

# 加强人工智能国际治理和合作

□蔡翠红

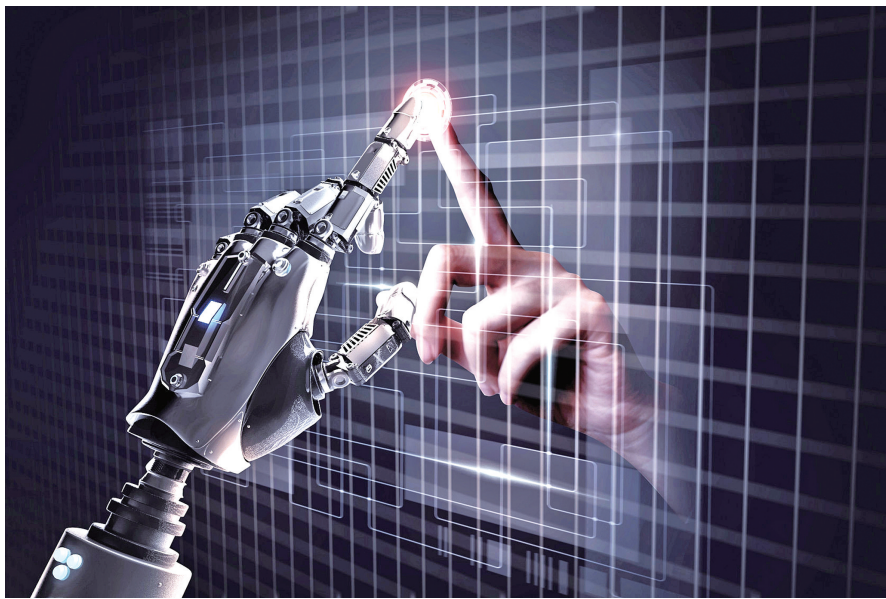
2024年11月,习近平主席在二十国集团领导人第十九次峰会上指出:“要引领数字化转型、数字经济和实体经济深度融合、新兴领域规则制定,加强人工智能国际治理和合作,确保人工智能向善、造福全人类。”人工智能技术的发展不仅为人类带来了前所未有的机遇,也带来了一系列治理挑战。特别是对于数据这个人工智能的核心要素来说,数据安全问题不仅涉及个人隐私还涉及国家安全、经济安全。此外,算法偏见可能加剧社会不公,影响决策透明度与公平性;技术垄断使发展中国家难以公平参与人工智能竞争,导致数字鸿沟扩大;人工智能军事化风险则可能威胁全球安全;人工智能自动化冲击就业,影响社会稳定;等等。如何平衡技术创新与治理规范,已成为全球共同面对的课题。

## 全球人工智能治理的现状

人工智能的发展正在重塑全球治理体系,多个国际组织和国家间合作平台已开始推动人工智能治理框架的建立。然而,目前人工智能治理总体仍处于碎片化、阵营化等滞后状态,缺乏统一的全球协调机制。

多边治理机制初步建立,但缺乏统一全球框架。近年来,联合国、二十国集团、经济合作与发展组织等国际组织相继推动人工智能治理。联合国发布《以人为本的人工智能治理》报告,强调人工智能伦理原则,倡导全球合作;二十国集团提出人工智能原则,鼓励以人为本的技术发展。此外,欧盟实施的《人工智能法案》,也旨在强化对高风险人工智能系统的监管。尽管多边治理机制不断建立,各国在人工智能安全、隐私保护、伦理道德、军事应用等领域的政策仍存在重大分歧,导致人工智能治理框架碎片化。在数据安全与隐私保护方面,欧盟严格执行《通用数据保护条例》,限制数据跨境流动;美国数据政策相对宽松,但在国家安全领域设立壁垒。在2025年法国巴黎召开的人工智能行动峰会上,欧盟代表强调要强化人工智能监管框架,而美国则更倾向于市场驱动模式,这加剧了全球治理的分歧。在人工智能军事化风险治理上,联合国虽就“致命性自主武器系统”(LAWS)展开讨论,但未能形成约束性协议。

全球人工智能标准竞争激烈,规则制定呈现阵营化。人工智能标准的制定决定了技术的全球适用性,各主要经济体在该领域展开激烈竞争。比如,美国强调市场驱动,主张企业主导,推动谷歌、微软等科技公司制定行业标准;欧盟强调严格监管,设立高风险人工智能应用的法律框架。然而,各国在技术标准上的立



场不同,可能导致人工智能市场碎片化,全球技术合作受阻。

人工智能治理面临技术治理与地缘政治的双重影响。在技术方面,人工智能治理涵盖算法透明度、数据安全、伦理原则等多个层面,并对就业市场、社会公平、经济结构带来深远影响。在地缘政治方面,人工智能已成为国际科技竞争的核心,例如,个别国家采取“技术封锁”战略,限制中国获取高端芯片、先进算法和云计算资源,试图在全球科技竞争中占据优势。这种做法可能导致全球人工智能供应链碎片化,影响技术创新的全球合作,甚至可能形成多个封闭的人工智能生态系统。

