

## 软件级别计量：概念辨析与应用途径

陆彩女 顾立平

**摘要** 在开放科学与开放源代码运动下,大多数科研人员都认识到了软件及软件源代码在科学研究中的作用。软件不只是研究工具,用于支持研究结果与研究数据的重现与验证,还是一种创造性研究产出,其学术价值得到了部分研究资助机构的承认。用量化的方式评价软件的学术影响力并将软件纳入学术评价体系,愈发重要和紧迫。软件级别计量就是通过量化指标来评价软件的学术影响力,旨在识别软件和软件源代码以及软件开发人员的学术贡献。软件级别计量指标包括但不限于软件引用次数、软件提及次数、软件下载次数、Altmetrics等,其中软件引用次数、软件提及次数是软件级别计量的两个重要方面。笔者剖析了当前软件级别计量测度指标存在的问题,包括数据获取难、无统一标准、易被操纵等;简要提出了未来软件级别计量的可能应用途径,如辅助研究软件发现、完善研究评估与贡献识别、促进软件长期保存等。图书馆可充分发挥知识服务中心的角色,在做好学术软件长期保存的同时,提供新型知识服务。图2。表5。参考文献83。

**关键词** 软件级别计量 软件引用 软件提及 Altmetrics 开放科学 开源软件 学术评价

**分类号** G250.7

## Software Level Metrics: Concepts and Future Applications

LU Cainü & GU Liping

### ABSTRACT

In the open science and open source environment, the importance of research data and pieces of software and source code which are used to create research data in scientific research has been increasingly recognized and valued by many scholars and organizations. However, the visibility problem of software and source code in academic publications is severe. For example, software haven't been cited or referenced or styles of software references or citations aren't standardized, resulting in that the roles of software developers, optimizers, and maintainers haven't been attached great importance in scientific research. Therefore, it is important and urgent to evaluate the academic impact of software and source code in a quantitative way and incorporate them into the academic assessment system. This study proposes the concept of software level metrics and gives the definition of software level metrics to evaluate the academic influence of software and source code quantitatively to identify the contribution of software developers and maintainers in scientific research.

Based on known theories and practices, this study combs several measurement methods and indicators and their inherent problems for software level metrics and divides the indicators into 3 types: usage statistics indicators, bibliometric indicators and supplementary measurement indicators. The usage statistics indicators

通信作者:顾立平,Email:gulp@mail.las.ac.cn,ORCID:0000-0002-2284-3856(Correspondence should be addressed to GU Liping,Email:gulp@mail.las.ac.cn,ORCID:0000-0002-2284-3856)

include the number of downloads, the number of registered users, the number of reuses and degree of dependence. The bibliometric indicators include the number of citations and the number of mentions in academic publications. Supplementary measurement indicators include the number of comments and the number of mentions in social network systems, etc. Among them, software citation and software mention in academic publications are two more important indicators in software level metrics. This study briefly summarizes four refined indicators of software citation and software mention, which are: 1) software mention #1-name mention; 2) software mention #2-URL mention; 3) software citation #1-citing papers; 4) software citation #2-citing URL or DOI or RRID. Then, this paper presents a practical example of software level metrics on three types of software: the proprietary software, open source software and non-commercial software. It is confirmed through the example that software citation #1-citing papers is more common in academic publications than the other 3 indicators of software citation and software mention. Finally, this study simply outlines the possible application scenarios of software level metrics, including: 1) Assisting researchers in academic software search and discovery; 2) Assisting funding agencies or reviewers in project funding assessment and project conclusion review in a more comprehensive manner; 3) Promoting code development and optimization and scholarly awarding all of the persons and websites in the lifecycle of software; 4) For libraries and librarians to preserve software and develop new types of knowledge services, such as innovating academic assessment services; 5) Effectively carrying out personal study for citizens. 2 figs. 5 tabs. 83 refs.

## KEY WORDS

Software level metrics. Software citation. Software mentions. Altmetrics. Open science.  
Open source software. Academic evaluation.

当前,开放科学运动如火如荼,将对科学研究行为带来重大变化和深远影响<sup>[1]</sup>。开放科学不仅涉及开放获取论文,一般来说还涉及其他科学研究成果,包括数据、模型、算法、软件、工具、样本等<sup>[2]</sup>。同时,软件在科学研究中的作用也日益凸显。据调查,90%的研究人员在研究过程中使用过软件,89%的研究人员表示如果没有软件会增加其研究难度或根本不可能完成研究<sup>[3]</sup>。软件是一种研究工具,其开放共享与可发现可以更快地重现研究结果与研究数据,它还是一种可计量的研究成果,但它的价值却经常被忽视<sup>[4]</sup>。因此,寻找可量化的方式将软件的贡献纳入学术评价体系中,不仅可以帮助科研资助机构和科研人员更好地认识并理解软件的价值,识别和发现软件开发者的贡献,而且可以在学术评价和学术交流中直接体现出这种价值,扩大学术影响力,增强学术成果类型多样化,促进学术生态体系均衡发展。

近年来,众多学者开始关注软件计量评价相关研究,同时从软件引用、软件存储与可持续性等领域出发,共同推动软件级别计量的发展。为了响应“破四唯”和学术评价改革要求,笔者长期对一线科研人员进行跟踪调查,得知此需求,故撰文介绍软件级别计量。

## 1 开源软件的作用日益凸显

### 1.1 开源社群成为推动软件发展的核心力量

开源(Open Source)即开放源代码,是一种软件源代码开放共享的开发模式<sup>[5]</sup>。在这种模式下,任何人可以不受限制地使用软件的全部功能,也可以根据自己的需求修改源代码,甚至将其编制成衍生品再次发布出去<sup>[6]</sup>。开放源代码的软件就是开源软件(Open Source Software, OSS),开源软件的版权持有人在软件协议的规定下,只保留一部分权利<sup>[7]</sup>。典型的开源软件包括 Linux、Apache、

MySQL、Firefox、Google Chrome、Java、Python、Android等,这些开源软件已经被广泛应用在人们生产和生活的各个场景中。

开源社群(或称为开源社区,本文统称为开源社群,旨在强调概念中所包含的人与人之间的社交关系与群体属性这一特征)又被称为开放源代码社群,一般由拥有共同兴趣爱好的人所组成,他们以开放软件源代码为核心,依托发布开源软件的网络平台,开展交流学习、合作开发与维护开源软件<sup>[8,9]</sup>。开源社群实际上就是依托互联网平台,汇集、累积大规模群体智慧,以实现技术创新。开源社群聚集了一批开源产品与开源技术的拥护者,目前在全球范围内比较知名的开源社群有 Apache、Debian、Linux 及 Mozilla<sup>[8]</sup>。

