

# 中国教育财政

怀仁怀朴 唯真唯实

北京大学中国教育财政科学研究所

2026年第4期(总第308期)

2026年3月4日

---

---

## 西方国家科技创新定量评价体系的演变 及其对中国大学创新评价的启示

闵亦杰 周奕多\*

建设针对科技创新活动的定量指标体系，是增强科技创新活动跨国家/跨组织可比性的基础，进而又能够在决策中发挥评价（“量尺”）与引导（“指挥棒”）的双重作用。在历史上，部分西方国家积累了设计科技创新定量指标体系的经验，在60年的时间尺度上形成了一幅连续性演变图景，集中体现在从《弗拉斯卡蒂手册》到《奥斯陆手册》再到欧盟创新记分牌的演变之中。本文旨在总结这一演变过程背后的规律，并分析其对于中国大学创新评价与定量指标体系设计的启示。

### 一、从《弗拉斯卡蒂手册》到《奥斯陆手册》：评价理念的演变

第二次世界大战后，世界各国逐渐意识到技术创新对于国家安全与发展的重

---

\* 闵亦杰，北京大学中国教育财政科学研究所助理研究员；周奕多，北京大学中国教育财政科学研究所博士研究生。

要意义并普遍加大研发投入，但缺乏可用作跨国家比较的研发（R&D）定义与归集/统计准则。针对这一问题，OECD 于 1963 年发布了首版 R&D 数据收集和报告指南性质的《弗拉斯卡蒂手册》（Frascati Manual）并经过多次迭代<sup>①</sup>，其核心逻辑是先定义 R&D 活动（符合新颖性、创造性、不确定性、系统性、可转移性/可复制性特征的活动），然后将用于这些活动的投入归集为 R&D 支出。最新版《弗拉斯卡蒂手册》将“高等教育部门 R&D”设为单独一章，详细说明了高等教育部门的范围、其 R&D 活动识别及特殊性（例如与教学活动之间的界限）、研发支出与人员的测度等。可以认为，目前中国大学通用的“项目-论文-专利”三元素定量评价体系背后的理念，就出自《弗拉斯卡蒂手册》。

相比于《弗拉斯卡蒂手册》，1992 年开始发布的《奥斯陆手册》（Oslo Manual）将自身定位为一个创新数据的收集、报告与使用指南，并自初版起就明确了“研发≠创新”的理念<sup>②</sup>。具体地，《奥斯陆手册》有以下特点：（1）将创新理解为一个“环境→投入→活动→产出→影响”的系统而非将创新约等于 R&D 这一个环节；（2）将因存在概念或测量模糊性而难以纳入 R&D 归集范畴、但对全面创新刻画十分重要的指标纳入评价体系；（3）将除创新者自身之外的外部社会主体纳入指标体系，并分析这些外部主体与创新者在投入、活动与产出等环节的互动以及创新成果对非直接相关外部社会主体的影响。《奥斯陆手册》以“欧盟创新记分牌”为最典型实践，我们将在下一节中详细讨论。

显然，《弗拉斯卡蒂手册》长于对 R&D 的精准识别，能够很好地服务于统计、会计、税务等部门的“计数”或“算账”需要。但如果混淆 R&D 与创新概念，以《弗拉斯卡蒂手册》理念设计针对大学科技创新的定量评价体系，则天然带有窄化乃至矮化大学创新的倾向。具体而言：第一，大学用于营造良好创新生态而做的努力（比如建立以人才发展为目的的预聘长聘制度）会因无法建立与 R&D 活动的直接关系而无法计入评价体系，反向导致论文、专利、项目指标在评价中占比过大，这是大学创新评价短期化倾向的一个理念根源；第二，大学创新以服务社会技术进步为追求，但这种从大学到社会的溢出效应无法计入评价体系（例如，代表着论文面向产业需求的论文被专利引用现象在各种体系中都不算

<sup>①</sup> 官方文件见 [https://www.oecd.org/en/publications/2015/10/frascati-manual-2015\\_g1g57dcb.html](https://www.oecd.org/en/publications/2015/10/frascati-manual-2015_g1g57dcb.html)。

