

AI科教革命，从精准考核、主题PI制 与教学系统重构开始

刘益东

(中国科学院自然科学史研究所, 北京 100190)

[摘要]当前,人工智能(AI)正以前所未有的速度渗透至人类社会生产生活与认知的各领域,DeepSeek的横空出世更让国人深切感受到AI的巨大潜力及其对中国科教事业带来的巨大挑战和机遇。AI正在快速重构科教领域的范式规则,包括考核精准化、PI精英化与教学系统重构。以往形成的“管理者通吃”“五唯”盛行的科教体系日益成为科技自立自强的阻碍。顺应AI快速发展趋势,以AI赋能中国科教改革,建设“一流研究主题”,实施主题PI制,凸显一流学者的重要性,建立起“点”(研究主题)、“线”(学科专业)、“面”(大学)三位一体的新型世界一流高教体系,有助于破解科教组织改革所急需的巨量经费投入、甄选一流人才担任学术带头人、确立教学系统重构原则、培养AI无法替代的人才等难题,有力推动我国科教事业从跟踪跟进阶段向追求原始创新和科技自立自强阶段的转型、跃升。

[关键词]DeepSeek 一流主题 主题名师 AI for PI 寄生式AI科研 高仿论文 AI之上

[中图分类号]G521; TP18 **[文献标识码]**A **[文章编号]**2096-983X(2025)02-0104-11

一、问题的提出：AI时代中国 科教领域面临的巨大挑战

目前,中国科教界正面临前所未有的双重挑战。一方面是人工智能(AI)大爆发,对各行各业产生巨大影响,尤其是起源于2023年,并于2025年横空出世的智能搜索与分析系统DeepSeek,让国人深切感受到AI的巨大潜力。目前,AI程序员的水平已经超过99.9%的人类程序员,在律师、医生、教师、设计、金融、科研、销售等领域,AI的工作能力正以极快的速

度接近并超过人类。可以预测,通用人工智能(AGI)在不远的将来就极有可能实现,智能机器人替代人类从事各种工作的情形迫近。人类大规模失业难以避免,形势极为严峻、紧迫。另一方面,我国科教事业正在从引进吸收、跟踪跟进阶段向原始创新、自立自强阶段转型、跃升,科研难度陡然提高。这种难度的提高不是成十倍八倍,而是成一百倍、五百倍,甚至更高。过去形成的“管理者通吃”“五唯”^①的科教体系难以胜任。近十年来,“双一流”^②建设成效显著,研究型大学是基础研究与原始创新的主力

收稿日期: 2025-02-22

作者简介: 刘益东, 研究员, 博士研究生导师, 主要从事科技战略、科技与社会、人才评价、科技史、未来学研究。

①“五唯”即唯分数、唯升学、唯文凭、唯论文、唯帽子的评价标准。

②“双一流”是指建设世界一流大学和一流学科。

军。但从总体上看,“双一流”建设宏观有余、微观不足,行政有余、学术不足。《统筹推进世界一流大学和一流学科建设总体方案》(国发〔2015〕64号)提出的世界一流学科的中国标准是“建设一流师资队伍、培养拔尖创新人才、提升科学研究水平、传承创新优秀文化、着力推进成果转化”。以上五项标准系统全面地概括了研究型大学的学科建设目标,但未能突出世界一流学科的核心特征应该是在原始创新方面取得突破,由此在世界前沿引领学科发展。

为此,笔者提出增加建设“一流研究主题”(简称“一流主题”)为评价标准之一,即该主题的研究水平为世界一流。国际学术前沿的突破与引领都是源于在某一研究主题领域的原始创新,没有一流研究主题就没有一流学科,没有一流学科就没有一流大学,只有“点”(研究主题)、“线”(学科专业)、“面”(大学)三位一体、多管齐下,才能尽快建成新型的世界一流高等教育体系。由此,“双一流”建设可升级为“叁一流”建设或“多一流”“全一流”建设。

面对史无前例的双重挑战,我国科教领域需要进行一系列深刻变革,急需解决三大难题,一是需要巨量科研经费投入;二是需要破除“五唯”评价标准、改革科研组织方式,甄选一流人才担任学术带头人;三是需要教育改革,培养AI无法替代的人才。为此,笔者认为当务之急是利用AI快速重构科教领域的范式规则,实现考核精准化、PI精英化与教学系统重构。其中,考核精准化可快速清除垃圾论文等学术泡沫,释放出巨量科研经费,大幅降本增效;PI精英化破除唯“帽子”、唯“头衔”,让真正的一流人才担任学术带头人,实现“一流研究主题”,加强原始创新,抢占科技制高点,快速提升我国的科技实力与智库实力;教学系统重构则是主动应对AI快速发展、传统STEM学科逐渐被AI替代的挑战,填补AI盲区,构建与AI互补、与AI共生的知识体系与实践智慧,培养AI无法替代的人才。^①

二、AI评价实现精准考核,清除垃圾论文,遏制学术平庸,大幅降本增效

提高科研效率与品质的关键是加强科研产出的考核。在基础研究、人文社会科学、交叉学科等学术领域,论文是主要的科研产出形式,结果以刊评文,成为科研产出考核的主要方式,造就了庞大的、等级分明的期刊产业,在国内形成了以发表SCI论文、SSCI论文、国际顶刊论文为导向的科教体系,以次充好的垃圾论文、泡沫论文与平庸学术吞噬了巨量科研经费,借助AI在评价领域的应用(简称AI评价),实施科研产出的精准考核迫在眉睫。

