

颠覆性研究文献计量识别方法述评

郭丽娜 李星琛 左文革 周群
(中国农业大学图书馆, 北京 100193)

摘要: 颠覆性研究是指挑战或颠覆原有研究范式, 能够创造新范式或新领域的研究。归纳梳理颠覆性研究的内涵、特征及其文献计量学识别方法, 为颠覆性研究的识别方法和指标优化研究提供参考与借鉴。利用文献调研方法对相关文献进行系统梳理, 将其分为基于引用曲线、引文网络的识别方法和基于多源数据的组合识别方法, 指出现有方法的局限性及未来发展趋势。颠覆性研究的文献计量识别方法除存在引文分析本身的时滞性外, 其在测度维度、数据源交叉融合等方面仍存在诸多挑战, 未来可考虑在充分调研颠覆性研究历史案例的基础上, 总结潜在颠覆性研究的特征, 结合内容分析、大数据或专家知识等手段, 以提高识别的准确性和前瞻性。

关键词: 颠覆性研究; 文献计量; 识别方法

中图分类号: G358

DOI: 10.3772/j.issn.1673-2286.2020.03.003

当前, 全球科学格局变化, 基础研究与应用研究交叉渗透, 科学研究与技术创新相互促进, 颠覆性研究为政治、经济、科技等领域带来革命性影响, 推进颠覆性研究对实现国家科技创新跨越式发展意义重大。基于此, 各国竞相占据新一轮科技创新博弈的制高点, 先后制定了推进颠覆性研究的相关政策, 为提升国家软实力、重塑世界竞争格局奠定基础。

颠覆性研究的识别和探测已成为科技创新领域的热点和前沿议题, 不仅有助于为国家科技管理部门和基金资助机构早期识别、及早部署具有颠覆性创新的基础研究提供决策支持, 也有助于在科学发展过程中占据先发优势, 提高综合竞争实力。在情报学领域, 颠覆性研究及其一系列相关概念的识别研究均在于对“颠覆性”这一特征的计量学信号的捕捉、探测和判断, 往往具有较高的不确定性和突变性, 也是颠覆性研究识别的重点和难点。鉴于此, 本文简要梳理颠覆性研究的相关概念及特征, 进而归纳其文献计量学识别方法, 指出该方法的适用性与局限性, 为颠覆性研究的识别及其方法指标的改进提供参考。

1 颠覆性研究的概念及特征

对于早期的科学社会学家而言, 颠覆性研究或科学突破是科学奖励体系中的概念, 其重点在于研究结果的新颖性或原创性, 它可以灵活地指代新理论的创建、现有理论的改进或对已知现象的更好描述。但Kuhn^[1]认为颠覆性研究是科学范式的根本突破, 与科学革命有关, 导致范式转变的突破和革命在科学体系内具有重要且相互补充的作用。受Kuhn的启发, 后来的科学社会学家将颠覆性研究(根本性突破)视为对科学权力结构的挑战, 即颠覆性研究的新颖性令人恐惧, 并在一定程度上被某个领域的“统治阶层”以同行评议等形式所抵制^[2]。

Kuhn指出, 颠覆性研究和渐进性研究是科学研究进程中的两种发展状态, 二者相互作用, 交叉渗透, 形成科学发展的动态结构^[1]。渐进性研究指在现有范式或理论框架内对知识和实践作增量改进, 促成科学的累积式发展, 而颠覆性研究属于具有革命性的科学突破或技术突破, 代表着对现有研究范式或技术产品范式的颠覆, 导致新的范式或领域的产生, 推动科学技术革命的发生^[2]。在以往的研究中, 颠覆性研究存在许

多与其相近或相似的概念,如变革性研究、颠覆性创新^[3]、破坏性创新^[4]、科学突破、里程碑论文或突破性研究等,但其内涵基本上继承了Kuhn所赋予的概念和特征。有学者对上述不同概念进行了界定、比较或辨析,如将变革性研究、颠覆性技术和颠覆式创新分别对应基础研究(science)、技术创新(technology)和产品市场^[5](innovation)3个不同角度的概念。Kostoff等^[6]认为颠覆性技术可以是现有技术的组合,也可以是一项新技术,可能导致其应用领域技术产品范式发生颠覆性转变。有学者认为重大创新突破在被科学领域认可之前,往往先在技术领域被发现或被承认,强调了颠覆性研究与颠覆性技术的相关性,Christensen^[3]提出颠覆性创新这一概念时,不仅强调技术本身的颠覆性,更注重技术应用所带来的颠覆性效应。因此,与颠覆性技术相比,颠覆性研究更容易通过文献计量特征表现出来,基于此,本文暂且将上述类似概念统称为颠覆性研究,通过梳理前人研究可以发现,颠覆性研究通常具有6个方面的特性。

