

文章编号:1004-115X(2015)02-0016-04

# 科研众包——科研合作的新模式

卫垆圻<sup>1,2,3</sup>,姜涛<sup>1</sup>,陶斯宇<sup>1</sup>,谢光锋<sup>4</sup>,谭宗颖<sup>1</sup>

(1. 中国科学院文献情报中心,北京 100190;

2. 中国科学院大学,北京 100049;

3. 中国科学院心理研究所,北京 100101;

4. 中国科学院院士工作局,北京 100190)

**摘要:**将众包作为一种新型的科技合作形态,首先论述了它的出现及其必然性,然后详细分解了这种新型科研合作形态的运行模式,并论及它的影响、优势和目前存在的主要问题。最后,结合我国的具体科研环境,提出4点具体的政策建议。

**关键词:**科研众包;科研合作;公民科学计划

**中图分类号:**G311 **文献标识码:**A

## Science Sourcing——A New model of Scientific Cooperation

WEI Tong—qi<sup>1,2,3</sup>,JIANG Tao<sup>1</sup>,Tao Si—yu<sup>1</sup>,XIE Guang—feng<sup>4</sup>,TAN Zong—ying<sup>1</sup>

(1. National Science Library, Chinese Academy of Sciences, Beijing 100190, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China;

4. Academician Bureau of the Chinese Academy of Sciences, Beijing 100190, China)

**Abstract:**As a new model of scientific cooperation, the present study reviewed the appearance of science sourcing, and then systematically analyzed its mode of operation. At the third part, the present study discussed the influences, advantages and current problems of science sourcing. Finally, four policy advices were put forward according to the scientific environment of our country.

**Key words:**Science Sourcing; Scientific Cooperation; Citizen Science Project

### 1 科研众包的兴起

“众包是指由个体、机构、非盈利组织和公司向具有不同知识背景的一组个体公开发起、自愿参与的在线活动。”<sup>[1]</sup>Jeff Howe于2006年6月在《连线》杂志上发表了“众包的兴起”一文,总结了网络时代经济模式的4大特点,并为这种新的经济模式创造了新的称谓——“众包”。商业项目采用众包模式,

已经基本成熟,有非常多的成功案例。近些年,当众包项目以科学发现和解决科技面临的挑战问题为目的,并且发起者为科研机构或科学家时,众包便从一种商业模式拓展为一种新型的科研合作模式,其应用范围极大地拓宽,也改变了传统上对科研合作的理解。这样的众包即为“科研众包”,参与众包科研项目的志愿者往往被称为“公民科学家”。

收稿日期:2014—11—18

作者简介:卫垆圻(1982—),男,河北人,中国科学院文献情报中心博士研究生,主要研究方向:学科情报,科学计量;

通讯作者:谭宗颖(1957—),女,四川人,博士生导师,主要研究方向:科技发展战略研究、学科战略情报研究、国际科技竞争力研究。

科研众包是应时而兴和科技发展的产物。科研众包形成的前提是互联网的普及。互联网,作为可以进行管理、分配和聚集大量分布式资源的信息系统,能低成本地将地理上分布广泛、文化上差别巨大的不同个体汇聚在同一个开放的环境下,同步或异步地彼此充分交换想法,达到使科研众包项目足以顺利完成所需要的人数阈值,在科研项目发起方的需求和公民科学家的行动之间建立快速有效的响应。这是科研众包得以形成的重要支柱<sup>[2]</sup>。

科研众包兴起的另一个重要动力源于科研活动产生的海量数据与数据处理能力之间的矛盾,即挖掘数据的能力始终落后于科学家们收集或生成数据的能力。Zooniverse 众包项目就是一个典型的例子,该项目的主要工作是利用公众的力量对天文图像进行分类和标识,以寻找行星或者查看天文物体。宇航员在过去数十年的工作成就巨大,产生了海量数据和图像,但是天文学家目前尚无能力在短时间内处理完这些洪水般的数据。数以百万计的星系图片完全需要手工来识别和分类。恰恰是众包模式来自全球各地的 85 万名志愿聚集组织起来者参与其中。Zooniverse 已经获得了很多成果,比如无核球螺旋星系的发现<sup>[3]</sup>。

推动科研众包成功开展还需要公民科学家的广泛自愿参与。能够吸引广泛、多样的公民科学家,并维持他们的参与热情对科研众包项目至关重要。公民科学家参与众包项目的动机不一。年纪较大的志愿者希望积极地利用好自己的休闲时光,年纪较轻的志愿者更多是希望融入一个新的集体。还有一些参与者的动机是可以到户外去,或者获得科学发现的成就感,当然还有一些人是为了物质奖励<sup>[4]</sup>。