“以刊评文”和“五唯”,容易导致鱼目混珠、以次充好现象,拉低考核标准。王志珍院士指出,“如果科研人员60%的精力都用在写论文应付考核上,哪有时间做真正有创造性的工作?这些‘泡沫论文’该挤一挤了!”^[1]施一公院士则直言:“‘垃圾文章’,就是纯粹为了发文而发文,这种情况太多了。”钟世镇院士指出,“我们发表在SCI上有用的论文却寥寥无几,15%都不够,另外85%都是垃圾论文”。^[2]知名学者熊丙奇也谈到,政府部门官员与两院院士多次批评我国的“科研成果”有90%的论文是垃圾。^[3]平庸论文同样危害巨大,中国人民大学教授欧阳志远在《光明日报》发文强调中国科学的真正危机不是作伪性论文的频现,而是平庸性论文的泛滥。《新华文摘》原总编辑张耀铭则指出,“学术平庸对学术的蛀蚀,比学术不端更为普遍、更为严重,……‘劣币驱逐良币’的现象大量发生”。^[4]

以刊评文的依据是论文引用率及其派生的期刊影响因子,引用率的局限性众所周知,这是信息技术欠发达时代的产物,当博士级智能、专家级智能的AI出现之后,对论文评价的方式应该随之改变。笔者提出双管齐下,一是从以论文为科研产出精准到以论文学术结论/学术

①本文所述一流人才、一流学者、一流成果、一流学科、一流主题等均指世界一流。

观点为科研产出,从“期刊-论文”模式精准化为“数据库-知识点”模式(论文作为必备的附件);二是利用自然语言理解、相似度分析等先进的AI技术开展甄别与评价。论文三要素包括研究问题-学术结论-接续比较。研究问题,通常是解释性问题而非描述性问题。有些论文用研究目的或任务代替研究问题,实际上也隐含了研究问题。新的学术结论,也是论文新的主要学术观点,包括论文新的研究视角、方法、理论、创见、数据等知识贡献与增量。接续比较,体现出研究是在他人工作基础上的推进或另辟蹊径,属于学术结论与流行学术观点(主流或前沿的学术结论)相比的优势和知识增量。

本文强调,学术结论/学术观点(并称学术结论)与结论、观点(并称结论)不同,后者只是就事论事的看法,而学术结论是从具体到一般的概括与升华。以气候变化研究为例,结论(就事论事)是“2023年夏季北京地区平均气温比过去30年同期升高1.2℃”,学术结论(理论升华)则为“温室气体通过增强大气辐射强迫效应,导致中纬度城市群热岛效应与全球变暖产生叠加共振,这一机制可解释21世纪东亚季风区城市极端高温事件频发的现象”。由此可见,学术结论通过温度数据发现规律→建立“辐射强迫-热岛效应”理论模型→提出可验证的“叠加共振”机制→推导出具有预测价值的区域性气候变化范式。学术结论是“know-why”层面的知识生产,结论则只是“know-what”层面的信息积累。

研究新案例可产生新结论,但不一定能提炼、升华出新学术结论。因为即使该案例没有被研究过,但已被流行学术结论所覆盖,这样的研究也不能形成学术论文。可以说,研撰合格的学术论文并非易事,笔者提出简明的“评文三问”法即可做出判断,即探究并展示出论文的学术结论是什么,是否新颖,与流行学术结论相比有何优势^[5]。新优学术结论是新颖且优于(至少一种)流行学术结论的学术结论,论文具有新优学术结论才是合格的论文。可进一步

简化,只要产出新的学术结论,就可视为合格论文,视为有效科研产出,因为即使不正确的新学术结论也有启发作用(常言道“有创意的错误胜过一打老生常谈”)。具有自然语言理解能力的AI评价可以确认论文是否有一般化的、新颖的学术结论,据此判断是否为合格论文。例如IBM Watson Discovery(提取论文核心主张并对比文献库)、Semantic Scholar(提取论文核心观点,生成简洁摘要)、Dimensions(整合文献、专利、基金等多种数据,提供学术观点查新服务)等,可识别论文是否产出学术结论及其是否新颖独特。持续改进的AI评价功能更强。

由此可见,当我们确立“新颖的学术结论是有效科研产出”的标准,AI评价就可清除垃圾论文、泡沫论文,遏制平庸学术。当然,随着AI智能水平的提高,不仅学校里面要杜绝利用AI代写作业,更要警惕、杜绝利用AI代为完成课题/项目,为此应该做专项审查。利用AI做科研分为两类,一类是超越式AI科研(AI for Science),借助AI做出明显高于AI自行完成的科研;另一类是寄生式AI科研,研究者水平平庸或疏于钻研,其水平低于AI(例如低于DeepSeek,低于不久将问世的GPT-5)的科研水平,借助AI科研掩盖自己的平庸,其科研成果与AI自行完成的成果,水平相差无几。如果涉及一些通过实验、调研、采访获得的数据开展研究,只要输入同样的数据,AI自行研究的结果与平庸学者借助AI开展科研的结果大同小异,同样属于寄生式AI科研。这类利用AI完成的科研成果仿冒高水平成果,同样浪费科研经费,败坏学术风气,可以称之为“寄生式AI的高仿论文、高仿成果”,简称“高仿论文”“高仿成果”。

