

论推动替代计量学发展的若干基本问题^{*}

邱均平 余厚强

摘要 替代计量学快速发展,同时也造成了一些困扰,主要表现在替代计量学的研究内涵存在争议,替代计量指标内在价值不够明朗,替代计量指标可信度遭到质疑。本文对近年来替代计量学的研究论著进行追踪和梳理,利用狭义、广义二分法归纳替代计量学的研究内涵,明确提出狭义的替代计量学专门研究相对传统引文指标的在线新型计量指标及其应用,广义的替代计量学研究在线新型科学交流体系和面向学术成果的全面影响力评价指标体系。若根据应用情境对替代计量指标和数据源进行细分,既能反映学术成果的社会影响力,也能反映学术成果的学术影响力。从规避数据操纵、数据严谨性和数据一致性三个方面,论证了替代计量指标具备实用的可信度。替代计量学研究有利于增加发展中国家科学家的话语权,对创新型科学交流机制起到促进作用。对应 Altmetrics 这个英文术语,“替代计量学”是最合适的中文译名。表 1。参考文献 60。

关键词 替代计量学 科学评价 科学交流 引文指标 可信度

分类号 G350

Some Basic Problems in Advancing the Development of Altmetrics

Qiu Junping & Yu Houqiang

ABSTRACT

With systematic survey, induction and analysis of altmetrics literature, the study aimed to answer the following three questions: 1) What does altmetrics study? What is it alternative for? 2) What is the relationship between altmetrics indicators and traditional citation-based indicators? And why? 3) Are altmetrics indicators valid? Are altmetrics indicators vulnerable and easy to manipulate because they are based on online data?

Results showed that the meaning of altmetrics evolved through three stages. In the first stage, the proposer of the term altmetrics made the initial definition. Altmetrics was deemed to possess the potential of replacing the traditional citation-based literature filter mechanism and scientific evaluation mechanism. It is defined as “the creation and study of new metrics based on the Social Web for analyzing, and informing scholarship”; In the second stage, researchers had defined altmetrics in various ways. To sum up, these definitions are different in the coverage. The most manifest definition is that “altmetrics is social media metrics”, the most general definition is that “altmetrics combined traditional bibliometrics indicators and novel online indicators to give alternative solution for measuring the complete impact of academic products”; In the third stage, altmetrics research community had come to agreement. Altmetrics in the broad sense emphasizes the change of research perspective, and aims to replace the quantitative scientific evaluation system relying solely on traditional citation-based indicators with the evaluation system of full impact oriented to academic products, and simultaneously advance the development of open science and online scientific communication. Altmetrics in the narrow sense specifically studies novel online indicators and their application compared to traditional citation-based indicators, especially indicators

^{*} 本文系中央高校基本科研业务费专项资金资助重点项目“替代计量学理论与实证研究”(编号:2014104010201)的研究成果之一。

通信作者:余厚强,Email:yuhouq@yeah.net

based on social media data.

The meanings of altmetrics indicators are gradually being revealed. 1) Studies showed certain types of the altmetrics indicators cannot reflect academic impact. Altmetrics indicators reflect different meaning with citation-based indicators, thus should be combined to use in evaluation. 2) Altmetrics indicators reflect visibility, popularity and societal impact. Altmetrics cover more general scientific product users and reflect societal impact. 3) Altmetrics indicators based on data source from the academia reflect more academic impact. Academic impact should undoubtedly be limited to academic users, because users outside the academia cannot contribute citations. To sum up, the key to properly use altmetrics indicators is to distinguish the application context.

Criticism to the validity of altmetrics can be summarized into three aspects. 1) altmetrics indicators bear the risk of being manipulated in large scale; 2) altmetrics data sources are not scientific and rigorous enough; 3) part of altmetrics data source cannot be traced. Despite the criticism, it must be made clear that these problems can be solved and avoided. 1) the risk of manipulating altmetrics indicators in large scale can be avoided; 2) every altmetrics indicator reflects users' value judgment; 3) the establishment of altmetrics infrastructure makes the raw altmetrics data easy to store and reuse. 1 tab. 60 refs.

KEY WORDS

Altmetrics. Scientific evaluation. Scientific communication. Citation indicators. Robustness.

替代计量学是近年国际科学计量学领域的热点研究主题,逐渐引起国内学者的关注。替代计量学在我国最早是由刘春丽等人^[1]引介,但是因为译名、时机、内容等问题,没有引起国内学者注意。自邱均平等^[2]系统梳理替代计量学的产生背景,并分阶段对替代计量学研究进展做详细分析以后,国内相关研究开始兴起。首先表现在引介性论文的大量涌现,例如:由庆斌等人^[3]介绍了替代计量学的指标、主要研究内容和分类情况;崔宇红^[4]介绍了替代计量学的定义、特征和数据源,综述其发展演变和研究主题;陈铭^[5]对替代计量学的起源、发展、定义进行了详细的阐述,并从研究对象、对待影响因子的态度、数据获取、发生时间、产生的影响力和同行评审六个方面分析其与传统计量学的不同。邻近学科也开始引介替代计量学,例如,赵文华等人^[6]从编辑的视角归纳了替代计量学公正、客观、透明的属性。

除了引介性论文外,还出现了实证性和理论性论文。宋丽萍等人^[7]将 F1000 与 WoS 引文指标做相关性分析,发现 F1000 因子与统计区间内的被引频次呈显著正相关,但同时有些 F1000 因子很高的论文并没有高频被引。刘春