随着深度信息技术(机器学习、人工智能、云计算、物联网、智慧城市、区块链、大数据、神经网络、量子计算等)的创新发展,传统商业软件公司(或IT公司)逐渐走向计算资源公司,加大在数据计算和云计算基础设施方面的投资与营收,而把代码开发和优化交给开源社群。商业软件公司通过开源软件和开源社群快速抢占市场,形成下游应用粘性,从而掌握软件产业生态的主导权,争取更大的商业利益<sup>[10]</sup>。开源是促进信息技术创新的重要途径<sup>[5]</sup>,开源社群已成为推动信息技术行业(尤其是软件与数字化应用行业)发展的核心力量。开源社群作为开放创新的代表性甚至极致性呈现形式,已经引领开放创新和数字创新实践前沿近三十年,它与其所代表的开源经济受到了多国政府、企业、高校的高度重视<sup>[8]</sup>。

## 1.2 “收编”开源软件社群是开放科学运动的方向和趋势

与开放获取、开放数据一样,开放源代码也被纳入开放科学的范畴内,并逐渐成为了开放科学的一个支柱(见图1)<sup>[11]</sup>。全部欧洲科学院(All European Academies, ALLEA)在2012年发布的《21世纪的开放科学——ALL欧洲科学院的声明》中提及,开放科学内容包括了出版物、

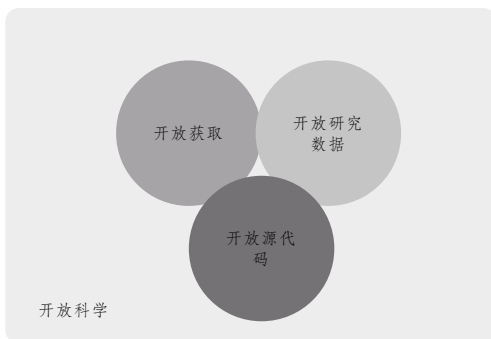


图1 开放科学的三大支柱

教育资源和软件<sup>[12]</sup>。欧洲开放科学云(European Open Science Cloud, EOSC)执行委员会提出,支持学术型/研究型软件相关学术基础设施的开发,并将其作为EOSC一部分<sup>[13]</sup>。联合国教科文组织2021年发布的《开放科学建议书》草案(现已由UNESCO第41届大会审议通过)明确指出,鼓励会员国将公共资金资助的研究所取得的科学知识(包括科学出版物、开放研究数据、开源软件、源代码和开放硬件)进行开放获取<sup>[14]</sup>。未来,开源软件、云服务将对开放科学产生巨大影响,很有可能出现基于云的开放科学,从而在云上构建具有创新力的开放社群,推动科学研究的发展<sup>[15]</sup>。

## 1.3 我国需大力扶持自主知识产权的软件研发工程

在中美贸易摩擦的背景下,由于外部力量的技术封锁,我国的技术突破和产业升级受到明显阻滞<sup>[16]</sup>。我国高端工业软件市场约有80%被国外垄断,中低端市场的自主率也不超过50%,目前还没有一家中国软件公司具备与世界级工业软件公司“掰手腕”的核心能力<sup>[17]</sup>。为尽早解决“卡脖子”问题,我国科技工作还需大力扶持自主知识产权的软件研发工程,把技术和发展的主动权牢牢掌握在自己手里。2020年,国务院印发《新时期促进集成电路产业和软件产业高质量发展若干政策的通知》<sup>[18]</sup>,支持软件关键核心技术研发与攻关,严格落实知识产权保护制度,努力打破国外核心技术封锁。但总的来看,我国软件

产业尤其是开源软件产业,正面临着开源软件断供、自主创新不足和知识产权保护等风险<sup>[10]</sup>。

## 2 软件级别计量的概念

软件级别计量这一概念继承了以往各种计量学概念的要义,本文在此简要回顾软件级别计量的历史脉络。一般认为,最早的文献计量研究始于20世纪初<sup>[19]</sup>,Cole和Eales于1917年开始进行文献统计方面的研究工作<sup>[20]</sup>。1922年,英国学者Hulme首次使用了术语“统计书目学”(Statistical Bibliography)<sup>[21]</sup>。1969年,英国学者Pritchard首次提出用Bibliometrics取代Statistical Bibliography,并将其解释为“数学和统计学方法在图书和其他学术交流媒介中的应用”<sup>[22]</sup>。同年,前苏联学者Nalimov和Mulechenko在合著中首次使用了“科学计量学”(Наукометрия)这一专业名词,其英文译名也很快被确定为“Scientometrics”<sup>[23]</sup>。至此,文献计量学作为一门学科正式诞生。1979年,德国学者Nacke提出了信息计量学(Informetrics)概念<sup>[24]</sup>。1994年,美国学者Narin依据自己在20世纪70年代对专利文献的计量分析,提出了专利计量(Patent Bibliometrics)术语,英文术语后演变成Patentometrics<sup>[25]</sup>。随着互联网的发展,文献计量学的研究对象也在不断拓展。1995年,Bossy使用术语“网络计量学”(Netometrics)来描述以互联网为媒介的科学交流<sup>[26]</sup>,随后被Almind和Ingwerson在1997年用Webometrics取代<sup>[27]</sup>,由于Cyber一词与以其作为词根的有关术语的广泛使用,Cybermetrics一词被频繁使用,也被认作是Webometrics的近义词。文献计量学在不断发展的过程中,涌现出了诸如对期刊作评价的期刊影响因子、对作者进行评价的H指数、G指数等各种指标和评价方法。随着越来越多的学者认为用期刊影响因子来评价学术成果存在问题和不合理性,2009年3月开放获取出版商PLoS提出了Article-Level Metrics(ALM)概念,并为其出版的每篇文献提供评价数据<sup>[28]</sup>。

2010年,Priem创造性地用Altmetrics(或Alternative Metrics)来指代传统期刊影响因子以外的替代性计量指标<sup>[29]</sup>。替代性计量指标不仅用于单篇论文的评价,还可用于学者、期刊、图书等的评价<sup>[30]</sup>。随着研究数据作为学术成果的价值逐渐被认识,数据(集)的质量评估也成了一个研究主题。2013年,Knowledge Exchange在一份报告中提出了数据集级别计量(Dataset-Level Metrics)<sup>[31]</sup>概念。同年,我国学者顾立平对数据级别计量(Data-Level Metrics)的概念和实践进展做了详细介绍,他认为数据级别计量是推进科学数据开放共享的一种途径,旨在识别科研人员对数据工作的贡献,主要是和数据发布与数据引用相关的计量,包括但不限于Altmetrics和论文级别计量(ALM)<sup>[32]</sup>。此外,还包括“实体计量学”(Entitymetrics)<sup>[33]</sup>等概念。