需要强调的是,存在笔者称之为“AI限高法则”,包括对齐AI和超越AI两种情况。对齐AI是指当使用者的水平低于或等于AI水平时,其使用AI产出的成果水平只相当于AI自行研究的水平或略高;超越AI是指当使用者水平高于AI自行研究的水平时,则使用者如虎添翼,不仅效率倍增而且其成果水平可以屡创新高。因此,

平庸的使用者并不能因为AI赋能而产生高于AI的有价值成果。也就是说,寄生式AI科研产生的论文、报告等成果的水平虽然高于以往的垃圾论文、垃圾报告,但是仍然没有价值,不应发表、不应作报告。因为AI照样可以无限生产出这种水平的“高仿论文”“高仿成果”,鱼目混珠,可以大规模生成SCI论文、SSCI论文、CSSCI论文、学术图书等,这是学术出版界面临的新挑战。结果是基于国际论文形成的大学国际排名、学科国际排名(如ESI全球前1%、全球前1%)将毫无意义,必将退出历史舞台。

按照目前AI的发展速度,可以预测,一两年内AI自行研究的水平或可超过多数国际顶刊论文的水平,大规模高仿论文成为国际顶刊论文导致依靠国际顶刊论文的国际大学排名体系将会崩溃。实际上,国内以发表国际顶刊论文为导向已经使国际大学排名沦为笑谈。如表1所示,全球自然指数排名前10的大学,中国占据9席,力压麻省理工、斯坦福、普林斯顿、伯克利、耶鲁、牛津、剑桥等一众世界名校,中国大学在世界前50所大学中的数量更是碾压美国大学,排名第1。但这与实际严重不符,充分表明按照国际顶刊论文数量根本无法正确反映大学科研的真实水平与前沿竞争力,自然指数全球大

学排行榜已被证伪,信誉崩塌^[6]。鉴于目前国际大学排名乱象丛生,加之高仿论文行将泛滥,其可信度堪忧,我国大学应该退出国际大学排名。目前,国内已有多所著名大学宣布退出,实际上并未因此影响招生和科研。建议教育部明确,考核、评估等不再参考大学国际排名。

同时,长期以来我国的项目/课题验收标准低,几乎都是“通过”“合格”。建议提高验收标准,严格按照任务书目标验收。AI评价可以有效甄别垃圾论文、泡沫论文和寄生AI的高仿论文、高仿成果等平庸之作。如此不仅节约巨量科研经费、大幅降本增效,而且避免劣币驱逐良币,实现“高标准、优胜出”。随着AI智能水平的提升,AI评价越发高效合理,明效大验。

三、一流人才评价改革:甄选一流人才担任学术带头人,实施主题PI制

目前AI大爆发,在博士级智能、专家级智能的AI(GPT-o3、DeepSeek等)横空出世的今天,人才战略升级为一流人才战略是必然选择。就目前智能水平而言,AI对一流(杰出、顶级)人才是助手,对二流(优秀、高级)人才是威胁,对三流(普通)人才是平替。随着AI智能水平提高,其替代性会逐级上升。因此,要使AI辅助人类而非替代人类,需要一方面改变AI发展模式^[7],另一方面要重视一流人才,使其充分发挥作用,确保人类社会安全与可持续发展。因此,提出一流人才标准和评价方法是当务之急。

(一) 五步识才法:以新标准识才、以“尖”识才、以长期反馈识才、以同主题比较识才、以用户需求识才

甄选一流人才绝非易事,唯“帽子”、唯头衔、唯国际顶刊论文、同行评议的主观性等,都是巨大阻碍。以原创成果论英雄,却存在着徐匡迪院士指出的“中国颠覆性技术被专家投票‘投’没了”的困境。甄选一流人才并充分发挥作用的问题不解决,实施一流人才战略就是一

表1 2024世界大学自然指数前10排名

排名	机构	2024世界大学自然指数排名(top10)		*诺贝尔奖数 (1901-2024)
		COUNT	SHARE	
1	哈佛大学	3660	1109.11	152
2	中国科学技术大学	2232	753.44	0
3	北京大学	2763	729.22	0
4	中国科学院大学	3603	722.93	0
5	浙江大学	1852	710.89	0
6	南京大学	1678	681.24	0
7	清华大学	2239	662.18	1
8	上海交通大学	1728	615.77	0
9	中山大学	1507	601.91	0
10	复旦大学	1595	572.30	0

注:*诺贝尔奖学科统计范围:物理、化学、生理或医学、经济学
Nature Index统计时间:2023年8月1日-2024年7月31日

数据来源:自然指数官网<https://www.nature.com/nature-index/>。
笔者将有关数据译为中文并标注各大学所获得诺贝尔奖的数量,其差距巨大,一目了然。

句空话。为此,笔者提出基于代表作的“五步识才”法(简称“五步识才”法),代表作等级与人才等级对应。例如,代表作是优秀成果则用于识别优秀人才,代表作是原始创新则用于甄选一流人才,此法高效合理、易学易用。

第一,确立一流人才标准。一流人才是做出一流成果的杰出人才,一流成果是原始创新成果,也可包括学术上的系统精深、奇艺绝学,两者都需要非常突出,如三农问题专家温铁军、古文字通才李学勤。简言之,一流人才是做出原始创新的杰出人才,厘清、明确“原始创新”的定义至关重要。目前流行的定义是,原始创新意为前所未有的重大科学发现、技术发明、原理性主导技术等创新成果。显然该定义过于模糊、笼统,许多国家自然科学奖二等奖、技术发明奖二等奖都被认为或自认为是原始创新,这显然与事实不符,因此原始创新需要准确的定义。973计划基础研究共性重大问题战略研究组认为:原始创新是指开拓新领域、引领新方向和孕育新学科的重大创新。^[8]后者是开辟新领域的一种高端形态。由此可知,原始创新可定义为:因做出突破性进展而开辟新领域、引领新方向的重大创新。从0到1的“1”是新领域、新方向。是否发生突破性进展,可用突破点四要素(突破什么、如何突破、突破开辟的新领域新方向、一句话概括核心贡献)加以识别和确认。