丽等人^[8]研究了 Mendeley、F1000 和 Google Scholar 与引文的关联性,发现 Mendeley 读者人数指标与 Google Scholar 的被引次数指标在论文评价结果中的相关程度相对较高。宋丽萍等^[9]对 F1000、Mendeley 和 WoS、Google Scholar 做相关分析,发现同行评议结果、传统引文分析指标以及以 Mendeley 为代表的影响计量指标具有低相关性。理论方面,余厚强等人^[10]阐述替代计量学的生态体系,据此构建在线科学交流新模式。

然而,与国外的替代计量学研究热潮相比,国内的相关研究还很薄弱。笔者认为有两点主要原因。①国内同行,包括期刊主编、专家学者,对替代计量指标持怀疑态度,认为下载、分享、阅读等作为评价指标,缺乏严肃性和可靠性。②我国学者对包括学术网络平台在内的在线科研工具利用率较低,使得替代计量数据积累不充分,可用于实证的替代计量数据源较稀缺,研究难度较大。第二点原因是我国客观科研环境造成的,这里先不做讨论,但是第一点原因却与缺乏对替代计量学的正确认识有关。诚然,替代计量学的快速发展,曾经带来一些困扰^[11],但是随着更多实证研究的推进,这些

问题日益清晰,本文对替代计量学最容易受到质疑的若干核心问题进行梳理,并对笔者^[12]之前提出的替代计量指标分层模型进行补充阐释。本文主要研究三个问题。①替代计量学到底研究什么?它要替代什么?②替代计量指标和传统引文指标之间是否存在关联关系?及其原因。③替代计量学的指标可信吗?基于网络数据的替代计量指标是否脆弱,又易受人为操控?

通过对这些问题的凝练和解答,能消除某些疑虑,为我国替代计量学研究铺开道路,吸引更多的学者投身到替代计量学的研究中去。尽管由于指标种类的多样性等原因,替代计量指标的内涵还没有完全明确,但是已经达成初步共识。有些研究人员也认识到,基于替代计量指标的服务是未来的潮流。

1 替代计量学内涵的演变

尽管 Priem 等人提出了“替代计量学”的术语,但是显然他们并不能决定它将向哪个方向发展,不同的利益相关者和研究人员,根据自己的理解对它进行阐释并开展研究。当我们认真审视替代计量学的发展过程时,不难发现最热衷于替代计量研究的群体是那些关注创新科学交流模式的组织或个人,例如致力于推动开放科学进程的 PLOS,专注于在线文献管理的 Mendeley 等。但是基金会资助政策的变革,出版巨头的青睐,和瞄准商机的创业公司,无疑起到推波助澜的作用。而支撑这些行为的源动力,来自对未来科研活动趋势的基本预测:科研活动将越来越开放,越来越多地转移到网上,与之相应的数据,也就越来越有价值。通过对所有替代计量学文献的回顾,笔者将替代计量学内涵的演变过程总结为三个阶段。

第一阶段,“替代计量学”提出者的最初界定。Priem 等人觉察到社会网络等在线交流渠道对科学交流产生了影响,相关的观察结果零星地散落在他们的早期文献里,随着他们对“科

学计量学 2.0”理念的提出和推广,他们日益意识到自己的想法可能是革命性的。怀着无限热忱,Priem 等人^[13]在 *Nature* 上发文描绘了相对传统期刊时代的另一个全新的在线科学交流时代,比起科学计量学 2.0,对这种在线科学交流新时代的计量研究需要更专业的术语来进行概括,于是 Altmetrics 得以提出。起初模仿 20 世纪 70 年代医学领域兴起的“Alternative Medicine”(替代医学),叫做“Alternative Metrics”,后来为了方便改称为“Alt-metrics”,最后从计量学构词角度出发,将横线也去掉了。不得不承认,作为科学术语,“Altmetrics”并不严谨,它既没有揭示出研究的对象,也没有明确的界限,但是在当时的情境下,要提出结构合理、含义清晰的术语,确实难度太大。这一方面和 Priem 等人的学术功底有关,另一方面,也和该研究领域的重要创新性有关。在这一阶段,替代计量学被认为具备替代传统基于引文的文献过滤机制和科研评价机制的潜力,Priem 等人^[14]将其界定为“基于社会网络的新计量指标的建立和研究,旨在分析和全面衡量学术”。

第二阶段,“替代计量学”研究人员理解纷纭。“替代计量学”术语在得到推广的同时,也日益受到严厉的批判,一个几乎相同的论断是这一新指标无法替代传统引文指标,充其量是既有体系的补充。Priem 等人认识到替代计量学还在发展初期,不宜挑战现有体系,所以做出了让步,认为现阶段替代计量新指标可以作为既有体系的有力补充^[15]。当然,从最开始,引文分析的价值就从没有被否认过,替代计量指标只是换一种角度来测度。从历史的角度看,新生事物在初期都会饱受争议,引文检索就是一例^[16]。然而,新生事物要获得认可,例如在线科学交流新体系,除了要以愿景激励人心以外,还要在实际操作上具备切实可行的方案。根据现实的需求,Priem 等人创立 ImpactStory (www.impactstory.org),而替代计量的商业价值也造就了众多创业公司,其服务为出版商、机构知识库和高校所采用。替代计量学在迅速发展,但是其

内涵诠释的多样性也给新进入的研究者造成一定困扰。美国信息标准化组织(National Information Standards Organization, NISO)甚至考虑要换掉“Altmetrics”这个术语,但是因为并没有提出

