技术知识基础跨度对专利颠覆性的影响研究*

吴海婷[1] 董克[1,2] 贾呈华[1] 吴佳纯[1]
（1.南京大学数据管理创新研究中心，苏州 215163；2.南京大学图书馆，南京 210023）

摘要：专利的技术知识基础跨度指专利所吸收技术知识的时间跨度与技术领域覆盖程度。为探究技术知识基础跨度对专利颠覆性的影响，本研究以美国专利商标局授权发明专利为样本，采用D指数测度专利颠覆性，从时间和领域两个维度构建专利的技术知识基础跨度测度指标，并综合运用描述性统计、逻辑回归和最小二乘回归进行实证分析。结果表明，技术知识基础跨度在时间维度与领域维度上均对专利颠覆性存在显著影响：在时间维度上，引证专利的时效性越高、平均引用年龄越大、时间覆盖广度越小，专利颠覆性越高；在领域维度上，引证专利的领域丰富性越低、均衡性越高、差异性越大，专利颠覆性越高。

关键词：技术知识基础跨度；专利颠覆性；D指数；专利评价
中图分类号：G255.53 DOI：10.3772/j.issn.1673-2286.2026.03.007
引文格式：吴海婷，董克，贾呈华，等. 技术知识基础跨度对专利颠覆性的影响研究[J]. 数字图书馆论坛，2026，22（3）：56-65.

颠覆性创新是一种能开创新范式，从根本上变革原有理论、方法或技术体系，并最终引发科技革命与产业变革的科技创新模式。任何颠覆性创新都不是凭空产生，而是建立在对已有研究成果的延伸和拓展之上。在规范的知识生态系统中，已有知识是颠覆性创新的基础，引用反映了颠覆性创新对现有知识的继承与依赖。颠覆性创新的知识基础往往具有跨度特征，例如，引文年龄分布反映知识基础的时间跨度[1]、跨学科性则反映其学科领域跨度[2-4]等。这些跨度特征反映了创新主体对已有知识的整合范围，会对创新的颠覆性程度产生影响。有研究[5]发现，在科技成果数量指数级增长态势下，创新主体过度依赖较狭隘范围的知识成果会对颠覆性创新程度产生负面作用。

专利是科技创新重要成果的产出形式，分析专利颠覆性的影响因素有助于深化对颠覆性创新内在机理的认知。在技术创新系统中，专利的知识基础主要由科学知识基础和技术知识基础两部分构成，前者体现为对科学文献的引用，后者反映在对专利技术的引用上[6]。现实中，作为技术创新系统的重要组成，专利创新对已有专利的依赖程度远高于科学知识，表现为引证专利文献数量远高于被引科学文献数量[7]；此外，专利具有明显的排他性特征，其引用行为受到法律法规的严格约束，使得专利与引证专利之间实质性相关程度较高，利用引证专利来反映技术知识的交流可靠性也较高[8]。因此，本研究以专利为研究对象，从时间和领域两个维度切入，构建专利的技术知识基础跨度测度指标体系，分析技术知识基础跨度对专利颠覆性创新的影响，旨在揭示专利颠覆性创新形成机制，为识别潜在的颠覆性创新专利提供新视角。

收稿日期：2026-01-20
*本研究得到国家社会科学基金重大项目“基于数智融合的信息分析方法创新与应用研究”（编号：22&ZD326）资助。

1　相关研究

1.1　知识基础跨度研究

引文知识能较准确反映科技成果对已有知识的吸收与继承特征[9]，因此可通过引文特征来定位科技成果的知识基础。现有知识基础跨度的研究主要集中在时间维度和领域维度。在时间维度上，引文的时间特征能反映某个领域知识的扩散和继承特点，以及知识的增长与老化程度，从而揭示领域知识整体的演化特征。具体分析中，引文的时间分布特点可以通过多种方式进行测度：四分位距即引文年份的第3四分位数与第1四分位数之差，通过中间50%引文的时间跨度来反映引文时间分布的集中范围[6]；普赖斯指数等指标能够表征引文的时效性[10]。

在领域维度上，引文的领域分布能反映知识基础的交叉、整合程度等特征。现有研究[11−12]分别从引文的丰富性、均衡性和差异性等角度进行分析，发现引文所涉及学科领域的跨度越大、分布越均衡、学科间差异越显著，其知识基础的整合程度往往越高。此外，还有研究[7]引入产业经济学中的赫芬达尔指数来刻画引文的领域分布特点。该指数原用于衡量市场集中度，因引文在各学科领域的分布结构与市场份额的分布结构具有同构性，故可迁移用于测度引文领域分布的集中或分散程度。

上述研究为知识基础跨度测度提供了多种指标，但主要针对科学论文，其指标设计也以科学知识基础的测度为应用情境。少量与专利相关的研究主要借鉴或应用科学论文知识基础跨度研究的方法，而专利技术与科学论文在知识结构与引用动机上存在差异，导致这些指标在专利情境中的适用性缺乏验证，这也是本研究的出发点之一。