(1) 高风险性。颠覆性研究可能带来基础性新发现,本质上是具有风险性的尝试,加之早期在同行评议时未被看好^[7],难以获得经费资助,进一步增加了研究的难度和风险,也因此,颠覆性研究往往由更敢于尝试和冒险的小团队完成^[8]。

(2) 偶然性。科学发展的历程和经验显示,颠覆性研究通常具有较高的偶然性,甚至完全出乎研究者意料。美国国家科学院在《促进地理科学的变革性研究》报告中分析了过去65年地理科学领域的主要颠覆性研究案例,认为传统的审查方式往往会忽略颠覆性研究,需要一种新的审查方法来加强对颠覆性研究的支持^[9]。

(3) 学科综合交叉特征明显。颠覆性研究的产生与不同学科间的交叉和依赖有关,随着学科界限越来越模糊,跨学科研究更易形成,许多重大基础科学问题已超出单一学科的研究范畴,需要借助不同的学科背景、理论方法,以及多种技术、设备等共同完成。

(4) 具有全球性共识并得到科学家的公认。尽管有些颠覆性研究因在早期无法得到认可而沉寂多年,但一旦被识别或认可,便会迅速获得科学界的高度关注与评价。如将诺贝尔奖获奖论文作为颠覆性研究案例或衡量颠覆性研究的标准^[10]。

(5) 基础性和长期性。与不断迁移演变的研究热点不同,颠覆性研究立足于创造新知识和发展新认知,

拓展科学的边界,进而改变我们对世界的认识。同时对基础科学发展进程有着革命性的颠覆和创造,能够开创新领域,并随时间推移产生广泛深远的影响^[11]。

(6) 潜在技术与应用属性。颠覆性研究多为应用导向型研究且被专利大量引用,Winnink等^[12]以首次分离出石墨烯的里程碑式论文作为颠覆性研究案例,发现石墨烯相关专利的申请量随该论文引用量增加而呈现滞后式增长,即颠覆性研究能够刺激相应技术研究的产生。

2 颠覆性研究的文献计量识别方法

如上所述,颠覆性技术多强调技术优势,而颠覆性研究更多地侧重于基础研究,而在学术网络中,论文和专利数据是基础研究的主要产出形式,颠覆性成果产生的影响可以通过文献计量数据和专利分析追踪及预测^[13]。学术论文和专利作为科学研究活动的主要产出形式,蕴含着先进的科学创新思想和丰富可挖掘的创新信息,是颠覆性研究识别方法的主要数据源,其文献计量识别方法主要是围绕引文特征,如引文曲线、引文网络以及不同数据源的组合展开。

2.1 基于引文曲线的识别方法

论文被引频次是衡量论文学术影响力最具代表性的指标,被引频次反映在时间上的曲线被称为引文曲线,引文曲线从引用的时间分布入手,分析引用的历史过程,以此反映论文影响力。引文曲线可分为经典引文曲线、指数增长引文曲线、睡美人引文曲线、双峰引文曲线和波型引文曲线等^[14]。根据颠覆性研究的引文曲线特征,该方法主要围绕高被引论文(指数增长引文曲线)和睡美人文献(睡美人引文曲线)两种类型展开。

2.1.1 基于高被引论文的识别方法

引文曲线呈指数增长的论文通常在发表之初就受到广泛认可,并迅速成为高被引论文。许多研究以高被引论文为研究案例或数据集,探索颠覆性研究的特征及识别方法。早在1992年,Garfield等^[15]曾基于高被引作者的排名预测诺贝尔奖得主,随后,科睿唯安基于该方法每年推出“引文桂冠奖”用来识别具有最重大影响的科研人员,用于预测可能的诺贝尔奖得主。

Winnink等^[12]以报道石墨烯的里程碑式论文作为颠覆性研究案例,指出高被引论文是识别潜在颠覆性研究的重要指标,进而提出颠覆性研究的引文曲线具有指数增长及高被引特征,有助于在科学和技术层面上识别和表征早期潜在的颠覆性研究。Bornmann等^[16]分析了25年内500多万篇文献的引文曲线,发现其中有1 013篇已知颠覆性论文在发表之初就被大量引用,且长期保持高被引状态。研究基于典型案例或对历史数据的回顾,指出了颠覆性研究的高被引特征,为颠覆性研究的识别提供了一定参考,但其准确性仍待考量。

