

有限与张力：人工智能时代科研规范的挑战与重塑

马光选（云南大学马克思主义学院）*

吴晓文（云南大学马克思主义学院）**

摘要：当前，人工智能正引发一场触及科学认识论核心的深层次变革，对科研规范构成系统性挑战。认识论层面，算法驱动的新“关系性知识”逐渐取代传统的因果分析模式，动摇了科学知识生产的合法性根基；方法论层面，过分看重效率使得科研进程中的偶然性与创造性要素被排挤；价值论层面，学术成果的所有权不明晰、资源分配不均衡以及学术诚信缺失等状况不断涌现，加重了伦理秩序的混乱状况。基于此，本文提议以自然辩证法为理论依据，塑造“人机协同”的动态治理架构，在认识论层面，创建涵盖研究主体、智能工具以及认知环境的多维互动模型，确立“偶然性守恒”的基本准则；在方法论层面，采用分层次的管理策略；在制度创建方面，通过教育改革，评价体系改良以及全球资源共享机制的完善，营造出公平有序的发展环境。本文着重剖析中国路径的独特内涵，把“道器相融”的传统哲学思想创造性地融入人工智能伦理规范体系当中；把“伦理引领—制度弹性—组织赋能”协同治理框架革新性地构建起来；依靠“包容性创新”理念促使全球伦理对话深入发展。

关键词：科研伦理；人工智能；伦理失范；中国路径；科学哲学

基金项目：国家社科基金项目“中越老缅边境农村社会风险治理研究”（项目编号：24XSH026）的阶段性成果。

* 马光选，云南大学马克思主义学院副教授，博士，博士生导师，研究方向为高等教育管理，E-mail: yunda125@163.com。

** 吴晓文，云南大学马克思主义学院研究助理，硕士，研究方向为数字时代科研规范的体系构建，E-mail: 2216499143@qq.com。

I. 引言

人工智能带来的“科学革命”已经深入到认识论的核心区域，正在改变着人类“探究自然”与“获取知识”的基本模式，这种颠覆性技术不但极大地改变了科研活动的形式，而且引发了对科学本质及其正当性的哲学探讨。其影响早已超出单纯的科技应用范围，直接渗入到科学实践的关键逻辑环节当中，促使学术界重新审视并反思现代科学体系赖以存在的理论框架和价值评判准则。

数字时代，算法能够从海量的数据和信息中挖掘出无数复杂变量间的关联，生成“关系性知识”^[1]，并输出为结构化的结果。这就使得传统意义上基于因果关系的逻辑推理受到极大冲击，进而引发认知范式的重大变革。在“数据统治”^[2]思维的影响之下，直觉判断、理性分析与偶然顿悟等曾经支撑科学研究的重要因素正被逐步削弱，甚至随时有可能被取代或被淘汰。这也就意味着，科学研究从以往的人类理性探索模式渐渐转变为算法主导模式，后者更加重视效率改善与预测能力的加强。这种转变既凸显出技术革新的重要性，也反映出科学规范体系的深层次问题和潜在风险。

这一情形并不只是理论上的推断，更多的是工具理性扩张超越科学精神核心价值的现实反应，人工智能依靠数据处理及模式识别方面的特长，正在重塑着“发现”与“理解”的内涵和外延。在此过程中，知识生产过度集中于数据相关性分析，忽略了对自然规律本质的研究。科研步骤被高度标准的操作流程主宰，令研究者主体思维与创新力遭到遏制，科学研究慢慢变为效率竞争的道具，背离了服务于人类利益、探寻客观真理的根本目的。这大概会对作为人类认识活动重要范式的科学规范产生严峻考验。该议题的解决迫切需要哲学社会科学领域进行深入的理论探究并给出系统回答，从技术角度讲要弄清人工智能的功能界限，还要深入分析其潜藏的风险，站在哲学角度就要全面考量人工智能给科学认识方式，研究范式以及价值取向这些关键领域带来的影响，而且要努力塑造符合人工智能时代的新型科研体系，促使人机协同创新，还要加大伦理规范研究力度以保证科技发展稳步前行，这些都是学术界亟须冲破的重要方向，本文也将就此展开深入探讨。

II. 挑战：科研规范的结构性危机

人工智能在科学研究领域中的应用已经不再局限于“辅助工具”这一基本定位，而是逐渐成为知识生产体系中不可或缺的知识生产构成要素。人工智能在提高科研效率的同时，也对长期以来所形成的稳定的知识结构造成了系统性冲击，并最终引发深层次的“结构性危机”^[3]。这种危机源自于人工智能高速发展同传统科学研究认识论、方法论以及价值论之间的矛盾。具体而言，随着算法冲破数据处理边界，直接介入知识生产过程，使得科学研究中依赖“认知主体”与“客观对象”之间二元对立的传统科学研究范式，首次遭受到哲学层面的质疑，并由此暴露出现有科学研究范式所存在的缺陷和不足。人工智能技术的迅猛发展使得科学知识的基本属性、生成过程和验证方式等诸