1.2　颠覆性创新与引用的关系研究

颠覆性创新研究已涵盖测度方法、影响因素、技术识别与预测等多个维度[13−15]，其中引用行为作为揭示创新颠覆性的重要切入点，日益受到学界关注。一方面，颠覆性创新往往并非凭空产生，而是源于对已有知识的创造性组合。研究[16]发现，引文知识的非常规性组合是颠覆性创新的重要来源，引文数量也被证实是影响科技成果颠覆性的重要因素之一[17]。另一方面，颠覆性成果往往对现有知识体系产生巨大的冲击，在引文网络中表现为局部结构的中断和重构[18]。由于颠覆性成果的高创新价值，施引成果在引用该成果时，往往不再引用其参考文献，从而导致先前知识基础被直接利用的程度下降。这种现象本质上反映了成果颠覆性对现有知识体系的影响能力，表现为后续施引文献的引用模式变革，学者们据此设计了CD指数[19]、D指数[20]等指标用于颠覆性测度。其中，D指数综合考虑了引用行为的模式和连续性，在海量数据集中的反复应用进一步验证了其可靠性和有效性[21−22]，在科技创新成果的颠覆性计算方面具有明显优势。

上述研究表明，颠覆性与引用之间存在关联，引文特征已成为揭示颠覆性形成机理的重要线索。然而，现有研究在追溯颠覆性的知识来源时，多以科学论文的参考文献为分析对象，对专利引用所承载的技术知识基础关注不足；专利作为技术创新的直接载体，其知识吸收与重组主要依托专利间的引用关系来实现，引证专利在时间与领域上的分布特征如何影响专利颠覆性，仍有待系统分析。因此，本研究聚焦引证专利的时间特征和领域特征，刻画技术知识基础在这两个维度上的跨度特征，从而揭示技术知识基础跨度对专利颠覆性的影响。

2　研究设计

2.1　研究思路

引文是专利知识基础的直接载体，其分布特征可从时间和领域两个维度来刻画：时间维度能够反映技术知识基础的演化过程，揭示了专利在创新过程中对先前知识的纵向整合能力；领域维度能够反映技术知识基础的领域覆盖面，揭示了专利在创新过程中对先前知识的横向整合能力。

从时间维度来看，如果专利引用了较长时间范围内的专利文献，往往表明专利所承载的技术创新对技术的发展脉络和演化轨迹有较为深刻的把握[6]。然而，不连续的非线性的技术轨道跃迁是颠覆性创新的重要特征之一[23]，因此，如果专利大量引用已内化为领域常识或基础的早期技术知识，就有可能导致专利难以突破传统技术框架，进而陷入技术连续性的路径依赖困境[24]。反之，如果专利大量引用近期专利，虽能更及时地吸收最新技术知识，但其能否带来颠覆性提高，尚无定论。

从领域维度来看，引用多领域的技术知识，能打破单一领域的思维局限和技术路径依赖，为专利创新提供更为丰富的知识素材和创新视角，有助于催生颠覆

性创新成果[25]。然而，跨领域知识重组的效益并非线性递增。当引用领域过于分散时，发明人有限的认知资源可能被稀释，难以对知识形成深度理解和根本性重组，知识整合也可能流于表面化和碎片化[26]。因此，领域维度的技术知识基础跨度对专利颠覆性的影响并非单向作用，而是同时受到知识吸收效益与知识整合成本的共同驱动。

基于以上分析，本研究将专利的技术知识基础跨度界定为专利所吸收技术知识的时间跨度与技术领域覆盖程度，用以衡量专利在创新过程中对先前知识的纵向与横向整合能力。

2.2 研究框架

本研究运用回归分析方法，探究技术知识基础跨度对专利颠覆性的影响，研究框架如图1所示。第一，数据获取与处理。以美国专利商标局（United States Patent and Trademark Office，USPTO）专利数据库为数据来源，经采集、筛选得到样本专利集，并进一步获取其引证专利集和施引专利集，为后续的研究提供数据基础。第二，变量选择与测度。以专利颠覆性指数为因变量，从时间和领域两个维度构建技术知识基础跨度测度指标体系。其中，时间维度以普赖斯指数、平均引用年龄和引证时间跨度来刻画，领域维度则采用丰富性、均衡性和差异性进行测度，以全面捕捉技术知识基础的分布特征。第三，结果与分析。考虑到专利颠覆性兼具定性与定量双重属性，先采用二元逻辑回归模型判断各指标对专利是否具有颠覆性的影响，再采用最小二乘回归模型分析各指标对颠覆性强度的作用大小，从而得出研究结论。

