中国引领人工智能全球治理的三重路径

陈志敏

习近平主席在2026世界人工智能大会暨人工智能全球治理高级别会议开幕式上发表的主旨讲话中提出“坚持开放共赢，驱动创新发展”“强化风险意识，确保安全可控”“鼓励包容并蓄，促进文明互鉴”“倡导和衷共济，完善全球治理”，从发展、安全、文明、治理四个维度，系统回答了构建公正合理的全球人工智能治理体系的一系列重大问题。

中国积极参与和引领人工智能全球治理进程，同各方一道共同推动全球人工智能健康、安全、公平、有序发展，体现负责任大国的担当。我们要深入学习贯彻习近平主席重要讲话精神，进一步增强人工智能治理理念的感召力、治理制度的有效性和治理行动的赋能力，汇聚全球共治合力，有力推动全球人工智能包容普惠发展，为推动构建公正合理的全球人工智能治理体系作出更大贡献。

理念引领：把准人工智能全球治理的正确方向

“促进全球人工智能朝着向上向善、造福人类的方向发展”，习近平主席这一重要论述为世界各国在人工智能领域跨越利益分歧、凝聚最大共识提供价值引领。中国的人工智能治理理念以增进人类共同福祉为目标，以保障社会安全、尊重人类权益为前提，把人的主体地位、各国共同发展、普遍安全保障、文明多样共存置于同一治理框架，确保人工智能始终朝着有利于人类文明进步的方向发展。

坚持以人为本、向上向善，把增进人类共同福祉作为人工智能全球治理的根本目标。技术进步能够延伸人的能力、拓展发展空间，但技术能力的增长不会自动转化为普遍福祉。人工智能如何研发、部署和使用，归根到底取决于人类作出怎样的价值选择和制度安排。习近平主席强调：“人工智能应当成为人类的可信工具”“确保人工智能始终处于人类控制之下”。这就要求尊重人的主体地位，把维护人的尊严、权利和安全贯穿研发应用全过程，不能让算法逻辑取代人的价值判断，更不能让技术成为操纵社会、侵害权益的工具。中国倡导向上向善，是为创新划定正确方向，使技术进步同人的全面发展、人类文明进步相统一。

坚持开放共赢、创新发展，把共同繁荣置于全球人工智能治理议程的重要位置。人工智能蕴含促进经济增长、改善民生福祉、推动可持续发展的巨大潜力。当前，全球南方国家在人工智能方面迫切的发展诉求和能力短板需得到更加充分的关注和回应。习近平主席强调：“鼓励开源开放、合作共享”“让人人智能赋能千行百业”。中国把发展权、使用权和受益权纳入全球人工智能治理议程，推动全球人工智能治理议程在重视风险防范的同时，充分关注和回应世界各国特别是全球南方国家如何共享发展机遇的问题，为全球南方国家平等进入智能时代提供了重要价值支撑。

坚持公平普惠、包容并蓄，把平等参与和文明多样性作为人工智能全球治理正当性的重要来源。各国不论大小强弱，都有平等发展和利用人工智能的权利，也应当平等参与议程设置、规则制定和治理决策。人工智能处理的不只是数据，还承载着不同国家的语言、文化、伦理和社会经验。习近平主席指出：“人工智能发展和应用不应该侵蚀和损害世界文明多样性和各国文化独特性。要用全人类共同价值塑造人工智能的价值观，善用人工智能技术增进不同文明的理解和包容，促进不同文明交流互鉴，精心培育各美其美、美美与共的文明百花园。”中国主张尊重各国国情、发展阶段和治理选择，既推动弥合基础设施、技术资源和人才能力差距，也重视增强发展中国家在全球治理中的代表性和发言权，避免在人工智能领域造成新的不公。这种兼顾权利公平、机会公平、规则公平和文明多样性的价值立场，是构建人类命运共同体理念在引领人工智能全球治理上的具体体现。

制度引领：推动人工智能全球治理由分散探索走向协同共治

习近平主席强调：“人工智能发展不应该是某个国家的独奏，而应当是全球合作的交响。”人工智能影响跨越国界、风险相互传导，任何国家都难以独善其身、独自应对。同时，人工智能带来的发展机遇前所未有，需要各国合作共享。目前，国际社会已经建立起人工智能领域的一系列治理机制，但仍存在治理机制碎片化和治理能力严重不足的突出问题。中国始终坚持真正的多边主义，通过一系列举措，积极推动联合国体系内的人工智能治理机制建设，倡导推动成立世界人工智能合作组织，以常态化合作机制凝聚治理共识，推动形成联合国发挥重要作用、全球和区域机制相互补充、多元主体协同参与的治理格局。