更好的,所以未能实行^[17]。在这一阶段,各个研究人员根据自己对替代计量学的理解,对其进行诠释,其中代表性的理解如表1所示。

表1 不同学者对替代计量学内涵的理解

序号	替代计量学的内涵	作者
1	替代计量学等同于社交媒体计量(Social Media Metrics),是利用社会网络来衡量学术成果更广泛影响力的替代性方法。	Haustein ^[18] 、Fenner等 ^[19]
2	替代计量学就是论文层次计量(Article-Level Metrics)。	Lin ^[20]
3	替代计量学是在论文层次计量之上更广泛计量指标研究的代称。	Lin ^[21]
4	替代计量学研究基于网络的替代性指标,与网络计量学是同义语。	Roemer等人 ^[22]
5	替代计量学对在线分享、保存、评论、采用和社交媒体使用进行计分,全面反映学术成果的影响力。	Konkiel ^[23]
6	替代计量学是一种研究在线空间,捕捉学者交互行为以反映学术影响力的新计量指标。	Lapinski等人 ^[24]
7	替代计量学衡量单篇论文获得的在线关注度。	Mounce等人 ^[25]
8	替代计量学要结合传统文献计量评价方法,给出学术影响力的替代性的、多维度的更加广泛的视角。	Rasmussen等人 ^[26]
9	新的学术成果和科学交流渠道需要新工具来衡量这些产出的影响力,替代计量学正是赋予这些新工具集合的名称。	Sutton等人 ^[27]
10	替代计量学将重构影响力的概念,正在挑战影响因子的地位。	Neylon等人 ^[28]
11	替代计量学还没有被广泛接受的定义。	Holmberg等人 ^[29]
12	替代计量学就是通过多种渠道获取学术成果的相关数据,包括主流新闻、博客、社交媒体、参考文献管理网站、同行评议网站和政策文件等。	Liu ^[30]
13	替代计量学是相对文献计量学而言的,是基于网络的指标,揭示文献和其他学术成果的影响力。	Bormann ^[31]
14	替代计量学是研究学术成果网络影响力的指标,并且侧重以社交媒体作为数据源。	Shema等人 ^[32]
15	替代计量学的“替代”是相对传统同行评议和文献计量学而言的,替代计量学来自两个背景,分别是论文层次计量和替代性计量指标,所以既可以用于反映单篇论文的影响力,也可以在更广泛的情境下替代传统评价方法,而替代计量学研究网络环境下的学术成果,应当视为网络计量学的子集。	Tammaro ^[33]

从表1可以看出,在这一阶段,学者对替代计量学内涵的认识不尽相同,甚至同一学者在不同的论文中,对替代计量学内涵的理解都在发生变化。归纳起来,这些理解主要在研究对象范围的界定上存在差异,最明确而狭义的理

解是“替代计量学就是社交媒体计量”,最笼统而广义的理解是“替代计量学结合传统文献计量指标与在线新指标,给出学术成果全面的、有影响力的替代性测度方案”,这些理解是学者根据研究需要做的设定,同时,这种广泛讨论也为

统一对替代计量学内涵的认识奠定了基础。

第三阶段,替代计量学界逐渐达成共识。替代计量的研究经过沉淀,内涵日益清晰,形成广义和狭义的区别,其学术地位得以确立,吸引越来越多的学者对此进行研究。国际学术界对替代计量学的内涵基本达成共识。广义的替代计量学强调研究视角的变化,旨在用面向学术成果全面影响力的评价指标体系,替代传统的片面依靠引文指标的定量科研评价体系,同时促进开放科学和在线科学交流的全面发展;狭义的替代计量学专门研究相对传统引文指标的在线新型计量指标及其应用,尤其重视基于社交网络数据的计量指标。所以说,替代计量学并非对既有引文指标的纯粹补充,因为替代计量指标能测度引文指标触及不到的领域,例如数据集的复用率、学术视频的影响力、学术博客的社会影响力等;替代计量学也并非全盘否定基于引文的传统指标,它要替代的是唯引文至上的学术评价体系,所以可以看到 Plum X 对替代计量指标的分类中,引文仍然是重要的一大类。替代计量学的内涵得到明确后,更多的精力将转向相关的深层研究。诚如 Priem^[34]所言,现在果树低处的果实快被摘完了,还有大量高处的果实等待采摘。

2 替代计量指标的内涵

与替代计量学内涵的演变相同,对替代计量指标内涵的认识,也经历了逐步深化的过程。最初,替代计量学者预期替代计量指标和传统引文指标呈正相关性,计划效仿加菲尔德(Eugene Garfield)建立引文指标的权威性那样,建立替代计量指标的权威性。引文指标的权威性建立在加菲尔德的研究团队进行的大量实证研究基础上,其主要方法是将引文分析结果和各种主流同行评议结果做关联分析,发现引文指标与同行评议结果呈高度正相关。加菲尔德领导的美国科技信息研究所甚至据此预测诺贝尔奖得主,准确率很高。于是,人们逐渐利用基

于引文指标的评价结果来近似替代耗时耗力的同行评议。久而久之,形成了今天以引文指标为黄金法则(Golden Rules)的局面。沿着相似的思路,替代计量学者展开了广泛研究,逐渐揭示了替代计量指标的内涵。