基于此,Schneider等^[17]提出颠覆性研究应既具有高被引特征又具备原创性,同时其重要引文在本领域和其他科学领域均有分布,进而通过特征分数和尺度方法(Characteristics Scores and Scales Approach)划分引文分布,结合特定算法对施引论文中的高被引论文进行分类和筛选,通过调用阈值和过滤识别高被引论文集中的颠覆性研究。Ponomarev等^[18]将设置高被引阈值与引文曲线拟合相结合,首先筛选发表5年后超过高被引阈值的论文作为候选的颠覆性论文;然后根据论文早期引文数据外推后期引文曲线,构建线性拟合模型和非线性拟合模型两种定量预测模型,并以已知颠覆性论文证实了方法具有较高精确度和召回率;在此基础上,又从多个维度出发对比已知颠覆性论文在被引频次、引文地理多样性(所有施引文献作者所属国家的数量)和跨学科性3个指标上的排名^[19],发现引文地理多样性与被引频次在排名上具有较高的一致性,认为可将引文地理多样性作为补充指标以提高预测的准确性。

2.1.2 基于睡美人文献的识别方法

睡美人文献是指发表初期零被引或低被引、一段时间后突然高被引的论文,属于科学领域的延迟承认现象,其引文曲线被称为睡美人引文曲线。有学者指出,睡美人文献的本质是科学研究的颠覆性或超前性^[20]。而对颠覆性研究案例的分析表明,许多颠覆性研究超出人们当前的认知范畴,易被科学共同体低估或抵制,需要经过较长时间才能被认可并获得引文影响力^[13],这也从另一角度表明睡美人文献是颠覆性研究的重要来源。因此,有研究通过发掘睡美人文献或延迟承认现象探测颠覆性研究。Du等^[21]指出睡美人文献与颠覆性研究存在极高的关联度,并基于引文曲线提出识别

睡美人文献的无参数指标,对睡美人文献及其典型案例的特征进行深入挖掘,通过对睡美人文献内容分析发现,其表征的创新成果多是颠覆性研究。

基于引文曲线的颠覆性研究识别和模型构建通常是回顾性的,属于基于历史数据的后验方法,仍需要借助引文分析以外的其他方法确定其颠覆性,进而根据案例中的引文模式进行比较和预测。

2.2 基于引文网络的识别方法

文献之间通过直接引用、共被引、引文耦合及互引等方式构成各种引用关系,这种由相互引证形成的引用关系结构称为引文网络。引文网络能够体现文献研究主题间的相关性,引文网络的结构能反映文献研究主题的演变和引发研究范式转变的颠覆性研究。

2.2.1 基于知识组合的新颖性识别方法

Uzzi等^[22]将科学研究视为知识组合的过程,认为颠覆性研究可能是多个领域已有知识以前所未有的方式进行重组的结果,同时将参考文献视为论文知识来源,利用参考文献所在期刊两两组合的新颖性来测度知识组合的创新程度。基于随机引文网络对WoS所有论文从期刊共被引角度进行分析,在将所有参考文献映射为相应期刊后,计算所有期刊对的Z分数,Z分数为负表示组合具有新颖性,反之为常规知识组合,结果发现同时具有高度新颖性和常规性知识组合的研究更易获得高影响力。Lee等^[23]继承了Uzzi等的思路,以论文参考文献(先前研究成果)两两组合的稀缺性来评价论文的创新性程度,并应用于测度科研团队的创造力。Wang等^[24]也从期刊组合层面入手,通过测度论文参考文献所在期刊的组合是否为首次来衡量研究颠覆性,并基于期刊相似度构建新颖性指标,发现颠覆性研究可能由于尝试新方法而具有跨学科性,往往在外领域被高度引用,建议使用学科标准化的指标进行测度。

一般来说,知识组合的新颖性是颠覆性研究的基本特征,但新颖性的产生源于已有知识,因此颠覆性知识与传统知识之间的平衡对于创新和影响之间的联系至关重要。也因此,Uzzi等^[22]认为高影响力研究通常是在已有研究的常规组合基础上,又具有新颖性组合的不断出现,同时极易成为高被引论文。