多根本性问题具有了前所未见的动态演化特征。这充分表明科学研究当下所遭遇的规范危机已然超越了传统意义上的技术伦理，触及到了科学知识生产的合法性根基、认知过程有效性的边界以及学术评价体系的深层架构。其冲击范围并不仅仅局限于实证科学内部，更扩展到了更为深刻的科学哲学范畴。那么，在人工智能时代，要想想准确把握这一危机的总体态势并深究其根源，就不得不站在三个相关联的重要维度上展开分析：一是知识生产过程中的本体论困境；二是研究行为开展中的方法论消解；三是学术伦理秩序的价值论失效。

2.1. 知识生产合法性面临本体论挑战

人工智能时代科研规范所面临的首要挑战是科学知识生产合法性的本体论挑战。传统的科学研究范式将严密的因果推理作为核心，意图找到各种自然现象之间的本质规律，而深度学习等现代人工智能技术仅仅依靠处理海量复杂数据，便能找出变量之间的统计学关联和因果关联，生产出“关系性知识”。然而，这种依靠相关性的认知方式存在着明显的“黑箱”^[4]特征，容易脱离“可证伪性”^[5]的科学核心原则。也就是说，由于算法决策过程中存在着内在隐蔽的特征，使得科研工作者难以追溯其推理过程，因此也就难以进行结论的逻辑性分析和实证性验证。一个常见的情况是，同样的人工智能算法在不同的计算机环境下处理不同的数据集时，可能会输出差距较大的几种结果，这样便会严重冲击科学研究中“可检验性”与“可重复性”^[6]的基本原则。同时，人工智能技术的发展也催生出算法的“幻象”^[7]这种新型本体论问题。人工智能通过高精度大语言模型，往往能够生产出逻辑严密、格式规范但却严重脱离现实情境的“科研成果”。例如，在部分专业领域中伪造包含虚假实验数据或信息的学术论文，类似的情况会给科学研究的真实性和客观性带来一系列冲击。在这种状况下，科学研究也许会日渐背离探寻自然法则以得到客观认识的核心价值，最终转变为纯粹的数据处理和模式生成工具。这种本体论上的改变，清楚地体现出人工智能对知识生产合法性的重大影响和潜在风险。

2.2. 科学研究偶然性的方法论湮灭

人工智能时代科研规范所面临的第二重挑战是科学研究中的偶然性将在人工智能的精准计算中趋于消散。纵观科学研究发展历程，会发现其间蕴含了无数的意外发现与灵光乍现。科学的进步往往也是依靠那些非线性过程偶然的交汇，或是依靠科学家本人深厚的积淀与敏锐的洞察力所带来的顿悟。当前，虽说人工智能支撑下的新研究范式在运行速率与性能改进等方面具有无可比拟的显著优势，但同时也削弱了上述那些科学研究发展中偶然性的独特魅力。人工智能算法，尤其是那些需要大量数据参与训练的大模型，大部分只是想要在某个特定领域内做到“局部最优解”^[8]，而且整个过程就像是不断扩大的工具理性，在穷尽所有可能的过程中来接近最完美的结果。这种方法虽然

极大地提高了工作效率，但是却难以避免地压制了科学家们从事科学研究时最为重要的一些偶然性因素、不确定性以及探索未知事物的精神。

诚然，人工智能在应对高度结构化且具有明确指向性的工作时有着明显的优势，但它在诱发创新思维或者促使理论取得突破等方面仍然有着明显的不足。算法虽然可以高效地完成某些任务，但是很难自动地去冲破既有的研究框架并质疑这种框架的有效性。例如，在那些需要人工智能技术参与的科研项目中，出现意外成果的频率变得很低，这就同科学哲学家托马斯·库恩所提出的“范式变革”^[9]理论形成了鲜明反差。根据库恩的观点，“反常现象”通常是促使科学范式发生改变的主要力量，可是当下的以人工智能为基础的科学已经在逐渐减弱这种机制的效果。这种现象显示出技术伦理方面的深层次矛盾，虽然人工智能在某些情形下有着明显的效率改善潜力，但是它潜藏的负面效应也许会损耗人类探索未知领域的潜力。在这个过程中，科研人员不可避免地从事知识生产者、变量调节者和结果验证者等角色向着数据标注者的角色转变，而后者往往是完全“工具化”的。特别是在问题提出与概念界定等关键环节，科研人员的功能定位已经产生了比较明显的偏离。这是因为深度学习算法具有归纳推理的能力，相较于科研人员，其内部逻辑更加适配于解决那些结构非常清楚的任务。因此，便很难再推动原创性思考或是引领革命性的研究方向，最终，就导致科学革新所依赖的动力源泉慢慢开始衰弱。

