

开放科学环境下负责任评价实现路径研究

宋丽萍 王建芳

摘 要 开放科学的发展正在改变学术交流生态,同时对传统科研评价产生挑战,对于科研评价创新的呼声也日益高涨。负责任评价(RRA)应时而生并成为开放科学政策框架的重要组成部分,而如何实现 RRA 成为关键问题。本文解析开放科学环境下的 RRA 内涵,总结国内外相关理论与实践进展,在此基础上尝试探索 RRA 的实现路径,以期“破四唯”指导下的我国科研评价提供借鉴与参考。以典型内容文本为数据源,通过关键词提取和共现分析得出结论,即以研究质量和影响力为主维度,以知情同行评议为主基调,构成国际 RRA 改革的主要思路。在分析开放科学背景下我国学术交流新业态与 RRA 的基础上,提出我国 RRA 的三层结构框架。其中以识别科研新范式为情境,以改善宏观政策为支撑,以改革评价方式为实操,以评价主体多元化、评价内容多维化、评价对象精准化、评价方法循证化为路径,建立有利于潜心研究和创新的我国科研评价制度,并回答科研评价改革的必答题:“谁来评”“评什么”“如何评”“怎么用”。研究认为,建立质量意识、培育“影响素养”、达成评价系统变革的国际共识,从而提高 RRA 的互操作性和可移植性,避免相关理论与实践的碎片化是保证 RRA 顺利实施的关键;将负责任评价理念与实现路径本土化,营造有利于开放科学发展的学术交流环境,将成为我国图书情报工作的新使命。图 2。参考文献 43。

关键词 开放科学 负责任评价 研究质量 社会影响力

分类号 G311

The Realization Path of Responsible Research Assessment in the Context of Open Science

SONG Liping, WANG Jianfang

ABSTRACT

The development of open science is changing the ecology of scholarly communication as well as challenging the traditional research evaluation. Therefore, a call for innovation in research assessment is growing. Responsible Research Assessment (RRA) emerges and becomes an important part of the open science policy framework. However, how to realize RRA has become the key issue. Through analyzing the fundamentals of RRA in the context of open science, and summarizing the progress of local and international theories and practices in relevant fields, this paper aims to explore the realization path of RRA to shed light on China's research evaluation under the guidance of breaking the "Siwei".

First of all, taking the typical relevant content texts as the data source, combined with the keyword extraction and co-occurrence analysis, it shows that the approach of international RRA takes influence and research quality as the main dimension, and informed peer review as the main keynote as well.

Secondly, not only adhering the concept of RRA to top-level design to embody the excellence and influence but also taking root in Chinese context and showing local characteristics, this paper puts forward

通信作者:王建芳,Email:wangjianfang@casisd.cn,ORCID:0000-0002-9950-7818(Correspondence should be addressed to WANG Jianfang,Email:wangjianfang@casisd.cn,ORCID:0000-0002-9950-7818)

the three-layer structured framework of China's RRA process path after analyzing the new situation of scholarly communication in China. The identification of the new paradigm of scientific research is employed as the research context, the improvement of macro policies (scientific research funding, management and publication policies) is defined as the research support, and the reform of the evaluation method is used as the research practice. Meanwhile, the diversification of the evaluation subject, the multi-dimensionality of the evaluation content, the precision of the evaluation object, and the evidence-based basis of the evaluation method are considered to be the paths, and then China's evaluation system, which is conducive to dedicated research and innovation, is established so as to answer to the mandatory questions of the scientific research evaluation reform: "who to evaluate", "what to evaluate", "how to evaluate" and "how to use".

This paper concludes that the reform and reconstruction of the research evaluation system is a gradual and continuous optimization process. Thus, establishing the quality awareness, cultivating "impact literacy", and investigating an international consensus on the reform of the evaluation system, thereby improving the interoperability and portability of RRA in parallel to avoid the fragmentation of related theories and practices have become the key to future development. Furthermore, it points out that localizing the concept and realization path of RRA and creating scholarly communication conditions to foster open science may become a new mission of library and information service in China. 2 figs. 43 refs.

KEY WORDS

Open science. Responsible research assessment. Research quality. Social impact.

0 引言

近二十年来开放科学迅速发展,开放论文、开放图书、开放存储、开放论坛、开放会议等新的理念与方法正在改变学术交流生态,同时对传统科研管理评价产生挑战。在学术交流、图书馆资源建设、用户推荐等实践中,对科研评价创新的呼声越来越高。负责任评价(Responsible Research Assessment, RRA)应运而生。RRA 源于欧盟“地平线 2020”计划中负责任科研与创新(Responsible Research and Innovation, RRI)理念。2015 年,英国专家小组发布《计量潮:研究评价和管理中指标作用的独立审查报告》(以下简称《计量潮报告》),将 RRI 延伸到科研评价领域进而提出负责任计量,并将其定义为研究治理、管理和评估中合理并适当使用定量指标的一种方式,同时将以精确数据为基础的稳健性、体现学科间差异的多样性、评价方法的多维性、数据收集和分析过程的透明性、不断更新的自反性界定为其基本特征^[1]。在此基础上,负责任计量拓展为更为宏观的负责任评价。全球研究理事会(Global Research Council, GRC)将负责任评价定义为:一类评估方法的总称,将反映并激励高质量研究,支持多元、包容的研究文化^[2]。

2021 年 11 月,欧盟发布《科研评价改革指南》,将科研评价改革列入优先发展领域,并指出其与开放科学原则的一致性^[3]。至此,RRA 成为开放科学的支点之一。然而,目前对于 RRA 如何实现莫衷一是。本文在总结国内外相关理论研究与实践进展的基础上,尝试探索开放科学环境下 RRA 的实现路径,以期为我国科研评价实践提供借鉴与参考。