(1)研究表明,部分替代计量指标不能反映学术影响力。替代计量学者以引文指标为参照,做了大量关联分析,得到的结论却各不相同,仅以 Twitter 提及量(Twitter Mentions)指标为例,既有 Twitter 提及量与传统引文指标呈正相关的研究结果。例如,Shuai 等人^[35]发现预印本论文的 Twitter 提及量与早期引文量显著相关,Eysenbach 等人^[36]计算得出 Twitter 提及量和引文指标有中等到显著的正相关关系,皮尔逊系数数值在 0.42 到 0.72 之间。也有研究得出两者关联度不高的研究结论,例如,Priem 等人^[37]做的实证分析得出 Twitter 提及量与引文量的相关程度仅为 0.1,Thelwall 等人^[38]的分析结果表明两者之间甚至呈负相关。不过,针对 Twitter 提及量这一指标而言,目前已经有了定论,Haustein 等人^[39]的大规模研究证明, Twitter 提及量和后期引文量没有任何关系,从而彻底结束了这场辩论,其研究成果被 *Nature* 网站报道^[40]。另外两个与引文指标做关联分析最多的替代计量指标是,博文提及量(Blog Counts)与在线参考文献管理网站标签量(Bookmark Counts of Online Reference Managers),Bormann^[41]的一项元分析表明博文提及量与传统引文量关联度只有 0.12,而在线参考文献管理网站标签量的关联度则在 0.23 到 0.51 之间。此外,还有将下载量、阅读量等替代计量指标与引文量做关联分析的研究^[42-43]。这些研究的共同结论是,替代计量指标能反映与引文指标不同的内涵,所以在评价时要结合起来使用。

(2)替代计量指标反映可见度、知名度和社会影响力。Altmetric.com 创始人 Euan^[40]表示对 Haustein 等人的研究结论并不意外,认为替代计量指标反映知名度,是网络上对科研成果的关注度。Altmetric.com 针对每个科研成果发布

衡量其受关注程度的“Altmetric 指数”，并早在网站公告中声明该指数的正确用法是，避免根据该指数来评价科研成果质量或对科研成果进行排序，Altmetric 指数仅反映科研成果在线受关注程度，研究表明 Altmetric 指数高的论文被引量不一定高，反之亦然^[44]。Altmetric.com 每年还会推出最受社交媒体关注的前 100 名论文，受到各界普遍关注。一般而言，Altmetric 指数最高的论文趣味性较强，易理解并与公众切身利益相关^[45-46]。Bormann^[47]则另辟蹊径，结合基金会日益要求科研人员证明其成果社会影响力的背景，认为替代计量指标能覆盖更加广泛的科研成果使用者（不仅限于读者），反映社会影响力，为解决衡量科研成果社会影响力的困难提供了一种简便的方法。

(3) 相比而言，基于学术界数据源的替代计量指标更能反映学术影响力。替代计量指标是否真的和引文指标没有关联，不能反映学术影响力呢？这种论断无疑是片面的，因为替代计量指标具有多样性。Haustein 等人^[39]的研究仅证明了单纯的 Twitter 提及量不能反映学术影响力，事实上，许多深层次因素在该研究中没有被考虑进去，例如 Twitter 的用户群体，以及 Twitter 提及的情境和向引文的转化率等因素。在线参考文献网站标签量与引文指标相关性更高，其内在原因之一是 Mendeley、CiteULike 等工具的使用者大多为学者，数据主要源于学术交流，学术影响力理所应当要限定在学术界用户，因为学术界之外的用户无法贡献引文，所以数据源为学术界的替代计量指标与反映学术影响力的引文指标的相关性更高。此外，还有许多其他替代计量指标与引文指标的关联性有待检验，例如对学术成果的评级（打分）、点赞量等。Beltrao^[48]曾将 Connotea 中标签为“Evolution”的论文抽取出来，结果发现这些论文的平均引文量比 *Nature* 和 *Science* 等顶级期刊论文的还要高。所以，在考察替代计量指标与学术影响力关系时，要更加关注学术界的替代计量数据源。

综上所述，替代计量指标反映科研成果获

得在线关注的强度，能反映社会影响力，若限定在学术界的数据源，若干指标如标签量、评级等也能反映学术影响力，关键要根据应用情境进行细分。

3 替代计量指标的可信度

除了揭示替代计量指标的内涵，保证指标的可信度对于建立威信同样重要。替代计量指标的内涵说明其有用性，而可信度则说明其可用性，可信度（或者健壮性）是作为指标的必备条件。替代计量指标受到的质疑可以归纳为三个方面。①替代计量指标数据源存在被规模化操控的风险。网络公关公司和水军的存在，让基于网络数据的指标天然受到更多的质疑，而替代计量学的很多指标正是基于网络数据，这些网络数据存在由机器大规模自动生成的可能，丧失了原本应有的含义。此外，替代计量指标具备透明性，更容易受到黑客攻击，例如一旦 PLOS 完全公布数据采集算法，将面临数据被篡改的风险。②替代计量指标数据源的科学严谨性受到质疑。科研活动的整个过程都要求严谨，网络尤其是社交网络被赋予的娱乐化印象，使其可信度大打折扣。例如对学术研究知之甚少的公众可能凭直觉对科研成果进行评价和传播，造成“民主暴力”和谬传。③替代计量指标的部分数据源无法回溯。例如，微博被作者删除的现象十分普遍，后期难以回溯和查证；碎片化的成果难以管理，例如数据集（Dataset）会被分割成若干单元，分布在不同的应用里，难以追踪。

尽管替代计量指标的可信度受到上述三个方面的质疑，但是质疑者存在的最大误区是，无形中夸大了保障替代计量指标可信度的困难。必须明确的是，这些威胁能够消除，并且，从替代计量学研究中获得的价值，要远大于消除这些威胁的成本。下面分三个维度进行阐述。