第二,以“尖”识才,而非以国际顶刊论文数量和人才“帽子”识才。人才评价不仅分类也要分层,拔尖人才、顶尖人才是因为有“尖”(原始创新),对于一流人才的评价无需面面俱到,只需要评价其代表作是否为原始创新等一流成果,一项足以认定。原始创新具有结构特征、过程特征、行为特征、综合特征等,可用原创成果四要点展示(突破点四要素、同题成果清单、学界好评、国际国内同类成果盘点比较及优势),用长期的同行评议替代一次性、一两次性同行评议^[9]。特斯拉CEO马斯克表示,他招聘面试的秘诀是让应聘者讲讲解决过哪些最困难的问题,具体是怎么解决的。真正解决过

难题的人会确切知道是如何解决的,他们知道过程细节,而假装解决过难题者只能说到表面一层,很难深及细节。^[10]马斯克的面试秘诀与笔者的思路不谋而合。突破点四要素是将如何解决难题的过程和细节格式化,并有助于利用AI实现人才、特别是一流人才的甄选与评价。

第三,以长期反馈识才。用长期的同行评议替代一次性或一两次的同行评议,确保客观公正,可克服同行相轻的弊端。长期反馈是指对代表作(原始创新等)的反馈,不是对该人才众多成果的反馈,只有一流成果才能长期积累出较多好评,普通成果则会淹没于时间长河。

第四,以同主题比较识才。是指研究同一主题的学者们以各自的代表作相互比较,如果甲的代表作是原始创新,乙没做出过原始创新(即使乙有许多其他成果和“帽子”),则甲胜出;如果两人都有原始创新,则均可胜出,因为都很杰出,均应认定为“主题学术带头人(主题PI)”“主题名师”,即在该研究主题领域,达到国际上与众不同、国内首屈一指或最好之一,实施主题PI制,使其充分发挥作用。在以“尖”识才、以长期反馈识才的基础上,进行同主题的横向比较,降低了判断的难度,代表作是否为原始创新一目了然,作为外行的用户也能够识别。例如,费孝通是农村社会学的主题PI、主题名师,其著作《江村经济》奠定了这样的学术地位。

第五,以用户需求识才。基于上述四步识才,已基本可以认定一流人才。那么,谁需要一流人才呢?同行未必需要,但用户需要。用户不是同行,不会同行相轻、不会嫉贤妒能,因此对接需要一流人才的用户至关重要。特别是当用户知道,一流人才与二流人才的区别不是程度上的,而是方向、层次、维度上的区别,一流知识服务可让用户后来居上、转危为安、以弱胜强、另辟蹊径。一流人才特征鲜明,在以“尖”识才、长期反馈识才、同主题比较识才的前提下,用户或猎头公司可自主甄选一流人才。开发AI评价智能体(AI评价专家、AI猎头)落实“五步识才”法,甄选一流人才更加客观高效。

(二)同主题横向比较胜出法:破除唯“帽子”的新胜出机制

与一流人才甄选同等重要的是设立、实施使一流人才充分发挥作用的体制机制。唯“帽子”是人才竞争的胜出机制,学者一旦戴上高端人才“帽子”,就备受重视、万事大吉,从学术地位、工作条件、科研经费、博士招生到薪酬收入及福利待遇乃至子女上学等,都得到优待。这种对人才的厚遇没错,怕就怕选错了对象。长期以来我国的一流人才标准错位,误以国际顶刊论文数量作为一流人才标准,结果是我国大学国际排名不断攀升,而原始创新成果总体较少,与巨额科研投入相比极不匹配。实际上国际顶刊论文不等于国际顶级成果,正如诺贝尔奖得主本庶佑所说,真正一流的工作往往没有在顶级刊物上发表,这是因为一流的工作往往推翻了已有的定论,评审员会给你提很多负面意见^[11]。破除唯“帽子”就是破除错误的胜出机制,代之以合理的胜出机制,使真正的一流人才及时胜出。科技史和国际学术界的共识是,一流人才是做出原始创新的杰出人才,凭借一项原始创新足以认定。破除唯“帽子”的关键有二,一是以合理的标准、有效的方法甄选出真正的一流人才,二是使其替代“高帽”人才,担任学术带头人,得到优厚的礼遇。“五步识才”法解决了前者,后者则需要新的胜出机制。为此,笔者提出“同主题横向比较胜出法”,简称“同主题比较法”。

一流人才受到重视分两种情况,一是政府机制,二是市场机制。企业高薪揽才已经比较普遍,使用“五步识才”法即可,需要解决的是政府机制问题。在唯“论文”、唯“帽子”的环境下,高校的动力在于追求大学与学科国际排名的提升,这是高校领导层最显见的“政绩”。延揽“高帽”人才一举两得,一是他们有话语权,可以获取更多科研经费等学术资源,二是他们依靠多发国际论文、国际顶刊论文胜出,这种优势继续发挥,会直接助力大学、学科提升国际排名。“高帽”人才与大学各得其所,却可能