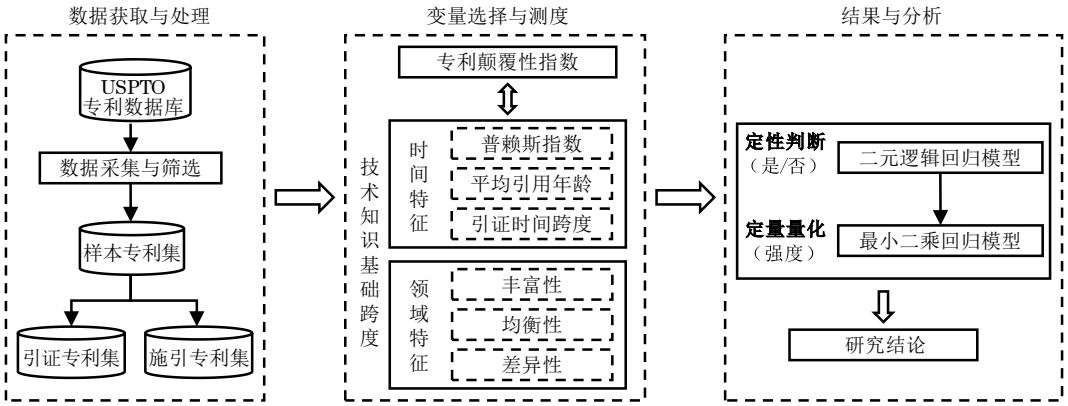

图1 研究框架

3 数据与方法

3.1 数据获取与处理

本研究以USPTO专利数据库为数据源。该库完整收录了1976年以来在美国获得授权的专利数据[27]，且专利引文制度相对完善[24]。其中专利引文包括申请人标注的引文和审查员补充的引文两类：前者是申请人依据现有技术披露义务主动标注的引证专利，直接反映了申请人的知识搜索行为；后者是审查员在审查过程中补充的申请人“当引未引”的相关引证专利[28]。尽管审查员补充的引文并非申请人主动选择，但根据美国专利商标局审查制度，此类引文须与专利的技术方案存在实质性关联，同样构成技术知识基础的重要组成部分。为此，本研究将两类引文合并使用。

为避免因专利类型差异而产生影响，本研究剔除外观设计专利、植物专利、再公告专利等其他类型专利数据，仅保留授权发明专利。美国发明专利的法定保护期为20年，在保护期内专利技术具有排他性，引用行为能更真实地反映技术知识的依赖关系；而专利失效后，它所蕴含的技术知识可能逐渐成为领域共识，影响后续引用。为此，本研究选取2005年为样本专利授权年份的起点，以保证样本专利在后续整个观测期内均处于有效保护状态。同时，为保证样本具有充分的时间跨度与代表性，本研究以10年为样本时间窗口。本研究在USPTO数据库中获取2005—2014年授权发明专利，共2 315 265项。

一方面，专利颠覆性指数的计算受引文时间窗口长

度的影响，当专利的被引量相对稳定后，才能计算出更接近专利真实创新水平的颠覆性指数，而专利从申请到授权公开通常需要经历较长时间，引用相对滞后，因此，本研究借鉴丁乐蓉等[29]所做研究，将引证时间窗口设定为5年，即考察样本专利自授权当年起5年内的被引情况。另一方面，专利颠覆性指数的计算还需有足够的引证专利数量和施引专利数量[30]，且两者的阈值通常设置为不少于5项[31−32]，为此，本研究进一步筛选出引证专利数量不少于5项且在5年时间窗口内获得至少5次引用的专利作为研究样本。最终，从USPTO数据库中获得308 351项授权发明专利作为样本专利，并提取其引证专利共12 994 742项，以及截至2018年引用样本专利的施引专利共4 576 218项。

3.2　变量选择与测度

3.2.1　因变量

专利颠覆性是本研究的因变量，指专利创造性破坏与变革现有技术体系或技术范式的程度。本研究关注技术创新系统中技术知识基础跨度对专利颠覆性的影响，因此重点从技术知识空间改变的视角来定义专利颠覆性，并采用D指数[33]进行测度，计算公式如式（1）所示。

$$D=\frac{n_i-n_j}{n_i+n_j+n_k} \tag{1}$$

其中，i表示仅引用样本专利的施引专利，j表示同时引用样本专利及其引证专利的施引专利，k表示仅引用样本专利的引证专利而未引用样本专利的施引专利；n_i、n_j、n_k分别对应i类、j类、k类专利的数量。D指数的取值范围为[−1，1]，当D指数≥0时，$n_i \geqslant n_j$，表明后续施引专利更倾向于引用样本专利而非其引证专利，说明样本专利的技术轨道可能实现了突破，体现出颠覆性特征；当D指数＜0时，$n_i < n_j$，表明后续施引专利仍倾向于引用样本专利的引证专利，说明样本专利更多延续了先前专利的技术路径，属于渐进性创新。

3.2.2　自变量

技术知识基础跨度是本研究的自变量，主要从时间和领域两个维度进行界定。在时间维度上，技术知识基础跨度主要体现专利对技术前沿的敏锐度、所吸收技术知识的新旧程度和历史纵深。因此，本研究分别采用普赖斯指数[10]、平均引用年龄[34]和引证时间跨度3个指标进行测度。其中，普赖斯指数用于计算授权年份与样本专利相差不超过5年的引证专利数量在全部引证专利中所占的比例，以此反映专利引用的时效性；平均引用年龄以样本专利与其各引证专利授权年份之差的均值来衡量，反映专利引用的整体新旧程度和时间分布重心；引证时间跨度借鉴了跨时域性测度框架中“时域差异性”指标的思想[4]，以样本专利的全部引证专利中授权年份的最大值与最小值之差来衡量，反映专利引用的时间分布广度。

在领域维度上，技术知识基础跨度表现为专利引用的领域分布特征，主要体现为引证专利所涉及技术领域类别数量、领域分布的集中与分散程度、领域的异质性与知识距离3个方面。因此，本研究以世界知识产权组织的专利技术领域分类体系为依据，基于各专利主国际专利分类（International Patent Classification，IPC）与专利技术领域之间的映射关系，确定专利的所属领域[35−36]。该对照表将IPC分类号（以小类为映射单位）以多对一的方式映射至35个技术领域类别。首先，提取每项专利的主IPC分类号，获取其小类信息；其次，根据对照表将其映射至对应的技术领域；最后，依据跨学科性三维测度框架[37]，从丰富性、均衡性和差异性测度领域维度的技术知识基础跨度。