发挥联合国主渠道作用，维护人工智能全球治理的普遍性、代表性和权威性。联合国是最具普遍性、代表性和权威性的政府间国际组织。支持联合国发挥统筹协调作用，有利于保障各国特别是广大发展中国家平等参与议程设置和规则讨论，有利于维护国际规则的普遍适用。自《全球人工智能治理倡议》提出以来，中国推动联合国大会协商一致通过加强人工智能能力建设国际合作决议，支持在联合国框架下建立“人工智能独立国际科学小组”和“人工智能治理全球对话”机制，并积极参与。建设多边合作平台，正是为了把对话共识转化为长期机制，把共同愿景转化为务实成果。

推动世界人工智能合作组织建设，搭建人工智能全球治理制度平台。世界人工智能合作组织作为全球首个人工智能政府间国际组织，遵循《联合国宪章》宗旨和原则，坚持向善普惠原则，秉持共商共建共享理念，向所有国家开放。其重要意义在于把阶段性会议和政策对话转化为常态化制度合作，把发展战略、治理规则、技术标准、能力建设和联合研究纳入相对稳定的合作框架，为各国特别是全球南方国家持续参与治理、对接资源、开展项目提供新的制度平台。同时，建设世界人工智能合作组织，也将对联合国框架下的现有国际机制形成有益补充，提供有力支持。

推动发展战略、治理规则和技术标准对接协调，增强人工智能全球治理的整体效能。《2026世界人工智能大会暨人工智能全球治理高级别会议主席声明》提出：“建立具有广泛共识、科学、透明、包容的标准规范框架，促进各国间标准互认，降低人工智能国际合作的成本和门槛。”当前，各国人工智能治理仍处于快速探索阶段，制度规范、监管路径和技术标准存在差异，需要针对不同区域和国家的区情国情强化对接的精准性。中国大力推进面向不同区域和国家的人工智能合作机制建设，2024年以来相继成立了中国—金砖国家人工智能发展与合作中心、中国—上合组织国家人工智能应用合作中心和中国—东盟国家人工智能应用合作中心。面向阿盟、非盟、拉共体的国际人工智能应用合作中心也已纳入规划。

行动引领：把中国主张转化为务实治理成果

习近平主席指出：“作为负责任大国，中国在人工智能领域始终致力于做国际公共产品的提供者。”理念的生命力在于实践，治理的有效性在于行动。中国以自身创新发展和治理实践不断丰富人工智能全球治理的经验，又围绕国际社会特别是全球南方国家的现实需求提供技术、人才、平台和应用方案，推动中国主张由原则倡议转化为可感知、可参与、可持续的合作成果。

强化创新发展和开放生态，拓展全球人工智能发展机遇。近年来，中国坚持有效市场和有为政府相结合，加强人工智能科技创新，积极推进“人工智能+”行动，培育各类主体共生共荣的健康生态。“十五五”时期，中国将加快突破人工智能基础理论和核心技术，促进模型算法、智能芯片、云平台和场景应用贯通发展；推进行业高质量数据集建设，完善数据供给、流通、利用和安全治理机制；发展开源开放生态，降低中小企业和广大用户使用人工智能技术的门槛；推进绿色算力基础设施建设，发挥人工智能在能源调度、节能降碳和绿色转型中的赋能作用；促进教育科技人才一体发展，为持续创新夯实人才基础。中国以完整产业体系、超大规模市场和丰富应用场景，为全球人工智能创新发展源源不断提供新的开源技术、市场空间和合作机遇。

统筹发展和安全，为人工智能全球治理探寻平衡路径。习近平主席高度重视统筹人工智能发展和安全，强调：“要高度重视人工智能引发的各类内生和衍生风险，推动构建法律法规、技术监测、风险预警、应急响应体系，筑牢安全底线，防范滥用恶用，确保人工智能始终处于人类控制之下。同时，应当共同反对在人工智能领域泛化国家安全概念、把本国安全凌驾于他国安全之上的做法。”中国围绕算法推荐、深度合成、生成式人工智能等新技术新应用，持续健全风险评估、算法备案、内容标识、技术监测和应急响应等制度，探索形成分类分级、全生命周期、敏捷治理等实践，为国际社会识别风险、明确责任、完善标准提供可资借鉴的经验和工具。