2.3. 学术伦理秩序的价值失范

人工智能时代科研规范所面临的第三重挑战在于学术伦理秩序的价值失范。首先，传统意义上的学术伦理框架大多是以人类为研究主体展开，但是，随着人工智能深度嵌入知识生产过程，许多新的伦理问题由此浮出水面。其中，学术成果署名权及其归属界定不清的问题尤其明显，这表明当前的学术评价体系 and 知识产权保护机制仍面临着诸多考验。例如，当人工智能技术被用于撰写学术论文或分析核心数据时，有关人工智能主体性地位的议题便不再只是技术层面的问题，而是变成了一个包含学术伦理和哲学思辨在内的深层议题。这种现象既对传统认知论框架下的“主体性”概念提出了疑问，也促使学者们开始重新思考知识生产过程里主体所扮演的角色、独有的价值以及应该承担的责任范围。同时，资源分配的非均衡性极大地加重了全球科研领域的公平性困境。例如，获取先进人工智能大模型所必需的关键计算能力、优质数据集以及专业技术人才的成本极高，而且这类资源在全球范围内存在着明显的地域分化特征。这就导致那些身处资源稀缺地区的研究人员在获取并高效利用这些核心要素时往往遇到更多阻碍，这不仅妨碍了他们参与到以人工智能推动的前沿性科学研究中去，也进一步巩固了历史上由殖民主义知识生产模式留下的不平等架构，与联合国教科文组织所提倡的“科学应当服务于全人类共同利益”这一宗旨背道而驰。

此外，随着人工智能技术的发展，学术诚信建设面临着前所未有的挑战。人工智能领域不断出现的深度伪造等新技术，使单纯的依靠技术手段去解决伦理问题变得捉襟见肘。这类技术手段不但无法有效的防止学术不端事件的发生，反而有可能使学术界对于技术的依赖增加，学术界的学者们

可能会降低对于道德规范重视的程度。在这种情况下，对于学术评价的标准逐渐模糊化，学术共同体的合作信任和友好氛围会逐渐瓦解，这种现象很好的说明了现代性批判理论中的核心思想，即“流变性约束”无论形式上看起来如何的严密，本质上缺乏一种内在的稳定性，无法根除根本性的价值冲突问题。人工智能领域学术伦理失范现象产生的根源，可以归结为价值观念模糊和重塑的需求，要想有效解决这种深层次的问题，就必须冲破技术上的限制，在哲学思辨、制度设计与跨学科整合等方面展开系统性的探索，从而创建起全面而可持续的治理架构。

III. 重塑：人机协同的辩证进路

人工智能领域的伦理困境，实际上折射出现代性进程中工具理性与价值理性之间深层的张力。要解决这一问题，并非单纯依靠技术创新或者复古思潮，而是在自然辩证法理论框架之内，对矛盾对立的本质特征展开深入剖析，去寻找更高层次的辩证统一之路，进而塑造起人机协同创新的实践范式，这牵涉到技术层面的优化升级，还是一项包含认知范式的重塑、研究方法的革新以及制度文化生态系统重构在内的综合系统工程。在这个过程中，要冲破“人类中心主义”^[10]和“技术决定论”^[11]这两种对立思维模式的束缚，按照马克思有关“人的对象性活动”的理论指导，把技术体系重新融入人类社会实践的历史长河之中，把它当作彰显人类主体性对象化的一种重要途径。只有从认识论的角度进行哲学思考，在方法论上使用实践智慧，在制度文化上形成系统的保障机制，才能超越技术发展本身所存在的局限，为人类的精神创新留下足够的空间，从而保证科技进步真正为人的全面发展和自由解放服务。

3.1. 认识论的重构：从主客二分到辩证统一

自然辩证法所蕴含的哲理，为重塑人机关系给予了重要的理论支持。传统科学认识论当中，主体与客体二元对立的认知模式，在人工智能深度介入知识生产场景的今天，暴露出固有的不足，难以适应复杂多变的发展环境中，急需构建一种更为包容并能够动态演化的“研究者—智能体—认知环境”三元辩证架构，来冲破既有范式所带来的束缚。该模型融合中国传统哲学中“天人合一”的整体宇宙观，重新解读了人工智能的本质特性，冲破了将其当作独立于自然或者人类的外界因素的传统观念，把它归位为自然演化进程中的一种特别的存在形式，表现出人类智能在特定历史阶段的具体化表现。