2.2.2 基于引文结构的识别方法

在引文分析视角下,科学范式从不断的科学进步到发生科学革命,发生了各种规模的引用级联,通过分析颠覆性研究论文的“增强效应”,能够揭示颠覆性思想如何实现突破,并在现有科学范式中占据优势。Mazlounian等^[25]通过挖掘诺贝尔奖论文数百万次引用的基础上,定量分析了推动科学范式转变的过程,他认为范式的转变是由于作者的某项颠覆性研究增强了作者先前研究的影响力,使得作者的突破性成果能够与现有范式进行竞争,提出能够监测引文结构发生变化的增强因子(boost factor),该指标是用作者发表某节点论文前后其所有论文的篇均被引量比值来表示,其值越大颠覆性越强。随后,Huang等^[26]认为颠覆性研究将导致现有范式下由关键文献组成的引文路径发生“破裂”,提出破裂分数(disruption score),即颠覆性研究的出现破坏了现有的引文路径,逐渐代替现有研究范

式的代表性论文被后续论文引用,并形成新的研究范式。与传统引文分析方法相比,潜在“助推因素”的挖掘可以更早地发现颠覆性研究和团队。

Funk等^[27]基于专利的动态引文网络提出测度技术颠覆性的指标(见表1,公式1,公式2),该指标通过测度新专利的出现对已有引文网络造成的巩固或破坏,来表征新专利对现有技术稳定性的影响,进而评估其颠覆性;Wu等^[8]将指标改进后提出颠覆因子(Disruption)用于识别颠覆性研究(见表1,公式3),并以获得诺贝尔奖的论文(高颠覆因子)、综述论文(低颠覆因子)、专家调查、关键词汇的使用来验证颠覆因子。计算颠覆因子的基本思路:若焦点文献(待测文献)的施引文献较多地引用其本身而非参考文献,则这篇论文是对先前成果的颠覆,颠覆因子越接近于1研究越具有颠覆性。研究发现,在科学和技术领域原创性或颠覆性研究往往是小团队的研究成果,且颠覆性研究多存在引用延迟现象。

表1 颠覆因子的不同计算方法

计算原理	基本公式	参数含义
描述一项新专利 <i>i</i> 对已有引文网络造成的影响。若引入新专利使得焦点专利(待测专利)的后向引用专利被引量减少,则说明焦点专利具有颠覆性 ^[28]	公式1: $D = \frac{1}{n} \sum_{i=1}^n (-2f_i b_i + f_i)$	<i>i</i> 为引用焦点专利或其后续引用专利的专利, <i>n</i> 为其数量。 <i>i</i> 引用焦点专利时 $f_i=1$,否则 $f_i=0$; <i>i</i> 引用焦点专利的后向引用专利时 $b_i=1$,否则 $b_i=0$
公式1仅能反映焦点专利对其后向引用专利影响方向,故引入参数 <i>m</i> 以反映其影响程度 ^[28]	公式2: $mD = \frac{m}{n} \sum_{i=1}^n (-2f_i b_i + f_i)$	参数 <i>m</i> 为只引用焦点专利的专利数,其他变量含义不变
计算只引用焦点文献(待测文献)和既引用焦点文献又引用其参考文献的论文的概率之差,即 $p_F - p_B$ ^[8]	公式3: $D = p_F - p_B = \frac{N_F - N_B}{N_F + N_B + N_R}$	N_F 代表只引用焦点文献的论文数, N_B 代表同时引用焦点文献及其参考文献的论文数, N_R 代表只引用其参考文献的论文数

颠覆因子的提出引起学界的广泛关注和讨论,Bornmann等^[28]发现某些论文在发表之初并未显示其颠覆性,随着时间的推移颠覆因子趋于稳定,研究的颠覆性才显现出来,指出颠覆因子的值与论文引用时间窗的长度有关,建议使用至少为期3年的引用时间窗。随后,Bornmann等^[29]又计算了2000—2010年《科学计量学》上所有论文的颠覆因子,对高颠覆因子论文(颠覆因子阈值设置为Top1%)的相关信息进行了统计分析,发现颠覆因子由焦点文献的被引文献主导,使用时必须考虑引用行为的影响。