丰富性以引证专利所涉及技术领域的类别数量来衡量，数量越多，表明专利所吸收知识的领域覆盖面越广。均衡性用以反映引证专利的技术领域分布均匀程度。本研究采用基尼不纯度指标[38]进行测度，其值等于1与各技术领域占比平方和之差，取值范围为[0，1），数值越大，领域分布越均匀；当引证专利同属一个技术领域时，基尼不纯度为0；反之，当引证专利分布越均匀且领域类别数越多，基尼不纯度越接近于1。丰富性和均衡性均只关注引证专利的内部分布特征，差异性则需要引入样本专利作为参照，衡量二者在技术领域上的语义差异。在综合对比多种语义抽取方法后，本研究主要采用Word2Vec[39]进行专利语义识别。首先，以专利摘要作为训练数据集，调用Word2Vec的CBOW模型进行训练；其次，利用训练后的Word2Vec模型将专利技术领域集合映射为词向量；最后，计算样本专利词向量与引证专利词向量间的平均余弦相似度作为差异性的测度结果。

3.2.3 控制变量

本研究选取被引频次、引证专利数以及独立权利要求数作为控制变量。①从颠覆性的机理来看，被引频次则反映了该专利在后续创新活动中的受关注程度和技术贡献度[40]，高被引专利通常在技术发展中具有较强的知识溢出效应，可能因其技术领先性而对后续技术演化路径产生影响；此外，引证专利数反映了专利所依赖的技术基础来源及其对现有技术的整合程度，较多的引证专利数意味着对现有技术的集成度更高，可能会催生出复杂而具有颠覆性的技术[12]。②从计算角度来看，D指数在计算过程中的取值大小与引证专利数、被引频次密切相关，也表明上述两个指标均应作为控制变量。③从技术复杂性角度来看，较多的独立权利要求数通常意味着更广泛的技术保护范围和更高的技术方案复杂度[41-42]，这可能增强专利的技术壁垒和市场影响力，进而影响其颠覆性程度。

本研究的全部变量及测度如表1所示。

表1 变量及测度

变量类型	变量编号	变量名称	变量测度
因变量	Y	D指数	基于样本专利前后向引用关系计算的颠覆性指数
自变量	X1	普赖斯指数	引证专利中近5年授权的数量占引证专利总数的比例
	X2	平均引用年龄	样本专利授权年份与各引证专利授权年份差值的均值
	X3	引证时间跨度	引证专利中最晚授权年份与最早授权年份之差
	X4	丰富性	引证专利所涉及的技术领域类别数量
	X5	均衡性	引证专利技术领域分布的基尼不纯度
	X6	差异性	样本专利与引证专利技术领域词向量的平均余弦相似度
控制变量	X7	被引频次	样本专利被后续专利引用的总次数
	X8	引证专利数	样本专利的引证专利总数
	X9	独立权利要求数	样本专利权利要求书中独立权利要求的条数

4 结果与分析

4.1 描述性统计

各变量的描述性统计结果如表2所示。在自变量方面，从时间维度来看，引证专利的普赖斯指数、平均引用年龄与引证时间跨度的平均值分别为0.279、11.686和29.620，标准差分别为0.243、6.647和27.044；从领域维度来看，丰富性、均衡性和差异性的均值分别为4.150、0.389和0.381，标准差分别为3.322、0.253和0.315。总体而言，技术知识基础跨度在时间和领域两个维度上的标准差均小于平均值，表明数据分布相对较为集中，不存在过度分散问题。在因变量方面，D指数的平均值为0.000，标准差为0.085，整体显现出略微的右偏分布（偏度为0.737），说明多数样本的D指数趋近于0，仅少量样本具备较高D指数水平。

表2 变量的描述性统计结果

变量编号	平均值	最大值	最小值	标准差
Y	0.000	1.000	−1.000	0.085
X1	0.279	1.000	0.000	0.243
X2	11.686	94.793	0.000	6.647
X3	29.620	232.000	0.000	27.044
X4	4.150	33.000	1.000	3.322
X5	0.389	0.931	0.000	0.253
X6	0.381	1.000	0.000	0.315
X7	14.840	1 432.000	5.000	23.718
X8	42.140	5 786.000	5.000	83.336
X9	22.030	803.000	1.000	16.511

在控制变量方面，样本专利的平均被引频次为14.840，但最大被引频次高达1 432次，标准差为23.718且大于平均值，表明专利的被引频次分布不均，离散程度较高。引证专利数的均值为42.140，最大值高达5 786次，标准差为83.336且远高于平均值，反映出样本专利的引证专利数量分布高度不均。样本专利的平均独立权利要求数为22.030条，最大值为803条，标准差为16.511，表明样本专利的独立权利要求数整体上数据分布集中度相对较好，但存在部分极端异常值。

4.2 相关性分析

由于各变量的极差、平均值、方差差异明显，且部分变量的分布并不符合正态分布，各变量与因变量之间也可能并非一般线性关系，因此本研究选择斯皮尔曼相关系数来检验变量相关性，结果如表3所示。从整体上看，除独立权利要求数（X9）外，其他自变量和控制变量均与D指数存在显著的相关关系，表明这些变量与专利颠覆性在统计上具有相关性。