(1) 替代计量指标的规模化操控风险完全可以规避。首先，替代计量指标的数据源，是主流媒体和平台，这些实体本身有数据质量保障

机制;其次,在汇聚数据时,替代计量指标集成者有数据一致性检验机制。Lin^[20]介绍了 PLOS 的数据可信度控制机制,在政策方面定义了四种类型的数据行为,在技术方面开发了数据审计系统 DataTrust,能监测数据异常变动。需要指出的是,谷歌公司、信用卡公司、银行等都面临着强有力的规模化操控的风险,但是这些公司已经有了充足的经验和坚实的技术支撑,数据质量能够得到充分保障。替代计量指标集成者一方面可以借鉴和利用这些经验与技术,另一方面替代计量指标面临的操控动机要远小于上述实体。所以说,替代计量指标规模化操控的风险可以规避。Priem^[13]认为,替代计量指标并不像表面上那样容易被操纵,这得益于替代计量指标的多样性和大规模特性,操纵某单项指标的数值难度可能不大,但要连动操纵系列指标所耗费的代价就相当昂贵。Fenner 等人^[19]比较了 PLOS、Altmetric.com 和 Mendeley 三个替代计量指标提供方的数据一致性,发现对于相同数据集,这三者提供的数据存在较大差异,并从三者数据收集方法上做了解释。

(2) 替代计量指标反映了使用者的价值判断。虽然大部分替代计量数据是在非正式学术交流中产生的,但是使用者生产替代计量数据存在生成成本(Generation Cost),包括时间成本、精力成本和声誉成本,使用者的价值判断正是反映在这些成本里。不同层次替代计量指标的生成成本是不同的。例如:相对浏览摘要,阅读全文的时间成本和精力成本要更高;相比个人收藏,推荐到好友圈或进行分享的声誉成本要更大,因为推荐或分享的科研成果往往被认为反映该使用者的水准。同时,不论哪种替代计量指标,其对象都是科研成果,很少具备娱乐价值,严肃性有所保证。在科研成果的传播上,学术界始终拥有高度自治权,不受公众参与科研成果传播的影响,与此同时,公众的教育和知识水平不断提高,他们既有利用科研成果的权利,也有促进科研进步的能力。

(3) 替代计量基础设施的建立让相关原始

数据得到较好保存和复用。在网络数据保存方面,维基百科、ArXiv 等组织积累了经验,可以实现数据的版本管理和长期保存。ImpactStory、Altmetric.com 等机构均致力于替代计量原始数据的积累和保存,凡是涉及替代计量指标的原始数据,都有望得到备份,作为替代计量数据库中的一部分,即使作者删除原文,仍然可以回溯,因为仅限科研用途,不会侵犯用户隐私。科研成果管理也逐渐规范化,例如 DOI 等组织^[49]的设立,让任意碎片化的成果都可以被引用。

虽然包括引文指标、链接指标等在内的所有指标,在其发展初期都受到可信度的质疑,替代计量指标也不能例外,但是网络巨头公司业已积累的丰富经验,能被充分借鉴,用于保证替代计量的数据质量,而替代计量指标的多维属性,也保证其免受人为操控,从而最大限度地确立替代计量指标的可信度。此外,还能通过使人为操控成本大于操控者的收益,来规避人为操控行为。总而言之,替代计量指标在实际应用中具备可信度。

4 替代计量学的其他问题

除了上述最受关注的核心问题外,替代计量研究是多角度的,笔者认为以下两个方面值得关注。

第一,替代计量学的学科外意义。诚如 Alperin^[50]所言,传统基于引文的学术影响力评价体系严重偏向北美和欧洲,对发展中国家不利,因为现有顶尖题录数据库(如 WoS、Scopus)都偏袒欧美期刊,以及和这些地区密切关联的研究主题,发展中国家的优秀科研成果甚至要付费才能纳入这种体系,而且其所有权还归发达国家所有。替代计量指标能更准确地衡量学术成果的影响力,并培养为本国发展目标服务的科研文化,让长期被主流平台屏蔽的发展中国家的期刊和学者的研究成果,在国际舞台上得到更好的展示,在繁荣科学交流的同时,扩大发展中国家科学家的话语权。Alperin^[51]还研究了

发展中国家情境下期刊与论文的替代计量指标,地区不限于北美,语言也不限于英语,指出了其与发达国家情境下研究结果的差异。

第二,替代计量新数据源不断涌现。替代计量学的提出,切实为许多新生科研交流模式和科研交流平台提供了契机。任何可能的科学交流创新活动都得到关注,既有的特色在线科学交流模式都可被挖掘,激发了学者的创造力,这一切部分归功于替代计量学这个学术平台的出现。例如, Lin 等人^[52]研究了维基百科的引文; Kaur 等人^[53-54]研究了 Scholarometer 的学术分类模式,提出了众源学科; Kousha 等人研究了学者网络简历链接分享对学术交流的作用^[55],还研究了 Youtube 视频^[56]在学术交流和教学中的作用,对被学术成果引用最多的视频做了内容分析; Thelwall 等人^[57]比较了基于 ResearchGate 使用与出版数据的大学排名,发现和已有的学术排名高度相关; Shema 等人^[58-59]研究了 Research Blogging.org 博客引文用于衡量学术影响力的潜在价值。此外,还有 Zotero、CiteULike、Academia.edu、Lnyrd.com、Slideshare、GitHub 等众多新数据源。中国也出现了数据堂、新科学网、科学网、万方学术圈等数据源,尽管其他社会网络用于学术用途的情况还有待调研,但是完全可以大胆地去探索和尝试。

5 结论与讨论

替代计量学在中国的研究刚开始兴起,厘清涉及替代计量学研究的核心内涵问题,不仅对相关研究人员来说十分必要,对领域外众多利益相关者理解替代计量学的研究内容、思路和目标,从而支持和利用替代计量学研究成果也至关重要。本文的主要结论有三点。