<sup>②</sup> 《奥斯陆手册》将创新的范围做了扩大化的定义，纳入了非技术的创新类型；这部分内容与本文主旨无关，故仅在脚注中做解释。官方文件见 [https://www.oecd.org/en/publications/oslo-manual-2018\\_9789264304604-cn.html](https://www.oecd.org/en/publications/oslo-manual-2018_9789264304604-cn.html)。

数),这是大学创新“闭门造车”倾向的一个理念根源;第三,大学创新活动(如科学家创办衍生企业)常跨越大学围墙内外,现有评价体系对此做生硬截断、只计入墙内部分,这又是说不清大学创新之于科技全局贡献的一个理念根源。前述倾向具有全球范围的普适性,在中国大学尤其明显。简而言之,我们需要明确《弗拉斯卡蒂手册》的本意是 R&D 归集、不应将其用于以“创新”为口径的数据收集与指标体系构建之中。

## 二、《奥斯陆手册》理念与欧盟创新记分牌

我们认为,基于《奥斯陆手册》理念设计的欧盟创新记分牌对于中国大学创新定量指标体系建设具有突出借鉴价值。为推动区域创新发展、提升整体竞争力,欧盟于 2000 年推出里斯本战略,明确提出建成全球最具竞争力知识经济体的目标。为保障该战略有效落地,欧盟以《奥斯陆手册》核心理念为基础,创设欧盟创新记分牌(以下简称记分牌)这一工具。其核心目的在于全面和标准化地收集各国创新相关数据,开展国家间创新水平比较,最终对各成员国战略实施进度与成效进行监督。最新版欧盟创新记分牌(2025)的定量指标体系被设计为一个“4 大模块—12 个维度—32 个指标”的层级结构(完整体系详见附件)。

欧盟创新记分牌在指标设置上延续了部分经典的 R&D 相关指标,体现出与传统做法之间的连贯性。例如,传统评价中核心的研发投入、学术论文数量、专利数量指标,分别被纳入投资模块下的企业投资维度、框架条件模块下的具有吸引力的研究体系维度,以及创新活动模块下的知识产权维度。此类传统指标虽得以保留,但在整个体系中的占比相对较低(约占四分之一左右)。关于记分牌指标体系的具体内容,分模块说明如下。

### (一) 框架条件模块

该模块聚焦支撑创新活动开展的基础环境建设,涵盖人力资源供给、科研体系效能、数字化基础设施保障等核心支撑要素,具体划分为人力资源、具有吸引力的研究体系与数字化三个维度。记分牌对框架条件的关注,反映出其设计的审慎性。若评价体系仅关注产出指标,就容易引导被评价对象忽视长期投资与能力积累,转而采取短期化策略以快速提升产出数据;而将环境构建和长期能力积累

等框架条件与产出指标共同纳入评价,就可相对平衡短期产出与长期发展之间的关系,缓解被评价对象的短期化倾向。此外,数字化维度的设置具有鲜明的时代性,该维度并非记分牌初始版本即有,是欧盟根据创新所嵌入基础环境的变化而后来增设,体现出记分牌指标体系与时俱进的特点。

## **(二) 投资模块**

该模块主要关注创新活动开展所需的资源投入情况,包含资金与支持以及企业投资两个维度。其中,资金与支持维度下的风险投资支出指标,集中体现了《奥斯陆手册》的宽口径理念。从功能定位来看,风险投资市场是初创型企业与新兴技术领域获取资金支持的主要渠道,但企业获得风投资金后并非专用于创新活动,也可能用于一般性的经营之中,由此在全面性与精准性之间形成了一个权衡关系:纳入该指标可更全面地反映企业创新资源的实际规模与来源构成,符合《奥斯陆手册》理念;但其固有的模糊性或不精确性导致难以被精确归集,与《弗拉斯卡蒂手册》的理念相冲突。换言之,是否选择此类指标本质上体现了定量指标体系在技术性(准确)与战略性(全面)上的取舍。