1 开放科学环境下 RRA 的内涵

2021 年,联合国教科文组织在《开放科学建议书》中提出了开放科学架构,特别包括了向传统科学界以外的社会行为者开放科学知识的创造、评估和传播进程,并强调营造有利于开放科学的政策环境,其中包括鼓励负责任的研究及研究人员评价^[4]。因此,可以认为 RRA 是开放科学政策框架中的重要组成部分。2012—2015 年先后形成的《关于科研评价的旧金山宣言》(以下简称《旧金山宣言》)、《科学评价十原则——莱顿宣言》(以下简称《莱顿宣言》)和《计量潮报告》共同构筑了负责任计量与评价的主体框架^[5]。2017 年,欧盟委员会和加拿大人文与社会科学联合会分别发布《开放科学背景下的新一代计量》和《关于人文社会科学评价指标的建议》,提出负责任使用指标的数据基础架构^[6,7]。上述行动响应了开放科学时代科学范式转变的需求,诠释了 RRA 评价目的与评价内容的新内涵。

首先,RRA 强调研究的卓越性与影响力。在 RRA 发展进程中,评价目的逐渐明朗化,即体现研究的卓越性与影响力,激励并认可高质量研究,同时进一步加强社会对研究和创新系统及其产出的信任,而“责任”成为其中支点。《计量潮报告》援引英国帝国理工学院 Grimm 事件指出,“错误的评价标准扭曲了我们的行为,决定了我们的职业生涯……”,由此强调评价主体的责任^[1]。同时鉴于科学研究的宗旨,即在于最大限度地弥合与社会需求之间的距离,研究主体(机构、项目、研究人员)需保证研究的卓越性,因此,作为 RRA 评价对象的研究主体同样蕴含着责任要素。从这一意义上,本文将 RRA 定义为:评价主体以反映并激励高质量研究为目标,切实反映研究主体能力和贡献、研究过程和研究成果质量及价值的评价方法的总称。

其次,RRA 体现了多元化要素。2018 年,科研管理学会国际联盟提出 RRA 的 SCOPE(Start-Context-Options-Probe-Evaluation)原则,即“目标、情境、方法、调查、元评价(对评价进行评价)”五步法^[8];2021 年 5 月,GRC 发布报告,明晰了 RRA 的全球化合作及其与研究文化、研究生态之间的关系^[9];2021 年 11 月,欧盟在《科研评价改革指南》中强调了科研评价改革的国际一致性,并将研究与创新、研究文化、研究生态纳入 RRA 框架中^[3]。因此,以纠偏期刊影响因子(JIF)滥用为起点的负责任计量逐渐拓展为 RRA,而研究与创新、研究质量、社会影响力、研究诚信、研究文化、研究生态成为 RRA 的要素。

最后,开放科学为 RRA 提出新的情境与方向。依据 GRC 的观点,对环境的敏感性,即对于科学研究过程及其方式的响应是 RRA 的重要特征^[2]。目前,研究与创新系统正经历重要变革,以研究过程的数字化、研究方式的协作与开放化、研究产出的多样化为主要特征。上述变化致使基于 JIF 的狭隘评价模式与科学研究的不匹配性更为凸显。为此,西方七国集团将研究评估改革列入《开放科学合作契约》^[10];欧盟开发了开放科学监测器(Open Science Monitor)以评估科学研究的质量、贡献及其影响,并实施了开放科学背景下的多维职业评估框架(OS-CAM)^[11];美国国家科学院将开放科学实践纳入评价体系^[2];英国卓越研究框架(Research Excellence Framework, REF)2021 包含了跨学科研究评审小组,并将数据、展览纳入研究成果的范畴^[12]。可见,国际 RRA 的进路在将研究的卓越性与影响力划为重点的同时,以科学研究的新范式为目标进行校准,从而实现与开放科学、跨学科研究相匹配。

2 开放科学环境下 RRA 的理论与实践

在 RRA 理念与框架引导下,国际上对开放科学环境下的机构、项目和人员评价进行了多方面理论研究,同时开展了相关实践探索。

2.1 国际 RRA 改革的主要思路

为理清国际 RRA 研究和实践的发展脉络和方向,本研究对典型内容文本进行了关键词提取和共现分析,结果显示:以研究质量和影响力为主维度,以知情同行评议(informed peer review)为主基调,共同构成国际 RRA 改革的主要思路。本研究选取的案例样本包括:《2021 年 RRA 报告》^[9]、荷兰《战略评估协议 2021—2027》(Strategy Evaluation Protocol 2021–2027, SEP 2021–2027)^[13]、英国 REF 2021^[12]、《科研人员评价香港原则:促进研究诚信》(以下简称《科研人员评价香港原则》)^[14]等。首先,利用自然语言处理开源软件 AntConc^①,以英国国家语料库为参考,通过 N-GRAM 模型设置词组长度的,抽取其中关键词及词丛并体现文档的内容特征;继而,经筛选、清洗,利用科学文献可视化分析软件 VOSviewer 生成上述关键词共现关系可视化图谱(见图 1),以呈现各国 RRA 实施路径的共性及重心;最后,经综合分析认为,RRA 发展进程中的重点要素如下:开放科学(open science)、研究质量(research quality)、社会影响力(social impact)、研究诚信(research integrity)、研究文化(research culture)、团队合作(research groups)、案例研究(case studies)、叙述式简历(narrative CV)、自我评价(self assessment)。