表3　变量相关性结果

组别	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9
Y	1									
X1	0.159***	1								
X2	−0.269***	−0.728***	1							
X3	−0.356***	−0.393***	0.802***	1						
X4	−0.212***	−0.206***	0.254***	0.413***	1					
X5	−0.014***	−0.067***	0.106***	0.145***	0.724***	1				
X6	0.015***	−0.058***	0.058***	0.064***	0.567***	0.821***	1			
X7	−0.136***	−0.001	−0.040***	0.034***	0.140***	0.001	−0.008**	1		
X8	−0.364***	−0.268***	0.292***	0.526***	0.652***	0.171***	0.091***	0.222***	1	
X9	0.004	−0.010***	−0.033***	0.010***	0.127***	0.058***	0.053***	0.115***	0.159***	1
样本量	308 351	308 351	308 351	308 351	308 351	308 351	308 351	308 351	308 351	308 351
Tolerance		0.581	0.254	0.294	0.238	0.291	0.432	0.939	0.353	0.974
VIF		1.722	3.944	3.402	4.202	3.433	2.314	1.065	2.833	1.027

注：“***”表示$p<0.001$，“**”表示$p<0.01$。

为避免变量间存在多重共线性而导致模型失真，进而出现“伪回归”现象，本研究在回归分析之前对各变量进行多重共线性检验。结果表明，本研究的各变量间的容忍度（Tolerance）均大于0.100，方差膨胀因子（VIF）均小于10，因此可以认为各变量间不存在严格意义上的多重共线性问题，满足回归分析的前提条件。

4.3　回归分析

本研究以零为分界点将D指数转换为二分类变量，将样本划分为颠覆性创新组（D指数≥0）与渐进性创新组（D指数<0）。此时，因变量为“专利是否属于颠覆性创新”，属于二分类变量。二元逻辑回归模型是处理此类因变量仅为0和1情形的常用方法，故本研究采用该模型估计技术知识基础跨度各维度对专利成为颠覆性创新概率的影响方向与程度。

为检验自变量影响结果的稳健性，本研究采用分层回归策略，逐步纳入不同组别的变量，观察各变量系数的显著性与方向是否保持稳定，判断结果是否受变量组合的影响。模型1为基准模型，仅纳入控制变量；模型2和模型3中在模型1的基础上分别加入时间特征变量和领域特征变量；模型4纳入所有控制变量、时间特征与领域特征变量，结果如表4所示。模型2与模型3中新引入的变量均通过显著性检验；模型4在同时纳入两类特征变量后，各自变量依然显著，且回归系数方向与模型2、模型3保持一致，表明结果具有良好的稳健性。从模型拟合效果来看，模型4的Nagelkerke R^2最高，拟合效果最好，模型2次之，模型3虽低于模型2，但较模型1仍有提升。这表明专利技术知识基础跨度的时间特征和领域特征对专利颠覆性均具有独立的解释贡献，二者共同纳入时模型整体解释力最强。

表4　逻辑回归结果

组别	模型1	模型2	模型3	模型4
X1		0.474*** （0.023）		0.455*** （0.023）
X2		0.019*** （0.001）		0.017*** （0.001）
X3		−0.025*** （0.000）		−0.024*** （0.000）
X4			−0.155*** （0.003）	−0.093*** （0.003）
X5			0.843*** （0.030）	0.690*** （0.031）
X6			0.299*** （0.019）	0.209*** （0.019）
X7	−0.007*** （0.000）	−0.010*** （0.000）	−0.007*** （0.000）	−0.009*** （0.000）
X8	−0.015*** （0.000）	−0.009*** （0.000）	−0.009*** （0.000）	−0.006*** （0.000）
X9	0.005*** （0.000）	0.003*** （0.000）	0.005*** （0.000）	0.003*** （0.000）
常量	0.307*** （0.007）	0.529 （0.018）	0.264*** （0.009）	0.437*** （0.019）
Nagelkerke R^2	0.123	0.184	0.135	0.190
样本量	308 351	308 351	308 351	308 351

注：“***”表示$p<0.001$；括号内为标准误。

逻辑回归仅能识别各自变量对“专利是否属于颠覆性创新”这一二分类结果的影响，而无法反映其对颠覆性程度的影响。考虑到D指数本身为连续变量，仅以二分类结果作为因变量会损失大量信息。为进一步考察技术知识基础跨度各维度变量对专利颠覆性程度的影响方向与强度，本研究以D指数作为连续因变量，构建最小二乘回归模型。由于D指数高度集中于零值附近且原始取值范围为[−1，1]，直接进行最小二乘回归容易受到分布偏态的影响，导致模型拟合效果不佳。为此，本研究对D指数进行双曲正弦变换后再进行最小二乘回归构建模型5，回归结果如表5所示。

表5　最小二乘回归结果

组别	模型5
X1	0.014***
X2	0.000*
X3	−0.001***
X4	−0.004***
X5	0.015***
X6	0.033***
X7	0.000***
X8	0.000***
X9	-1.657×10^{-5}
常量	0.031***
R^2	0.014
样本量	308 351

注：“***”表示$p<0.001$，“*”表示$p<0.05$。

可以发现，在控制了被引频次、引证专利数、独立权利要求数等因素后，各自变量均能显著影响专利颠覆性程度，其影响方向与显著性与模型4整体一致，表明这些自变量不仅影响专利成为颠覆性创新的概率，同样能显著影响其颠覆性程度的高低。由于因变量D指数经双曲正弦变换（asinh）处理，回归系数反映的是各自变量对asinh(D)的边际效应。具体而言，普赖斯指数每增加一个单位，asinh(D)提高0.014；引证时间跨度每增加一个单位，asinh(D)降低0.001；丰富性每增加一个单位，asinh(D)降低0.004；均衡性每增加一个单位，asinh(D)提高0.015；差异性每增加一个单位，asinh(D)提高0.033。平均引用年龄的系数极小，虽具有统计显著性，但对颠覆性程度的实质影响微乎其微。综合来看，差异性对颠覆性程度的正向促进效应最强，均衡性与普赖斯指数次之，引证时间跨度和丰富性则表现出抑制效应。