与加强原始创新事与愿违。加强原始创新、实现科技自立自强,需要原创导向的学术评价体系,上述“五步识才”法即可甄选做出原始创新的一流人才,“同主题比较法”则进一步使得一流人才担任学术带头人。考虑目前国际环境和国情,“同主题比较法”是对作为学术带头人的候选人之间进行公开比较,包括四项内容。

第一,确立学术带头人的标准。分两个等级,A级PI是世界一流学者,是双一流高校一流学科的学术带头人,即A级PI是在所在学科领域做出原始创新的一流学者;B级PI是国内一流学者,要求其代表作做到国际上与众不同、国内首屈一指或最好之一。也就是说,A级PI以做出原始创新而达到国际上与众不同、国内首屈一指或最好之一,是一流学科必备的学术带头人,如果暂时缺位,则一流学科暂时摘牌、降格;B级PI是国内相对水平高的学者,至少做到国际与众不同(凭借创新性成果可以做到)、国内首屈一指或最好之一。简单地说,A级PI是世界一流学者,B级PI是国内一流学者,两者都担任学术带头人,都称之为“主题名师”,A级PI还称之为“主题PI”,是一流学科的学术带头人。

第二,明确研究主题。研究主题有四个来源,一是国家标准(GB/T 13745《学科分类与代码》涵盖约60个一级学科、600个二级学科)、教育部目录(111个一级学科,部分学科下设二级学科);二是学术数据库主题词(CNKI、Web of Science等数据库中的高频主题词与CSSCI、SCI等的关键词共现分析识别高频研究主题);三是新出现的高频主题词(如“新质生产力”);四是咨询DeepSeek等高水平人工智能获得的结果。例如科技伦理(学科归属:哲学、社会学、法学等)、科技风险(学科归属:哲学、管理学、社会学、系统科学等)都是公认的研究主题。在研究主题领域做出原始创新,铸就一流研究主题,在国际学术前沿抢占研究主题的制高点。

第三,明确比较内容。A级PI与B级PI两者的共同之处是做到国际与众不同、国内首屈一

指或最好之一；不同之处是前者做出了原始创新成果。“同主题比较法”的比较内容为不同学者在该研究主题领域比较“1+N”：“1”是其一项标志性成果，包括原始创新成果或重大创新成果（共两类，如图1所示）；“N”是该研究主题领域的其他研究成果、获奖、项目资助、学术团体任职、人才荣衔及头衔等，以辅助证明学者在该研究主题的学术成就、地位与影响力。以“1”为主、以“N”为辅，如果比较的各方的标志性成果是原始创新成果，则全部胜出；如果参与比较方的“1”不是原始创新成果，则无论其“N”内容如何丰富，也不是世界一流学者。

	自由探索类	目标导向类
从0到1	例如： 石墨烯的发现。	例如： 晶体管的发明
从1到N	例如： 石墨烯的制备方法	例如： 硅平面型晶体管

图1 基础研究重大成果分类

第四，揭榜挂帅，公开招聘学术带头人。学术带头人无需面面俱到，在符合政治标准前提下，应以学术能力为首要标准，原始创新能力为首选。“双一流”建设的一流学科和争取入选一流学科的高校学术带头人，应予以公开招聘，可毛遂自荐、揭榜挂帅。高校或科研院所的学术带头人，也应该在一定范围内公开竞聘，揭榜挂帅。俗话说，不怕不识货，就怕货比货。以“1”为主、以“N”为辅，可实施有效比较。

同主题比较法有助于落实揭榜挂帅，有助于使得一流人才脱颖而出，避免名不符实者担任学术带头人，劣币驱逐良币。学术带头人不需要负责全面工作，专门负责组建科教团队，在相关研究主题做出原始创新成果，铸就一流研究主题，引领发展。一流研究主题是指在该研究主题（如科技风险）领域做出原始创新，取得世界一流成果，该研究主题即称之为“一流研究

主题”或“一流主题”，显示出其在国际学术前沿的领先地位。同样也有教学主题上的创新与引领，研究主题、教学主题，统简称为主题。建成一流学科的首要标准是在该学科领域，至少有一个研究主题是一流研究主题，可引领学科前沿发展。一流主题建设、一流学科建设、一流大学建设，点、线、面三位一体，构成世界新型一流高教体系。

教育评价、科技评价是指挥棒，教育评价改革、科技评价改革的关键是一流人才评价改革，不仅自下而上实施新的一流人才标准和甄选方法（例如行之有效的“五步识才”法），而且需要自上而下在对高校，特别是对双一流高校的评估及学科评估中，增加两项核心指标。其一，凡是一流学科，至少需要一名一流人才作为学术带头人，否则一流学科称号暂时取消。一流学科的学术带头人应该至少做出一项原始创新成果，在国际学术前沿引领发展。教育部学科评估应设立一流人才评价和原始创新评价的专项，以原创评价为导向，对于全国双一流高校约500个一流学科逐一评估，确认有没有做出过原始创新，其完成人是否为学术带头人，评估公开透明，通过公开实现公平公正。其二，盘点双一流大学的一流研究主题数量。即盘点该大学内，在哪些研究主题领域做出过原始创新而在国际前沿居于领先地位，以及“准一流研究主题”，即做出国际上与众不同、国内首屈一指或最好之一的重大创新成果。机构评估的重要内容之一就是盘点、查验该科研机构在哪些研究主题领域做出过原始创新，在哪些研究主题领域做到国际上与众不同、国内首屈一指或最好之一。经过评估，该研究机构的科研实力与前沿水平一目了然。