这一理论框架的关键含义在于确立并守住“偶然性守恒”这一基本原理。尽管算法优化在提升必然性方面有着明显的优势，但是它的发展进程无法避免地缩减了科学探究过程中的偶然性、不确定性以及灵感乍现的空间，这种现象是由于技术演进的必然规律造成的。“偶然性守恒”的提出，并不是要否定算法的效率，而是提倡应当在准确规划和标准化流程的支持下，给科研人员的直觉判

断、创新构想以及非线性思维提供充分的条件和灵活的操作环境。例如，在高度自动化的智能系统当中，恰当地加入随机性或者非确定性因素，可以有效地促使超越常规模式的突破性成果产生，进而推动深层次知识发现和技术革新向前发展。这种设计理念蕴含着过程哲学的主要内涵，现实世界的发展和存在，从本质上讲，就是确定性与不确定性、必然性与偶然性相互作用并不断生成的过程。其深层的理论价值在于，这一设计试图冲破海德格尔所指责的“集置”^[12]限制，重新体现它的“解蔽”^[13]本质，也就是成为探寻世界奥秘，拓宽认知边界的有力工具，给解决技术异化问题赋予了实际操作的方法。

3.2. 方法论创新：分层治理的实践探索

在人工智能技术发展所带来的科研伦理挑战面前，要创建起一套系统化且层次清晰的理论框架，从而对技术复杂性实施精准控制。这个框架要整合多种治理理念和更新要素，全面覆盖科研活动的各个部分，真正提升其实际操作指导意义。

从技术角度来讲，保障人工智能系统拥有可解释性与可追溯性特点，这是重塑科学知识正当性的关键要素。促使可解释型人工智能技术发展并得到广泛应用有着重要意义，其核心目的在于破解算法决策的“黑箱”难题，把隐藏起来的推理过程转变为明晰，直观且容易领悟的因果联系模型或者决策支撑架构，这种更新有着重要的哲学价值。它旨在把科学哲学里的“可证伪性”理念从理论层面落实到实际操作环节当中，以此提升知识生产过程中对理性的反思与批判能力。在实际运用时，加强对科研全过程透明度管理和数据品质把控的做法成为守护学术诚信的一道重要防线。采用区块链等前沿技术对科研数据以及操作流程实施存证，可以完整地记录整个实验设计、数据收集、算法参数调整以及结果分析等核心环节的信息，依靠去中心化机制达成分散式存储，为科研进程创建起透明的监督架构，保证各个阶段都会遭受潜在的伦理检查，进而加大学术不端行为的成本，守护研究的真实可信程度。在成果评定方面，迫切需要形成科学并可衡量的评定指标体系，准确界定人类和人工智能在知识更新方面的贡献比例。目前国际学界普遍赞同采用人工参与度分级声明制度。成果发布的时候，得系统讲述人工智能在数据处理、建模、成果展示以及理论演绎这些不同环节中的实际运行流程，用标准符号做到准确标示和科学说明。这种体系既是评判人机合作速率的现实凭据，又是站在现象学层面仔细探究人类创造性思维在知识产生过程中的特殊功能和无法代替的作用的深入探索。这有利于科研评价朝着越发看重创新内涵方面进步，这种分层管理的形式进一步丰富了交往行为理论在学术界的应用广度，目的是创建一套可以看清、有架构且互相作用力很强的科研准则，让学者群体产生一致看法并且形成新的相互依靠关系。

3.3. 制度保障：文化赋能与生态重建

要推动人机协同研究范式不断向前发展并取得更好的效果，就急需形成一套科学而完备的制度保障体系以及文化支持环境。这个过程非常复杂，综合程度很高，而且具有重大的战略价值和发展前景。

教育体系深层次变革要被当作塑造创新文化的重点途径，科研伦理和科技哲学教育不能仅仅依靠理论讲述，而应该深入融合技术实践情境之中。利用基因编辑、脑机接口以及算法决策等领域内也许会引发伦理争论的案例，借由沉浸式“实践教学”，激发学生开展价值辨别，加强其伦理判断能力，把抽象的伦理认识转变成技术决策里的价值指向和实践智慧。这样的教育方式更加符合实用主义哲学观念，看重通过实际操作或者模仿环境来提升应对复杂问题的能力。学术评价体系必须达成范式转变，当下过于依赖成果数量和影响因子等量化指标的评价手段，实际上体现出工具理性占据主导地位，严重脱离科学研究的根本目的。当下迫切要推进科研评判体系由单纯关注“成果数量”转而侧重“过程创新性”的根本性转变，创建起含有研究假说的前瞻性、试验设想的独特之处以及识别难题的敏感度等诸多方面标准的革新判定框架。这有益于引领科研活动回到探索未知事物与打破常规界限的本质特性上，促进研究者大胆尝试非主流走向，在制度范围内冲破工具理性桎梏，并给予马尔库塞所说的融合了感觉和道理的“新感性”^[14]以实际操作空间。这种转向是激发科研活力的主要制度杠杆之一。在全球治理层面，“数字鸿沟”^[15]正在逐渐变大，做到科技资源的公平安排与恰当分配是创建“包容性科研生态”的关键步骤。这就要求各个国家间必须展开跨国合作，做出制度上的改动举措。通过依靠多方力量共同搭建算力、数据与算法模型之类的关键要素共享平台，尤其着重于帮助欠发达地区和发展中国家能够以合理成本获取先进科研硬件设施。这样既能有力阻止殖民主义在知识制造领域潜藏的风险，并且也是践行全球科研公平观念，践行“科技为人类全体利益效力”这一宗旨的具体途径。这种全球合作的实践，能够很好地体现出全球正义理论在科技治理方面的实际价值。要让人机协同研究模式持续发展下去，就要从教育引导、绩效评定以及资源分配等方面形成起完备的制度体系，还要促使文化更新，营造出充满生机的生态氛围。