## 2 科研众包的运作模式

### 2.1 公民科学家

科研众包的重要行动要素是公民科学家。他们自愿投入金钱、计算机资源、人力、时间等来参与科研活动。科研众包项目的成功与否很大程度上取决于是否能吸引足够多的公民科学家参与,并激发他们的热情与能力。每个众包科学计划都需要细分成具有执行度的小任务,反映了“参与项目后个体所必须做出的最小付出”。依据细分程度的不同,公民科学家参与科研众包项目的形式和程度有区别。

有些科研众包项目对参与者的要求比较简单,不需要投入脑力活动,只需利用其计算设备的空闲时间参与科学计算,分布式计算项目便属于此类,这

类众包方式已应用于搜索梅森素数,搜索外星文明发出的无线电信号,蛋白质结构预测与设计,研究气候变化的趋势等等<sup>[6]</sup>。

当众包科研项目的细分任务比较复杂时,对公民科学家的专业水平要求就相应提高。比如,美国怀俄明大学教授 Bruce Parkinson 希望合成一种合适的催化剂来驱动生成氢燃料,因为 60 多种金属之间的可能组合是多种多样的,所以,他将催化剂设计成廉价、快速、简单组合的化学试剂盒。利用这些小套件,该合成项目吸引了 70 所大学和高中的 500 多名学生参与,很快创造出了独特的金属氧化物,一些学生还将他们的研究成果提交到了一本科学杂志上<sup>[6]</sup>。

个体无法完成的、更为复杂的众包科研合作需要机构组织来承接。针对出血性大肠杆菌菌株 O104:H4 的研究<sup>[7]</sup>就是这类科研众包的典型案例。先后有多个国家的科研机构 and 大学参与,在几天内,以科研众包的形式,便完成了分离暴发菌株、公开发布数据、获得基因组测序、描述病原体特点等各种工作,其效率令人印象深刻。这些实践表明,世界各地的专家可以通过互联网迅速组织起来,以科研众包的方式工作,产生高质量的分析成果,从而解决科学问题和紧急的实际问题(如疫情爆发等)。

### 2.2 众包项目的发起者与设计者

众包项目的发起者和设计者是科研众包项目的主要获益者,他们常常是某个学科领域的专职科学家或专门的科研机构。众包项目是否能获得成功,众包方案的设计至关重要,这是众包项目前期的核心工作。不同类型的科研众包项目的设计方案是不同的。

2.2.1 集成型科研众包旨在收集更多更广泛的分布式信息与数据,其成功的关键在于吸引广泛多样的个体的热情参与。这类项目需要重点考虑如何在兼顾任务的执行环境、手段与工具,以及公民科学家的专业水平等因素的情况下,将总体需求分解为单个可执行的、有趣的简单任务,同时考虑响应回馈,给予参与者有价值的激励。项目设计者还需要考虑如何整合与分析广泛的分布式信息<sup>[8]</sup>。Fold it 蛋白质折叠游戏便是其中的典型案例。它设计用于解析蛋白质酶的结构,玩家可以通过在屏幕上随机折叠分子塑造新的蛋白质,它的辉煌成绩是在 3 周内解开了困惑科学家 20 年的一个艾滋病相关酶的蛋白质结构<sup>[9]</sup>。Fold it 对每位玩家的专业水平几乎没有要求,玩家只需要 1 台联网的电脑,在简单的指

导下便可以用鼠标随意创造和修改蛋白质,正是因为对玩家的专业要求低,任务细分明确,游戏设计有趣,才激励了大量用户参与,从而解决了复杂的科学问题。

2.2.2 选择型科研众包采用了“赢者通吃”的模式,唯一的目的是找到最优的解决方案。这类众包不需要设计细分任务,但需要提出明确的需求和目标,并在最优解决方案的选择和激励上精心设计<sup>[8]</sup>。典型案例来自 Netflix 公司。该公司公开悬赏征集提高影片推荐准确性的解决方案,快速吸引全球 186 个国家的上千个研发小组参与了推荐算法的改进,并成功实现了将推荐准确性提高 10% 的目标,而花费仅为 10 万美元<sup>[10]</sup>。类似模式对解决目标明确的复杂问题非常有用,在重要的科研众包平台 innocentive.com 上,可以看到大量这类研究项目,比如 NASA 要求寻找一种新的方法使用、可视化和分析气候和地球科学数据,美国打击恐怖主义技术办公室的项目希望设计一套信息管理和分析的手机系统等等。