模型R^2为0.014，表明本研究所构建的技术知识基础跨度指标体系仅能解释专利颠覆性程度变异的极小部分，其原因主要是：一方面，D指数高度集中于0值附近，绝大多数专利的颠覆性程度差异极小，该分布特征本身会导致模型拟合较为困难；另一方面，专利颠覆性受到发明人特征、研发团队构成、制度环境等多元因素的影响，这些因素在本研究中尚未纳入考察范围。

4.4　稳健性检验

考虑到部分专利的颠覆效应可能需要更长的观测时间才能充分显现，本研究将引证时间窗口调整为10年重新计算D指数，经双曲正弦变换后构建模型6，以检验结果对引证时间窗口选择的敏感性，结果如表6所示。

表6　稳健性检验结果

组别	模型6
X1	0.045***
X2	0.001***
X3	−0.002***
X4	−0.006***
X5	0.037***
X6	0.036***
X7	−0.001***
X8	0.000***
X9	1.038×10^{-5}
常量	0.043***
R^2	0.022
样本量	308 351

注：“***”表示$p<0.001$。

从回归系数的方向来看，普赖斯指数、平均引用年龄、引文领域均衡性和引文领域差异性对D指数的影响均显著为正，引证时间跨度和引文领域丰富性的影响均显著为负，与模型5的结果一致；从显著性水平来看，6个自变量均在$p<0.001$水平上显著；从系数量级来看，模型6中各自变量的回归系数普遍高于模型5，表明扩大时间窗口后，技术知识基础跨度对专利颠覆性的影响效应进一步放大，符合“颠覆效应需要较长时间才能充分显现”的预期。此外，模型6的R^2由模型5的0.014上升至0.022，模型解释力进一步提升。综上，技术知识基础跨度对专利颠覆性的影响在不同引证时间窗口中均保持稳健。

5　结论与启示

5.1　研究结论

本研究以USPTO数据库中的发明专利为研究对象，从引证专利的时间和领域两个维度测度专利的技术知识基础跨度，并采用D指数量化专利的颠覆性创新程度，探讨了技术知识基础跨度对专利颠覆性的影响。结果表明，在时间维度上，技术知识基础跨度对颠覆性具有显著影响。具体而言，普赖斯指数与专利颠覆性正相关，即引证专利中近5年内授权的专利占比越高，专利颠覆性越强，表明对最新技术前沿的敏锐追踪有助于及时吸收尚未被广泛整合的新兴技术知识，加速前沿知识的创造性组合，从而激发出更具颠覆性的技术创新。平均引用年龄与专利颠覆性正相关，即引证专利的平均引用年龄越大，专利颠覆性越高，这说明深度回溯早期奠基性技术成果有助于发明人跳出当前主流技术路径，重新审视技术本质与演化脉络，从被既有范式取代或忽略的历史知识中发掘新的创新可能，进而实现对现有技术范式的突破。高普赖斯指数反映的是引文中近期文献的比例优势，而高平均引用年龄则表明引文中同时包含了年代久远的早期专利，二者共同表明，兼顾“追踪前沿”与“回溯源头”的双向知识搜索策略更有利于颠覆性创新的产生。此外，引证时间跨度与专利颠覆性负相关，当引证时间跨度较小时，专利引文的时间分布较为集中，有助于形成连贯、系统的技术知识体系，使发明人能对特定时期的技术进展形成深入透彻的理解，从而为专利颠覆性创新提供支持。

在领域维度上，引证专利的领域跨度对颠覆性具有显著影响，表现为当引文领域丰富性越低、均衡性越高、差异性越大时，专利颠覆性越高。跨领域的技术知识整合能够产生协同效应，激发深层次的创新潜力，催生高颠覆性成果。然而，过多领域的技术知识吸收也会导致知识重组的表面化和碎片化，稀释知识重组的深度和均衡性。颠覆性创新往往要求发明人对关键技术领域形成深入透彻的理解，在此基础上实现知识的根本性重组，而非对大量领域进行浅层的信息采集。当专利引用的技术领域数量过多时，可能意味着发明人采取了以广度为主的搜索策略，虽接触了更多领域的知识来源，但也会增加知识整合的复杂性，稀释核心技术方向上的创新聚焦，这一发现与“过度搜索会抑制技术创新”的观点相一致[11,43]。均衡性反映了专利技术知识基础的领域分布均匀程度，高均衡性越能有效规避对某一领域的过度依赖及其可能带来的创新局限。此外，技术知识基础的领域差异性较大，即样本专利与引证专利所属领域间的技术距离较远，有助于打破传统技术边界，为技术发展开辟全新的创新路径，从而激发颠覆性创新的产生。

5.2　启示与展望

本研究揭示了技术知识基础跨度与颠覆性创新之间的关系，为专利研发活动中的知识搜索和整合提供了策略指导。在时间维度上，专利发明人在知识搜索过程中，既应敏锐追踪近期前沿技术进展以捕捉尚未被充分整合的新兴知识，也应主动回溯早期奠基性技术成果以跳出当前主流路径的认知局限。在领域维度上，专利发明人应注重跨领域知识的均衡整合，优先考虑技术距离较远的领域知识，以激发非连续性的技术突破。同时，需警惕领域多样性过高所带来的整合复杂性，在知识来源的广度与融合深度之间寻求合理平衡，避免陷入“广而不精”的知识搜索陷阱。