IV. 中国路径：传统智慧与科技治理的融合共生

在全球人工智能技术飞速发展的大环境下，中国凭借自身的文化根基和治理智慧，慢慢摸索出一条把传统文化元素同现代科技特点相融合的更新治理之路。这种模式既不是照搬西方经验，也不是单纯沿用旧有想法，而是以五千年的中华文明为根基，在继承发扬古圣先贤理念的基础上，创造性地建立起带有时代印记的制度架构，做到了技术效能与人文价值的有机整合，达成本土实践和全球使命的辩证统一。其根本动因是把“天人合一”“仁民爱物”和“经世致用”这些传统思想要素融入到人工智能等现代科技领域的创新发展之中，以此来体现东方哲学在数字文明创建中的独特价

值，利用中国智慧推动人工智能伦理规范体系的形成，进而守护人类尊严和文明本质的延续。

4.1. 传统创新基因的当代转化

中国技术伦理的思想要义之一，便是要做到“道器相融”^[16]，即技术的发展应该和自然法则、社会规范以及人文价值相协调。这种源远流长的智慧，在人工智能时代焕发出新的生机与活力，而且展现出十分关键的引领作用。先秦典籍里对造物法则展开全面叙述，重点提及顺应天时、体察地气、选取上乘材料并加以改善工艺等等，搭建了包含“天一地一人一技”四方面要素的理论体系，给现代人工智能研发和应用赋予生态伦理准则。此观念要求算法设计必须超越纯粹工具理性约束，将自然生态的可延续性、资源的有限性以及人类深层次的情感 and 价值等要素内化到技术逻辑当中，让代码变得更加温暖，更有人情味一些。

类似的“以自然之道驭技术之术”的中国式哲学思维，在现代科技界焕发出了强大的生命力。将中国传统金属热处理中所蕴含的“渐进”“均衡”和“顺应自然之道”等思想融入人工智能的关键算法之中，用以改善神经网络的学习和训练过程。如此，既能明显改进大模型的性能，也能大量缩减能耗，做到技术效能和资源利用效率之间的辩证统一。这种“师法自然”的创新实践很好地体现出中华传统文化当中“观象制器”的思想，其在数字时代的新发展走向，清楚显示了古圣先贤的智慧对于解决当下技术难题的重要意义。

儒家“格物致知”^[17]思想提倡通过系统探究事物本质来达成真知获取，其中所蕴含的“渐进性”“实践性”和“整体性”思维给人工智能认知模型的建构提供了重要参考和启示。现代认知科学和多模态学习技术的持续创新和突破，使得将“物格而后知至”的认知发展路径可以变成机器学习里的层级化和递进式认识机制。一些学者试图塑造包含东方智慧的认知架构，想要模仿传统心学里“事上磨炼”的实践形式，看重知识要在某种情形之下经过实践加深与改进的过程。这种技术架构依靠具身化学习范式，促使人工智能系统冲破传统数据驱动的模式，慢慢在复杂场景里开展价值评判和伦理推断。这是一种融合本土哲学思想的更新尝试，表现出中华工匠精神由传统物质生产范畴向数字化智能创造范畴的深入转变与发展路线。

4.2. 科技治理模式的制度创新

中国在科技伦理治理方面存在明显优势，这是由中国打造的“伦理价值引领、制度规范弹性、组织效能提升”三位一体协同模式所带来的。这一模式冲破了西方治理架构里常见程序化趋向和形式主义局限，而且有效地防止了“技术乌托邦”^[18]观念也许会引发的价值虚无危险，做到了原则性和灵活性的完美结合。给应对技术快速进步带来的伦理难题给予了强有力的制度保证和操作支持。