基于本文所提出的原始创新定义、原创成果四要点展示与“五步识才”法、“同主题比较”法，笔者为教育评价改革提出关键一招，一举多得：一是甄选一流人才，作为一流学科的学术带头人；二是破除唯“帽子”，认定一流人才不看“帽子”、头衔，不看国际顶刊论文数量，

而以原始创新论英雄,同主题横向比较,公开透明,“是骡子是马拉出来遛遛”;三是甄选主题PI,实施主题PI制,聚焦国际学术前沿;四是因为一流学科至少需要一名一流人才作为学术带头人,一流人才将备受重视,为双一流高校和重点高校争先重金礼聘,以此破除唯“帽子”顽瘴痼疾;五是铸就一流研究主题,抢占科技制高点。研究主题是知识体系的基本单元,一流研究主题是指在该研究主题领域做出原始创新,在国际学术前沿居于领先地位。

(三) AI for PI的主题PI制:原创科教团队是抢占科技制高点的决定性力量

AI for Science不仅助力科研、开展科研,同时引发科研组织变革、科研管理变革、科教融合变革以及双一流建设的变革。互联网、数字技术赋能,使得组织扁平化,AI赋能使得组织精约化,借助AI等信息技术运行“班长的战争”“一个人的公司”等等。在科研上,组织小型化趋势同样明显,“小团队、大能量”“小团队、大创新”,Nature杂志官网在线发表一项研究,发现小团队比大团队更能做出颠覆式的创新成果^[12]。实施PI制,特别是实施主题PI制至关重要,主题PI制强调一流学者担任PI,是国际上与众不同、国内首屈一指或最好之一,是在全国范围内某一研究主题的学术带头人(当然,根据研究主题的热门程度,可以有若干学术带头人,即该学者的研究水平达到国际上与众不同、国内最好之一。通过同主题比较法即可确认),他们或来自大学、科研院所或工商企业,英雄不问出处,都是该研究主题的全局领军人才。主题PI或称主题名师、超级PI,在全国范围内一视同仁,体现出国家提倡的新型科研机构、民营企业与大院名校等公平竞争,以原始创新论英雄,而非以“帽子”、以顶刊论文数量、以科教机构的等级论资排辈。主题PI是机构评估的首要指标,其数量是该机构的核心竞争力,反映出在国际学术前沿的地位。例如,一个智库机构,有多少智库领域的主题PI,直接决定其智库的竞争力。在全国范围,主题PI的数量、专

业分布决定了国家的科技竞争力和国际地位。

主题PI可直接接受国家层面的资助与考核,可以选择继续在原单位供职,也可以带着经费和团队自由选择其他科教单位或企业或落户卓越科教特区。5年考核一次,如果不能继续在该研究主题的国际前沿保持领先,则失去主题PI的称号及经费特供等权益。也可以放宽到只要保持在国际前沿继续占有一席之地,仍然可以延长5年。主题PI或主题名师更加明确是在哪个研究主题(术业有专攻,通常是二级主题或三级主题的专家)领域一流的学术带头人,例如组织创新主题学术带头人(组织创新PI),更便于对接用户。通过“五步识才”法和同主题比较法,公开透明、公平公正选出主题PI,一目了然,接受监督。一所大学或科研院所或科技企业拥有主题PI、主题名师的数量与学科分布,最能反映其科教实力与前沿领导力。主题PI以产出原始创新为标志、为学术品牌,两者绑定,就像高锟-光纤通信、谢赫特曼-准晶体、波兰尼-默会知识、费孝通-差序格局,易于识别,易于接受监督,这与虚化的人才帽子完全不同。

AI for PI包括三点:一是助力甄选一流人才担任PI,并配以AI助理;二是赋能PI打造以一流人才为中心的卓越科教组织,使得主题PI团队、超级PI团队成为一流学科的核心、成为科研国家队的组元;三是科教融合,主题PI团队、超级PI团队也是原创科教团队,前瞻引领,抢占科技制高点,同时自主培养一流人才和领军人才。AI for PI形成一个人才微特区,营造原创小环境。将原创科教团队作为一流学科评估的重点指标,是破“五唯”、立新标的重要举措。

做成事的关键是选对人,实施一流人才评价改革迫在眉睫。利用“五步识才”法和“同主题比较”法可快速甄选一流人才,使其担任学术带头人,成立原创科教团队,抢占科技制高点,加快实现科技自立自强。

一部科技史,就是创造者创造的历史,一流人才的作用至关重要。科研,特别是基础研究、理论研究、原始创新、人文社科、智库咨

询研究,一流人才的作用更是决定性的。正如杨文采院士指出,“从0到1”的产生来源于极少数特殊人才,他们能思考“将要到来的关键问题”。^[13]施一公院士指出,“一个国家的科技实力,并不是看所有科研人员的平均水平,而是看顶尖科学家的水平”。^[14]在博士级智能AI、专家级智能AI问世并快速普及的今天,一流人才的作用更加突出,以一流人才为中心构建卓越科教体系迫在眉睫^[15],特别是在智库咨询领域应该对接一流人才(主题PI),而非行政化地对接智库机构,事关决策,唯有杰出。