### 2.1 区分数据和代码

随着以数据密集型计算为特征的科学研究第四范式的兴起,科学(研究)数据的价值得到了越来越广泛的认可<sup>[34]</sup>。科学研究中科学(研究)数据的产出离不开相应软件或代码的支撑,没有软件的支持,大数据集的获取、处理、分析和可视化等工作则无法完成<sup>[35]</sup>。但是软件往往被认为是科学研究的副产品或被降级为二等地位,其学术价值一直被低估甚至被忽略<sup>[36,37]</sup>。近年来,软件在科学研究中的作用逐渐得到了认识,并逐渐与数据区分开来,一些资助机构(如美国NSF)也开始将软件认定为有效的科研成果<sup>[38-40]</sup>。Katz等详述了软件与数据的差别:①软件是可执行的,而数据不是;②数据提供证据,而软件提供工具;③软件是创造性的工作,科学数据是事实或观察;④相较于数据,软件要遭受不同类型的位元衰减(Software/bit rot,如软件腐败、代码腐败、比特腐败、软件腐蚀、软件衰变、软件熵),即软件开发时通常要使用其他的软件,相互之间有复杂的依存关系,而这些具有依赖性的软件包也经常变化;⑤软件的生命周期通常比数据要短<sup>[41]</sup>。



## 2.2 建立合理评估软件代码的工作机制

越来越多的研究人员将软件开发作为他们自己研究成果的一部分<sup>[42]</sup>,将其产出软件在互联网上公开共享<sup>[43]</sup>,也非常重视软件对科学研究的作用<sup>[44,45]</sup>。但是,与研究数据、研究论文不同的是,绝大多数的软件源代码是在学术界之外的行业和开发社群中被开发和使用<sup>[46]</sup>,而且在发表后开发人员还需要不断地对软件进行维护、更新、升级以保证其持续可用<sup>[47]</sup>。软件没有得到正式识别或根本没有被引用或标注,或者就算被标注,其方法也是完全不同于学术出版物的引用方式<sup>[46]</sup>。软件可见性问题,尤其是在学术出版物中的可见性问题,也引起了一系列担忧<sup>[48]</sup>。不同软件开发人员在软件优化过程中所发挥的作用也没有得到充分认识,这影响了他们对软件开发和维护的热情<sup>[36]</sup>。另外,由于研究数据通常依赖于软件生成,因而软件(精确到版本)对于研究的可重复性至关重要<sup>[49]</sup>。因此,建立一套合理评估软件代码的工作机制刻不容缓。

## 2.3 软件级别计量的定义

尽管国内外已有研究开展了软件影响力测量工作,但是目前还没有学者明确提出软件级别计量(Software-Level Metrics)这一概念,与之接近的术语有软件计量(Software Metrics)、软件测量指标(Software Measurement)、代码级别计量(Code-Level Metrics)或(源)代码计量(Source Code Metrics)。但是,这些术语概念都是从软件源代码的质量入手,旨在发现或预测软件源代码中可能出现的问题或改进点,属于软件开发与测试领域的质量控制研究,而不是对软件的影响力(尤其是学术影响力)进行评估,因而不是本文的研究范畴。

基于前文对软件级别计量概念发展脉络的介绍,本文认为软件级别计量就是对已知确定软件的影响力(以学术影响力为主)开展评价,旨在识别软件开发与维护人员的贡献,计量指标与方法包括但不限于在学术论文中提及和(或)引用软件频次、软件使用频次、Altmetrics等指标。软

件级别计量中的软件也包括科学研究工作中所用的代码、脚本、算法等工具型实体。

从软件生命周期来看,软件经历了从开发到发布再到使用和维护的一个过程,在这个过程中还会出现二次开发或版本更新。基于此,笔者简单归纳了软件级别计量可能的适用对象与作用。①对于软件开发人员而言,开展软件级别计量的目的是识别开发人员的贡献,尤其是学术贡献,也包括二次开发人员或后续开发人员的贡献。②对于发布者/发布机构/网站/数据库来说,软件级别计量可以扩大他们的影响力,实现软件在更大范围内的扩散、交流与共享。③对于使用人员来说,软件级别计量可以帮助他们识别、发现真正有用的软件,可以辅助研究工作,提高研究可重复性的可能。④对于维护人员(或机构)/网站/数据库来说,软件级别计量可以引起学术界对软件维护与保存的重视,提高维护机构/网站/数据库的影响力。

## 3 现有测量方法、标准和指标

与数据集的计量评价指标<sup>[50]</sup>类似,目前国内外关于软件级别计量的评价指标大多基于频次,通常是学术论文中软件的引用次数、提及次数、使用次数等(见表1)。

Crowston等在2003年首先提出,软件项目流行度的一个简单测量方式是下载次数,同时也提出了用户数、代码重用、信息页浏览次数、用户评级等测量方式<sup>[51]</sup>。Howison等于2015年提出了可用来测度软件影响力的多个指标,如下载次数、注册人数、用户评论数量等,同时也指出了使用这些指标可能会存在的问题<sup>[52]</sup>。2015年,我国学者Pan等根据软件在论文全文中的引用次数和提及次数来评估软件的学术影响力<sup>[37]</sup>。Thelwall等在2016年证实了下载次数可用作开源软件学术价值的一个评价指标<sup>[53]</sup>。Priem和Piwowar基于此结论发起了开源项目Depsy(<http://depsy.org>),以发现软件在开源社区中的影响力,指标包括下载次数(CRAN和

表 1 软件级别计量现有研究人员与评价指标

提出者	具体评价指标	年份
Crowston 等	下载次数、用户数、代码重用次数、信息页浏览次数、用户评级	2003
Howison 等	下载次数、被引次数、注册人数、邮件列表订阅人数、用户评论数	2015
Pan 等	全文提及次数、被引次数	
Thelwall 等	下载次数、被引次数	2016
杨波	提及软件的文献数量、软件提及文献的总被引次数、提及软件的网页数量	
潘雪莲	使用次数	
Priem 等	Desy: 下载次数、被引次数、被其他软件使用次数	2017
赵蓉英等	一级指标: 下载量、被引指标、复用指标 二级指标: 总下载量、近六个月的平均下载量、全文数据库中的提及次数、参考文献中的引用数量、软件的顶点度、软件的入度、PageRank 指数	
Zhao 等	下载量、引用量	
王玉琢等	提及论文数、提及总次数、提及位置	
Li 等	被引次数	
Pan 等	提及次数(标题、摘要、关键词)、对核心论文的引用次数、扩散深度与速度(文章、期刊、学科层面)	2018
于晓彤	软件提及、被引用情况	2019
He 等	被引次数	
魏瑞斌	在论文中的提及次数、使用次数	2021

PyPI 收集下载统计数据)、学术论文引用次数(软件论文+DOI+RRID)、被其他软件使用次数(代码再利用的 PageRank 算法)等<sup>[54]</sup>,但项目已于 2018 年 3 月结束并停止提供服务和数据。2017 年,赵蓉英等基于下载、引用和复用,提出了开源软件的学术影响力评价指标体系<sup>[55]</sup>。此外,国内外其他学者也在积极探索使用学术论文的引用次数和(或)提及次数来简要评价软件的学术影响力<sup>[56-67]</sup>。除了文献计量学指标,其他替代性指标也被用来评价知识实体的影响力,例如投票数、被讨论次数、被点赞数量等。IEEE 国际数据挖掘会议的组织者曾通过专家投票来评价算法的影响力<sup>[68]</sup>。