从制度建设角度来讲，我国正在形成起多层次、多维度的复合型治理体系。在宏观层面，通过

提倡性规范文件来确定人工智能发展的价值导向和基本准则，促使社会达成伦理共识；在微观层面，依靠具备法律效力的法规草案，明晰技术研发以及应用过程里各个主体的行为边界和责任划分。这种“软法引领方向，硬法守住底线”的制度安排，体现出中国传统“礼法融合”^[19]治理理念的现代转变，目的是在技术飞速改变之时，永远守住人的主体地位和尊严。特定区域的试点经验体现出组织力量深入技术治理的特别效果。以党组织为主心骨塑造起来的科技伦理审查体制，它的功能超越了传统的行政监督范畴，渐渐变成具有“价值引导”和“共识形成”两种职能的重要体系。依靠开放性、民主化的决策平台，把群众路线观念融入到算法风险评判和伦理审查全过程中，大量吸收不同主体的意见，在技术应用之初就全面识别并规避可能存在的伦理危险，特别在个人隐私保护之类的领域收获了明显成果。依靠系统化的组织架构设计，促使企业、高校和科研机构等各类主体加深伦理治理合作，引领创新型组织创建符合技术风险特点的伦理影响评价机制以及防控体系，这种以价值引领为主导，重视多方协同参与的治理模式冲破了西方传统“政府规制”与“市场自律”二元对立的既有模式，在技术创新过程中重新塑造了公众对于“公共福祉”的认知框架和行为准则。

4.3. 全球伦理对话的中国方案

全球化进程中，西方国家在技术伦理等方面占据着主导地位，并表现出文化霸权的倾向。“人类文明新形态”^[20]这一概念为中国提出的新型文明对话机制给予了理论支撑。它鲜明地反对单一文化或伦理标准的绝对化和普适化主张，在人工智能等前沿科技的价值判断及规制过程中，要充分顾及全球文明的多样特性，创建起以多元文明价值观为根基的包容性评价体系。

这一理念将儒家的“仁爱”观念同“差序格局”^[21]的特点融合起来，又吸收了非洲社群本位观所包含的集体主义精神，还把伊斯兰教义里关于诚信和善治的主要成分纳入进来，从而形成了一套具有多维度特征的技术伦理评判架构。这个架构没有采取固定的或者标准的形式，而是设计出一种适应于跨文化情境下的技术应用调适机制。这种既有普遍适用性又带有地方特色的“和而不同”伦理实践范式，正在逐渐冲破技术殖民主义的知识障碍，给全球技术治理增添多元对话和包容共生的价值理念。有开创性意义的实践范例是新兴国家围绕人工智能伦理构建的协作平台，它打破了传统的自上而下的标准化制定模式，提倡以平等对话为基础的“文明互鉴”理念，其研发出的伦理评价工具有着非常新颖的运用情境化方法论框架，充分尊重各个文化社群独特的价值取向以及敏感领域。该工具通过深入剖析并融合当地主流宗教哲学里有关公共福祉与神圣空间的关键观念，重新界定了针对宗教场所开展人脸识别等技术应用时的伦理标准，从而化解了社区成员对于技术会触碰宗教禁忌的忧虑。国际实践表明，非西方伦理体系与价值观念不是技术发展的桎梏，而是解决全球科技治理困境的一笔重要财富，依靠多文明传统的智慧，“旧瓶”中也蕴含着解决现代价值相对主义与伦理分歧、以及多样化创新路径的可能。而中国模式给予的主要启示是：伴随人类科技文明从单一种子现代性迈向多样性现代性转型，建立更开放包容的文化对话机制，在相互尊重协商的基础上达成共识并共同塑造未来的方向，已成为需要解决的时代课题。

V. 结论：在技术边界开拓人文领域

人工智能技术飞速发展，促使人类文明迈向崭新历史阶段。以算力为根基、数据为核心、算法为动力的认知革命，其影响早已超出传统技术演进范畴，彻底改变人类生活方式的本质特征。在人工智能时代，技术效能与人文价值相互交融，形成协同共生且动态平衡的关系，“技术阈值”成为文明演进过程中的关键议题和人文精神复苏的契机。面对算法逻辑给科研范式和伦理秩序带来的系统性冲击，要冲破工具理性的桎梏，用辩证唯物主义方法论保证技术创新进程中人的主体地位，在数字化进程里塑造文明发展的价值导向。