### 2.3 中介平台

中介平台是科研众包运作模式的第 3 个重要环节。科研众包合作的顺利开展非常依赖能够聚集大量用户的网络平台或社区,这样才能将专职科学家、公民科学家和合作伙伴等各方力量连接和组织起来。Innocentive.com 就是典型例子。美国主要的政府机构、非盈利组织和商业企业都在 Innocentive.com 网站上提交困扰他们的技术难题来寻求最佳的解决方案和途径,并为最佳解决方案提供者支付 1~10 万美元不等的奖金。

## 3 科研众包的影响和优势

科研众包突破了不同学科间的界限。公开的数据或任务需求,通过网络和众包平台,聚集各领域的科学家和技术人才,使跨学科研究更为便利。科研众包突破了专职科学家和业余科学家的界限,使有一技之长的任何个人都可以投入自己的智力和时间参与到科学发现和知识创造中,并获得回报<sup>[11]</sup>。科研专职科学家的工作模式也在发生改变,越来越多的科学家意识到公众可以作为劳动力、技能、计算能力甚至是财政资助的重要来源<sup>[12]</sup>,充分地利用公众,他们将有更多时间和精力投入到对专业水平要求更高的研究创新中,同时,他们与公众之间借由众包项目加强联系;科研众包让科学发现走出实验室,在减少了研发操作成本<sup>[8]</sup>,优化了科研资源配置的

同时,还提高了科研成果的影响力与普通公众的认知度和参与度。因此,科研众包在作为一种科学教育和科学传播形式,促进公众对科学的理解的作用上也颇值得重视<sup>[13]</sup>。科学家必须也有责任让公众理解科学研究的价值,而众包科研项目可以让公众亲身参与研究,无疑是达成这一目标的良好手段之一<sup>[14]</sup>。实践证明,众包科学计划经过巧妙的在线设计,在良好的社区环境下,可以发挥规模效应,产生高质量的科学数据和可靠有效的科研产出,以及出乎意料的洞见和创新<sup>[15]</sup>,集体智慧的成果比这个集体中最优秀的个人的创见更为出色。

## 4 促进科研众包的政策建议

研究机构应积极鼓励研究人员尝试众包模式来解决科学问题,对有此意向的科学家给予资金、技术和政策倾斜。科研众包项目在取得新发现的同时,还节约了宝贵的科研经费,优化了科研资源的配置和人员结构,增强了专业科学家与公众的互动交流,扩大了研究机构的社会影响力,以及提高了公众对科研活动的认知度和参与度。这些积极效益在研究机构对研究人员的科研评价中应有所显现。

在众包项目实施前,应该预判科研众包的风险及知识产权等问题,以有效规避和降低众包项目的风险性。在众包项目开始前,应成立负责风险管理和知识产权管理的专门委员会,要求主要负责人提交具体的操作计划,在此基础上,制定项目成果的版权归属规章和约定,评估操作的可行性和风险。对不适用于众包模式的研究项目,要防范其未经许可地采用众包模式后所导致的负面影响。

研究资助机构应增设专门的公民科学项目。这样不仅可以推进科学发现,还将促进培养更具科学素养的公众,并激励下一代科学家。甚至,科研资助本身也可采用众包形式。众包科研资助已成为一种获取科研经费的可能路径。科学家可将研究项目在众包科研资助平台上公布,寻找有热情的公众给予资助。

### 参考文献:

- [1] Estelles-Arolas E., Gonzalez-Ladron-de-Guevara F. Towards an integrated crowdsourcing definition [J]. Journal of Information Science, 2012, 38(2):189-200.
- [2] Nov O, Arazy O, Anderson D. Crowdsourcing for science: Understanding and enhancing scisourcing contribution [C]. ACM CSCW 2010 Workshop on the Changing Dynamics of Scientific Collaborations. Savannah, Georgia, USA, 2010.
- [3] Simmons B D., Lintott C., Schawinski K., et al