笔者根据已有研究文献,将评价指标简单归纳为三种(见表 2):①用量统计类指标,如下载量、注册数、用户数、复用数等;②文献计量学指标,如引用次数、提及次数;③其他补充指标,如订阅用户

数、用户评论/点赞数、社交网络提及数等。但上述指标都或多或少地存在一些问题。用量统计类指标(包括使用和复用)比较难追踪,存在非官网下载的使用行为,下载后是否使用也是一个问题<sup>[67]</sup>。在软件引用或提及方面,在学术论文中普遍存在软件引用(或提及)缺失或软件引用随意、不规则<sup>[48,60,61,69]</sup>等问题,而且软件被引用或被提及不等于被使用,软件提及次数也难以获取。其他补充指标可能存在的问题是数据来源标准不一致、易被“有心人”引导而造成指标缺乏公信力,只被用作补充评判指标而不是主要评价指标。在这三大类指标中,被引次数和被提及次数依然是当前国内外评价软件学术影响力的主流指标。

### 3.1 软件引用

国际学术团体和资助机构一直致力于通过

表2 软件级别计量指标及其存在的问题

类别	具体指标	存在问题
用量统计类指标	下载量、注册数、用户数、复用数或依存度等	①数据比较难获取或准确统计,如存在非官网下载的使用行为、不同版本的下载次数统计、多数软件官网不提供下载次数等; ②下载后是否使用也是一个问题。
文献计量学指标	引用次数、提及次数	①软件引用(或提及)缺失或软件引用随意、不规则; ②软件被引用或被提及不等于被使用; ③软件提及与引用次数难以获取,需从全文中挖掘。
其他补充指标	订阅用户数、用户评论/点赞数、浏览次数、社交网络提及数等	①来源平台众多且标准不一致; ②易被“有心人”引导而造成指标缺乏公信力; ③只被用作补充评判指标而不是主要评价指标。

软件引用来评价软件,其前提是软件引用规范化。目前,学术论文中软件引用的方式和内容比较随意和多样化。简单来看,软件引用方式可分为两种,一种是直接引用,一种是间接引用。直接引用的内容可能包括:软件论文(Software paper)、软件指南或说明书、软件/项目网址、软件存储/下载网址、软件存储位置提供的DOI号或RRID(Research Resource Identifier)号等永久标识符;间接引用则是引用其他使用或介绍该软件的文献(非软件论文或软件指南与说明书)。软件论文是一种详细描述某个软件的论文,也是一种经同行评议的正式出版物,通常由软件开发人员撰写,类似数据论文<sup>[70]</sup>。由于主流学术奖励体系依然以出版物作为主要评价方式,所以对开发者而言,他们的学术贡献一直难被认可。为了解决这一问题,软件论文在几年前应运而生。据统计,目前全球已有80多种同行评议期刊接收软件论文<sup>[71]</sup>。软件论文出现后,人们对软件本身的引用自然都转到软件论文上<sup>[72]</sup>,但是软件论文中的一些理论、方法等内容可能才是其被引用的原因,而不是软件自身<sup>[70]</sup>,同一个软件也可能被多篇软件论文介绍。

FORCE11<sup>①</sup>在2015年专门成立了软件引用工作组(FORCE11 Software Citation Working Group),该工作组在2016年发布了软件引用六大原则<sup>[73]</sup>:重要性、信誉和归属、唯一标识、持久性、可访问性和特异性。原则发布后,软件引用工作组解散,又成立了新的软件引用实施推进工作组(Software Citation Implementation Working Group),该组于2021年发布了软件引用指南,建议作者依据软件在论文中的提及位置选择不同的引用方式,同时建议将软件论文和软件本身同时引用而不是只引用软件论文<sup>[74]</sup>。国际研究软件联盟(Research Software Alliance, ReSA)、FORCE11和研究数据联盟(Research Data Alliance, RDA)也在联合推进研究软件的FAIR原则<sup>②</sup>(FAIR Principles for Research Software, FAIR4RS Principles)<sup>[75]</sup>。实践方面, Du等人开发了CiteAs平台(<https://citeas.org>),目的是:①鼓励软件开发人员明确提出对他们最有价值的引文方式,②鼓励软件用户更多地引用软件并遵循开发者要求的引用方式<sup>[76]</sup>。虽然目前来看软件引用规范化还有很多挑战和问题<sup>[77]</sup>,如论文元数据、学术基础设施(软件发现系统)如

① FORCE11的全称是The Future of Research Communications and e-Scholarship,是一个由研究人员、图书馆员、出版商和研究资助者组成的国际联盟,致力于改革或加强研究出版和传播系统。

② FAIR是Findable、Accessible、Interoperable、Reusable四个单词的首字母缩写,主要是强调研究数据的可发现、可获取、可互操作和可重用。

何保障,如何让科研人员接纳并做好软件引用,如何做好软件长期保存与维护等,但上述软件引用原则和指南及相关工作为学术界的软件引用制定了统一的原则和标准,相信未来软件引用的形式会更加规范。

### 3.2 软件提及

根据潘雪莲整理的出版物中软件提及和引用特征编码表<sup>[78]</sup>,学术文献中软件提及的信息通常可能包括软件名称、创建者/开发者、版本号、软件或项目的下载或存储网址或 DOI 号、RRID 号等。目前从软件提及角度对软件进行评价主要使用了机器学习算法,从文献中自动提取提及的软件名称等信息,辅之以人工判别。但是,这些提取工作都是在特定文献范围内开展,若想大范围推广以作为软件级别计量的指标之一,不仅需要公开此类算法,还需要有人员、机构或平台设立专门的项目,自动抓取全网

论文全文并用算法提取其中可能提及的软件。同时,提及软件的不同位置,提及是否真正使用也需要经人工判别后赋分计算。这是软件级别计量的一大难点。目前美国 SciCrunch.org 可以提供经人工判别后的引用或提及软件的文献,个人用户也可以登录后添加某软件的引用或提及文献,这为未来构建依赖学界、用户参与的软件学术贡献发现生态提供了一种实践可能和方向。

由于软件提及和软件引用都可能会涉及软件/项目网址、软件 DOI 号、RRID 号等,为了有效计算,本文提出软件引用的标准是:在参考文献中出现软件相关信息,如软件名称、软件论文、说明书或指南、其他介绍文档、下载网址、DOI 号或 RRID 号等;在论文中提到这些信息,但未写成参考文献的格式,则为软件提及。因此,笔者将软件提及方式、软件引用方式简要归为以下四种(见表 3),其中软件引用方式一的论文占比最多,其次是软件提及方式一<sup>[48]</sup>。

表 3 软件提及、软件引用方式说明及案例展示

类别	方式	说明	案例
软件提及	提及方式一——名称提及	提及软件名称(可能包括版本号等)	使用 Matlab 软件分析
	提及方式二——URL 提及	直接给出软件/项目下载网址	在 <a href="https://www.python.org/downloads">https://www.python.org/downloads</a> 处可免费下载
软件引用	引用方式一——引用文件	引用软件论文、指南或说明等资料性文件	使用 biosys (Swofford Selander, 1981) 计算
	引用方式二——引用地址或号码	引用、下载或访问 URL、DOI 号或 RRID 号等	使用 R <sup>[1]</sup> 绘制统计图 [1] The R project for statistical computing [EB/OL]. [2021-12-20]. <a href="https://www.r-project.org">https://www.r-project.org</a> .