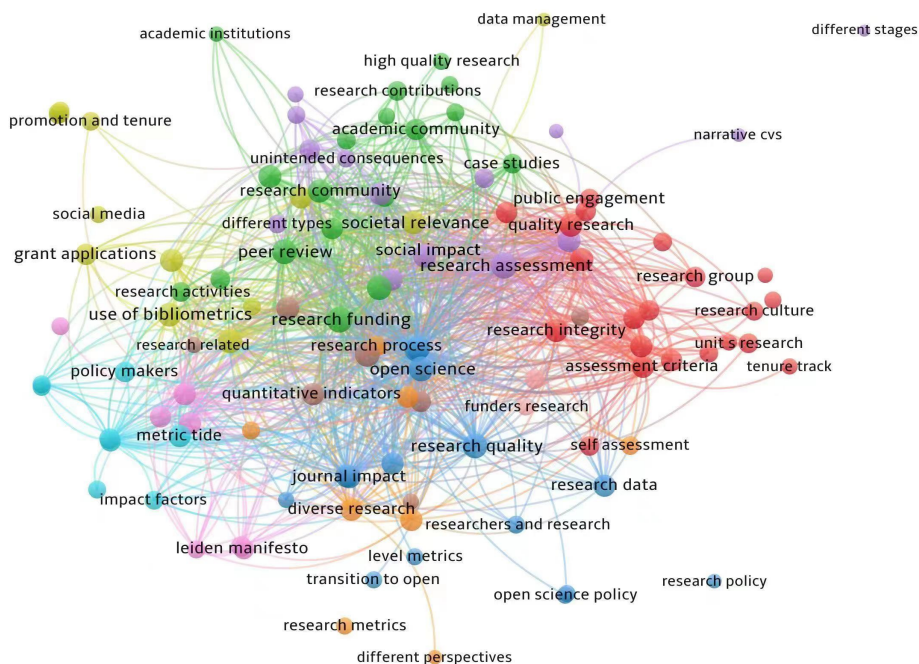


图 1 关键词共现关系

① <http://www.laurenceanthony.net/software/antconc>

2.2 机构评价:社会影响力成为主维度

在机构评价中,RRA 以强调研究质量与社会影响力为路径,实现激励措施与科研使命的契合,以支撑“从社会中的科学到科学为社会”的转变^[15]。以泰晤士高等教育(Times Higher Education,THE)世界大学排名为例,2019 年起增加了社会影响力评价框架,以联合国 17 项可持续发展目标为依据,报告大学应对贫困、饥饿、性别平等、清洁饮水、空气和气候变化以及提供高质量教育等问题的水平^[16]。REF 以研究质量、在学术界以外的影响(社会影响力)作为评价的两个要素,其中社会影响力权重从 REF 2014 的 20%调增到 REF 2021 的 25%,并从研究对人类健康、公共政策、法律与服务的影响等九个维度进行考量^[12]。SEP 2021–2027 中将研究质量及社会相关性,即研究给学术界以外的经济、社会、文化、公共政策、服务、健康、环境、生活质量带来的影响和收益作为评价的主维度^[13]。就此,兰德欧洲前任总裁 Grant 在题为《学术激励与研究影响力》的报告中,以 THE、英国 REF 2021、荷兰 SEP 2021–2027 为例探索了以社会影响力为重心的激励方式的变革,并提出“影响素养”的概念,进而明确了科学研究承担社会责任、解决社会问题的核心使命,同时标志着 RRA 在对卓越研究的重新定义中,将社会影响力纳入其中^[16]。

对于社会影响力的评价,当前主要以实证方法——案例研究进行^[17]。尽管人们一直寄希望于利用 Twitter、Facebook 等第二代社交网络及信息交流的数字痕迹表征研究的社会影响力,然而 Bornmann、Thelwall 等计量学家的研究表明,Twitter 等替代计量指标主要侧重于关注度,并非 REF 定义的研究成果对学术界以外的经济、社会的贡献^[17,18]。同时,ArXiv 创建者、康奈尔大学物理学家 Ginsparg 的文本挖掘结果表明,案例研究与同行评议结果显著正相关^[19]。由此,1870 年由哈佛大学提出的案例研究成为 RRA 中社会影响力评价的主流方法。英国 REF、荷兰 SEP、澳大利亚卓越科研框架均予以采用,其中专家依据案例的定性描述对研究的社会影响力进行遴选与甄别,因此其被认为是社会影响力评价中的同行评议方法。

2.3 项目评价:从基于产出的总结性评估转向形成性评价

在项目评价中,RRA 以减少对于产出及其载体的关注、强调研究过程为路径,加强对研究过程的监督、对研究行为的规范,进而识别并激励高质量研究。受重视产出的传统科学价值观的影响,Leydesdorff 所谓指标利用中的“层次谬误”盛行,即用一种高层次分析单位(如期刊)做调查,却用另一种低层次分析单位(如论文)做结论,研究质量及其影响被狭隘地定义为被引次数或 JIF^[20]。指标的误用、扭曲激励方式下的 PoP(Publish or Perish,出版或者出局)文化,意味着以牺牲质量、完整性和研究诚信为代价而优先考虑产出,由此带来的系统性偏见抑制了创新与研究的多元化,削弱了研究的完整性和可重复性,并损害了社会福祉。

为此,欧盟《科研评价改革指南》规定:根据研究的内在特征和表现而非载体进行评价,并将研究质量与学术影响力确定为共同的价值观^[3]。《科研人员评价香港原则》将研究实践的可靠性、研究报告的完整性列为五大原则之二,并着重分析了发表偏差(选择性发表阳性结果)带来的科学证据偏移^[14]。英国惠康以及盖茨开放出版平台均实行实验程序报告制度;加拿大国际发展研究中心开发了项目研究质量评价的 RQ+框架,其中包括基于量化的专家判断,以及以研究诚信、重要性、应用性、完整性、合法性、包容性为子维度的质量考量^[21];荷兰乌得勒支大学医学中心新评估框架的目标可谓是对上述操作最好的诠释,即摆脱基于产出的总结性评估而转向过程性或形成性评价^[22]。

2.4 研究人员评价:走向自我评价与指标精准化

在研究人员评价中,RRA以自我评价为路径,通过叙述性简历彰显人才的不同类型、突出其学术贡献,从而弱化文献计量指标效应并引导年轻的研究人员专注于科学研究。斯坦福大学元研究创新中心提出的《科研人员招聘、晋升和任期评价六原则》^[23]、荷兰发布的《为每个人提供施展才华的空间》的立场文件^[24],表达了对于早期职业研究人员(Early Career Researcher, ECR)的关注,催化了叙述性简历的流行,并在开放科学的情境下定义了研究的卓越与影响力,从而将预印本、数据等广泛的研究产出以及对研究社区、研究文化的贡献罗列其中,因此降低了JIF、H指数的影响。