(1) 替代计量学有狭义和广义之分,狭义的替代计量学专门研究相对传统引文指标的在线新型计量指标及其应用,尤其重视基于社交网络数据的计量指标,广义的替代计量学研究新型在线科学交流体系,强调研究视角的变化,旨

在利用面向学术成果的全面影响力评价指标体系,替代传统片面依靠引文指标的定量科研评价体系。所以说,替代计量学并非对既有引文指标的纯粹补充,也并非全盘否定传统的引文指标。国内对“Altmetrics”的翻译曾经产生过困惑,主要因为“Alternative”这个词的意思很微妙,译为“选择性”、“综合”则不知所云,译为“社媒影响”、“网媒影响”则过于狭隘,译为“另类”在大陆语境有贬义,译为“新型”则与原词偏离较远,有学者将其意译为“补充”,实则其内涵并非“补充”。尽管 Rousseau 教授^[60]撰文认为这个术语提得不好,因为出现下一代指标就会存在命名困难,但是从文中可知他也是理解为“替代”的。其实,医学领域早在 20 世纪 70 年就有了替代医学(Alternative Medicine)和补充医学(Complementary Medicine)的先例,可见“Alternative”这个词的翻译并没有想象中那样复杂,“替代”是比较合适的译名。

(2) 替代计量指标直观上反映科研成果获得的在线关注强度,若根据应用情境对替代计量指标和数据源进行细分,既能反映学术成果的社会影响力,也能反映学术成果的学术影响力。例如,学术成果的微博提及量、新闻报道量等可以反映社会影响力,而学术界在线参考文献管理平台中的阅读量、评级和标签等可以反映其学术影响力。同时也需注意,替代计量指标可以反映社会影响力绝不意味着它是衡量社会影响力的唯一标准,社会影响力的体现形式多样,除了在网络中得以体现之外,还有科技成果转化、开设讲座、项目合作等许多线下形式,当然,这些线下形式也会逐渐在网络上有迹可寻。

(3) 替代计量指标是可信的,通过充分的政策保障和技术支撑,替代计量数据将得到严格的审计、保存和利用。互联网公司(如谷歌)对数据控制的丰富经验,以及替代计量指标自身的多样性和大规模特性,都有力支撑了替代计量指标的可信度。同时,笔者认为很多时候,替代计量数据的情境对用户更加重要。例如,同

样是被微博提及,该科研成果是被谁的微博提及和以怎样的方式被提及,比起单纯的数字更加有价值,学者将对个人学术成果被谁利用、怎样被利用有更好的认识,这样的价值同样可以被传递给基金决策者和政策制定者。

替代计量学蓬勃发展之际,有学者认为它是让发展中国家科学家获得平等学术地位的契机,在线科学交流让发展中国家有同样的机会获得学术成果的可见度。但是,我们也要注意,当前国际主流科学交流平台都是发达国家

在开发和维护,而不同国家之间网络并非完全畅通,这是现实中要考虑的因素。此外,除了发现新的替代计量数据源,也要积极开发具备创新科学交流模式的新数据源,在这个过程中,不应过分夸大语言的障碍。多语言平台技术已经比较成熟,例如腾讯公司的微信在国外同样有市场。我国学者应当切实利用好替代计量学这个学术平台,创新和改革中国的科学交流与科研定量评价体系。

参考文献

- [1] 刘春丽. Web 2.0环境下的科学计量学:选择性计量学[J].图书情报工作,2012(14):52-56, 92. (Liu Chunli. Scientometrics in the Web 2.0 environment: Altmetrics[J]. Library and Information Service, 2012(14): 52-56, 92.)
- [2] 邱均平, 余厚强. 替代计量学的提出过程与研究进展[J].图书情报工作,2013(19):5-12. (Qiu Junping, Yu Houqiang. The putting forward process and research progress of Altmetrics[J]. Library and Information Service, 2013(19):5-12.)
- [3] 由庆斌, 汤珊红. 补充计量学及应用前景[J].情报理论与实践,2013(12):6-10. (You Qingbin, Tang Shan-hong. Altmetrics and its potential applications[J]. Information Studies: Theory & Application, 2013(12):6-10.)
- [4] 崔宇红. 从文献计量学到 Altmetrics: 基于社会网络的学术影响力评价研究[J].情报理论与实践,2013(12):17-20. (Cui Yuhong. From bibliometrics to altmetrics: evaluation research on academic impact based on social network[J]. Information Studies: Theory & Application, 2013(12):17-20.)
- [5] 陈铭. 期刊利用统计与 Altmetrics 的兴起[J].图书与情报,2014(1):12-17. (Chen Ming. The utilization statistics of journals and the rise of Altmetrics[J]. Library and Information, 2014(1):12-17.)
- [6] 张侃, 赵文华, 孙保存. 从编辑的视点探讨科技期刊 Altmetrics 的重要性[J].编辑之友,2013(9):41-43. (Zhang Min, Zhao Wenhua, Sun Baocun. Investigating the value of Altmetrics in scientific journals from perspective of editors[J]. Editorial Friend, 2013(9):41-43.)
- [7] 宋丽萍, 王建芳. 基于 F1000 与 WoS 的同行评议与文献计量相关性研究[J].中国图书馆学报,2012(2):62-69. (Song Liping, Wang Jianfang. The correlation between peer review and bibliometric analysis in evaluating scientific publication outputs: a case study of F1000 and WoS[J]. Journal of Library Science in China, 2012(2): 62-69.)
- [8] 刘春丽, 何钦成. 不同类型选择性计量指标评价论文相关性研究——基于 Mendeley、F1000 和 Google Scholar 三种学术社交网络工具[J].情报学报,2013,32(2):206-212. (Liu Chunli, He Qingcheng. Study on correlation of different altmetrics indicators for paper evaluation based on three academic social networking tools: mendeley, F1000 and google scholar[J]. Journal of the China Society for Scientific and Technical Information, 2013, 32(2):206-212.)
- [9] 宋丽萍, 王建芳, 王树义. 科学评价视角下 F1000、Mendeley 与传统文献计量指标的比较[J].中国图书馆学报,2014(4):48-54. (Song Liping, Wang Jianfang, Wang Shuyi. Observation of F1000, Mendeley and traditional bibliometric indicators from the perspective of scientific assessment[J]. Journal of Library Science in

China, 2014(4):48-54.)