在库恩理论中,科学的转折点通常表现为科学新发现、经典著作或知识突破点等关键节点的出现。陈超美等^[30]以此为基础设计了CiteSpace,在对知识突破点进行探测时,利用渐进式可视化反映共被引网络的演变过程,通过识别潜在重要节点来探测里程碑式的颠覆性研究论文;后又提出结构变异理论(Theory of Structural Variation)来探测文献的潜在颠覆性,其基本思想是科学发现或创新会在差异性较大的网络间产生新颖性连接,并且会快速吸引大量的引用,使网络整体结构发生较大变化。同样,也有研究通过比较不同时期专利引文网络主路径中重要新节点的出现来识别潜在颠覆性研究^[31]。Foster等^[32]通过识别提高网络连通性的节点或连线来测度颠覆性研究,认为如果一项研究能将两个引文网络上的文献联通,或者能在两个未联

2.2.3 基于关键节点的识别方法

通过测度科学结构的变化可以探测颠覆性研究,

通的研究间建立联系,使原本松散的网络结构变得紧密,则这项研究具备颠覆性。

此外,还有研究在对论文引文网络进行拓扑测量及可视化的基础上,利用节点的类内聚集度和参与系数来识别颠覆性研究^[33]。某节点的类内聚集度表示其与类内其他节点的连接程度,参与系数指其所有连接在其他类内的分布状况,颠覆性研究通常类内聚集度逐渐增高、参与系数逐渐降低。

2.3 基于多源数据的组合识别方法

数据源的可靠性与充分程度在较大程度上影响识别颠覆性研究的准确率,单一数据源用于信息分析的有效性往往有限,同时单一指标也无法全面测度研究的颠覆性,许多研究人员尝试构建多源数据融合组合识别方法,在融合论文与专利数据的基础上,从颠覆性研究的典型特征入手,构建多维指标进行分析。

知识突变是颠覆性研究导致研究范式转变的特征之一,基于多源数据对其定量分析可以捕捉和识别早期颠覆性研究。张金柱等^[34]以专利直接和间接引用的科学论文关键词、学科分类为研究对象,在对比不同时间段关键词簇或学科分类簇、研究主题和学科分类组合的差异程度基础上,计算突变程度指标,预测可能产生颠覆性技术的时间、研究主题和学科分类组合。白光祖等^[35]从文献知识的内外外部关联入手,对核心专利、核心论文分别引用论文产生的知识突变进行分析,识别内外外部潜在的颠覆性研究主题。此外,从科学与技术的交叉

中识别创新前沿,可以揭示和判别可能发生突破性或颠覆性研究的领域和主题,杜建等^[36]从科学与技术交叉渗透特征入手,建立科学与技术交叉象限模型,分别将引用专利的ESI高被引论文共被引聚类、引用基础科学论文的三方专利耦合聚类作为识别科学维与技术维颠覆性研究的指标,通过论文共被引和专利耦合分析聚类识别并命名创新前沿。

多维测度指标的组合运用是目前情报学领域的趋势,Ehrnberg等^[37]将文献计量指标与工业指标进行组合,基于相关专利及文献数量的激增来预测和判断颠覆性研究。还有学者通过被专利引用和延迟承认指数高2个指标识别科学与技术交叉处的研究内容,经专家鉴定这些研究主题与颠覆性研究密切相关性^[8]。此外,Jones等^[38]以论文的参考文献为切入点,认为高被引论文的研究基础时间跨度与研究类型有关,并基于后向引用建立测度参考文献平均年龄的BCiteAge指标,用于表征当前工作与先前工作的时间距离及对先前工作的借鉴程度,研究表明颠覆性研究倾向于引用较年轻的参考文献。进而Staudt等^[39]基于文本和引文构建识别颠覆性研究的指标体系,并将11个指标依照颠覆性研究的特点分为破坏性、高风险性、多学科性等7种类型,利用因子分析选出代表性指标用于识别颠覆性研究(见表2),并且发现各指标大多存在正相关关系,部分弱相关或不相关,这表明颠覆性研究的各个特征存在一定独立性,且在不同时期和不同领域,颠覆性和影响力均呈正相关。

表2 用于识别颠覆性研究的主要指标

指标类型	代表性指标	计算原理
破坏性	BCiteAge	范式转变与参考文献平均年龄的相关性
由革命性思想驱动	Concepts	利用N-Grams模型测度颠覆性思想的诞生
高风险性	FCiteVar	利用前向引用的方差来衡量文献的风险性
多学科性	BHerfCite	通过后向引用反映某文献的被引文献在多个学科中分布的离散程度
影响的广泛性	FHerfCite	通过前向引用反映某文献的施引文献在多个学科中分布的离散程度
影响的深远性	FCiteAge	利用文献发表年与被引用年差值变化的趋势来衡量影响的变化程度
影响力	FCiteMean	计算平均前向引用数以测度文献影响力