2020年GRC调查表明,包括爱尔兰科学基金会、奥地利科学基金会、芬兰科学院、挪威大学联盟在内的近60%的受访机构正在试用或实施叙述性简历。就内容而言,以英国国家科研与创新署为例,要求以不超过1000字叙述研究与创新、个人与团队发展、对研究社区的贡献以及对社会的贡献^[2]。可见,以自我评述全面展示研究产出、研究质量、研究活动以及对研究社区的贡献是叙述性简历的主要特征。而荷兰研究委员会Veni计划中ECR叙述性简历关于个人研究愿景的附加项(即欧盟《科研评价改革指南》中“前瞻性叙述”)进一步引导年轻研究人员投身创新^[25];专家则基于标准化、结构化的简历识别出卓越的研究人员;加州大学伯克利分校神经科学家Llorens认为“这在研究人员评价方式上迈出了一大步”^[2]。

与此同时,以RRA指标开发与利用为路径的研究人员评价方式的探索也在进行中。2021年3月,在荷兰莱顿大学科学评价中心宣布停用H指数的同时,科睿唯安发布科研人员绩效分析新工具——射束图,以科研人员的出版物集合及其科研经历的全面画像破解H指数等单点指标,以学科领域百分位进行论文在同一出版年、同一研究领域影响的标准化,进而实现跨学科比较^[26]。Bornmann与Leydesdorff的分析表明,与其他文献计量指标相比,百分位与同行评议结果契合度最高,因此,射束图对于研究质量具有一定的表征意义^[27];科睿唯安科学信息研究所总监Szomszor认为,“射束图适宜以负责任的方式对单个科研人员的研究水平进行评估”^[26]。从这一意义上说,叙述性简历与射束图为RRA提供了具有可操作性的新工具和新路径,代表了研究人员评价的新取向。

3 开放科学环境下我国学术交流新业态与RRA

在全球开放转型的浪潮下,我国积极探索开放科学的中国路径,在其影响下的学术交流呈现新业态,这一嬗变亦扩散到科研评价领域。

3.1 开放科学环境下我国学术交流新业态

开放科学旨在建立一种促进科学创新的学术生态系统,具体表现为科技论文的开放获取、科学数据的开放共享、科技平台的知识融合、科研群体的广泛参与以及不断涌现的新型学术交流模式^[28]。《中国科技期刊发展蓝皮书(2021)》指出,中国开放获取出版、开放数据、开放平台和参与式科研等实践经历了从无到有、从小到大、从少到多的发展历程^[29]。首先,开放出版开启学术交流新模式。例如,国家新闻出版署融合发展重点实验室发起的开放科学计划(Open Science Identity, OSID),以全媒体思维实现轻量化的期刊获取;《心理学报》公开审稿意见,进行开放同行评议的尝试;中国电子科技大学与约翰威立国际出版公司联合创办世界一流高影响力开放获取期刊——InfoMat^[30];中国科学院打造科技论文预发布平台——ChinaXiv。其次,开放平台构筑了学术成果交流新机制。例如,国家自

然科学基金委发布大数据知识管理服务平台,提供基金项目科研成果的开放共享。再次,开放数据拓展学术交流与合作新渠道。例如,中国科技云实现与国际开放科学基础设施互联互通;中国国家基因库开展与欧洲分子生物学实验室的数据共享、与斯瓦尔巴全球种子库的合作。最后,开放社区形成学术交流新途径。诸如科学网、小木虫等学术社交网络在扩散研究成果、促进沟通合作的同时,以其跟踪测度功能为学术成果的利用和评价提供依据;OSID 在线实时评论便利了交流,并为出版后评价提供了辅助工具。以上种种实践表明,开放科学环境下,我国学术交流呈现出新的业态。

3.2 开放科学环境下我国 RRA 与“破四唯”

开放科学环境下的学术交流新业态为 RRA 提供了新的内容、思路与方法,而我国科研评价体系改革可谓是具有中国特色的 RRA。科学计量被定义为“对学术交流进行量化的科学”,即科研评价以学术交流为媒介对科研过程进行描述^[31]。因此,开放科学作为科学研究的新范式,不仅带来整个学术交流生态的变革,而且助推了以 RRA 为代表的科研评价方法的重塑。在 RRA 发展进程中,近几年我国也启动了科研评价改革,先后出台《关于深化项目评审、人才评价、机构评估改革的意见》(以下简称《改革意见》)、《关于规范高等学校 SCI 论文相关指标使用 树立正确评价导向的若干意见》《关于破除高校哲学社会科学研究评价中“唯论文”不良导向的若干意见》《关于完善科技成果评价机制的指导意见》。上述文件直指科研评价中的“唯论文、唯职称、唯学历、唯奖项”以及“SCI 至上”不良导向,同时明确了“建立以创新质量和贡献为导向的绩效评价体系、完善有利于创新的评价机制”的发展目标。可见,“破四唯”与纠偏“以刊评文”“以刊评人”“量化至上”的《旧金山宣言》《莱顿宣言》具有高度的一致性,而我国科研评价体系改革可以视作具有中国特色的 RRA。事实上,GRC 已将中国的 RRA 视为国际 RRA 的重要组成部分^[2]。目前,在“破四唯”导向下,我国科研评价改革正在进行中,而“破四唯”后如何“立新标”仍在探索中,下文将就开放科学环境下我国 RRA 实现路径进行分析,以利于新的评价标准的建立。

4 开放科学环境下我国 RRA 的实现路径

基于对我国政府文件与《2021 年 RRA 报告》、SEP 2021-2027 等文件的关键词与词丛分析的对,秉承 RRA 体现卓越性与影响力的宗旨,并结合坚持中国特色、扎根中国、融通中外、立足时代的指导原则,本文提出我国推进 RRA 的三层结构框架,如图 2 所示。

4.1 开放科学环境下我国 RRA 实施框架构建

我国推进 RRA 的三层结构框架以识别科研新范式为情境,以改善宏观政策(科研资助和管理政策、科学出版政策)为支撑,以改革评价方式为实操。其中以研究质量和社会影响力作为科研评价的重要标准,政策完善与评价改革并重,建立有利于潜心研究和创新的评价制度,从而回答科研评价改革的必答题:“谁来评”“评什么”“如何评”“怎么用”。

