator[J]. Behaviour & Information Technology, 1997, 16(4-5) :232-245.
- [ 9 ] Azarian A, Siadat A. A synthesis of software evaluation methodologies and the proposal of a new practical approach [J]. Journal of Software, 2011, 6(11) :2271-2281.

- [10] Dujmović J, Kadaster M. A technique and tool for software evaluation[J]. *Evolution*, 2002, 374: 246.
- [11] Priem J. Scholarship: beyond the paper[J]. *Nature*, 2013, 495(7442): 437-440.
- [12] Sud P, Thelwall M. Evaluating altmetrics[J]. *Scientometrics*, 2014, 98(2): 1131-1143.
- [13] Colquhoun D P A. Why you should ignore altmetrics and other bibliometric nightmares[EB/OL]. [2016-10-21]. <http://www.dscience.net/?p=6369>.
- [14] Yew B C, Foon J, Wee J. Altmetrics is an indication of quality research or just hot topics[C]//Proceedings of International Association of Scientific and Technological University Libraries. Finland, Alto University, 2014: 89-102.
- [15] Bornmann L. Alternative metrics in scientometrics: a meta-analysis of research into three altmetrics[J]. *Scientometrics*, 2015, 103(3): 1123-1144.
- [16] Priem J, Piwowar H A, Hemminger B M. Altmetrics in the wild: using social media to explore scholarly impact[J]. arXiv preprint arXiv:1203.4745, 2012.
- [17] Haustein S, Peters I, Sugimoto C R, et al. Tweeting biomedicine: an analysis of tweets and citations in the biomedical literature[J]. *Journal of the Association for Information Science and Technology*, 2014, 65(4): 656, 669.
- [18] Shuai X, Pepe A, Bollen J. How the scientific community reacts to newly submitted preprints: article downloads, Twitter mentions, and citations[J]. *Plos One*, 2012, 7(11): e47523.
- [19] Thelwall M, Haustein S, Larivière V, et al. Do altmetrics work? Twitter and ten other social Web services[J]. *Plos One*, 2013, 8(5): e64841.
- [20] Thomas K, Grier C, Paxson V, et al. Suspended accounts in retrospect: an analysis of Twitter spam[C]//ACM SIGCOMM Conference on Internet Measurement (IMC), 2011: 243-258.
- [21] Sud P, Thelwall M. Not all international collaboration is beneficial: the Mendeley readership and citation impact of biochemical research collaboration[J]. *Journal of the Association for Information Science and Technology*, 2016, 385(8): S99.
- [22] Li X, Thelwall M. F1000, Mendeley and traditional bibliometric indicators[C]//Proceedings of the 17th International Conference on Science and Technology Indicators. 2012: 451-551.
- [23] Torressalinas D, Cabezasclavijo A, Jimenezcontreras E. Altmetrics: new indicators for scientific communication in Web 2.0[J]. *Comunicar*, 2013, 41(41): 53-60.
- [24] Mohammadi E, Thelwall M. Mendeley readership altmetrics for the social sciences and humanities: research evaluation and knowledge flows[J]. *Journal of the Association for Information Science and Technology*, 2014, 65(8): 1627-1638.
- [25] 赵蓉英, 郭凤娇, 谭洁. 基于 Altmetrics 的学术论文影响力评价研究——以汉语言文学学科为例[J]. *中国图书馆学报*, 2016, 42(1): 96-108. (Zhao Rongying, Guo Fengjiao, Tan Jie. Evaluation of academic papers impact based on Altmetrics: a case study of Chinese Language and Literature[J]. *Journal of Library Science in China*. 2016, 42(1): 96-108.)
- [26] 王睿, 胡文静, 郭玮. 高 Altmetrics 指标科技论文学术影响力研究[J]. *图书情报工作*, 2014, 58(21): 92-98. (Wang Rui, Hu Wenjing, Guo Wei. Research on academic influence of high Altmetrics Sci-tech papers[J]. *Li-*

- brary and Information Service, 2014, 58(21):92-98.)
- [27] 郭飞, 游滨, 薛婧媛. Altmetrics 热点论文传播特性及影响力分析[J]. 图书情报工作, 2016(15):86-93.  
(Guo Fei, You Bin, Xue Jingyuan. Analysis on transmission characteristics and influence of Altmetrics hot papers  
[J]. Library and Information Service, 2016(15):86-93.)
- [28] 刘晓娟, 宋婉姿. 基于 PLOS ALM 的 altmetrics 指标可用性分析[J]. 图书情报工作, 2016, 60(4):93-101.  
(Liu Xiaojuan, Song Wanzi. Analysis on the usability of Altmetrics indicators based on PLOS ALM[J]. Library  
and Information Service, 2016, 60(4):93-101.)
- [29] 叶文芳, 于洪洋, 王春霞. 基于社交媒体的学术影响力计量方法研究——以 altmetric.com 为例[J]. 科技与  
出版, 2015(9):81-85.(Ye Wenfang, Yu Hongyang, Wang Chunxia. Research on the measurement of academic  
influence based on social media: a case study of Altmetric.com [J]. Science-Technology and Publication, 2015  
(9):81-85.)
- [30] 邱均平, 张心源, 董克, 等. Altmetrics 指标在机构知识库中的应用研究[J]. 图书情报工作, 2015(2):100-  
105.(Qiu Junping, Zhang Xinyuan, Dong ke. A research of application of Altmetrics in institutional repository[J].  
Library and Information Service, 2015(2):100-105.)
- [31] Zuccala A A, Verleysen F T, Cornacchia R, et al. Altmetrics for the humanities comparing Goodreads reader ratings  
with citations to history books[J]. Aslib Journal of Information Management, 2015, 67(3):320-336.
- [32] Adie E. The grey literature from an Altmetrics perspective-opportunity and challenges[J]. Research Trends, 2014,  
37:23-25
- [33] Kousha K, Thelwall M, Abdoli M. The role of online videos in research communication: a content analysis of You-  
Tube videos cited in academic publications[J]. Journal of the Association for Information Science and Technology,  
2012, 63(9):1710-1727.
- [34] Depsy [EB/OL]. [2016-10-09]. <http://depsy.org>.
- [35] Depsy: valuing the software that powers science [EB/OL]. [2016-10-09]. <http://blog.impactstory.org>.
- [36] Shema H, Bar-Ilan J, Thelwall M. Research blogs and the discussion of scholarly information[J]. Plos One, 2012,  
7(5):e35869.
- [37] Impactstory. Let's value the software that powers science: Introducing Depsy[EB/OL]. [2016-10-19]. [http://  
blog.impactstory.org](http://blog.impactstory.org).
- [38] Singh Chawla D. Researchers argue for standard format to cite lab resources[EB/OL]. [2015-05-29]. [http://  
doi.org/10.1038/nature.2015.17652](http://doi.org/10.1038/nature.2015.17652).
- [39] Bressert E. SciPy and NumPy: an overview for developers[M]. O'Reilly Media. 2012.
- [40] Abe H, Tsumoto S. Analyzing behavior of objective rule evaluation indices based on pearson product-moment cor-  
relation coefficient[M]. Springer Berlin Heidelberg, 2008:84-89.

赵蓉英 武汉大学信息管理学院教授, 博士生导师。湖北 武汉 430072。

魏明坤 武汉大学信息管理学院博士研究生。湖北 武汉 430072。

汪少震 武汉大学信息管理学院硕士研究生。湖北 武汉 430072。

(收稿日期:2016-12-01; 修回日期:2017-01-17)