

人工智能时代大国竞逐下国家间关系应对策略

张子宁

扬州职业技术大学，江苏扬州，221000；

摘要：人工智能技术的加速扩散正在深度重构大国竞逐的底层逻辑，技术霸权争夺从硅基算力延伸至算法疆域，数据主权博弈催生新型地缘政治断层，由此引发技术民族主义阵营化、智能军备互信侵蚀与全球治理规则供给滞后等多维冲击。应对上述挑战，需在非对称创新、数据流通制度设计、安全对话机制更新与规范共建四个维度协同推进，以功能性议题的渐进式制度积累对冲地缘政治博弈对国际合作秩序的结构性侵蚀，在竞争与合作并行共存的现实约束下寻求可落地的治理路径。

关键词：人工智能治理；技术民族主义；数据主权；智能军备控制

DOI：10.69979/3029-2735.26.04.099

前言

2025年是新一轮人工智能（AI）技术与应用进程中浓墨重彩的一年，AI从实验室的算法迭代到千行百业的场景落地，正以不可逆之势深刻重塑经济形态、社会生活乃至地缘格局。这一年，技术突破与产业应用深度交织，大国博弈与全球治理并行推进，压注与质疑同频共振，勾勒出一幅复杂而深刻的智能时代新图景。

1 人工智能赋能大国竞逐的核心态势

1.1 技术霸权争夺从硅基延伸至算法疆域

当前大国科技竞争的逻辑已经发生了根本性转变，竞争的焦点也不再是芯片产能和制造工艺的物理层面的竞争，是向模型架构设计、知识表征体系和算法生态主导权的深层延伸。实质上是算力的稀缺性可以通过产业政策来弥补，而算法范式一旦形成路径依赖，后者会陷入“在他人定义的规则框架内追赶”的结构性困境^[1]。美国对华出口管制的深层次目的，不是单纯地阻断硬件供应，是制造算力瓶颈，迫使中国大模型研发在数据规模、训练效率、泛化能力上同步承受压力，从而拉大两国在基础模型质量上的代际差距。但封锁逻辑有内在的矛盾，算法创新具有很强的路径可替代性，稀疏激活、状态空间模型等非主流架构的异军突起，说明技术霸权的边界比半导体供应链更难控制。决定算法疆域归属的，是哪一个国家最早把模型能力转化为行业标准、开发者生态、国际规范，而不是一时的参数规模领先。

1.2 数据主权博弈催生新型地缘政治断层

数据要素的政治化进程，正在全世界范围内制造出

新地缘裂痕，其深浅已远远超出了传统意义上的贸易壁垒或者技术封锁。裂痕的产生机制是各个主要大国就数据跨境流动分别创建起排他性的法律管辖框架，把数据主权从技术治理问题变成国家间政治博弈的关键变量^[2]。欧盟用隐私保护之名推行数据本地化，美国用国家安全之名扩大域外司法管辖，中国用数据安全审查机制强化出境管控，三种逻辑在价值取向上互相抵牾，在制度设计上彼此不兼容，共同推动全球数字空间向封闭化、圈层化方向发展。碎片化趋势对于人工智能发展所造成的影响是结构性的，AI模型的能力边界很大程度上是由训练数据的大小、多样性和跨域覆盖情况决定的，而数据主权壁垒的加强，使得每一个治理体系内部的数据积累都会陷入边际收益递减的境地，各方为了争夺数据控制权，实际上也在系统性地削弱全球AI协同创新的要素基础，这个悖论至今没有在国际上得到有效的制度性回应^[3]。

2 AI竞赛对国家间关系的多维冲击

2.1 技术民族主义思潮下阵营化趋势显现

技术民族主义兴起不是孤立的政策转向，是大国战略竞争向科技领域深入渗透的阶段性产物。当技术能力被纳入国家安全的核心评价维度之后，“技术自主”就从产业政策目标变成了政治话语的重要架构，促使各国在半导体、AI基础设施以及数据生态方面加快自主化推进。美国主导的“芯片四方联盟”、印太经济框架中的数字贸易条款，共同勾画出以技术兼容性为纽带的合作边界，而这一边界并不是封闭固化，东南亚、中东等

中间地带国家普遍保持战略弹性,在不同技术体系之间寻找利益平衡,客观上为跨阵营的技术对话留有制度性空间。阵营化趋势所造成的真挑战,就是技术标准若沿地缘政治的界限持续分化,会从底层架构方面加大跨体系互操作的协调成本,不过这同压力也有同轴线的现实动力,国际标准化组织框架内的多边技术协调进程因此有了新的战略意义。