第一层为框架的基础层,主要用于对科学研究及其范式的识别,目的在于理解作为评价对象的科学研究以及作为科学研究与 RRA 媒介的学术交流,因此,开放科学和跨学科研究等科学研究的新范式、预印本等学术交流的新形式受到关注。第二层为框架的支撑层,既依托于科研环境,又指引着科研评价实践,因此,不仅需要评价管理政策,而且需要科研资助政策、出版政策与之呼应,以及相关培

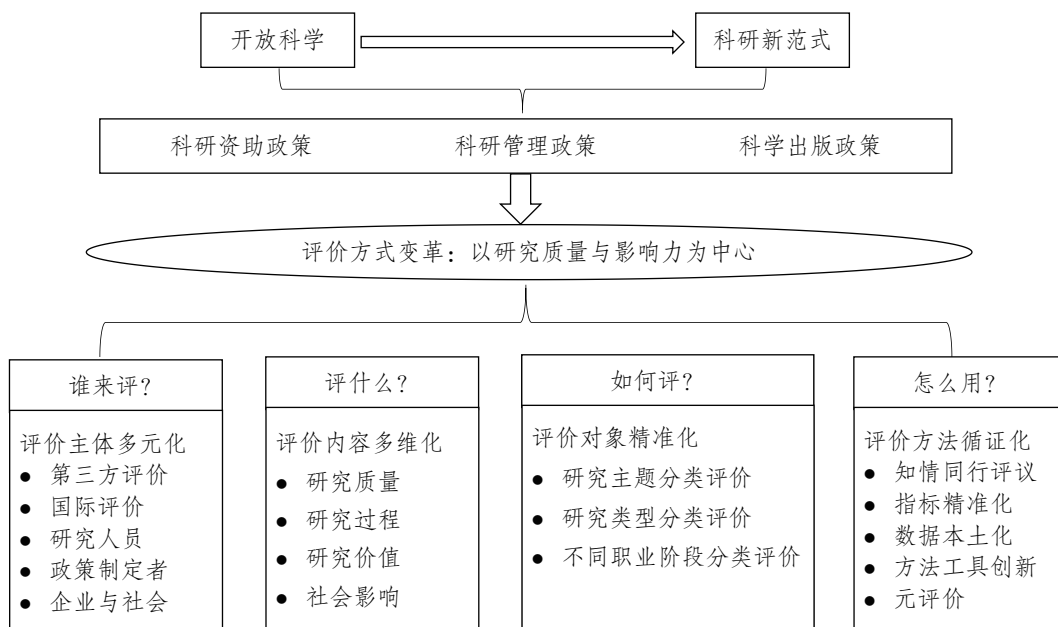


图2 我国推进 RRA 的结构框架

训的引导,以通过政策导向、最佳实践案例传播、培训指导共同促进 RRA 实施。第三层为操作层,以研究质量与影响力为核心开展评价,以体现我国 RRA 对研究质量的敏感及其对于研究影响力的侧重,其中由评价主体、评价内容、评价对象、评价方法构成。

4.2 基于评价要素的 RRA 实施路径

4.2.1 评价主体多元化

评价主体的多元化,即采用多方评价与第三方评价。兰德公司在题为《开放科学——公民对科研的职责与贡献》的报告中强调了公民与开放科学的关系,意即科学共同体以及所有利益攸关方在开放科学中都拥有发言权^[32]。RRA 通过多方评价与第三方评价,实现与开放科学的匹配,并弥合科学研究、创新社区与整个社会之间的差距,同时避免评价过程局限于科学共同体内部而产生偏颇。REF 在社会影响力评估方法上做了重要改革,制定了辅助评价指标,并邀请工商企业、政府、学界用户和其他利益相关方代表参与^[12];芬兰坦佩雷大学在教师招聘、晋升、绩效考核中亦聘请了外部专家^[25]。中国科学院在此方面也进行了尝试,在评价基础研究成果时邀请国际、国内顶尖专家开展同行评议,在应用研究与技术研发成果方面又兼顾市场、社会和第三方评价^[33]。因此,我国推进 RRA 时可以评价主体的多元化为路径,实现激励措施与社会价值观的契合,并解决评价工具与评价目的的脱节带来的评价功能异化问题。

4.2.2 评价内容多维化

评价内容的多维化,即以研究质量、过程、价值、社会影响力等对研究进行多维度考量。首先,在我国“破四唯”主导精神指引下,依据《科研人员评价香港原则》,借鉴英国、荷兰等国 RRA 的经验,以研究及研究人员的社会责任为内嵌价值,加强对研究过程的监管及研究人员行为的规范,从而逐渐转向过程或形成性评估。其次,以破除“SCI 至上”为突破口,引入创新文化要素,突出科技创新的质量、贡献与绩效,以正确评价科技成果的科学、技术、经济、社会、文化五种价值,从而消除“唯论文”“唯数

量”“急功近利”等弊端。再次,建立我国的开放引文索引(OpenCitations)、开放科学监测器,以覆盖从假设提出、数据搜集与分析到知识发布出版、发布后审查与认可的科学研究全周期,评估科学研究的质量与影响,从而反映学术交流新业态,并驱动研究社区的合作共享与创新。借此达到两个目标:一是满足评价主体需求;二是通过对研究的可信性、严谨性、透明性、创新性的认可促进研究人员的职业发展,以激励研究主体创新。目前,科技部会同中国科协,以及中国数学会、中国物理学会、中国化学会、中国计算机学会、中华医学会开展“学科学术评价规范研究试点”,形成了《学术评价规范建议》和《科研活动行为规范建议》。