## **(三) 创新活动模块**

该模块聚焦创新实践的具体过程,包含中小企业创新、公私合作与知识产权三个维度,是《奥斯陆手册》关注全链条创新理念的集中体现。由于创新过程存在高度不确定性,部分符合创新规律的正确行为或决策,未必能产生预期的创新成果。基于此,增设针对创新行为特征(而非仅针对创新产出)的考核标准,可减轻那些开展了正确创新行为但未取得成效主体的评价压力。此类指标的纳入,本质上反映了欧盟创新记分牌不仅是为评价而评价,更是为促进各国更好响应里斯本战略目标而评价。例如,公私合作维度就关注大学与企业间合作,其目的既在于促进企业技术创新吸收大学的科学研究成果,还在于促进分属不同国家的大学与企业加强沟通,加速欧洲的一体化进程。在实操中,公私合作出版物与PCT专利申请量两个指标有着相同的权重,说明欧盟创新记分牌兼重创新的过程与产出维度,以全面数据收集最终服务于宏观战略决策。

## **(四) 影响模块**

该模块关注创新活动产生的最终成效，且不局限于对创新主体自身的影响，具体划分为经济影响、社会影响与环境的影响三个维度。该模块对社会与环境的影响维度的关注，实质是对“何为高质量创新”问题的系统性回答——尽管欧洲对创新价值标准的界定未必符合其他国家的实际，但该评价思路为衡量创新活动的综合价值提供了重要参考。从具体指标设计来看，社会影响维度下的高技能岗位占比、终身学习参与率这两个指标与教育、科技、人才一体化发展的整体要求高度契合。而创新成果在欠发达地区普及率指标则体现了公平原则：若某项创新成果对创新主体而言具有较高经济价值，但在欠发达地区的推广应用存在障碍，按照记分牌评价规则，其得分将相应下调，以此引导创新活动兼顾经济价值与社会普惠性。

### **三、西方国家科技创新定量指标体系演变的启示**

欧盟创新记分牌虽聚焦国家层面并侧重企业创新，但其核心思路仍多有借鉴价值。在总体层面，中国大学科技创新定量评价体系应超越仅聚焦创新投入与产出的狭义视角，转向“框架→投入→活动→产出→影响”的全链条设计。在指标体系设计层面，应考虑更多引入有助于全面刻画大学创新活动但存在模糊性的指标，以此更完整地呈现中国大学的创新全貌。在数据收集层面，在传统的归集方法之外，应更充分地利用大数据与文本分析方法从各类论文、专利、工商注册等数据库中直接挖掘并形成与大学创新相关的指标。以下具体讨论具有抛砖引玉的性质，并不全然正确或恰当。

#### **（一）框架条件：构建支撑长期可持续发展的大学内部科研环境**

在大学创新情境下，框架条件模块应聚焦大学科学家创新活动所嵌入的组织环境与支持条件，以响应“十五五”规划“投资于人”核心理念、支撑大学长期可持续发展的基础。这一模块可以考虑学术评价体系、内部资源配置机制与数字基础设施三个维度。第一，构建以人为中心的学术评价体系。摒弃将科研人员视为论文生产工具、单纯服务于学科评估或“双一流”建设的功利化导向，转而关注科研人员的长期成长与发展需求，通过科学制订人才引入启动支持与晋升考评标准，避免制度性压力迫使科研人员陷入短期化科研行为。第二，设计公平包容

的内部资源分配机制。重点向新实验室、青年科研团队等外部资源获取能力较弱的群体倾斜，同时对非共识性科学技术路径等给予宽容与支持，防止资源过度集中于少数主流领域与资深科研人员。第三，建设普惠性数字基础设施作为支撑。顺应大科学与 AI 时代发展趋势，由大学层面统筹建设算力、存储设备等核心数字基础设施，组建专业技术支持团队；面向人文社科、基础学科等筹资能力较弱的领域，强化学生数字化技能培养，确保各学科均有机会利用好数字基础设施，为大学全领域创新提供坚实基础支撑。