技术革命迅猛发展,但人工智能治理体系滞后。人工智能技术的突破性进展远超治理体系的适应速度,现有法律法规、技术监管和伦理框架难以有效应对新兴技术挑战。例如,大规模人工智能模型的快速迭代使监管者难以评估其社会影响,生成式人工智能(AIGC)已影响舆论传播、知识产权保护、就业市场等多个领域,而现行法规尚未建立完整的监管体系。治理滞后不仅影响技术发展,还可能导致技术滥用和社会风险扩大。

## 中国在人工智能国际治理中的责任与担当

2023年10月,习近平主席宣布提出《全球人工智能治理倡议》,推动各国协力绘就全球人工智能治理的未来蓝图。中国在人工智能国际治理中支持在充分尊重各国政策和实践基础上,推动形成具有广泛共识的全球人工智能治理框架和标准规范,并在规则制定、合作机制建设、实践推广等方面持续贡献力量。

作为责任践行者,中国在国内积极推进人工智能治理实践,为国际社会提供可借鉴的模式。中国实施了《互联网信息服务算法推荐管理规定》《网络信息内容生态治理规定》等法规,建立了

较为完善的算法监管、数据安全、伦理审查体系,为全球人工智能监管提供经验参考。此外,中国通过“一带一路”数字合作网络共享人工智能治理最佳实践。在数据治理方面,中国提出“数据主权”与跨境数据流动结合的治理模式,探索兼顾数据安全和全球数据共享的平衡方案,为国际社会提供了一种新的数据治理框架。

作为治理倡议者,中国积极参与并推动全球人工智能治理规则的制定。2024年,第七十八届联合国大会协商一致通过中国主提的“加强人工智能能力建设国际合作”决议,140多个国家参加决议联署,彰显了中国在全球治理中的影响力。与此同时,中国在二十国集团、联合国教科文组织等平台上提出人工智能安全、隐私保护、算法透明度等治理倡议,推动全球数据安全规则协调,呼吁各国开发和使用人工智能必须符合《联合国宪章》的宗旨和原则。此外,中国推动人工智能在可持续发展、气候变化、公共卫生等领域的规范应用,让人工智能真正服务于全球社会进步。

作为机制建设者,中国积极搭建全球人工智能合作平台,促进各国在技术、政策、产业等方面开展合作。中国倡导以联合国为核心的多边合作,推动国际标准协调。“人工智能全球治理论坛”,为各国政府、企业、学术机构提供交流平台,协调人工智能政策、规范行业发展。此外,中国推动“一带一路”数字经济国际合作,设立跨国人工智能实验室,实施“一带一路”科技创新行动计划,加强绿色发展、数字经济、人工智能等领域的多边合作平台建设。借助《全球人工智能治理倡议》文件,中国倡导“以人为本”“共商共建共享”的理念和“智能向善”的宗旨,强调人工智能应服务于人类福祉。

## 推动人工智能国际治理与合作的路径

习近平主席强调:“坚持以人为

本、智能向善,在联合国框架内加强人工智能规则治理”。具体而言,人工智能国际治理与合作需要在规则协调、技术普惠、安全可控三大方面取得突破,以确保人工智能的发展符合全人类的共同利益,而非成为技术垄断、数据壁垒和地缘政治竞争的工具。

加快建立全球人工智能治理框架,推动多边规则协调。在联合国、二十国集团等多边机制下,各国应围绕人工智能伦理、数据治理、算法透明度、跨境监管等关键议题达成共识,减少政策分歧,避免治理碎片化。同时,需推动全球法律和标准的对接,在世界贸易组织、二十国集团等平台上协商人工智能相关规则,避免个别国家垄断技术标准,确保全球人工智能产业的公平竞争。此外,还需鼓励区域与双边人工智能治理合作,探索与不同国家发展阶段相适应的人工智能监管模式,构建更加灵活、多层次的治理体系。

推动全球人工智能能力建设,促进技术普惠。加强发展中国家人工智能技术能力建设,通过联合国“加强人工智能能力建设国际合作”决议,推动人工智能技术、人才、基础设施的开放共享,缩小全球数字鸿沟。同时,建立国际人工智能技术共享平台,推动负责任技术流动,避免数据壁垒和技术垄断限制发展中国家的技术升级,确保发展中国家能公平利用数据资源。此外,全球人工智能合作也需要加强国内外企业和学术机构的协作,推动跨国企业共建人工智能产业生态,促进全球人工智能人才交流,避免“人才保护主义”阻碍技术进步。

强化全球人工智能风险治理,确保人工智能安全可控。人工智能技术的广泛应用带来了伦理风险、算法歧视、数据安全隐忧和军事化风险,国际社会需尽快建立人工智能风险评估与预警体系,在军事、金融、医疗等关键领域完善技术安全评估、算法透明度审查和责任追溯机制,确保人工智能发展符合全球社会责任。此外,各国需加强人工智能伦理与算法监管,健全全球范围内的算法歧视审查,制定人工智能透明度标准,以确保人工智能发展符合人类社会的共同价值观。

总之,人工智能的快速发展带来了全球性的机遇与挑战。人工智能治理的国际合作不仅关系到全球科技创新的方向,也关乎数字经济的健康发展 and 人类共同福祉。只有通过国际协调与合作,才能携手打造公平、安全、可持续的人工智能未来,让科技进步真正造福全人类。

(作者为复旦大学美国研究中心教授)  
来源:《人民日报》