4.2.3 评价对象精准化

评价对象的精准化是与多样性科学研究产出、活动的接合点,即像《开放科学建议书》中指出的:“鼓励负责任的研究与研究人员评价,以激励卓越科学,并承认研究产出、活动和任务的多样性。”^[4]分类评价是实现精准化的必要条件,同时也是构建科研评价体系的基础;分研究主题评价是分类评价的进一步细化;而与职业生涯不同发展阶段相契合的评价方式则成为研究人员 RRA 的新趋向。英国 REF 2021 成立跨学科评审小组,以保证对不同类型研究评价的公正性^[12];比利时根特大学为每位教授设立了个性化档案,以利于评价的准确性^[25];推进研究评估改革联盟(Coalition for Advancing Research Assessment, CoARA)成立职业早期、中期研究人员评价工作组。我国则以分类评价作为科研评价体制改革的重点,《改革意见》指出,“要建立健全分类评价体系,对不同类型的科研工作应分别建立各有侧重的评价路径。”据此,以评价对象的精准化为路径实现我国的 RRA,其中根据机构、人员、项目等不同研究主体,以及基础研究、应用研究、软科学研究等不同研究类型动态调整并合理配备定性评价方法与定量评价指标。我国的代表作制度即是评价对象精准化的一种尝试和体现,其实质在于通过评价与科学研究活动的切合,全方位认可研究成果,反映多样性的研究活动。

4.2.4 评价方法循证化

评价方法的循证化源于循证医学的概念。欧盟在《科研评价改革指南》中提到循证化路径^[3],具体到我国评价实践,包括以下三点。

首先,以知情同行评议为路径,即以计量支撑的同行评议取代对 SCI 指标的“专注”,以削弱马太效应的影响、提升同行评议的质量,从而实现从基于计量的科学价值观到基于质量与影响的价值观的转化。从国际 RRA 实践来看,存在两种情况,一方面,为提高 REF 排名,利物浦大学以教师在世界领先出版物上发文情况证明其实质性贡献的做法引起强烈反响^[34];另一方面,科学家对乌得勒支大学完全依赖叙述性简历进行招聘和晋升的举措忧心忡忡^[35]。可见,计量与同行评议二者缺一不可。因此,我国在 RRA 推进过程中,“SCI 至上”固然不行,但是贸然放弃具有公正形式的定量指标亦不可取^[36]。

其次,推进指标精准化、数据本土化,进行试点并开发工具包。《旧金山宣言》指导委员会致力于学术评估创新和变革成功案例的传播以及 RRA 工具的挖掘;欧洲大学协会与《旧金山宣言》指导委员会联合开发了案例库,就学术评价制度变革的关键要素进行探索^[37];莱顿大学科学评价中心与伊利诺伊理工学院合作开发研究评估工具以促进科研人员职业评估的新政策制定及其实践的发展^[38];中国科学院尝试了“破四唯”的实践之路,逐步实现从定量排名到以定性为主评价的演进^[39]。上述实践为我国 RRA 的进一步发展提供了基础与借鉴。

最后,在我国目前的科研评价体系中,逐步建立元评价机制^[40]。通过对于评价流程、评价方法、评价结果、评价目的的元评价,发现问题、改进评价技术和方法、改善评价结果的公布方式、加强对科

研评价自身的管理,从而提高评价质量^[41]。《计量潮报告》以及《斯特恩评论》即是对于 2014 年首次实施的 REF 的元评价^[42],从而为 REF 2021 奠定了基础,同时为我国提供了良好范例。

5 结语

通过上述分析,得出如下结论。

(1) 在环境扫描基础上,以研究质量和影响力为主维度,以同行评议为主导,以指标的负责任使用为依据,构成开放科学背景下国际 RRA 的进路。具体来说,在对开放科学、跨学科研究等科研环境认知的基础上,机构评价以社会影响力为主维度,项目评价从基于产出的总结性评估转向形成性评价,研究人员评价走向自我评价与指标精准化。

(2) 纵观国际 RRA 的理论与实践,基于研究卓越性(质量)及其影响力的科学价值观、知情同行评议的评价方法,已有效融入评估系统的渐进式改革之中^[43]。其中,在研究与创新、研究质量、社会影响力、研究诚信、研究文化、研究生态等多元氛围中,RRA 以责任为主线,以“体现卓越和影响,激励高质量研究”为宗旨,实现激励措施与科研使命的契合,与开放科学、跨学科研究相匹配。

(3) 缘起于欧洲的 RRA 已引起全世界的共鸣。欧盟调查表明,改革评价系统已被许多国家、地区、机构视作保证研究质量并提升其影响的优先选项^[3]。然而,从根本上说,RRA 属于方法创新,在荷兰心理学家 Groot 建立的观察、归纳、演绎、测试和评价实证研究周期审视下,当前 RRA 仍处于探索阶段,尚未就 RRA 中机构、项目、研究人员的评价标准和流程达成共识,而基于 JIF 的评价体系已根植于学术研究文化,按论文、专利数量等简单量化、重数量轻质的倾向依然存在。科研评价体系改革和重构是一个循序渐进、不断优化的过程,与其息息相关的研究文化、研究生态等系统性变革更是复杂且缓慢的历程。因此,一方面,需要建立研究的质量、影响意识;另一方面,需要加强各方交流,并达成评价系统变革的国际共识,形成行动路线进而提高评价系统的互操作性和可移植性,以避免 RRA 理论与实践的碎片化,并保证改革的有效实施。

(4) 开放科学环境下的学术交流新业态,对我国科研评价体系的变革提出了挑战。科研评价“破四唯”,推进我国科研评价体系改革创新是时代的新要求。我国不断增长的科学研究规模及国际影响力,以及图书馆发展对开放资源利用的迫切需求,为负责任的科研评价体系完善奠定了实践基础。对负责任评价的研究将为科研交流体系的完善和图书馆服务质量的提升提供参考。同时,科学共同体对于科研评价体系的“拟态”将进一步助推我国开放科学实践。因此,将负责任评价理念与实现路径本土化,将其数据获取、分析和集成的标准纵向延伸并嵌入文献采购资源价值评估、电子资源评估、学科服务评估、网络资源可信度评估等业务流程中,以形成富有时代特征、彰显中国特色、体现世界水平的评价体系,营造有利于开放科学发展的学术交流环境,或将成为图书情报工作的新使命。

致谢:本文系国家社会科学基金一般项目“负责任计量视角下科学评价方法与指标优化研究”(项目编号:18BTQ074)的研究成果。

参考文献

- [1] WILSDON J, ALLEN L, BELFIORE E, et al. The metric tide: report of the independent review of the role of metrics in research assessment and management[M]. London: SAGE Publications Ltd, 2015: 25-45.