推进能力建设和提供国际公共产品，推动全球南方国家平等受益。中国呼应全球南方国家的人工智能能力建设需求，在联合国相继提出《人工智能能力建设普惠计划》和《“人工智能+”国际合作倡议》，围绕基础设施、数据资源、人才培养、应用场景和治理能力形成系统合作框架；发起成立人工智能能力建设国际合作之友小组，汇聚政府、国际组织、企业和科研机构等多方资源；积极参与联合国支持下的全球人工智能能力发展网络建设，促进各国需求与技术、知识、人才和项目资源持续对接，常态化开展国际交流培训，推动气象智能预警方案“妈祖”在30个国家落地应用，助力广大发展中国家弥补短板。由议题倡导到方案设计，再到合作网络和具体项目，中国正在推动能力建设由原则倡议走向实践举措，以实际行动助力弥合智能鸿沟，彰显大国责任担当。

（作者为上海市习近平新时代中国特色社会主义思想研究中心复旦大学基地研究员）

落实四大全球倡议 推进人工智能全球治理

张薇薇

今年7月，习近平主席出席2026世界人工智能大会暨人工智能全球治理高级别会议开幕式并发表主旨讲话，全面系统阐述中国对人工智能发展和治理的政策立场、理念主张，指出：“各国应当秉持以人为本、向上向善理念，让人工智能成为促进共同繁荣、维护共同安全的一个重要动力源，携手构建公正合理的全球人工智能治理体系。”当前，全球人工智能技术创新正步入密集突破的上升期，各类内生和衍生风险亦随之而来。习近平主席直面全球人工智能发展与治理的痛点问题，从发展、安全、文明、治理维度提出4点意见，与四大全球倡议的核心要义高度契合，为推进人工智能全球治理贡献了中国智慧和中国方案。

人工智能全球治理亟待破局

习近平主席指出：“当前，世界百年变局加速演进，新一轮科技革命和产业变革加速突破，全球人工智能技术创新进入前所未有的活跃期，万物智联、人机协同、跨界融合、共创分享的智能技术，叠加释放巨大能量，既蕴含巨大机遇，也面临治理挑战。”人工智能是新一轮科技革命和产业变革的重要驱动力量。近年来，在算法、数据、算力全链条突破的推动下，人工智能各方面能力快速提升，推动经济社会发展各领域从数字化、网络化向智能化加速跃升。

随着人工智能技术发展和应用普及，相关风险愈发突出。一方面，模型幻觉、信息操控、就业冲击等各类风险持续向经济社会层面传导。如果不及时处理，相关风险挑战容易叠加扩散。这要求我们必须加强主动前瞻应对。另一方面，人工智能风险具有天然的跨国属性。在一国训练的大模型被另一国用户调用，风险可能随之穿越国界。一个恶意的深度伪造视频可能会在数小时内传遍全网，引发社会冲突。此外，不同国家人工智能发展差距持续拉大，还将加剧国际政治经济不平等。这就决定了人工智能治理不能局限在单一国家层面，必须在全球层面凝聚治理共识，形成治理合力。

当前，人工智能全球治理还面临多重挑战，具体体现在以下方面。一是治理碎片化。一些国际组织、地区组织将人工智能作为重点议题并形成相关机制，但这些机制的治理主题、治理理念、制度安排、政策节奏等存在不小差异，相互之间也没有层级关系、缺乏有效协调。二是治理行动力不足。一些关于人工智能全球治理的宣言、倡议等国际文件还停留在原则共识阶段，签署国自行决定落实方式和落实力度，执行和监督缺乏有效的制度保障。三是地缘政治博弈干扰。人工智能被视为影响国家综合国力的战略性技术。一些国家出于维护或扩大自身优势的考虑，或拒绝人工智能全球治理，或试图垄断治理规则制定权，侵蚀多边合作的信任基础，推高全球治理的协调难度。这些挑战意味着，推进人工智能全球治理，需要科学清晰的理念引领、稳定有效的制度支撑和切实有力的行动举措。