- Galaxy zoo: Bulgeless galaxies with growing black holes [J]. Monthly Notices of the Royal Astronomical Society, 2013, 429(3):2199-2211.
- [4] Robson C. Using mobile technology and social networking to crowdsource citizen science [D]. EECS Department, University of California, Berkeley, 2012.
- [5] Nouman Durrani M., Shamsi J. A. Volunteer computing: Requirements, challenges, and solutions [J]. Journal of Network and Computer Applications, 2014, 39:369-380.
- [6] Lockwood D. Crowdsourcing chemistry[EB/OL]. (2012) <http://cen.acs.org/articles/90/i46/Crowdsourcing-Chemistry.html>, 2014-08-04.
- [7] Outbreak genomics [J]. Nature Biotechnology, 2011, 29(9):769-769.
- [8] Schenk E, Guittard C. Towards a characterization of crowdsourcing practices [J]. Journal of Innovation Economics & Management, 2011,(1):93-107.
- [9] Khatib F, DiMaio F, Foldit Contenders Group, et al. Crystal structure of a monomeric retroviral protease solved by protein folding game players [J]. Nature Structural & Molecular Biology, 2011, 18(10):1175-1177.
- [10] Peng G. The preliminary settlement of crowdsourcing legal issues[C]. 2010 International Conference on E-Business and E-Government (ICEE). 2010. IEEE
- [11] 冉伟, 卢远添. 科研社交网络与众包平台研究报告[EB/OL]. (2013) [http://www.coinsay.com/article/coinsay\\_3165.html](http://www.coinsay.com/article/coinsay_3165.html), 2014-08-04.
- [12] Cohn J. P. Citizen science: Can volunteers do real research?[J]. Bioscience, 2008, 58(3):192-197.
- [13] Brossard D., Lewenstein B., Bonney R. Scientific knowledge and attitude change: The impact of a citizen science project [J]. International Journal of Science Education, 2005, 27(9):1099-1121.
- [14] Silvertown J. A new dawn for citizen science [J]. Trends in Ecology & Evolution, 2009, 24(9):467-471.
- [15] Trumbull D. J. Bonney R. Bascom D. et al. Thinking scientifically during participation in a citizen-science project [J]. Science Education, 2000, 84(2):265-275.

(上接第 11 页) 适应当前网络概念逐渐成熟的时代。跨组织横向协同创新是将产学研等组织中在同一产业中的能力与资源进行有效整合并创造更大效用的系统。随着网络概念的发展,跨组织横向创新系统将逐步完善,纳入更多的角色与元素,充分调用整合各种能力与资源,以创造更多的方便、解决更多的困难。

在研究方法上,目前跨组织横向协同创新的研究方法虽然整体和多数细节上能够有效地解决具体问题,但其样本的筛选过程由于为了减小误差带来的影响而混杂着主观性。如今是大数据时代,研究以全面多样的数据为根据,这要求研究者需要容忍误差,主要关注数据能够表达的事物间相关关系。跨组织横向协同创新本身就是注重要素之间的相关关系,其相关关系的强弱取决于诸多量化因素,因此其研究需要更多的量化数据,应建立长期的数据采集计划,将诸多软性因素以科学的方式量化。

#### 参考文献:

- [1] Yang Dongsheng, Zhang Yongan. Simulation Study on University-Industry Cooperative Innovation Based on Multi-agent Method[C]. Proceedings of the 2008 International Conference on Computer Science and Software Engineering, 2008:528-531.
- [2] Hongzhan Chen, Qiangqiang Zhao, Zhen Xinjin. Study on Grey Evolutionary Game of Industry-University-Institute Cooperative Innovation[C]. Proceedings of 2009 IEEE International Conference on Grey Systems and Intelligent Services, 2009:1120-1125.
- [3] 陈晓红, 解海涛. 基于“四主体动态模型”的中小企业协同创新体系研究[J]. 科学学与科学技术管理, 2006, (8):37-43.
- [4] 胡冬云, 陶丹. 面向行业产业的协同创新中心运行机制研究[J]. 中国高校科技, 2012, (11):22-24.
- [5] 辛冲. 企业组织与技术的协同创新研究[J]. 研究与发展管理, 2011, (23):37-43.
- [6] 杜兰英, 陈鑫. 政产学研用协同创新机理与模式研究——以中小企业为例[J]. 科技进步与对策, 2012, 29(22):103-107.
- [7] 李金海. 基于协同创新的概念性结构模型研究[J]. 河北工业大学学报, 2013, (42):112-118.
- [8] 解学梅. 中小企业协同创新网络与创新绩效的实证研究[J]. 管理科学学报, 2010, 13(8):51-64.
- [9] 解学梅, 左蕾蕾. 企业协同创新网络特征与创新绩效: 基于知识吸收能力的中介效应研究[J]. 南开管理评论, 2013, (3):47-56.
- [10] 甄晓非. 协同创新模式与管理机制研究[J]. 科学管理研究, 2013, 31(1):21-24.
- [11] Pan Jieyi, Wang Fen. Analysis and Evaluation of Knowledge Transfer Risks in Collaborative Innovation Based on Extension Method[C]. IEEE Xplore, 2008:1-4.
- [12] 李煜华, 武晓锋, 胡瑛瑛. 基于演化博弈的战略性新兴产业集群协同创新策略研究[J]. 科技进步与对策, 2013, 30(2):70-73.