-
- [2] CURRY S, RIJCKE S D, HATCH A, et al. The changing role of funders in responsible research assessment: progress, obstacles & the way ahead [R/OL]. (2020-11-18) [2023-02-13]. https://www.researchgate.net/publication/349762528_The_changing_role_of_funders_in_responsible_research_assessment_progress_obstacles_and_the_way_ahead.
- [3] European Commission. Towards a reform of the research assessment system scoping report [R/OL]. (2021-11-26) [2023-02-13]. <https://op.europa.eu/en/publication-detail/-/publication/36ebb96c-50c5-11ec-91ac-01aa75ed71a1/language-en>.
- [4] UNESCO. UNESCO recommendation on open science [EB/OL]. (2021-11) [2023-02-14]. <https://en.unesco.org/science-sustainable-future/open-science/recommendation>.
- [5] WILSDON J. Has the tide turned towards responsible metrics in research? [N/OL]. The Guardian, 2018-07-10 [2023-02-13]. <https://www.theguardian.com/science/political-science/2018/jul/10/has-the-tide-turned-towards-responsible-metrics-in-research>.
- [6] WILSDON J, BAR-ILAN J, FRODEMAN R, et al. Next-generation metrics: responsible metrics and evaluation for open science [R/OL]. (2017-03-20) [2023-02-13]. <https://op.europa.eu/en/publication-detail/-/publication/b858d952-0a19-11e7-8a35-01aa75ed71a1>.
- [7] Federation for the Humanities and Social Sciences. Approaches to assessing impacts in the humanities and social sciences [R/OL]. (2017-10-26) [2022-07-12]. <https://www.federationhss.ca/en/programs-policy/advocacy-research-impact>.
- [8] INORMS Research Evaluation Group. The SCOPE framework: a five-stage process for evaluating research responsibly [R/OL]. (2021-08) [2022-05-23]. https://inorms.net/wp-content/uploads/2021/11/21655-scope-guide-v9-1636013361_cc-by.pdf.
- [9] Global Research Council. Responsible research assessment: Global Research Council conference report 2021 [R/OL]. (2021-05) [2022-07-10]. https://www.globalresearchcouncil.org/fileadmin/documents/GRC_Publications/GRC_RRA_Conference_Summary_Report.pdf.
- [10] Cabinet Office. G7 2021 research compact [EB/OL]. (2021-07-12) [2022-07-10]. <https://www.gov.uk/government/publications/g7-2021-research-compact/g7-2021-research-compact>.
- [11] O'CARROLL C, ESPOSITO F, RENTIER B, et al. Evaluation of research careers fully acknowledging open science practices—rewards, incentives and/or recognition for researchers practicing open science [R/OL]. (2017-07) [2022-05-15]. <https://orbi.uliege.be/handle/2268/215460>.
- [12] Research England. Research excellence framework 2021 [EB/OL]. [2023-02-14]. <https://www.ref.ac.uk/>.
- [13] VSNU, KNAW, NWO. Strategy evaluation protocol 2021-2027 [EB/OL]. (2020-03) [2023-02-14]. https://storage.knaw.nl/2022-06/SEP_2021-2027.pdf.
- [14] MOHER D, BOUTER L, KLEINERT S, et al. The Hong Kong principles for assessing researchers: fostering research integrity [J/OL]. PLoS Biology, 2020, 18 (7) [2022-07-10]. <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000737>.
- [15] OWEN R, MACNAGHTEN P, STILGOE J. Responsible research and innovation: from science in society to science for society, with society [J]. Science & Public Policy, 2012, 39 (6) : 751-760.
- [16] GRANT J. Academic incentives and research impact: developing reward and recognition systems to better people's lives [R/OL]. (2021-02) [2022-01-26]. https://academyhealth.org/sites/default/files/publication/%5Bfield_date%3Acustom%3AY%5D-%5Bfield_date%3Acustom%3Am%5D/academicincentivesresearchimpact_feb2021.pdf.
-