### 3.3 Altmetrics

与出版物和数据集类似的是,软件在存储库中也可以通过类似 Altmetrics 的补充指标进行计量。以 Zenodo(<https://zenodo.org>)为例,软件级别计量指标有浏览次数、下载次数、唯一浏览次数(1 小时内的浏览次数算 1 次)、唯一下载次数(1 小时内的下载次数算 1 次)、被引用次数(计算来自文献、数据集、软件等的引用)。

GitHub(<https://github.com>)的评价指标包括跟踪(Watch)、点赞(Star)、使用(Fork)。PLoS(<https://plos.org/publish/metrics/>)也将代码作为与论文及数据同级别研究产出提供 Altmetrics 测量指标。另外,Altmetrics 还包括在社交软件中的讨论次数、网络中视频或教程视频数量及各视频的浏览次数与讨论次数等<sup>[79]</sup>。《荷兰研究软件可持续性——当前实践与建议》报告中

指出,软件引用和其他相关计量指标(如公开发布次数、贡献者群体大小)也要被考虑在研究评估和研究人员评价中<sup>[80]</sup>。但诸如 Altmetrics 的其他计量指标可能会因易被操控、标准不一致、暂无集成工具或指标等而有争议或可操作性较低,目前只能作为补充指标或参考指标。

#### 4 软件级别计量成为新的科研贡献承认机制的可能性

##### 4.1 软件级别计量具体示例

软件依据其可用性或可获得性分为以下四

表 4 软件类型及其访问权限<sup>[48]</sup>

软件类别	是否可被访问	是否需付费访问	是否提供源代码访问	是否提供源代码修改
不可访问软件 <sup>①</sup>	否	/	/	/
专有软件	是	是	否	否
开源软件	是	否	是	是
非商业软件	是	否	多数可以	否

① 由于软件无法获取就无法被使用和学术文献中体现,所以不可访问软件不在本文所述软件级别计量讨论范围中。

笔者将前文软件计量指标分别应用在后三种软件类型上,每种类型软件各选一个案例,以此来展示软件级别计量的应用探索(见表5)。表5证实了引用软件相关文献这种方式在学术

种:①不可访问软件(Not accessible);②专有软件(Proprietary),必须付费且源代码作为专有机密保存的软件;③开源软件(Open source),无需付费即可使用,提供对源代码的访问,并提供修改代码的明确许可;④非商业软件(Non-commercial),无需付费即可获得的软件,但不提供修改代码的明确许可,不过大多数(不是全部)这类软件都提供对源代码的访问<sup>[48]</sup>。笔者按照是否可被访问、是否需付费访问、是否提供源代码访问权限、是否提供源代码修改权限等将这四种软件加以区分(见表4)。

论文中较为普遍,引用的文献可能是软件说明书与指南、软件论文,也可能是其他人撰写的关于该软件的论文或资料等。其次较多的使用方式是软件名称提及。

表 5 不同类型的软件计量指标示例

软件计量指标类别	软件计量指标	专有软件(次) (如 GraphPad Prism)	开源软件(次) (如 Python)	非商业软件(次) (如 VOSviewer)
提及	提及方式一——名称提及	2 084	24 911	1 982
	提及方式二——URL 提及	8	38	3
引用	引用方式一——引用文件	3 937	173 027	4 019
	引用方式二——引用地址或号码	2 449	8 363	526
Altmetrics	社交网络讨论次数等	27 300	96 400 000	8 270

注:统计数据来源于 Scopus 和 Google 视频,检索日期为 2021 年 12 月 31 日。

##### 4.2 未来可能的应用情景

目前,国际上软件在学术中的作用和贡献

已经得到多方重视。部分研究资助机构已经认识到软件是一种有价值的研究产出,学术界正



在推动软件引用的规范化、标准化,出版商也在推出或更新各自的软件引用政策,要求在研究论文撰写中明确给出软件引用格式,如 Springer-Nature、Elsevier、AAAS 等国际知名出版商和期刊都发布了软件存储与引用相关政策<sup>[81]</sup>。软件级别计量在开放科学环境的助力下,将成为学术评价的下一片“蓝海”,它不仅可以帮助技术开发人员获得晋升机会,甚至是科研资助,还可以辅助用户识别、发现可靠软件,促进软件行业的繁荣发展。

软件级别计量可能会有以下应用场景(见图2):①对研究人员来说,可以辅助科学研究过

程中软件的查找与发现,甚至为研究结果复现提供最后一块“拼图”;②对资助机构或评审专家而言,可以全方位开展项目资助评估、项目结题审核与评价;③对软件从业人员/维护网站等来说,可以促进代码开发与优化,其学术贡献在研究评估中也会得到认可,并继续支撑科研;④对图书馆员来说,可以推动软件长期保存,开展新型知识评价服务以及相关科普工作;⑤对公众来说,可以帮助他们更高效便捷地开展个人学习。不过,应用前提是已经具备了相关政策、技术和基础设施。

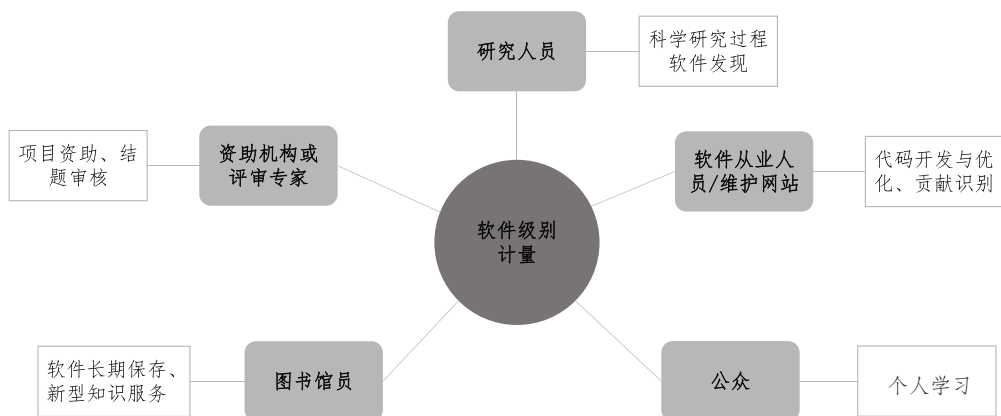


图2 软件级别计量未来可能的应用场景

## 5 结语

在开放科学环境下,图书馆将继续扮演知识服务中心的角色,推动知识可持续发展<sup>[82]</sup>。软件作为一种创造型知识产出,也引起了部分图书馆和馆员的注意。如美国国会图书馆在

2013年发布了软件长期保存国家战略的报告,提出要发现有价值且存在风险的软件,做好可执行软件的长期保存<sup>[83]</sup>。作为知识服务机构,图书馆可以面向科研人员 and 科研管理人员提供软件级别计量服务,并将其作为学术评价与评估中的一个方面加以广泛宣传和实践,再以此为契机嵌入到科学研究过程中。

## 参考文献

- [1] 易志军,庄岩,江丽辉. 拟定全球“开放科学”准则:促进后疫情时代的国际科学合作[J]. 科学观察,2020,15(5):63-68. (Yi Z J, Zhuang Y, Jiang L H. Draft global “open science” guidelines: promote international scientific cooperation in the post-epidemic era[J]. Science Focus, 2020, 15(5):63-68.)
- [2] Hong N P C, Hammitzsch M, Hufton A, et al. Open science goes Geo-part II: scientific software[C]//General Assembly European Geosciences Union, 2015.