## **（二）创新投资：强化创新投入的可持续性与战略能动性**

大学创新投资模块的核心在于保障大学创新活动的资金供给数量/质量与大学层面的战略能动性，建议考虑科技财政可持续性、多元化科技经费筹措能力与校级可支配科技经费三个维度。第一，科技财政可持续性。聚焦自然科学基金、国防军工项目等传统科技财政路径的经费支持，既关注经费规模的充足性以保障科研活动长期开展，也关注年度之间科技财政经费的稳定性，避免因资金波动过大影响科研布局的连续性。第二，多元化科技资金筹措。呼应《奥斯陆手册》重战略、宽口径的理念，将社会资本捐赠或者大学科学家创业企业融资等情况纳入评价范畴——后者情况下资金虽未直接进入高校账户，但大量用于支付研究生实习薪酬、购置科研装备与耗材等核心科研支出，且相关科学成果仍以学校为作者单位发表，本质上是大学创新资源池的重要组成部分。这部分统计相对复杂，可以考虑一手问卷与二手数据库形式相结合进行收集。第三，校级可支配科技经费。核心关注归大学层级支配的科技资金的绝对数量与占比，这类资金区别于记在学校名下但由科研人员自主支配的经费，其占比越高意味着学校越具备主动设计科技发展战略、布局重点领域与新兴方向的能力。

## **（三）创新活动与产出：聚焦有组织科研、新科研范式应用与高质量成果**

评价创新活动的一个关键在于正视其高不确定性特征——正确的创新决策与实践未必能收获预期成果。因此，为防止评价导向过度偏向产出导致抑制科研人员探索未知的积极性，建议专门设立创新活动模块，对符合导向的创新活动予

以认可而不论其产出如何。对创新活动的讨论可以包括两个维度。第一，需求导向的有组织科研。针对传统机构（学院、研究所）因跨学科性不足而难以承接大科学项目与大产业项目的问题，突破“按学科划分机构”的固有模式，增加具备行政统筹能力的、跨学科的、实体性质的有组织科研机构在大学中的占比，提升大学响应复杂技术创新需求的能力。第二，对采用新科研范式的研究给予更多包容与鼓励。新科研范式在起步阶段往往与传统范式存在冲突、面临争议，大学应建立容错机制，减少对使用非共识新范式进行的科学研究采用基于期刊级别的刚性考核；应采用专家评议的方式，合理评估其在研究方法与研究范式层面的探索价值，为创新突破预留空间。

在科研产出评价方面，可考虑沿用学术论文与专利两大核心评价对象，兼顾数量与质量双重维度。在学术论文方面，数量维度涵盖论文发表的绝对数量、不同学科领域的人均论文数等；质量维度则包括中科院分区（尤其是一区）论文占比、JCR 顶刊发表数量、高被引论文（如 ESI 高被引、热点论文）数量、论文总被引频次、篇均被引频次、他引率等。在专利方面，数量维度包含专利申请量、授权专利总量、年度新增专利数等绝对指标；质量维度则聚焦 PCT 专利申请/授权量、专利被引频次、专利实施转化率、专利维持年限（维持时间越长通常质量越高）等指标。

#### **（四）创新影响：从产业、战略与全球维度拓展对创新价值的理解**

创新影响模块核心在于突破学术共同体单一评价视角，勾勒中国大学科技创新在学术价值之外的实际应用价值与广泛影响力，全面回答社会（乃至国际社会）是否认可中国大学创新这一问题。该模块可分为中国大学创新的产业价值、战略价值以及全球价值三个维度，对应面向经济主战场、国家重大战略、人类命运共同体三个方面。具体而言：第一，大学创新的产业价值。基于科学论文被专利引用现象，讨论中国大学科学研究是否能够直接支撑产业技术创新；进一步分析引用中国大学论文的专利持有人的技术先进性与行业特征，考察中国大学科学研究是否支撑了先进技术以及是否具有赋能型知识特征。同时，考察大学师生创业企业在风险投资市场上的融资规模与频次，反映大学技术在进入产业化阶段后的被