3 讨论

综上所述,随着文献计量学理论和方法的不断拓展和深入,文献计量学应用研究在情报学和管理学领

域扮演着越来越重要的角色,在科学评价、研究前沿识别和学科服务实践中发挥重要作用。近年来,国内外学者不断尝试和探索将文献计量学应用于颠覆性研究的识别,但大多仍是基于已有指标和方法的组合,并未形

成独特的研究方法和范式。通过比较上述研究方法可以发现,除无法突破引文分析所固有的时滞性和数据源单一等缺陷,颠覆性研究的文献计量学识别方法还存在以下方面的问题。

(1) 颠覆性研究产生因素复杂,识别准确度不高。如前所述,颠覆性研究具有高风险性和偶然性等特征,在成长过程中受到社会、科学、政治和经济等各种环境因素的影响,早期文献计量特征并不明显。尽管在一定程度上能够识别并区分颠覆性研究和渐进性研究,但得到的潜在颠覆性研究数量仍然很大,对于其后期能够成长为真正的颠覆性研究仍缺乏更多依据支撑,导致分析结果的准确度和有效性降低。

(2) 数据源和方法难以交叉融合,识别维度单一。颠覆性研究的识别与研究前沿的识别有许多相似之处,相较而言,研究前沿的识别具有更可靠的文献计量学特征和理论基础,而颠覆性研究的识别仍处于尝试阶段。对颠覆性研究早期计量特征的捕捉和探测难度更大,需要突破已有的论文和专利数据等数据源限制。此外,现有的综合识别方法仍停留在对单一指标的简单组合,不能有效地将多种方法充分交叉融合,发挥互相补充和相互验证的作用。

结合以上分析,颠覆性研究识别方法的未来发展方向可从以下方面展开。

(1) 充分结合文献外在特征和内容特征,提高识别的时效性。尽管引文分析法不断改进,但仍难以避免时滞性,同时外在引用指标难以测算文献的内在研究本质。基于论文内容的识别方法,利用文本挖掘技术能够第一时间直接分析科技文本的内容特征,弥补引文信息的滞后性。未来可以考虑引文分析和内容分析的结合,基于文本挖掘等技术识别文献信息内容,同时注重在引文关系和主题内容中的“异常”,这些研究主题更可能成为潜在的颠覆性研究。

(2) 不断拓展和融合多源数据,增强识别的前瞻性。当前,颠覆性研究的文献计量学识别方法仍以学术论文和专利数据为主,识别结果的时效性和前瞻性有所欠缺。尽管已有研究以论文和专利的交叉作为科学与技术交叉的基础测度指标,但其他数据源的拓展应用并不多见。因此,未来应综合利用当前情报学领域的各种数据源(如科技媒体报道、科研项目数据和文本、网络信息和科学家个人主页等多源信息),借助大数据处理分析技术,以促进多源数据的交叉融合和相互验证。

(3) 介入专家知识研判研究内容,提高识别的准确度。将定性定量识别方法相结合,能够同时发挥专家判断、客观分析的优势。目前已出现部分综合识别方法,但并未将二者充分交叉渗透,未来研究应促进专家研判与计量分析的有机结合,进一步细化分析粒度。构建全方位的颠覆性研究识别指标体系,有助于专家掌握足够信息,进而在更高信息集成水平上形成更具权威性的意见。

4 结语

基于文献计量方法的颠覆性研究识别和预测仍处于摸索阶段,还没有形成成熟可靠的识别理论和方法,目前仍以传统引文分析方法为主,即以论文的外在特征代替论文的内在特征进行价值和质量判断,或通过分析公认颠覆性研究案例来研判颠覆性研究外在特征。但文献计量方法对于颠覆性研究的识别仍具有重要的理论意义和应用价值,尤其随着信息技术和文本处理技术的发展,颠覆性研究识别方法将会在识别的准确度、时效性和前瞻性等方面得到进一步提升和完善。当然,最终的分析结果必须获得领域专家认可,方能真正发挥其推动学科发展、提高国家综合竞争实力的作用和优势。