- [17] TAHAMTAN L, BORNMAN L. Altmetrics and societal impact measurements: match or mismatch? A literature review [J/OL]. *Profesional de la Información*, 2020, 29(1) [2022-07-10]. <https://revista.profesional-delainformacion.com/index.php/EPI/article/view/epi.2020.ene.02>.
- [18] THELWALL M. Measuring societal impacts of research with altmetrics? Common problems and mistakes [J]. *Journal of Economic Surveys*, 2021, 35(2): 1302-1314.
- [19] VAN NOORDEN R. Seven thousand stories capture impact of science [J]. *Nature*, 2015, 518(7538): 150-151.
- [20] LEYDESDORFF L, MILOJEVIC S. The citation impact of German sociology journals: some problems with the use of scientometric indicators in journal and research evaluations [J]. *Soziale Welt*, 2015(66): 193-204.
- [21] IDRC. Research quality plus: a holistic approach to evaluating research [EB/OL]. (2018-07-06) [2022-02-22]. <https://www.idrc.ca/en/stories/research-quality-plus-holistic-approach-evaluating-research>.
- [22] SCHUURMANS M, BENEDICTUS R. Guide for reviewers/evaluators that use the UMC Utrecht indicators for impact [R/OL]. (2020-11) [2022-07-12]. <https://assets-eu-01.kc-usercontent.com/546dd520-97db-01b7-154d-79bb6d950a2d/a2704152-2d16-4f40-9a4b-33db23d1353e/Format-Impact-indicator-evaluation-pilot-incl-introduction.pdf>.
- [23] NAUDET F, LOANNIDIS J P A, MIEDEMA F, et al. Six principles for assessing scientists for hiring, promotion, and tenure [EB/OL]. (2018-06) [2023-02-14]. <https://blogs.lse.ac.uk/impactofsocialsciences/2018/06/04/six-principles-for-assessing-scientists-for-hiring-promotion-and-tenure/>.
- [24] VSNU, NFU, KNAW, et al. Room for everyone's talent towards a new balance in the recognition and rewards of academics [EB/OL]. (2019-11) [2022-01-25]. <https://www.universiteitenvannederland.nl/recognitionandrewards/wp-content/uploads/2019/11/Position-paper-Room-for-everyone%E2%80%99s-talent.pdf>.
- [25] DORA. Case study [EB/OL]. [2023-02-14]. <https://sfdora.org/dora-case-studies/>.
- [26] 马丁·索姆佐尔, 大卫·彭德尔伯里. 通过射束图揭示科研人员的引文表现 [R/OL]. [2022-02-22]. https://discover.clarivate.com/wp_beamplots_CN. (SZOMSOY M, PENDLEBURY D A. Interpreting the citation performance of individual researchers with beamplots [R/OL]. [2022-02-22]. https://discover.clarivate.com/wp_beamplots_CN.)
- [27] BORNMAN L, LEYDESDORFF L. The validation of (advanced) bibliometric indicators through peer assessments: a comparative study using data from InCites and F1000 [J]. *Journal of Informetrics*, 2013, 7(2): 286-291.
- [28] 陆彩女, 顾立平. 开放科学与科技期刊的应对 [J]. *中国科技期刊研究*, 2022, 33(2): 183-191. (LU C N, GU L P. Open science and suggestions on scientific journals [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2022, 33(2): 183-191.)
- [29] 中国科学技术协会. 中国科技期刊发展蓝皮书(2021) [M]. 北京: 科学出版社, 2021: 117-135. (China Association for Science and Technology. Blue book on the development of Chinese science and technology periodicals(2021) [M]. Beijing: Science Press, 2021: 117-135.)
- [30] 李方芳, 秦青松, 杨文婷, 等. 开放科学: 构建开放创新生态 [R/OL]. (2021-11) [2022-02-22]. <https://www.wlaforum.com/res/journal/>. (LI F F, QIN Q S, YANG W T, et al. Open science: building an open and innovation ecosystem [R/OL]. (2021-11) [2022-02-22]. <https://www.wlaforum.com/res/journal/>.)
- [31] 宋丽萍. 数字环境下的学术交流与学术评价 [M]. 北京: 中国社会科学出版社, 2019: 24. (SONG L P. Scholarly communication and research assessment in the digital environment [M]. Beijing: China Social Sciences Press, 2019: 24.)
- [32] SMITH E, PARKS S, GUNASHEKAR S, et al. Open science: the citizen's role and contribution to research [R/OL].

- (2021-07-17) [2022-02-22]. https://www.rand.org/content/dam/rand/pubs/perspectives/PE200/PE246/RAND_PE246.pdf.
- [33] 李晓轩,徐芳. “四唯”如何破:中国科学院研究所评价的实践和启示[J]. 中国科学院院刊,2020,35(12):1431-1438. (LI X X,XU F. How to break the “Siwei”?—Practice and enlightenment based on research institute evaluation of Chinese Academy of Sciences[J]. Bulletin of Chinese Academy of Sciences,2020,35(12):1431-1438.)
- [34] Responsible research assessment faces the acid test[J]. Nature,2021,595(7868):471-472.
- [35] DALMEET S C. Scientists at odds on Utrecht University reforms to hiring and promotion criteria[EB/OL]. (2021-08-09) [2022-07-09]. <https://www.natureindex.com/news-blog/scientists-argue-over-use-of-impact-factors-for-evaluating-research>.
- [36] 邱均平,张裕晨,周子番. 新时代我国科研评价体系重构中必须处理好八大关系[J]. 中国图书馆学报,2021,47(1):47-60. (QIU J P,ZHANG Y C,ZHOU Z F. Eight relationships in the reconstruction of scientific research evaluation system in China in the new era[J]. Journal of Library Science in China,2021,47(1):47-60.)
- [37] RITA M,BREGT S,FEDERICA G,et al. From principles to practices:open science at Europe's universities[R/OL]. (2021-07) [2022-02-22]. <https://eua.eu/downloads/publications/2021%20os%20survey%20report.pdf>.
- [38] Tools to Advance Research Assessment (TARA) is a project to facilitate the development of new policies and practices for academic career assessment[EB/OL]. [2022-02-25]. <https://sfdora.org/project-tara>.
- [39] 徐芳,代涛. 全过程多主体责任评价的概念、内涵和关键问题[J]. 科学学与科学技术管理,2021,42(6):3-14. (XU F,DAI T. The concept,connotation and key issues of responsible evaluation based on multi-agent and whole-process[J]. Science of Science and Management of S. & T.,2021,42(6):3-14.)
- [40] 俞立平,潘云涛,武夷山. 元评价对提高科技评价质量的影响研究[J]. 科学学研究,2012,30(4):512-516. (YU L P,PAN Y T,WU Y S. A effective method to enhance science and technology evaluation quality:meta-evaluation[J]. Studies in Science of Science,2012,30(4):512-516.)
- [41] 叶继元. 学术“全评价”体系与中国特色哲学社会科学学术评价体系的构建与完善[J]. 社会科学文摘,2021(7):4-6. (YE J Y. The construction and improvement of the academic all-round evaluation system and the academic evaluation system of philosophy and social sciences with Chinese characteristics[J]. Social Sciences Digest,2021(7):4-6.)
- [42] STERN N. Lord STERN's review of the Research Excellence Framework:call for evidence[R/OL]. (2016-01) [2022-02-22]. <https://core.ac.uk/download/pdf/74380480.pdf>.
- [43] 张琳,SIVERTSEN G. 科学计量与同行评议相结合的科研评价——国际经验与启示[J]. 情报学报,2020,39(8):806-816. (ZHANG L,SIVERTSEN G. Combination of scientometrics and peer-review in research evaluation:international experiences and inspirations[J]. Journal of the China Society for Scientific and Technical Information,2020,39(8):806-816.)

宋丽萍 天津师范大学信息管理系教授。天津 300387。

王建芳 中国科学院科技战略咨询研究院副研究员。北京 100190。

(收稿日期:2022-07-11)