- [10] 余厚强, 邱均平. 替代计量学视角下的在线科学交流新模式[J]. 图书情报工作, 2014(15):42-47. (Yu Houqiang, Qiu Junping. Novel pattern of online scientific communication from perspective of altmetrics[J]. Library and Information Service, 2014(15):42-47.)
- [11] Adie E. Taking the alternative mainstream[J]. El Profesional De La Información, 2014, 23(4):349-351.
- [12] 余厚强, 邱均平. 替代计量指标分层与聚合的理论研究[J]. 图书馆杂志, 2014(10):13-19. (Yu Houqiang, Qiu Junping. Theoretical research on stratifying and aggregating altmetric indicators[J]. Journal of Library, 2014(10):13-19.)
- [13] Priem J. Scholarship: beyond the paper[J]. Nature, 2013, 495(7442):437-440.
- [14] Priem J, Taraborelli D, Groth P, et al. Altmetrics a manifesto[EB/OL]. [2014-08-20]. <http://altmetrics.org/manifesto>.
- [15] Priem J, Costello K L. How and why scholars cite on twitter[J]. Proceedings of the American Society for Information Science and Technology, 2010, 47(1):1-4.
- [16] Garfield E, Merton R K. Citation indexing: its theory and application in science, technology, and humanities[M]. Wiley New York, 1979.
- [17] NISO. NISO altmetrics standards project white paper[EB/OL]. [2014-09-19]. http://www.niso.org/apps/group_public/download.php/13295/niso_altmetrics_white_paper_draft_v4.pdf.
- [18] Hausteijn S, Peters I, Bar-Ilan J, et al. Coverage and adoption of altmetrics sources in the bibliometric community[J]. Scientometrics, 2014(1):1-19.
- [19] Fenner M, Costas R, Zahedi Z. How consistent are altmetrics providers: study of 1000 PLOS ONE publications using the PLOS ALM, Mendeley and Altmetric.com APIs[EB/OL]. [2014-09-14]. http://figshare.com/articles/How_consistent_are_altmetrics_providers_Study_of_1000_PLOS_ONE_publications_using_the_PLOS_ALM_Mendeley_and_Altmetric_com_APIs/1041821.
- [20] Lin J. A case study in anti-gaming mechanisms for altmetrics: PLOS ALMs and DataTrust[EB/OL]. [2012-07-21]. <http://altmetrics.org/altmetrics12/lin>.
- [21] Lin J, Fenner M. Altmetrics in evolution: defining & redefining the ontology of article-level metrics[J]. Information Standards Quarterly, 2013, 25(2):20.
- [22] Roemer R C, Borchardt R. From bibliometrics to altmetrics a changing scholarly landscape[J]. College & Research Libraries News, 2012, 73(10):596-600.
- [23] Konkiel S, Scherer D. New opportunities for repositories in the age of altmetrics[J]. Bulletin of the American Society for Information Science and Technology, 2013, 39(4):22-26.
- [24] Lapinski S, Piwowar H, Priem J. Riding the crest of the altmetrics wave how librarians can help prepare faculty for the next generation of research impact metrics[J]. College & Research Libraries News, 2013, 74(6):292-300.
- [25] Mounce R. Open access and altmetrics: distinct but complementary[J]. Bulletin of the American Society for Information Science and Technology, 2013, 39(4):14-17.
- [26] Rasmussen P G, Andersen J P. Altmetrics: an alternate perspective on research evaluation[J]. Sciecom Info, 2013, 9(2):28-39.
- [27] Sutton S W. Altmetrics: what good are they to academic libraries?[J]. Kansas Library Association College and University Libraries Section Proceedings, 2014, 4(2):1-11.
- [28] Neylon C, Willmers M, King T. Impact beyond citation: an introduction to altmetrics[J]. Scholarly Communication in Africa Programme, 2014(1):10-17.
- [29] Holmberg K. The meaning of altmetrics[C/OL]. Proceedings of the IATUL Conferences. 2014. Paper 1. <http://>

docs.lib.purdue.edu/iatul/2014/altmetrics/1.