参考文献

- [1] KUHN T S. The structure of scientific revolution: Second Edition [M]. Chicago: University of Chicago press, 1970.
- [2] SHATZ D. Peer review: A critical inquiry [M]. Lanham: Rowman & Littlefield, 2004.
- [3] CHRISTENSEN C M. The innovator's dilemma: The revolutionary book that will change the way you do business [M]. New York: Harper Business Essentials, 2003.
- [4] 张金柱, 张晓林. 基于科技资源的突破性创新指标及识别方法综述 [J]. 图书情报工作, 2012, 56 (22): 56-61.
- [5] 杜建, 孙轶楠, 张阳, 等. 变革性研究的科学计量学特征与早期识别方法 [J]. 中国科学基金, 2019, 33 (1): 88-98.
- [6] KOSTOFF R N, BOYLAN R, SIMONSG R. Disruptive technology roadmaps [J]. Technological Forecasting and Social Change, 2004, 71 (1/2): 141-159.
- [7] 杨卫, 郑永和, 董超. 如何评审具有颠覆性创新的基础研究 [J]. 中国科学基金, 2017, 31 (4): 313-315.

- [8] WU L, WANG D, EVANS J A. Large teams develop and small teams disrupt science and technology [J]. *Nature*, 2019, 566 (7744): 378.
- [9] National Academies of Sciences, Engineering, and Medicine. *Fostering transformative research in the geographical sciences* [M]. Pittsburgh: National Academies Press, 2015.
- [10] SANGWAL K. On the growth dynamics of citations of articles by some Nobel Prize winners [J]. *Journal of Informetrics*, 2015, 9 (3): 466-476.
- [11] 郑永和, 陈淮. 美国国家科学基金会加强支持变革性研究考察 [J]. *中国基础科学*, 2008 (4): 39-42.
- [12] WINNINK J J, TIJSSSEN R J W. Early stage identification of breakthroughs at the interface of science and technology: lessons drawn from a landmark publication [J]. *Scientometrics*, 2015, 102 (1): 113-134.
- [13] WANG D, SONG C, BARABÁSI A L. Quantifying long-term scientific impact [J]. *Science*, 2013, 342 (6154): 127-132.
- [14] 李江, 姜明利, 李玥婷. 引文曲线的分析框架研究——以诺贝尔奖得主的引文曲线为例 [J]. *中国图书馆学报*, 2014, 40 (2): 41-49.
- [15] GARFIELD E, WELLJAMS-DOROF A. Of Nobel class: a citation perspective on highimpact research authors [J]. *Theoretical Medicine*, 1992, 13 (2): 117-135.
- [16] BORNMAN L, YEA, YE F. Identifying landmark publications in the long run using field-normalized citation data [J]. *Journal of Documentation*, 2018, 74 (2): 278-288.
- [17] SCHNEIDERJ W, COSTAS R. Identifying potential “break through” publications using refined citation analyses: Three related explorative approaches [J]. *Journal of the Association for Information Science and Technology*, 2017, 68 (3): 709-723.
- [18] PONOMAREV I V, LAWTON B K, WILLIAMS D E, et al. Breakthrough paper indicator 2.0: can geographical diversity and interdisciplinarity improve the accuracy of outstanding papers prediction? [J]. *Scientometrics*, 2014, 100 (3): 755-765.
- [19] PONOMAREV I V, WILLIAMS D E, HACKETT C J, et al. Predicting highly cited papers: a method for early detection of candidate breakthroughs [J]. *Technological Forecasting and Social Change*, 2014, 81: 49-55.
- [20] PALOMERASN. Sleeping patents: any reason to wake up [EB/OL]. [2020-01-01]. <https://pdfs.semanticscholar.org/98b2/6d5977f1719690634ce0786cc7c9dce73be1.pdf>.
- [21] DU J, TANG X, WUY. The effects of research level and article type on the differences between citation metrics and F 1000 recommendations [J]. *Journal of the Association for Information Science and Technology*, 2016, 67 (12): 3008-3021.
- [22] UZZI B, MUKHERJEE S, STRINGER M, et al. A typical combinations and scientific impact [J]. *Science*, 2013, 342 (6157): 468-472.
- [23] LEE Y N, WALSHJ P, WANG J. Creativity in scientific teams: Unpacking novelty and impact [J]. *Research Policy*, 2015, 44 (3): 684-697.
- [24] WANG J, VEUGELERS R, STEPHAN P. Bias against novelty in science: a cautionary tale for users of bibliometric indicators [J]. *Research Policy*, 2017, 46 (8): 1416-1436.
- [25] MAZLOUMIAN A, EOM Y H, HELBING D, et al. How citation boosts promote scientific paradigm shifts and nobel prizes [J]. *PLOS ONE*, 2011, 6 (5): e18975.
- [26] HUANG Y H, HSUC N, LERMAN K. Identifying transformative scientific research [C] // 2013 IEEE 13th International Conference on Data Mining. IEEE, 2013: 291-300.
- [27] FUNKR J, OWEN-SMITH J. A dynamic network measure of technological change [J]. *Management Science*, 2016, 63 (3): 791-817.
- [28] BORNMAN L, TEKLES A. Disruption index depends on length of citation window [J]. *El Profesional De La Información*, 2019, 28 (2): 1-2.
- [29] BORNMAN L, TEKLES A. Disruptive papers published in *Scientometrics* [J]. *Scientometrics*, 2019, 120 (1): 331-336.
- [30] CHEN C. Searching for intellectual turning points: Progressive knowledge domain visualization [J]. *Proceedings of the National Academy of Sciences*, 2004, 101 (suppl): 5303-5310.
- [31] MARTINELLI A. *The dynamics of technological discontinuities: a patent citation network analysis of telecommunication switches* [D]. Eindhoven: Eindhoven University of Technology, 2010.
- [32] FOSTER J G, RZHETSKY A, EVANS J A. Tradition and innovation in scientists’ research strategies [J]. *American Sociological Review*, 2015, 80 (5): 875-908.
- [33] SHIBATA N, KAJIKAWA Y, TAKEDA Y, et al. Detecting emerging research fronts based on topological measures in