认可程度。讨论该指标的意义在于，社会观点普遍认为大学创新以基础研究为主业，而对大学能否在技术开发乃至产业化阶段仍发挥重要作用的认知不足。第二，大学创新的战略价值。一方面，以大学参与国家重点研发计划等的数量与实际贡献衡量其对国家重大战略的直接支撑；另一方面，通过大学科技成果被全球范围内重大项目（含国际顶尖科技计划与他国国家级重点项目等）引用的情况，考察中国大学创新的战略价值。第三，大学创新对全球南方国家的价值。兼顾“授人以鱼”与“授人以渔”衡量中国大学创新的全球影响力：“授人以鱼”体现为中国大学的科技成果被全球南方国家科技成果引用频次与占比，反映技术成果对发展中国家的实际赋能；“授人以渔”则聚焦全球南方国家学生来华就读博士研究生后的长期发展情况，特别是其中的归国人才在全球南方国家当地科技领域形成的同伴效应与辐射影响力，以此考察中国大学创新对全球科技普惠与共同发展的贡献。

## **（五）技术可行性：大数据技术、文本分析技术与公共数据库建设**

当前，我国高校科研数据的收集仍主要依赖“自下而上”归集，这种方式的全部过程限于教育行政体系之内，具有行政约束力强、执行效率高的优势，但在指标可得性与精细度上存在显著短板。以衡量论文影响力这一目标举例：现有实践往往依靠统计“论文被引次数”加以实现、无法区分引用主体的性质（诺贝尔奖团队或不知名科学团队），更难以判断引用意图（作为发展理论的核心基础或仅作为背景知识被提及），导致评价颗粒度粗糙，难以真实反映科研成果价值。近十年，随着大数据技术与文本分析技术的迭代、算力硬件的升级以及高质量公共数据库的完善，构建基于公共数据的收集与分析方法已成为可能。换言之，我们已具备深度挖掘海量异质性数据、设计与国家战略意图高度匹配的精细化指标的技术基础。未来的中国大学技术创新定量指标体系不应局限于传统行政归集方法，而应探索行政归集与大数据挖掘相结合的混合模式，从而极大地丰富数据来源，提升指标体系的全面性、准确性与可行性。

我们举两个具体例子详细说明。第一，在成果转化与产业化问题上，现有解决方案已经可以利用大语言模型进行跨库匹配，将工商注册信息中的企业创始人

与大学教师名录进行关联，从而构建出大学教师校外创业的全景数据库。这不仅能揭示大学科技成果转化的路径偏好，还能通过追踪企业的市场表现来回答“大学技术是否有广泛的转化潜力”。第二，在人才流动与全球影响方面，结合行政记录与 OpenAlex 等新一代学术出版物数据库，已经可以清晰追溯全球范围内的人才流动轨迹，特别是能够精准识别来自全球南方国家的学生在中国获得博士学位后返回母国的科研表现及其产生的同伴效应。这些具体实践充分证明，新技术与新数据库的结合，已使构建更全面、更深入、更具战略导向的定量指标体系成为可能。