四大全球倡议指引人工智能全球治理的正确方向

当前，国际社会对人工智能的关切涉及技术创新、安全防控、伦理规范、产业发展等诸多领域，议题多元、挑战交织，但归根结底集中于发展、安全、文明、治理四个方面。习近平主席在这次会议上提出“坚持开放共赢，驱动创新发展”“强化风险意识，确保安全可控”“鼓励包容并蓄，促进文明互鉴”“倡导和衷共济，完善全球治理”4点意见，与四大全球倡议一脉相承，与人工智能向上向善的现实要求深度契合，系统回答了人工智能全球治理的基本问题，为破除人工智能全球治理困局提供了重要思想引领。

“坚持开放共赢，驱动创新发展”与全球发展倡议一脉相承，为促进普惠发展、弥合智能鸿沟提供指引。全球发展倡议将发展置于全球宏观政策框架的突出位置，强调推动多边发展合作进程协同增效，做到发展为了人民、发展依靠人民、发展成果由人民共享，并呼吁抓住新一轮科技革命和产业变革的历史性机遇，携手实现跨越发展等。人工智能是一种可以广泛赋能、层层渗透的关键共性技术，如何将其精准嵌入生产生活各环节，成为促进普惠发展、提升民生福祉的利器，考验治理智慧。同时，发展中国家在数字基础设施、数据、人才等方面存在短板，人工智能鸿沟加深的风险始终存在。坚持开放共赢，驱动创新发展，有利于推动人工智能与经济社会发展深度融合，全面赋能农业、制造业、城市管理、防灾减灾、教育医疗等领域，让技术进步尽快转变为可感可及的民生温度。在这一过程中，坚持普惠包容，帮助发展中国家建设人工智能基础设施、培训人工智能人才，推动人工智能模型开源共享，支持发展中国家开发本土化的人工智能应用，有利于利用先进技术更好发挥各国比较优势，助力解决世界范围的发展不平衡问题。

“强化风险意识，确保安全可控”与全球安全倡议一脉相承，为化解技术风险、维护人类安全廓清路径。全球安全倡议倡导坚持共同、综合、合作、可持续的安全观，共同维护世界和平和安全；秉持安全不可分割原则，反对把本国安全建立在他国不安全的基础之上；等等。当前，人工智能发展带来的安全风险大致包括两类：一是由技术自身导致的内生风险，如算法黑箱、数据偏差以及突破人类管控边界、背离人类价值与指令等；二是技术被恶意使用导致的衍生风险，如被用于网络攻击、制造生物武器等。特别需要指出的是，人工智能被应用于军事领域，将增强态势感知能力、降低打击成本等；被应用于网络空间，可能极大提升相关国家的信息压制和认知塑造能力。一味追求通过技术带来不对称优势，将加重国际社会的安全困境，导致更大的不安全甚至全人类悲剧。面对这一情况，一国一域的孤岛式治理难以奏效。强化风险意识，确保安全可控，推动各国加强对话协作，就防范底层技术隐患、约束技术恶意滥用等开展规则共商、风险联防，才能筑牢人工智能的安全底线。

“鼓励包容并蓄，促进文明互鉴”与全球文明倡议一脉相承，为守护文明多样性、促进交流互鉴划定坐标。多样性是世界的基本特征，也是人类文明的魅力所在。当今世界有80多亿人口，200多个国家和地区，2500多个民族和多种宗教。但在世界历史长河中，曾出现“文明优劣论”“文明冲突论”等，甚至引发战争和屠杀，留下难以抚平的历史创伤。全球文明倡议倡导尊重世界文明多样性、弘扬全人类共同价值、重视文明传承和创新、加强国际人文交流合作。人工智能对推动人类文明发展进步具有双重作用。一方面，可以让讲不同语言的人们便捷对话，能对濒危语言和文化遗产进行系统性建模与数字化保存，也可使小众文化精准对接受众群体，帮助其传承。另一方面，若人工智能大模型训练长期依托少数语言，可能使其他语言和文化在数字空间中被边缘化，侵蚀世界文明多样性。鼓励包容并蓄，促进文明互鉴，为人工智能发展和治理引入文明视角，用全人类共同价值塑造人工智能的价值观，构建多语言、跨文化的训练数据集，让不同文明都能被算法“看见”和“表达”，人工智能就可以成为增进不同文明相互理解和包容的桥梁。