- citation networks of scientific publications [J]. Technovation, 2008, 28 (11): 758-775.
- [34] 张金柱, 张晓林. 利用引用科学知识突变识别突破性创新 [J]. 情报学报, 2014, 33 (3): 259-266.
- [35] 白光祖, 郑玉荣, 吴新年, 等. 基于文献知识关联的颠覆性技术预见方法研究与实证 [J]. 情报杂志, 2017, 36 (9): 38-44.
- [36] 杜建, 孙轶楠, 李永洁, 等. 从科学-技术交叉处识别创新前沿: 方法与实证 [J]. 情报理论与实践, 2019, 42 (1): 94-99.
- [37] EHRNBERG E, JACOBSSON S. Indicators of discontinuous technological change: an exploratory study of two discontinuities in the machine tool industry [J]. R&D Management, 1997, 27 (2): 107-126.
- [38] JONES B F, WEINBERG B A. Age dynamics in scientific creativity [J]. Proceedings of the National Academy of Sciences, 2011, 108 (47): 18910-18914.
- [39] STAUDT J, YU H, LIGHT R P, et al. Highimpact and Transformative Science (HITS) metrics: Definition, exemplification, and comparison [J]. PLOS ONE, 2018, 13 (7): e0200597.

作者简介

郭丽娜, 女, 1996年生, 硕士研究生, 研究方向: 科学计量与评价。

李星琛, 女, 1997年生, 硕士研究生, 研究方向: 科学计量与评价。

左文革, 女, 1966年生, 硕士, 研究馆员, 研究方向: 信息计量与科学评价。

周群, 男, 1982年生, 博士, 副研究馆员, 通信作者, 研究方向: 信息计量与情报分析方法, E-mail: qzhou@cau.edu.cn。

A Review of Disruptive Work and Its Bibliometric Identification Methods

GUO LiNa LI XingChen ZUO WenGe ZHOU Qun
(Library of China Agricultural University, Beijing 100193, China)

Abstract: Disruptive work refers to challenging or disrupting the original research paradigm and creating new paradigms or new areas of research. The connotation, characteristics and bibliometric identification methods of disruptive work provide reference for the research of the identification method and index optimization of disruptive work. The literature investigation method is used to systematically sort out related literatures, and it is divided into four categories based on citation curve, citation path, citation network identification method and multi-source data based combination identification method, pointing out the current challenges and development trends. At present, the citation analysis method of disruptive work still has problems such as time lag, difficulty in measuring dimension, and difficulty in cross-integration of data sources. In the future, based on the full investigation of historical cases of disruptive work, summarize the characteristics of potential disruptive work, and combine content analysis, big data and expert knowledge to improve the accuracy and forward-looking of identification.

Keywords: Disruptive Work; Bibliometrics; Identification Methods

(收稿日期: 2020-02-07)