本研究仍存在一定的局限。①D指数的计算依赖样本专利的被引与施引数据，需达到一定的数量规模门槛，因而可能将大量低可见度或新兴专利排除在外。因此，研究结论主要适用于已具备一定影响力和引用基础的专利群体，对早期或低引用专利的颠覆性特征外推能力有限。②当前对技术知识基础跨度的测度仅涵盖时间与技术领域两个维度，尚未深入至语义层面。未来研究可进一步探索技术知识基础在语义层面的跨度，分析样本专利与其引证专利在技术内容上的语义跨度，充分挖掘专利颠覆性创新的影响因素，为专利评估提供更加丰富的分析维度。③本研究以USPTO的授权发明专利为分析对象，未区分不同技术领域及发明人引用与审查员引用的差异。未来可进一步拓展数据来源、细分技术领域并区分引用类型，以更精确地揭示发明人真实的技术知识基础跨度对专利颠覆性的影响。④本研究关注专利知识属性层面的技术知识基础跨度，未纳入发明人特征、团队背景、机构类型及合作网络等创新主体层面变量。后续研究可将创新主体层面变量纳入分析，以更全面地揭示专利颠覆性的形成机理。

参考文献

[1] 刘小慧，朱曼曼. 科研论文颠覆性与参考文献年龄之间的关系研究[J]. 情报理论与实践，2023，46（7）：60–66.

[2] 丁乐蓉，杨欣谊，张靖雯. 不同合作模式下论文跨学科性对颠覆性创新的影响研究[J]. 现代情报，2025，45（2）：133–144.

[3] 吕冬晴，阮选敏，李江，等. 跨学科知识融合对D指数的影响[J]. 情报学报，2022，41（3）：263–274.

[4] 杨杰，王左戎，邓三鸿，等. 基于参考文献的论文跨学科性、跨时域性及其影响力研究[J]. 情报学报，2024，43（9）：1003–1014.

[5] PARK M，LEAHEY E，FUNK R J. Papers and patents are becoming less disruptive over time[J]. Nature，2023，613：138–144.

[6] 张佳锐，康乐乐，孙建军. 科学与技术知识融合中的新近性与时域性作用研究[J]. 信息资源管理学报，2024，14（4）：86–102.

[7] 孔嘉，邓三鸿，张佳锐，等. 技术创新中的知识融合：“科学–技术”二元知识对专利影响研究[J]. 情报学报，2022，41（11）：1161–1173.

[8] 张谙宁. 专利审查员引文的特点及其评价功能探析[J]. 中华医学图书情报杂志，2022，31（2）：36–41.

[9] CHEN L. Do patent citations indicate knowledge linkage? the evidence from text similarities between patents and their citations[J]. Journal of Informetrics，2017，11（1）：63–79.

[10] SHENG L B，LYU D，RUAN X M，et al. The association between prior knowledge and the disruption of an article[J]. Scientometrics，2023，128（8）：4731–4751.

[11] 侯剑华，邓先江，唐诗琪. 跨领域知识融合对高价值专利的影响研究[J]. 数据分析与知识发现，2025，9（3）：69–82.

[12] 方曦，彭康，刘云. 基于技术融合视角的颠覆性专利预测研究[J]. 情报杂志，2025，44（5）：139–146，138.

[13] 曹晓丽，于改红，吴振新. 颠覆性技术识别研究进展[J]. 国防科技，2024，45（6）：5–17，33.

[14] 唐旭丽，李信. 科研团队多样性对学术颠覆性创新的影响研究：以人工智能领域为例[J]. 情报学报，2023，42（1）：43–58.

[15] 石慧，潘云涛，赵筱媛，等. 基于文献挖掘的颠覆性技术早期识别研究[J]. 中国科技资源导刊，2019，51（4）：102–110.

[16] BORNMAN L，TEKLES A，ZHANG H H，et al. Do we measure novelty when we analyze unusual combinations of cited references? a validation study of bibliometric novelty indicators based on F1000Prime data[J]. Journal of Informetrics，2019，13（4）：100979.

[17] 李欣哲，鲁晓. 基于关键词特征和多元异构信息的论文颠覆性指数预测模型研究[J]. 情报探索，2025（2）：1–10.

[18] YANG A J，YAN X，HU H，et al. Are disruptive papers more likely to impact technology and society?[J]. Journal of the Association for Information Science and Technology，2025，76（3）：563–579.

[19] FUNK R J，OWEN-SMITH J. A dynamic network measure of technological change[J]. Management Science，2017，63（3）：791–817.

[20] 罗素平，寇翠翠，金金，等. 基于离群专利的颠覆性技术预测：以中药专利为例[J]. 情报理论与实践，2019，42（7）：165–170.

[21] WEI C，LI J，SHI D. Quantifying revolutionary discoveries: evidence from Nobel prize-winning papers[J]. Information Processing & Management，2023，60（3）：103252.

[22] 林婷，何郁冰. 互补者激励对企业颠覆性创新的影响研究[J]. 管理学报，2025，22（5）：887–896.

[23] 刘安蓉，李莉，曹晓阳，等. 颠覆性技术概念的战略内涵及政策启示[J]. 中国工程科学，2018，20（6）：7–13.