人工智能带动的认识改变根源要从现代性进程中的工具理性和价值理性的深层冲突找起头来，因为算法准确掌握并有效操作客观规则之后，人们靠直觉决定、偶尔发觉与因由推理的能力正逐步衰减掉，这不是文明倒退的象征意义表达出来，恰恰是黑格尔所说的“理性的狡计”^[20]在技术领域的生动展开。当机器的聪明才智不停地冲破常识范畴的时候，同时也在帮助人们反省自己的身份状况来重新界定它。就像恩格斯通过分析发现的辩证法则一样，在科技创新带来难题的情况下同样能为人文方面开创出改进进步的新途径。如此一来，整个社会的向前发展得以延续下来。中国科学家的研究成果显示，“道器相融”的东方哲学思想在量子通信、深空探测等前沿领域有着独特的理论意义和实践价值，它把技术伦理融入科研范式，既推动技术创新又兼顾人文关怀，这体现出真正的创新不是单纯的技术比赛，而是工具理性与价值理性相互融合的过程，这个过程通过技术应用来提升人类的精神境界，从而做到技术创新和人文发展相互推进。

当代中国现代化进程中，科技伦理建设既是塑造文化主体性的关键路径，也是构建全球技术治理体系的重要部分，这种主体性不是简单地继承传统文化的封闭形态，而是依靠五千年的文明积累，借助创造性转化和创新性发展传统智慧来达成。从《考工记》提出的“天时—地气—材美—工巧”生态造物观念，到宋应星《天工开物》体现的尊重自然规律的思想，中华文明技术哲学一直把器物创造放在“天人合一”的宏大宇宙观里做系统性思考。进入人工智能时代以后，这些深厚的民族文化基因重新焕发生机，中国科技治理革新形成起以“伦理引领、制度包容、政党统筹”为主干的多层级协同机制，从而避免了西方案程序化思维的局限性以及技术乌托邦思想的空泛化倾向。

《新一代人工智能伦理规范》和《人工智能法（草案）》共同构成了“柔性引导与刚性约束并存”的治理体系，很好地体现了“礼法相济”的传统治理智慧。浦东新区作为国家级人工智能创新发展试验区，依托这一制度框架具有明显优势，依靠党委主导的伦理审查机制，通过技术伦理“红色议事厅”吸纳社会公众参与算法评估，预防潜在风险；依托专业平台促进产学研协同创新，推动企业采取差异化应对措施，在技术研发环节重塑公共价值共识，为全球治理提供突破“国家—市场”二元模式的新方案。

通用人工智能技术也许会带来文明危机，在此情况下，形成超越民族国家界限的全球伦理共识

体系变得十分紧迫，中国方案以“增强型人类文明”作为核心范式革新，这既是马克思“人的自由全面发展”理念的现代诠释，也是《礼记·礼运》大同理想在当代的再构。该范式在国际治理范畴提出三项重要支柱：一是创建算法决策里的“人类否决权”伦理机制，保证技术发展一直服务于改善个体尊严与主体地位的目的，防止技术异化或者消解人的本质属性；二是规划代际公平的技术影响评判架构，把《吕氏春秋》中“长利”观念变成推进文明可持续发展的制度安排，慎重衡量短期利益和长远福祉之间的联系；三是推动创新成果的全球共享机制，通过实质性的技术转移和开放知识协作，彻底消除“数字鸿沟”造成的文化隔阂，践行联合国教科文组织提倡的“科学普惠”原则。“一带一路”倡议框架下算力共享网络的搭建，同南南人工智能伦理联盟提倡的“文明互鉴”理念非常吻合，这个联盟依靠“情境文化敏感性”，努力开发文化算法评价工具，这很好地体现了“包容性创新”的核心价值取向，这些实际操作有效地解决了技术殖民主义造成的知识壁垒问题，表现出非西方伦理思想既是一种阻碍，又可以成为解决全球治理难题的关键资源，一起绘制出“和而不同”的科技文明发展蓝图。

在智能文明发展的新阶段，人类要从《尚书》里“正德、利用、厚生、惟和”的理念当中去发掘其思想内涵，这个理念看重技术革新要以改良个体道德素质和生活品质为根本目的，依靠科学调配资源、依照自然规律、守护生态平衡和社会稳定来做到可持续发展，人工智能飞速发展不能仅仅局限于工具理性，而是变成推动人文精神加深的关键契机，针对技术引发的双重难题和机会，务必运用辩证思维去统筹工具效果和价值引领的关系，在算法逻辑之外拓展新的创意，在大数据分析时守住人性尊严和主体意识的根本价值，从而做到技术为人服务、人归于文明本质的历史性统一，进而塑造出一个兼具技术伦理关怀和人文复兴特点的新时代。这既是中国科技伦理建设的核心使命，也是人类文明在智能时代实现螺旋式发展的必由之路。

参考文献

- [1] Tara Safavi & Danai Koutra. (2021). Relational World Knowledge Representation in Contextual Language Models: A Review. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, pages, 1053 - 1067.
- [2] 梁颀. (2021). 字母表、数据统治与数字孪生：一切都是现在——对德里克·德克霍夫的访谈. 东南传播, (11):85-89.
- [3] 伊曼纽尔·沃勒斯坦 & 张发林. (2017). 结构性危机：一次迥异的危机. 北京大学学报(哲学社会科学版)54(01), 5-10.
- [4] 仇筠茜 & 陈昌凤. (2018). 黑箱：人工智能技术与新闻生产格局嬗变. 新闻界(01), 28-34.
- [5] 江天骥 & 段秀芳. (2007). 可证伪性、可批评性和科学方法. 世界哲学(02), 24-31.
- [6] 刘永谋. (2022). 可检验性与重复性危机：科学知识是绝对真理吗？. 民主与科学(03), 25-30.