- [ 3 ] Hettrick S, Antonioletti M, Carr L, et al. UK research software survey 2014[ EB/OL]. [ 2021-07-25]. <https://zenodo.org/record/14809#.Yc7nkmBBY3A>.
- [ 4 ] Kelley I. Publish or perish: the credit deficit to making software and generating data[ EB/OL]. [ 2021-07-25]. [https://figshare.com/articles/journal\\_contribution/Publish\\_or\\_perish\\_the\\_credit\\_deficit\\_to\\_making\\_software\\_and\\_generating\\_data/1112579/2](https://figshare.com/articles/journal_contribution/Publish_or_perish_the_credit_deficit_to_making_software_and_generating_data/1112579/2).
- [ 5 ] 中国开源软件推进联盟. 2021 中国开源发展蓝皮书[ EB/OL]. [ 2021-07-25]. <https://codechina.csdn.net/COPU/copu/-/blob/master/static/images/%E5%BC%80%E6%BA%90%E8%93%9D%E7%9A%AE%E4%B9%A6.pdf>. ( China OSS Promotion Union. China open source blue book ( 2021) [ EB/OL]. [ 2021-07-25]. <https://codechina.csdn.net/COPU/copu/-/blob/master/static/images/%E5%BC%80%E6%BA%90%E8%93%9D%E7%9A%AE%E4%B9%A6.pdf>.)
- [ 6 ] 开源软件是什么? 有哪些?[ EB/OL]. [ 2021-07-25]. <http://c.biancheng.net/view/2943.html>. ( What is open source software? What are they?[ EB/OL]. [ 2021-07-25]. <http://c.biancheng.net/view/2943.html>.)
- [ 7 ] Open source software[ EB/OL]. [ 2021-07-25]. [https://en.wikipedia.org/wiki/Open-source\\_software](https://en.wikipedia.org/wiki/Open-source_software).
- [ 8 ] 陈光沛, 魏江, 李拓宇. 开源社区: 研究脉络、知识框架和研究展望[ J]. 外国经济与管理, 2021, 43( 2) : 84-102. ( Chen G P, Wei J, Li T Y. Open source community: research context, knowledge framework and research prospects[ J]. Foreign Economics & Management, 2021, 43( 2) : 84-102. )
- [ 9 ] 开源社区[ EB/OL]. [ 2021-07-25]. <https://baike.baidu.com/item/%E5%BC%80%E6%BA%90%E7%A4%BE%E5%8C%BA/1373560>. ( Open source community [ EB/OL]. [ 2021-07-25]. <https://baike.baidu.com/item/%E5%BC%80%E6%BA%90%E7%A4%BE%E5%8C%BA/1373560>.)
- [ 10 ] 王晓冬. 我国开源软件产业面临的突出风险及对策研究[ J]. 信息安全研究, 2021, 7( 10) : 973-976. ( Wang X D. Research on outstanding risks and countermeasures faced by open source software industry of China [ J]. Journal of Information Security Research, 2021, 7( 10) : 973-976. )
- [ 11 ] Jomier J. Open science-towards reproducible research[ J]. Information Services & Use, 2017( 37) : 361-367.
- [ 12 ] All European Academies. Open science for the 21st century[ EB/OL]. [ 2021-07-27]. <https://allea.org/portfolio-item/open-science-for-the-21st-century>.
- [ 13 ] European Open Science Cloud. Scholarly infrastructures for research software[ EB/OL]. [ 2021-07-29]. <https://www.ouvri.lascience.fr/wp-content/uploads/2021/02/Scholarly-infrastructures-for-research-software.pdf>.
- [ 14 ] UNESCO. Final report on the draft text of the UNESCO recommendation on open science[ EB/OL]. [ 2021-07-28]. <https://unesdoc.unesco.org/ark:/48223/pf0000376130>.
- [ 15 ] Gentemann C L, Holdgraf C, Abernathy R, et al. Science storms the cloud[ J]. AGU Advances, 2021, 2( 2) : e2020AV000354.
- [ 16 ] 邢冬梅. “卡脖子”技术问题的成因与规避——技术轨道的分析视角[ J]. 国家治理, 2020( 45) : 21-25. ( Xing D M. The cause and avoidance of “stuck neck” technical problem: analysis perspective of technical track [ J]. Governance, 2020( 45) : 21-25. )
- [ 17 ] 何立胜. 中央经济工作会议两处提到“卡脖子”, 如何破解? 关键就靠这一招[ N/OL]. [ 2021-07-27]. [https://s.cyol.com/articles/2020-12/30/content\\_k6N5O5IG.html](https://s.cyol.com/articles/2020-12/30/content_k6N5O5IG.html). ( He L S. The central economic work conference mentioned “stuck neck” in two places. How to solve it? The key lies in this trick[ N/OL]. [ 2021-07-27]. [https://s.cyol.com/articles/2020-12/30/content\\_k6N5O5IG.html](https://s.cyol.com/articles/2020-12/30/content_k6N5O5IG.html).)
- [ 18 ] 国务院关于印发新时期促进集成电路产业和软件产业高质量发展若干政策的通知[ EB/OL]. [ 2021-07-27]. [http://www.gov.cn/zhengce/content/2020-08/04/content\\_5532370.htm](http://www.gov.cn/zhengce/content/2020-08/04/content_5532370.htm). ( The State Council issued the notice on several policies to promote the high-quality development of the integrated circuit industry and the software industry in the new period[ EB/OL]. [ 2021-07-27]. [http://www.gov.cn/zhengce/content/2020-08/04/content\\_5532370.htm](http://www.gov.cn/zhengce/content/2020-08/04/content_5532370.htm).)
- [ 19 ] 邱均平, 赵蓉英, 董克. 科学计量学[ M]. 北京: 科学出版社, 2016: 5. ( Qiu J P, Zhao R Y, Dong K. Scientometrics[ M]. Beijing: Science Press, 2016: 5. )
- [ 20 ] Cole F T, Eales N B. The history of comparative anatomy[ J]. Science Progress, 1917( 11) : 578-596.
- [ 21 ] Hulme E W. Statistical bibliography in relation to the growth of modern civilization[ M]. London: Grafton, 1923:

- 29-44.
- [22] Pritchard A. Statistical bibliography or bibliometrics[J]. *Journal of Documentation*, 1969, 25(4): 348-349.
- [23] Hood W W, Wilson C S. The literature of bibliometrics, scientometrics, and informetrics[J]. *Scientometrics*, 2001, 52(2): 291-314.
- [24] Nacke O. Informetrie; ein neuer Name für eine neue Disziplin[J]. *Nachrichten für Dokumentation*, 1979, 30(6): 219-226.
- [25] Narin F. Patent bibliometrics[J]. *Scientometrics*, 1994, 30(1): 147-155.
- [26] Bossy M J. The last of the litter; “Netometrics”[J]. *Solaris Information Communication*, 1995(2): 245-250.
- [27] Almind T C, Ingwersen P. Informetric analyses on the world wide web; methodological approaches to “webometrics”[J]. *Journal of Documentation*, 1997, 53(4): 404-426.
- [28] Article-level metrics[EB/OL]. [2021-07-29]. [https://en.wikipedia.org/wiki/Article-level\\_metrics](https://en.wikipedia.org/wiki/Article-level_metrics).
- [29] Tattersall A. Supporting the research feedback loop; why and how library and information professionals should engage with altmetrics to support research[J]. *Performance Measurement and Metrics*, 2017, 18(1): 28-37.
- [30] Dutta B. The journey from librmetry to altmetrics; a look back [EB/OL]. [2021-07-29]. <http://eprints.rclis.org/23665/2/B-Dutta-JU-Golden-Jubilee-Paper.pdf>.
- [31] Costas R, Meijer I, Zahedi Z, et al. The value of research data[R/OL]. [2021-07-29]. [https://repository.jisc.ac.uk/6205/1/Value\\_of\\_Research\\_Data.pdf](https://repository.jisc.ac.uk/6205/1/Value_of_Research_Data.pdf).
- [32] 顾立平. 数据级别计量——概念辨析与实践进展[J]. *中国图书馆学报*, 2015, 41(2): 56-71. (Gu L P. Data level metric: its concepts and progress[J]. *Journal of Library Science in China*, 2015, 41(2): 56-71.)
- [33] Ding Y, Song M, Han J, et al. Entitymetrics: measuring the impact of entities[J]. *PLoS One*, 2013, 8(8): e71416.
- [34] Hey T, Tansley S, Tolle K. The fourth paradigm: data-intensive scientific discovery[M]. Redmond, WA: Microsoft Research, 2009: xix.
- [35] Goble C. Better software, better research[J]. *IEEE Internet Computing*, 2014, 18(5): 4-8.
- [36] Hafer L, Kirkpatrick A E. Assessing open source software as a scholarly contribution[J]. *Communications of the ACM*, 2009, 52(12): 126-129.
- [37] Pan X, Yan E, Wang Q, et al. Assessing the impact of software on science; a bootstrapped learning of software entities in full-text papers[J]. *Journal of Informetrics*, 2015, 9(4): 860-871.
- [38] NSF. GPG summary of changes [EB/OL]. [2021-07-29]. [https://www.nsf.gov/pubs/policydocs/papguide/nsf13001/gpg\\_sigchanges.jsp](https://www.nsf.gov/pubs/policydocs/papguide/nsf13001/gpg_sigchanges.jsp).
- [39] Piwowar H. Value all research products[J]. *Nature*, 2013, 493(7431): 159.
- [40] NSF. Supporting scientific discovery through norms and practices for software and data citation and attribution [EB/OL]. [2021-07-29]. <https://www.nsf.gov/pubs/2014/nsf14059/nsf14059.jsp>.
- [41] Katz D S, Niemeyer K E, Smith A M, et al. Software vs. data in the context of citation[J]. *PeerJ Preprints*, 2016, 4: e2630v1.
- [42] Hannay J E, MacLeod C, Singer J, et al. How do scientists develop and use scientific software?[C]//2009 ICSE Workshop on Software Engineering for Computational Science and Engineering. IEEE, 2009: 1-8.
- [43] Kraker P, Leony D, Reinhardt W, et al. The case for an open science in technology enhanced learning[J]. *International Journal of Technology Enhanced Learning*, 2011, 3(6): 643-654.
- [44] Software Sustainability Institute. It's impossible to conduct research without software, say 7 out of 10 UK researchers[EB/OL]. [2021-07-30]. <https://www.software.ac.uk/blog/2014-12-04-its-impossible-conduct-research-without-software-say-7-out-10-uk-researchers>.
- [45] Nangia U, Katz D S. Track 1 paper: surveying the US National Postdoctoral Association regarding software use and training in research[C]//Workshop on Sustainable Software for Science: Practice and Experiences (WSSSPE 5.1). 2017.
- [46] Research Data Alliance. Use cases and identifier schemes for persistent software source code identification[EB/OL]. [2021-07-30]. <https://www.rd-alliance.org/sites/default/files/Card%20RDA%2020200910.pdf>.