四、AI教育革命,重构教学系统,培养AI无法替代的人才

AI大爆发,特别是以GPT-o3、DeepSeek等为代表的AI大模型已经达到博士级智能、某些领域达到教授级智能,而且AI前沿竞争激烈、迭代迅速,AI智能水平将持续提升。可以预测,比人类聪明的通用人工智能(AGI)在短期内实现的可能性将大大提高。这意味着现行高等教育面临前所未有的严峻挑战,其培养人才的最高学位是博士学位,理论上现行教学体系已经不能大量培养出超过AI的人才了,传统的STEM学科多数将逐渐被AI替代,AI可能大规模替代传统的学士、硕士、博士,高等教育正在进入“柯达时刻”。为了培养AI无法替代的人才,亟须重构教学系统,填补AI盲区,构建与AI互补、与AI共生的知识体系与实践智慧,这是世界各国高等教育的头等大事与当务之急。

熟练使用DeepSeek等高水平AI只能做到不落伍、不掉队,却不能拥有竞争优势,只有掌握前瞻、原创、跨界、融合、怡情、审度、自省、社交、领导等AI无法替代的能力,才有独特优势。但是现行的教学体系并没有侧重培养这些能力。逐一培养这些能力是不够的,应重构教学系统,大幅调整、补充新理科、新工科、新文科、新医科、新艺科、新农科等,既要利用AI赋能,又要填补AI盲区。

教学系统包括教师、教学内容、学习者、学习环境、评价等五个基本要素,这些是重构教学系统的主要内容。这里主要探讨前三项。关于教师,只有一流学者才胜任教师,上述“五步识才”法可有效识别一流学者,成立主题名师工作室,主持相应领域的教学与科研工作,建立组织保障体系,从集中式学校形态转变为分布式“主题名师+AI助教”网络形态与集中式学校形态并存,让学习者有更多选择,让一流教师充分发挥作用,乃至名校让位于名师^[16]。关于学习内容,创建内容的原则有三:一是AI赋能,AI+传统内容,比如AI+STEM,实现精英教育普惠化;二是填补AI盲区,培养AI不擅长的内容,包括前瞻、原创、跨界、融合、怡情、审度、自省、社交、领导等九种能力,由此培养出AI无法替代的人才;三是培养借助AI、借助开放教育资源等的自学能力、与AI互补能力、人机共生能力和自我修炼,根据需要进行自学,培养站在未来看今天的逆向思维、安全是发展的前提、致善才能致富、行稳致远等新思维、新价值观、新理论方法,培养审美、判断、权衡等审度能力^[17],这是急剧变化时代以不变应万变的法宝。关于学习者,根据AI智能水平和机器人智能水平的提升,把学习者分为两类,一类是可以做到“AI之上”“AI之外”(统称为“AI之上”),不被AI替代者,另一类是被AI替代者。对于后者要加强心理疏导,使其能在赋闲中找到生活的意义,而“AI之上”的学习者则成为高等教育培养的重中之重。一流人才培养一流人才,采取以主题名师网络培养人才的方式,N对1方式。O-Ring理论表明,一流人才的能力具有互补性,他们共同工作时,其整体生产力呈指数级提升^[18],因此多位一流学者培养一位学习者能产生奇效。以此可以培养AI无法替代的人才,为社会贡献“原理可行、技术可行、经济可行、社会可行、生态可行、人道可行”的成果、产品与服务,这六个“可行”简称六“可行”或6F(Six Feasibilities)。“人道可行”包括伦理可行与人文关怀,比如脑机接口技术,用于治疗疾病

可行,但是普遍用于健康人群则不符合“人道可行”原则,明显是短期行为,因为一旦实施,技术会不断升级换代,不同企业还会竞争,软件会有“漏洞(Bug)”“后门”“病毒”,使用者被企业或黑客控制,一旦使用这种智力兴奋剂,就会内卷,主动或被迫不断升级换代,脆弱的人脑必然被不断损害、永久损害,人类集体自残。脑机结合并不能优于AI机器人,因为人脑迟缓、脆弱,跟不上AI的速度和升级换代,与AI脑机对接只能拖累AI,就像阿尔法零与世界围棋冠军脑机对接,肯定变得迟缓而不敌阿尔法零。应对AI风险需要制度变革,发展只有致善才能致富的致善经济与致善制度,是当年只有守法才能致富的升级换代。创造六“可行”(6F)的成果、产品与服务,科技致善,致善创新,实现人类社会的安全与可持续发展。

五、结语

AI大爆发开启全新时代,人类面临前所未有的挑战与机遇。DeepSeek横空出世,华为、新疆等民营科技企业表现卓越,他们能在国际市场上与国际顶级科技企业一决高下,有赖于其运行与管理机制与时俱进,与国际先进管理体系对齐,加之聪明勤奋,结果必然如此。而我国的双一流高校与科研机构的运行与管理机制仍然较为守旧,行政化、“管理者通吃”“五唯”等现象较为突出,导致在原始创新、科技制高点、前沿引领方面与国际领先水平存在显著差距。因此,深化科技体制改革和教育改革,潜力巨大,任重道远。为此,本文在AI大爆发的时代背景下有感而发,进行上述探索。

博士级智能AI、专家级智能AI的出现,导致人才问题归要为一流人才问题,人才战略升级为一流人才战略。“双一流”建设,突出的是学校领导、学院领导的重要性,本文提出增设“一流研究主题”以突出一流学者的至关重要。“点”(研究主题)、“线”(学科专业)、“面”(大学)三位一体、多管齐下,才能培养AI无法