- [7] 韩志明. (2019). 技术治理的四重幻象——城市治理中的信息技术及其反思. 探索与争鸣 (06), 48-58+157+161.
- [8] 王冰. (2014). 基于局部最优解的改进人工蜂群算法[J]. 计算机应用研究 31 (04), 1023-1026.
- [9] 郑杭生 & 李霞. (2004). 关于库恩的“范式”——一种科学哲学与社会学交叉的视角. 广东社会科学 (02), 119-126.
- [10] W.H. 默迪 & 章建刚. (1999). 一种现代的人类中心主义. 哲学译丛 (02), 12-18+26.
- [11] 王建设. (2007). “技术决定论”与“社会建构论”：从分立到耦合. 自然辩证法研究 (05), 61-64+69.
- [12] 孙周兴. (2018). 技术统治与类人文明. 开放时代 (06), 24-30+5-6.
- [13] 薛永龙 & 汝倩倩. (2020). 遮蔽与解蔽：算法推荐场域中的意识形态危局. 自然辩证法研究 36 (01), 50-55.
- [14] 范晓丽. (2006). 马尔库塞批判的理性与新感性思想研究. 复旦大学.
- [15] 李哲罕. (2025). 社会再生产视角下的数字鸿沟问题——基于布尔迪厄理论的考察. 求索 (04), 34-40+206.
- [16] 杨羽 & 石连海. (2022). 道器相融：职业院校应用型人才培养机制探索. 职业技术教育 43 (18), 45-50.
- [17] 孙逸超. (2025). 北宋中期洛阳学者与二程的格物论——基于二程语录记录者的解析. 哲学动态 (06), 41-55.
- [18] 李筱. (2024). 恩斯特·布洛赫技术乌托邦思想研究. 山东大学.
- [19] 秦双星 & 薄振峰. (2011). 传统法律文化论——评《礼法融合与中国传统法律文化的历史演进》. 学术交流 (02), 207-209.
- [20] 韩喜平 & 郝婧智. (2021). 人类文明形态变革与中国式现代化道路. 当代世界与社会主义 (04), 49-56.
- [21] 费孝通. (2012). 乡土中国. 北京：北京大学出版社.
- [22] 俞吾金. (2010). 从康德的“理性恨”到黑格尔的“理性的狡计”. 哲学研究 (08), 70-76+128.

Limitedness and Tension: Challenges and Reforms of Research Norms in the Era of Artificial Intelligence

Ma Guangxuan (College of Marxism, Yunnan University)

Wu Xiaowen (College of Marxism, Yunnan University)

Abstract: Currently, artificial intelligence is triggering a profound transformation that reaches the core of scientific epistemology, posing a systematic challenge to research norms. At the epistemological level, the new “relational knowledge” driven by algorithms is gradually replacing the traditional causal analysis model, undermining the legitimacy foundation of scientific knowledge production. At the methodological level, an excessive focus on efficiency has led to the marginalization of the accidental and creative elements in the research process. At the value level, unclear ownership of academic achievements, unbalanced resource distribution, and the lack of academic integrity are constantly emerging, exacerbating the chaos in the ethical order. Based on this, this paper proposes to take dialectics of nature as the theoretical basis and

shape a dynamic governance framework of “human-machine collaboration”. At the epistemological level, a multi-dimensional interactive model covering research subjects, intelligent tools, and cognitive environments should be created, and the basic principle of “conservation of contingency” should be established. At the methodological level, a hierarchical management strategy should be adopted. In terms of institutional creation, a fair and orderly development environment should be created through educational reform, improvement of the evaluation system, and improvement of the global resource sharing mechanism. This paper focuses on analyzing the unique connotation of the Chinese path, creatively integrating the traditional philosophical thought of “the unity of Dao and Qi” into the ethical norms of artificial intelligence; innovatively constructing the collaborative governance framework of “ethical leadership-institutional flexibility-organizational empowerment”; and promoting the in-depth development of global ethical dialogue through the concept of “inclusive innovation”.

Keywords : Research ethics; Artificial intelligence; Ethical misconduct; Chinese path; Philosophy of science

Copyright ©2026 by Author(s). This article is open accessed under the CC-BY License
(<http://creativecommons.org/licenses/by/4.0/>)