## 附录

表 1 欧盟创新记分牌（EIS 2025）指标体系<sup>①</sup>

<p><b>框架条件模块</b>反映企业外部创新绩效的主要驱动因素，分为人力资源、有吸引力的研究体系和数字化三个维度：</p> <ul style="list-style-type: none"><li>● 人力资源维度包含 3 项指标，衡量高技能和受过良好教育的劳动力供给情况，具体指标为：（1）科学、技术、工程和数学（STEM）领域新增博士毕业生；（2）25-34 岁完成高等教育人口；（3）25-64 岁参与终身学习活动人口。</li><li>● 有吸引力的研究体系维度包含 3 项指标，衡量科研基础的国际竞争力，具体指标为：（1）国际科学合作出版物；（2）高被引出版物；（3）外国博士研究生。</li><li>● 数字化维度包含 2 项指标以衡量数字化技术水平，具体指标为：（1）企业宽带普及率；（2）具备基础以上综合数字技能的人口（供给情况）。</li></ul>
<p><b>投资模块</b>反映公共部门和企业部门的投资情况，分为资金与支持、企业投资和信息技 术使用三个维度：</p> <ul style="list-style-type: none"><li>● 资金与支持维度包含 3 项指标：（1）私人资金（风险资本投资）；（2）高校和政府研究机构的研发支出；（3）政府对企业研发的直接资助和税收支持。</li><li>● 企业投资维度衡量企业为产生创新而进行的研发投资和非研发投资，包含 3 项指标：（1）企业研发支出；（2）非研发创新支出；（3）创新活跃企业的人均创新支出。</li><li>● 信息技术使用维度衡量信息技术的使用情况，共 2 项指标：（1）积极提升员工信息通信技术（ICT）技能的企业；（2）受雇的信息通信技术专业人员。</li></ul>
<p><b>创新活动模块</b>反映企业部门创新的不同方面，分为创新者、合作关联与知识产权三个 维度：</p> <ul style="list-style-type: none"><li>● 创新者维度关注中小企业在市场或企业内部推出的创新情况，包含 2 项指标： （1）推出产品创新的中小型企业；（2）推出流程创新的中小型企业。</li><li>● 合作关联维度评估创新生态系统中的关联情况，包含 3 项指标：（1）创新企业间的 合作努力；（2）公私部门间的研究合作；（3）科技领域人力资源的职业流动。</li><li>● 知识产权维度衡量不同形式知识产权（IPR）的表现，具体指标为：（1）PCT 专利申 请；（2）商标申请；（3）外观设计申请。</li></ul>
<p><b>创新影响模块</b>反映企业创新活动产生的效果，分为三个维度：</p> <ul style="list-style-type: none"><li>● 就业影响维度基于 2 项指标衡量创新活动对就业的影响：（1）知识密集型活动中的 就业；（2）创新企业中的就业。</li><li>● 销售影响维度衡量创新的经济影响，包含 3 项指标：（1）中高技术产品出口；（2） 知识密集型服务出口；（3）创新产品带来的销售额。</li><li>● 环境可持续性维度基于 3 项指标反映在减少负面环境影响方面取得的进展：（1）资 源生产率；（2）细颗粒物（PM2.5）空气污染暴露程度；（3）环境相关技术的发展。</li></ul>

<sup>①</sup> David Rosenfeld, Owen McColgan, Paresa Markianidou, et al. European Innovation Scoreboard 2025 – Methodology Report. European Commission, 2025. [https://research-and-innovation.ec.europa.eu/document/download/6340fb44-1dad-4a11-b5d7-a298ad5922a5\\_en?filename=ec\\_rtd\\_eis-2025-methodology-report.pdf](https://research-and-innovation.ec.europa.eu/document/download/6340fb44-1dad-4a11-b5d7-a298ad5922a5_en?filename=ec_rtd_eis-2025-methodology-report.pdf).

## 上期回顾

2026 年第 3 期（总第 307 期）

《中国教育财政二十讲》成果系列摘录之二十：

中国全社会教育投入指数——基于中国教育财政家庭调查的估计

**摘要：**本文在 2023 年中国教育财政家庭调查数据的基础上，对全社会的教育投入情况进行估计。首先，本文估计了以家庭为单位的子女教育支出（家庭教育支出）和单个子女的教育支出（生均家庭教育支出）。其次，基于三个指数来衡量家庭教育支出的负担情况，分别是家庭教育支出占人均可支配收入的比例、家庭教育支出占家庭总消费支出的比例和家庭教育相对支出比。最后，基于三个指数来衡量全社会教育投入的情况，分别是国家财政性教育经费占 GDP 的比例、家庭教育支出占 GDP 的比例、全社会教育投入占 GDP 的比例。

《中国教育财政》由北京大学中国教育财政科学研究所主办；旨在反映本所最新的学术科研活动；相关内容仅体现作者本人观点，并不必然代表本所的立场。

文章内容仅供参考，如需转载须事先征得本研究所同意。

本期印发：2000 份

下载网址：<http://ciefr.pku.edu.cn>

---

**主办单位：**北京大学中国教育财政科学研究所

**邮箱：**[workingpaper@ciefr.pku.edu.cn](mailto:workingpaper@ciefr.pku.edu.cn)

**责任编辑：**毕建宏

**传真：**010-6275-6183

**地 址：**北京市海淀区颐和园路 5 号

**微信公众号：**中国教育财政

北京大学教育学院大楼四层（100871）