2.2 数据主权竞争重塑地缘政治互动逻辑

数据流动路径的政治化,把国家间关系的敏感节点从传统的领土、资源问题,转移到了数据管道走向、数据处理权归属上,外交谈判的议程也因此被深度技术化。中欧数据传输协议的阶段性推进、美欧《跨大西洋数据隐私框架》的反复磋商,均表明即便在同盟关系内部,数据主权的制度性张力也持续考验政治互信,但谈判本身的存续恰恰说明各方仍将规则协调视为可行路径而非零和博弈。发展中国家的本土数据通过跨国平台汇聚却无法分享 AI 增值收益的结构不平等,正在使它在全球数据治理谈判中产生更加主体性的利益表达,非盟数据政策框架的逐步完善是典型例证。数据主权博弈的深层含义是把地缘政治竞争逻辑嵌入数字基础设施的物理拓扑里,但各方对于数据互联互通的共同依赖,在博弈过程中不断产生推动规则协商的结构压力,竞争与合作的张力不是单向演进的,而是在具体议题上呈现出动态的利益交汇空间。

2.3 智能军备发展考验传统安全互信框架

自主武器系统和 AI 辅助决策在军事领域渗透,正从两个方面对既有安全互信架构造成实质性冲击。从技术透明度维度来看,传统的军备控制核查逻辑是以武器平台可见、数量可数为前提的,而算法能力的军事化应用具有很强的隐蔽性,现有的核查机制已经不能适应技术发展的需要了;从决策时效维度来说, AI 系统把响应窗口缩短到人类认知无法跟上的时间区间里,人在回路原则的落实遇到技术层面的现实困难,对于意外升级风险的管控提出了更高的要求。值得注意的是,中美两国学界和智库层面就 AI 军事伦理的对话渠道还在继续,对于致命性自主武器系统在《特定常规武器公约》框架下的讨论也在进行中,说明各方对于智能军备失控风险有共同的认识,为危机管控机制、事故归因程序等具体问题达成有限共识奠定了现实基础。

2.4 规则供给滞后制约全球治理共识凝聚

全球 AI 治理所面对的困境是规则供给的速度和质量跟不上技术扩散的广度,而且现有的规则倡议在地缘政治的张力下很难形成有效的制度聚合。欧盟的《人工智能法案》、美国行政命令体系、中国生成式 AI 管理规定,分别代表风险监管、创新优先、内容管控三种治理思想,三种制度在适用范围、执法标准等方面存在制度性的落差,给跨境 AI 应用造成多重合规摩擦。治理分歧不一定会导致规则碎片化的单一结果,在 AI 安全评估方法、高风险场景分类标准等技术性问题上,各方存在共同寻求共识的务实意图, G20、经合组织等平台也在逐渐形成渐进式规范共识。AI 治理的规则竞争虽已成为大国博弈的延伸场域,但技术风险的跨国溢出性使任何单一体系都难以实现自我封闭式的治理,这一共同脆弱性构成推动多边协调的内生动力,也为全球治理制度化进程保留了持续推进的现实空间。

3 大国竞逐下国家间关系的应对策略

3.1 以非对称创新突破技术封锁与阵营化困局

技术封锁的本质是控制关键节点的供给,迫使对手在既定的技术路径上陷入追赶的困境,而对称创新的反制逻辑恰恰在于绕开这个路径依赖,在对手还没有形成垄断优势的新兴技术方向上集中资源形成局部超越。受封锁方理性选择不是投入巨资复制对手已经成熟的成熟技术体系,而是主动在计算范式、模型架构、应用场景三个维度寻找颠覆性切入点,在计算范式上,光子计算、神经形态芯片等非冯·诺依曼架构路线还处在技术窗口期,各方都没有代际优势,具备弯道超车的现实条件,在模型架构上,以更算力实现更强泛化能力的轻量化推理路线,既是对硬件封锁的工程性应对,也可能成为重新定义行业标准的战略机遇。破解阵营化困局,就要把开源生态建设纳入国家技术外交的战略议程——开源模型的广泛传播天然具有跨越阵营边界的渗透力,通过主导具有国际影响力的开源项目,在不引发政治对抗的情况下形成技术生态的软性聚合,这条路径在标准竞争层面的战略价值,已经超出了单纯的技术能力输出的意义。