- [47] 于晓彤,潘雪莲,华薇娜. 基于内容分析法的开源软件使用与引用研究——以 Word2vec 为例[J]. 图书情报知识,2020,13(1): 101-107,121. (Yu X T, Pan X L, Hua W N. The use and citation of open source software based on content analysis: a case study of Word2vec[J]. Documentation, Information & Knowledge, 2020, 13(1): 101-107, 121.)
- [48] Howison J, Bullard J. Software in the scientific literature: problems with seeing, finding, and using software mentioned in the biology literature[J]. Journal of the Association for Information Science and Technology, 2016, 67(9): 2137-2155.
- [49] Niemeyer K E, Smith A M, Katz D S. The challenge and promise of software citation for credit, identification, discovery, and reuse[J]. Journal of Data and Information Quality (JDIQ), 2016, 7(4): 1-5.
- [50] Belter C W. Measuring the value of research data: a citation analysis of oceanographic data sets[J]. PLoS One, 2014, 9(3): e92590.
- [51] Crowston K, Annabi H, Howison J. Defining open source software project success[C]//Proceedings of the International Conference on Information Systems(ICIS), 2003:327-340.
- [52] Howison J, Deelman E, McLennan M J, et al. Understanding the scientific software ecosystem and its impact: current and future measures[J]. Research Evaluation, 2015, 24(4): 454-470.
- [53] Thelwall M, Kousha K. Academic software downloads from Google code: useful usage indicators?[J]. Information Research: An International Electronic Journal, 2016, 21(1): n1.
- [54] Singh C D. The unsung heroes of scientific software[J]. Nature News, 2016, 529(7584): 115.
- [55] 赵蓉英,魏明坤,汪少震. 基于 Altmetrics 的开源软件学术影响力评价研究[J]. 中国图书馆学报, 2017, 43(2): 80-95. (Zhao R Y, Wei M K, Wang S Z. Evaluation of academic impact of open source software based on altmetrics[J]. Journal of Library Science in China, 2017, 43(2): 80-95.)
- [56] Li K, Yan E, Feng Y. How is R cited in research outputs? Structure, impacts, and citation standard[J]. Journal of Informetrics, 2017, 11(4): 989-1002.
- [57] Zhao R, Wei M. Impact evaluation of open source software: an altmetrics perspective[J]. Scientometrics, 2017, 110(2): 1017-1033.
- [58] He J, Lou W, Li K. How were science mapping tools applied? The application of science mapping tools in LIS and non-LIS domains[J]. Proceedings of the Association for Information Science and Technology, 2019, 56(1): 404-408.
- [59] Pan X, Yan E, Hua W. Disciplinary differences of software use and impact in scientific literature[J]. Scientometrics, 2016, 109(3): 1593-1610.
- [60] Pan X, Yan E, Cui M, et al. How important is software to library and information science research? A content analysis of full-text publications[J]. Journal of Informetrics, 2019, 13(1): 397-406.
- [61] 杨波,王雪,余曾漂. 生物信息学文献中的科学软件利用行为研究[J]. 情报学报, 2016, 35(11): 1140-1147. (Yang B, Wang X, She Z L. Research on using behavior of scientific software in bioinformatics literature[J]. Journal of the China Society for Scientific and Technical Information, 2016, 35(11): 1140-1147.)
- [62] 杨波,王雪,苏娜. 不同文献集中中国学者引用软件和数据集的特征比较研究[J]. 图书情报工作, 2017, 61(14): 109-115. (Yang B, Wang X, Su N. A comparative study on characteristics of how Chinese scientists citing software and datasets in different publication communities[J]. Library and Information Service, 2017, 61(14): 109-115.)
- [63] 王玉琢,章成志. 考虑全文本内容的算法学术影响力分析研究[J]. 图书情报工作, 2017, 61(23): 6-14. (Wang Y Z, Zhang C Z. Using full-text to analyse academic impact of algorithms[J]. Library and Information Service, 2017, 61(23): 6-14.)
- [64] 崔明,潘雪莲,华薇娜. 我国图书情报领域的软件使用和引用研究[J]. 中国图书馆学报, 2018, 44(3): 66-78. (Cui M, Pan X L, Hua W N. Software usage and citation in the field of library and information science in China[J]. Journal of Library Science in China, 2018, 44(3): 66-78.)
- [65] 于晓彤,潘雪莲,华薇娜. 知识图谱研究中的软件引用和扩散分析[J]. 情报资料工作, 2019, 40(2): 19-29. (Yu X T, Pan X L, Hua W N. Analysis of software citation and diffusion in knowledge mapping research[J]. Information and Documentation Services, 2019, 40(2): 19-29.)

- [66] 魏瑞斌. 基于论文提及的科学软件在国内图书情报学领域的应用现状分析[J]. 情报杂志, 2021, 40(5): 165-170. (Wei R B. Analysis of the application status of scientific software mentioned in the paper in the field of library and information science in China[J]. Journal of Intelligence, 2021, 40(5): 165-170. )
- [67] 潘雪莲, 孙梦佳, 于晓彤, 等. 中国科研人员的科学软件使用和引用行为研究[J]. 现代情报, 2021, 41(8): 76-86. (Pan X L, Sun M J, Yu X T, et al. Study of Chinese researchers' scientific software usage and citation behavior[J]. Journal of Modern Information, 2021, 41(8): 76-86. )
- [68] Wu X, Kumar V, Quinlan J R, et al. Top 10 algorithms in data mining[J]. Knowledge and Information Systems, 2008, 14(1): 1-37.
- [69] Costa J, Meirelles P, Chavez C. On the sustainability of academic software; the case of static analysis tools[C]// Proceedings of the XXXII Brazilian Symposium on Software Engineering. 2018: 202-207.
- [70] Li K, Chen P Y, Yan E. Challenges of measuring software impact through citations; an examination of the lme 4 R package[J]. Journal of Informetrics, 2019, 13(1): 449-461.
- [71] Hong N P C. In which journals should I publish my software? [EB/OL]. [2021-08-24]. <https://www.software.ac.uk/which-journals-should-i-publish-my-software>.
- [72] Li K, Lin X, Greenberg J. Software citation, reuse and metadata considerations; an exploratory study examining LAMMPS[J]. Proceedings of the Association for Information Science and Technology, 2016, 53(1): 1-10.
- [73] Smith A M, Katz D S, Niemeyer K E. Software citation principles[J]. PeerJ Computer Science, 2016, 2: e86.
- [74] Katz D S, Hong N P C, Clark T, et al. Recognizing the value of software: a software citation guide[version 2; peer review: 2 approved][J]. F1000Research, 2021, 9: 1257.
- [75] Hong N P C, Katz D S, Barker M, et al. FAIR principles for research software (FAIR4RS principles)[EB/OL]. [2021-08-08]. [https://www.rd-alliance.org/system/files/FAIR4RS\\_Principles\\_v0.3\\_RDA-RFC.pdf](https://www.rd-alliance.org/system/files/FAIR4RS_Principles_v0.3_RDA-RFC.pdf).
- [76] Du C F, Cohoon J, Priem J, et al. CiteAs: better software through sociotechnical change for better software citation [C]// Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing. Association for Computing Machinery, New York, USA, 2021: 218-221.
- [77] Druskat S, Katz D S, Todorov I T. Research software sustainability and citation[J]. arXiv preprint arXiv: 2103.06681, 2021.
- [78] 潘雪莲. 软件实体的自动抽取和学术影响力研究[D]. 南京: 南京大学, 2016. (Pan X L. Analysing the academic impact of software: a bootstrapped learning of software entities in full-text papers[D]. Nanjing: Nanjing University, 2016. )
- [79] Osinska V, Klimas R. Mapping science: tools for bibliometric and altmetric studies[J]. Information Research, 2021, 26(4): 909.
- [80] Saskia V E, Tom B, Maria C, et al. Research software sustainability in the Netherlands: current practices and recommendations[EB/OL]. [2021-08-24]. <https://zenodo.org/record/4543569#.Yc7Zk2BBz30>.
- [81] FORCE11 Software Citation Implementation Working Group. Software citation policies index[EB/OL]. [2021-08-24]. <https://www.chorusaccess.org/resources/software-citation-policies-index/>.
- [82] 于永丽. 开放科学环境下高校图书馆科研支持服务模式研究[J]. 图书馆学报, 2019, 41(8): 87-91, 101. (Yu Y L. Research on the scientific research support service mode of university library in the open science environment[J]. Journal of Library Science, 2019, 41(8): 87-91, 101. )
- [83] National Digital Information Infrastructure and Preservation Program. Toward a national strategy for software preservation [R/OL]. [2021-08-26]. [https://www.digitalpreservation.gov/multimedia/documents/Preserving\\_EXE\\_report\\_final101813.pdf](https://www.digitalpreservation.gov/multimedia/documents/Preserving_EXE_report_final101813.pdf).

陆彩女 中国科学院上海药物研究所信息中心馆员。上海 201203。

顾立平 中国科学院文献情报中心研究员, 博士生导师。北京 100190。

(收稿日期: 2022-01-04)