替代的人才,快速建成新型的世界一流高等教育体系。唯“帽子”顽瘴痼疾是实现科技自立自强的最大障碍,需要自下而上、自上而下多管齐下加以解决。基于以上研究,建议教育部对双一流高校约500个一流学科进行学术带头人和原始创新的专项评估,每个一流学科至少有一位一流学者作为学术带头人,这是必备条件,否则暂时摘牌、降格。“五步识才”法、“同主题比较”法可以高效合理地甄选一流学者,盘点原始创新成果,普查其研究主题分布,根据国家战略需求与构建人类安全共同体的需要,扬长补短、优化布局;盘点主题PI、主题名师,实施主题PI制,加强一流学科建设、一流主题建设。同时实施揭榜挂帅,公开透明,符合条件的一流学者可以毛遂自荐,通过同主题横向比较,公平公正,揭榜挂帅。

以一流人才为本,双一流大学、科研院所和高端智库,通过同主题比较,尽快盘点在哪些所属研究主题领域拥有国内首屈一指或最好之一的学者,据此评估作为科教“国家队”的实力与发展态势。如此可破除唯“帽子”弊端,主题PI、主题名师等一流人才将备受重视和礼遇。建立主题PI、主题名师数据库,特别是主题PI是在该研究主题领域做出原始创新的一流学者,在国内数以百计,他们特征突出,易于甄选,是国家最宝贵的战略资源,他们是抢占科技制高点、抢占认知战制高点、抢占未来制高点、为决策者提供一流知识服务的核心力量。

参考文献:

- [1]王志珍院士:挤掉科研“泡沫”需要动真格[N].中国科学报,2014-03-13(01).
- [2]科技界声音[J].科技导报2012,30(26):10.
- [3]熊丙奇.警惕科研“举国体制”[J].民主与科学,2010(3):20-41.
- [4]学术期刊乱象诊断:假刊增刊 买卖版面 虚假引用[EB/OL].(2015-06-08)[2025-01-05].<http://www.chinaneews.com.cn/cul/2015/06-08/7328305.shtml>.
- [5]刘益东.以论文展评模板评价代表作与科技人才[N].中国科学报,2023-02-09(3).

- [6]刘益东. 是国际大学排名还是开国际玩笑? [EB/OL]. (2024-11-06)[2025-02-21]. <https://baijiahao.baidu.com/s?id=1814926575222410868&wfr=spider&for=pc>.
- [7]刘益东. 人工智能B模式: 下一个科技制高点[J]. 延边大学学报(社会科学版), 2024(5): 108-117.
- [8]973计划基础研究共性重大问题战略研究组. 对提升原始性创新能力的一些建议[J]. 中国基础科学, 2004(2): 7-12.
- [9]刘益东. 怎样让一流人才脱颖而出——对一流人才的及时甄选、自主培养与前沿学习革命[J]. 中国经济报告, 2022(3): 112-119.
- [10]马斯克: 招聘员工我通常只问2个问题[EB/OL]. (2024-07-21)[2025-02-15]https://ishare.ifeng.com/c/s/v002M9dupsIftH3E--TVgxFcjm2ReQY23jbYsYlDy1_nxhRw__.
- [11]本庶佑谈科研: 做第一个搭独木桥的人[EB/OL]. (2019-06-25)[2025-02-21]. https://www.sohu.com/a/322930084_773043.
- [12]WU L, WANG D, EVANS J A. Large teams develop and small teams disrupt science and technology[J]. Nature, 2019(566): 378-382.
- [13]杨文采院士: “从0到1”突破的起点在哪里? [EB/OL](2021-03-19)[2025-02-18] https://m.thepaper.cn/baijiahao_11797076.
- [14]施一公, 试答钱学森之问[J]. 中小学科学教育, 2024(1): 13-15.
- [15]刘益东. 粤港澳大湾区建设高水平人才高地的关键——以一流人才体系为中心[J]. 深圳社会科学, 2024(1): 22-32.
- [16]刘益东. AI教育革命加快拔尖创新人才培养与评价[J]. 中国考试, 2024(8): 1-11.
- [17]刘大椿, 等. 审度: 马克思科学技术观与当代科学技术论研究[M]. 北京: 中国人民大学出版社, 2017: 1-6.
- [18]KREMER M. The O-Ring Theory of economic development[J]. The Quarterly Journal of Economics, 1993, 108(3): 551-575.

【责任编辑 刘婉华】

AI-Enabled Paradigm Shift in Science and Education: Tripartite Implementation through Precision Assessment Mechanisms, the matic Principal Investigator Systems, and Pedagogical Architecture Reformation

LIU Yidong

Abstract: Currently, artificial intelligence (AI) is permeating all domains of human social production, daily life, and cognition at an unprecedented pace. The emergence of DeepSeek has particularly heightened Chinese scholars' awareness of AI's tremendous potential and the significant challenges/opportunities it presents for China's scientific and educational endeavors. AI is rapidly reshaping paradigms in education and research through four key dimensions: precision assessment, elite principal investigator (PI) systems, pedagogical restructuring, and talent cultivation. To address critical issues in scientific-educational reform—including massive funding requirements, selection of top-tier academic leaders, principles for instructional system redesign, and cultivation of AI-irreplaceable talents—we propose establishing “world-class research themes” through a thematic PI system. This approach emphasizes the pivotal role of distinguished scholars in constructing a trinity framework that integrates “points” (research themes), “lines” (academic disciplines), and “areas” (universities), thereby forming a new paradigm for world-class higher education systems.

Keywords: DeepSeek; world-class convergent research themes; convergent theme principal investigators; AI for PI; AI-augmented plagiarized manuscripts; high-fidelity AI-generated paper; Transcending AI