3.2 以数据流通合作弥合数字地缘政治分歧

数字地缘政治分歧的本质是各方对于数据要素控制权的争夺、对于数据流通收益的不均等分享所反映出

的矛盾,消除此矛盾的有效途径,并不是消灭主权诉求,是设计出既保留主权保留又能够使要素流通的制度。隐私计算技术的工程化成熟给工程化提供了技术基础,联邦学习框架可以实现数据在本地完成模型训练,不出境,可信执行环境保证数据在计算过程中对调用方不可见,使数据主权不让渡、数据价值可流通从理论构想变为可落地的技术方案。在制度设计上可以参考欧盟的“数据空间”,以跨境医疗数据共享、气候监测数据互认等具体的垂直领域为议题,建立以议题为领域的数据流通协议,用场景化的制度设计规避宏观层面的主权争议,用不断积累的合作实践形成可以向其他领域推广的制度模板。全球南方国家数据权益的制度性保障,是促使多边数据治理由大国博弈转向包容性合作的重要变量,在双多边数字经济协定中加入数据增值收益分配条款,是将其变成全球数据治理积极参与者的务实途径。

3.3 以安全对话机制管控智能军备竞争风险

智能军备风险管控的制度建设要从两个方面同时推进,一是技术透明度机制的建立,二是危机沟通程序的更新迭代。前者可操作路径是建立AI军事系统的行为边界申报制度,要求各方对自主武器系统决策授权层级、人工干预触发条件和交战规则设定定期通报,用信息共享的制度降低战略误判概率;后者在现有军事热线基础上嵌入AI系统异常行为专项通报程序,保证自主系统出现意外行为时可以立即启动沟通渠道。中美两国军事AI竞争的危險性在于双方的战略意图不能被对方准确感知,而算法行为的不透明性又进一步加大了这种感知偏差,因此将“算法行为可解释性”纳入双边安全对话的技术议程,是填补现有互信机制空白的优先事项。多边层面,《特定常规武器公约》框架为讨论致命性自主武器系统规范提供了现成平台,把“人类实质性控制”原则转化为具有可核查内容的具体操作标准,是当前多边安全规范建设中技术可行、政治可接受的务实选择。

3.4 以规范共建推动人工智能全球治理制度化

在承认治理哲学分歧客观存在的基础上,找到各方都有制度化动机的最小共识领域,用功能性议题逐步规范化代替整体框架构建的政治博弈。目前最具共识基础的领域是AI系统的安全测试标准和高风险场景的使用

限制,各国虽然在监管哲学上存在根本分歧,但是都对避免人工智能系统造成大规模不可控后果这一点有共同利益,为建立跨国AI安全评估互认机制提供了现实的政治基础。从规范推进的方法论角度来讲,应当先加强国际标准化组织、IEEE等技术性机构在AI治理中的规范中介作用,把原则性共识转化为技术可操作的标准文件,再经由各国国内立法程序实现规范的本土化落地,“技术标准先行、政治框架跟进”的路径,在绕开大国政治博弈的同时,能够在实践中积累出具有实质约束效果的治理经验。与此同时,建立一个具有普遍代表性的全球AI治理机制,必须把全球南方国家的能力建设诉求纳入制度设计的核心议程,防止治理规范再成为技术强国输出自身利益偏好的工具性安排。

4 结语

人工智能改变大国竞争格局的过程,本质上就是技术能力分布和制度供给能力之间的结构性竞赛,目前国际社会所面临的真实困境就是,前者的扩散速度已经远远超出了后者的响应边界。竞争与合作并存并不是外交辞令,它是技术风险跨国溢出的客观现实所造成的一种制度压力,任何一个国家都无法在封闭的体系中将智能技术风险全部内化,共同脆弱性是促使国际协调产生最低限度理性基础的原因。现实可行的路径不是要弥合治理哲学的根本分歧,而是在安全基准、标准互认、危机沟通等各方面的制度实践中不断积累可操作的制度实践,用渐进式的规范沉淀来对抗地缘政治博弈对全球治理进程的侵蚀,这是在大国竞逐格局短期内难以根本改变的情况下,国际社会维持最低限度秩序稳定的最可行的战略选择。

参考文献

- [1] 余南平. 通用人工智能时代的国际权力重塑 [J]. 中国社会科学, 2025, (04): 41-59+205.
- [2] 肖晞, 刘坤焯. 人工智能对国家间关系的冲击及未来选择 [J]. 教学与研究, 2024, (08): 68-81.
- [3] 唐诗. 人工智能对大国权力关系的影响机制研究 [D]. 南京大学, 2020.

作者简介: 张子宁(1996.10-), 男, 汉族, 江苏南京人, 硕士, 讲师, 研究方向: 国际关系 政治哲学。