[24] 逯万辉. 基于专利引文网络挖掘的技术研发路径识别与颠覆性创新信号探测研究[J]. 情报学报，2024，43（9）：1059–1069.

[25] MENGUC B，AUH S，YANNOPOLUS P. Customer and supplier involvement in design: the moderating role of incremental and radical innovation capability[J]. Journal of Product Innovation Management，2014，31（2）：313–328.

[26] 郑世鑫，汪雪锋，陈虹枢，等. 知识重组视角下突破性技术形成机理研究[J/OL]. 科学学研究，2025：1–37[2026–03–06]. https://doi.org/10.16192/j.cnki.1003–2053.20251204.003.

[27] 贺德方. 基于USPTO专利数据的全球技术创新态势分析及我国的对策研究[J]. 中国软科学，2012（8）：178–184.

[28] 肖冬梅，陈颖. 专利审查员引文的特征与价值[J]. 图书情报工作，2015，59（19）：6–14.

[29] 丁乐蓉，石静，吴柯烨，等. 技术团队跨学科性对突破性创新的影响研究[J]. 情报学报，2024，43（5）：503–515.

[30] BORNMAN L，DEVARAKONDA S，TEKLES A，et al. Disruptive papers published in Scientometrics: meaningful results by using an improved variant of the disruption index originally proposed by Wu，Wang，and Evans（2019）[J]. Scientometrics，2020，123：1149–1155.

[31] 王丹，周潇，赵捧未，等. 基于技术融合视角的颠覆性专利识别研究[J]. 图书情报工作，2024，68（15）：58–71.

[32] 潘一如，毛进，李纲. 基于引文网络的高颠覆性专利知识扩散特征研究[J]. 数据分析与知识发现，2023，7（10）：1–14.

[33] WU L，WANG D，EVANS J A. Large teams develop and small teams disrupt science and technology[J]. Nature，2019，566（7744）：378–382.

[34] MUKHERJEE S，ROMERO D M，JONES B F，et al. The nearly universal link between the age of past knowledge and tomorrow's breakthroughs in science and technology: the hotspot[J]. Science Advances，2017，3（4）：e1601315.

[35] 吴菲菲，李睿毓，黄鲁成，等. 跨产业技术溢出识别与效应测度研究：以无人机技术为例[J]. 科学学与科学技术管理，2018，39（6）：84–98.

[36] 王格格，刘树林. 国际专利分类号间的知识流动与技术间知识溢出测度：基于中国发明授权专利数据[J]. 情报学报，2020，39（11）：1162–1170.

[37] STIRLING A. A general framework for analyzing diversity in science, technology and society[J]. Journal of the Royal Society Interface，2007，4（15）：707–719.

[38] BREIMAN L，FRIEDMAN J H，OLSHEN R A，et al. Classification and regression trees (CART)[J]. Biometrics，1984，40（3）：358.

[39] 熊回香，李跃艳. 基于Word2vec的学者推荐与跨语言论文推荐模型研究[J]. 情报科学，2019，37（12）：19–26.

[40] 张玉蓉，祝靖雨. 基于中国专利金奖的高价值专利影响因素研究：模糊集定性比较分析[J]. 科学学与科学技术管理，2025，46（2）：109–119.

[41] LLANES G，TRENTO S. Patent policy, patent pools, and the accumulation of claims in sequential innovation[J]. Economic Theory，2012，50（3）：703–725.

[42] 奉国和，刘任铧，邓伟伟. 高价值专利影响因素分析及主题挖掘：以纳米药物领域为例[J]. 图书馆论坛，2024，44（7）：48–60.

[43] LAURSEN K，SALTER A. Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms[J]. Strategic Management Journal，2006，27（2）：131–150.

作者简介

吴海婷，女，博士研究生，研究方向：知识创新与信息计量。
董克，男，博士，研究员，博士生导师，通信作者，研究方向：信息计量与科学评价，E-mail：dongke@nju.edu.cn。
贾呈华，男，硕士研究生，研究方向：网络舆情与信息计量。
吴佳纯，女，博士研究生，研究方向：知识创新与信息计量。

Research on The Impact of Technological Knowledge Base Span on Patent Disruptiveness

WU HaiTing[1]　DONG Ke[1,2]　JIA ChengHua[1]　WU JiaChun[1]
(1. Research Institute for Data Management Innovation, Nanjing University, Suzhou 215163, P. R. China;
2. Nanjing University Library, Nanjing 210023, P. R. China)

Abstract: A patent's technological knowledge base span captures both the temporal range and the field breadth of the prior art it builds on. To explore how this span shapes patent disruptiveness, this study draws on a sample of USPTO granted invention patents and measures disruptiveness with the D index. Indicators of the technological knowledge base span are then constructed from the temporal and field dimensions based on backward citations. Descriptive statistics, logistic regression, and OLS regression are then employed for the empirical analysis. The results show that the technological knowledge base span exerts significant effects on patent disruptiveness in both dimensions. In the temporal dimension, patents that draw on more recent prior art, carry a larger mean citation age, and pull from a narrower time window tend to be more disruptive. In the field dimension, patents that cite fewer distinct fields, distribute their citations more evenly across those fields, and exhibit greater disparity among the fields they cite are associated with higher disruptiveness.

Keywords: Technological Knowledge Base Span; Patent Disruptiveness; D-index; Patent Evaluation

（责任编辑：管清漾）