- [30] Liu J. Bringing altmetrics to institutions [EB/OL]. [2014-09-20]. http://libraryconnect.elsevier.com/sites/default/files/Jean_Liu_June_2014.pdf.
- [31] Bornmann L. Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics [OL]. [2014-09-15]. <http://arxiv.org/abs/1406.7091>.
- [32] Shema H, Bar-ilan J, Thelwall M. Do blog citations correlate with a higher number of future citations? Research blogs as a potential source for alternative metrics [J]. *Journal of the Association for Information Science and Technology*, 2014, 65(5): 1018-1027.
- [33] Tammaro A M. Altmetrics in the humanities: perceptions of italian scholars [J]. *Libraries in the Digital Age (LIDA) Proceedings*, 2014, 13(2): 13-25.
- [34] Priem J, Piwowar H. Top 5 altmetrics trends to watch in 2014 [EB/OL]. [2014-10-08]. <http://blog.impactstory.org/top-5-altmetrics-trends-to-watch-2014>.
- [35] 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.
- [36] Eysenbach G. Can tweets predict citations? Metrics of social impact based on twitter and correlation with traditional metrics of scientific impact [J]. *Journal of medical Internet research*, 2011, 13(4): e123.
- [37] Priem J, Piwowar H A, Hemminger B M. Altmetrics in the wild: using social media to explore scholarly impact [OL]. [2014-09-17]. <http://arxiv.org/abs/1203.4745>.
- [38] 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.
- [39] 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.
- [40] Noorden R V. Twitter buzz about papers does not mean citations later [EB/OL]. [2014-08-19]. <http://www.nature.com/news/twitter-buzz-about-papers-does-not-mean-citations-later-1.14354>.
- [41] Bornmann L. Alternative metrics in scientometrics: a meta-analysis of research into three altmetrics [OL]. [2014-09-20]. <http://arxiv.org/abs/1407.8010>.
- [42] Bazrafshan A, Haghdoost A A, Zare M. A comparison of downloads, readership and citations data for the journal of medical hypotheses and ideas [J]. *Journal of Medical Hypotheses and Ideas*, 2015, 9(1): 1-4.
- [43] Schloegl C, Gorraiz J, Gumpenberger C, et al. Are downloads and readership data a substitute for citations? The case of a scholarly journal [J]. *Libraries in the Digital Age (LIDA) Proceedings*, 2014, 13(2): 38-46.
- [44] Costas R, Zahedi Z, Wouters P. Do altmetrics correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective [OL]. [2014-09-22]. <http://arxiv.org/abs/1401.4321>.
- [45] Bornmann L. Which kind of papers has higher or lower altmetric counts? A study using article-level metrics from PLOS and F1000Prime [OL]. [2014-09-25]. <http://arxiv.org/abs/1409.2863>.
- [46] Costas R, Zahedi Z, Wouters P F. Disentangling the meaning of 'altmetrics': content analysis of Web of Science scientific publications [EB/OL]. [2014-08-29]. http://figshare.com/articles/Disentangling_the_meaning_of_altmetrics_content_analysis_of_Web_of_Science_scientific_publications/1041770.
- [47] Bornmann L. Validity of altmetrics data for measuring societal impact: A study using data from Altmetric and F1000Prime [OL]. [2014-09-25]. <http://arxiv.org/abs/1406.7611>.
- [48] Beltrao P. Connotea tag: evolution citation report [EB/OL]. [2014-08-30]. <http://www.evocellnet.com/2006/11/connotea-tagevolution-citation-report.html>.
- [49] Bando K. Measuring impact of integration of orcid and altmetrics into open access repository upon scholarly commu-

- nication and research impact[EB/OL].[2014-09-03].http://figshare.com/articles/Measuring_Impact_of_Integration_of_ORCID_and_altmetrics_into_Open_Access_Repository_upon_Scholarly_communication_and_Research_Impact/1052995.
- [50] Alperin J P. Ask not what altmetrics can do for you, but what altmetrics can do for developing countries[J]. Bulletin of the American Society for Information Science and Technology, 2013, 39(4):18-21.
- [51] Alperin J P. Exploring altmetrics in an emerging country context[EB/OL].[2014-07-15].http://figshare.com/articles/Exploring_altmetrics_in_an_emerging_country_context/1041797.
- [52] Fenner M, Lin J. An analysis of Wikipedia references across PLOS publications[EB/OL].[2014-07-15].http://figshare.com/articles/An_analysis_of_Wikipedia_references_across_PLOS_publications/1048991.
- [53] Kaur J, JafariAsbagh M, Radicchi F, et al. Crowdsourced disciplines and universal impact[EB/OL].[2014-07-20].http://figshare.com/articles/Crowdsourced_disciplines_and_universal_impact/1037741.
- [54] Kaur J, Hoang D T, Sun X, et al. Scholarometer: a social framework for analyzing impact across disciplines[J]. PLOS ONE, 2012, 7(9):e43235.
- [55] Kousha K, Thelwall M. Disseminating research with web CV hyperlinks[J]. Journal of the Association for Information Science and Technology, 2014, 65(8):1615-1626.
- [56] Kousha K, Thelwall M, Abdoli M. The role of online videos in research communication: a content analysis of youtube videos cited in academic publications[J]. Journal of the American Society for Information Science and Technology, 2012, 63(9):1710-1727.
- [57] Thelwall M, Kousha K. ResearchGate: disseminating, communicating, and measuring scholarship?[J]. Journal of the Association for Information Science and Technology, 2014(1):1-14.
- [58] Shema H, Bar-Ilan J, Thelwall M. Research blogs and the discussion of scholarly information[J]. PLOS one, 2012, 7(5):e35869.
- [59] Shema H, Bar-Ilan J, Thelwall M. How is research blogged? a content analysis approach[J]. Journal of the Association for Information Science and Technology, 2014(1):25-39.
- [60] Rousseau R, Ye F Y. A multi-metric approach for research evaluation[J]. Chinese Science Bulletin, 2013(26):3288-3290.

邱均平 武汉大学中国科学评价研究中心主任, 博士生导师。

通信地址: 湖北省武汉市珞珈山。邮编: 430072。

余厚强 武汉大学信息管理学院博士研究生。通信地址同上。

(收稿日期: 2014-09-20; 修回日期: 2014-10-28)