“倡导和衷共济，完善全球治理”与全球治理倡议一脉相承，为推动加强人工智能全球治理确立遵循。当前，人工智能全球治理正处于起步阶段。全球治理倡议以奉行主权平等、遵守国际法治、践行多边主义、倡导以人为本、注重行动导向为核心理念，为推动加强人工智能全球治理确立了遵循。倡导和衷共济，完善全球治理，坚持共商共建共享的全球治理观，切实发挥联合国的重要作用，把人工智能全球治理规则建立在广泛的国际共识基础上，加强人工智能发展战略、治理规则、技术标准的对接协调，既有利于早日形成具有广泛共识的全球治理框架，提升发展中国家在人工智能全球治理中的代表性和发言权，保证各国在人工智能全球治理中权利平等、机会平等、规则平等，也有利于以提升人的获得感、幸福感、安全感为目标，让这一前沿技术更好造福人类社会，避免在人工智能领域造成新的不平等。

中国以实际行动为推进人工智能全球治理作出贡献

习近平主席指出：“作为负责任大国，中国在人工智能领域始终致力于做国际公共产品的提供者。”中国不仅是人工智能全球治理的积极倡导者，更是重要践行者。2025年，中国的人工智能核心产业规模已突破1.2万亿元，企业数量超过6200家。随着应用场景日益丰富，中国的人工智能治理也有序展开，为推进人工智能全球治理作出中国贡献。

治理实践为国际社会提供有益借鉴。中国的人工智能治理实践遵循以人为本、智能向善的价值导向，注重技术进步与社会公平、公共秩序、人的全面发展相协调，增进社会整体福祉。同时，坚持发展与安全并重，统筹推进技术创新与风险防控，既鼓励人工智能全面赋能经济社会高质量发展，又重视筑牢安全底线。在实践中，中国采取敏捷治理的方式，小步快跑、持续迭代，动态校准治理方案，建立起多层次的人工智能法律制度体系，形成政府主抓、企业主责、公众监督的多元协同治理格局。这些实践有力证明，人工智能发展与安全可以统筹兼顾，以人为本、安全可控、公平透明等原则可以被嵌入技术开发与应用的全生命周期，敏捷治理能够在构筑安全底线的同时给技术创新预留空间。

积极推动人工智能全球治理合作。2023年，中国提出《全球人工智能治理倡议》。2024年，中国推动联合国大会协商一致通过人工智能能力建设国际合作决议，发布《人工智能能力建设普惠计划》。2025年，中国推动2025世界人工智能大会暨人工智能全球治理高级别会议发表《人工智能全球治理行动计划》，倡议成立世界人工智能合作组织，在全球发展倡议高级别会议上提出“人工智能+”国际合作倡议。今年，29个创始成员国代表在上海签署《关于成立世界人工智能合作组织的协定》。该组织的成立将为人工智能全球治理合作搭建机制化平台，推动理念倡导与行动落地有机结合。这些举措和行动，为推动人工智能全球治理作出中国贡献。

用实际行动帮助发展中国家。中国始终将促进人工智能普惠包容发展放在人工智能全球治理的突出位置，同联合国共同举办人工智能能力建设研讨班，帮助世界各国特别是全球南方国家加强能力建设，积极协助发展中国家建设数字基础设施，帮助发展中国家构建自主可控的人工智能模型等。比如，中国帮助老挝研发的国家大模型在中文、老挝文、英文的互译，特别是老挝文的理解与内容生成，以及与当地文化相关的领域、本土化适应能力等方面表现卓越，目前已部署上线并开放使用。同时，中国的开源大模型是促进人工智能普惠包容发展的重要工具。开源大模型通常可以自主适配、本地化部署，使用成本较低，能够大幅降低人工智能的使用门槛，为发展中国家打造贴合自身实际需求的解决方案、提升自身人工智能能力等提供重要助力。凭借低成本、高性能、高可控的综合优势，开源大模型正为越来越多国家和人群带来福音，推动人工智能真正成为造福全人类的公共产品。

（作者为中国国际问题研究院世界和平与安全研究所副所长）

人民要论 REN MIN YAO LUN