

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主旨讲话,强调:“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。”习近平主席的重要讲话,着眼人类社会共同未来和人工智能发展大势,深刻阐明中国参与人工智能全球治理的价值立场和责任担当。人工智能作为引领新一轮科技革命与产业变革的战略性技术,既蕴含巨大发展机遇,也带来治理挑战。我们要深入学习贯彻习近平主席重要讲话精神,以人工智能大国的胸怀与担当,不断丰富人工智能国际公共产品供给,为推动构建公正合理的全球人工智能治理体系贡献中国智慧、中国方案、中国力量。

统筹发展和安全,以治理型公共产品贡献中国方案

习近平主席强调:“人工智能应当成为人类的可信工具。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。”人工智能在为经济社会发展注入强劲动能的同时,伴随着恶意使用、技术失灵、系统性风险等多重挑战。统筹发展和安全、平衡创新与治理,在推动人工智能广泛应用的同时,有效防范化解各类风险,成为国际社会共同面对的重大课题,也是人工智能全球治理的重要方向。特别是随着人工智能向通用化、自主化方向加速演进,其感知、推理、决策和行动能力不断增强,风险可能由单一环节向复杂系统传导,一旦系统目标偏离人类意图且无法被人类及时干预和有效控制,就可能引发难以预测的严重后果。当前,在特定实验环境中已发现人工智能系统存在策略性欺骗、规避干预和自我保护等倾向,存在隐瞒信息、误导操作者等策略性行为。尽管相关现象目前主要出现在模拟测试中,但其揭示的潜在风险不容忽视。

治理与创新相互促进、相辅相成。科学有效的治理能够明确规则边界,完善责任机制、强化风险防范,为技术创新营造稳定、透明、可预期的制度环境。人工智能技术迭代快、应用场景广、风险形态不断变化,客观上要求及时更新治理思路、重塑治理关系,推动治理体系与技术发展同向而行、协同演进。要加强前瞻研判和政策储备,针对新技术、新业态、新风险及时出台相应规则,并根据技术演进、产业发展和实践反馈进行动态评估和持续完善,推动治理由相对静态的规则管理转向适应技术变化的动态调适。摒弃监管者与被监管者之间简单对立的博弈思维,健全政府、企业、科研机构、行业组织和社会公众之间的沟通协作机制,推动风险信息共享、治理责任共担,形成多元参与、协同共治的治理格局。根据应用场景和风险等级实行分类分级治理,在工具选择上突出柔性原则,综合运用监管沙盒、试点示范、动态评估等治理工具,实现规范与激励的有机统一。

早在2014年6月,习近平总书记在中国科学院第十七次院士大会、中国工程院第十二次院士大会上发表重要讲话,对人工智能等前沿技术发展作出前瞻性谋划。10余年来,中国人工智能发展始终沿着习近平总书记擘画的蓝图稳步前行,实现了从“跟跑”“并跑”到多个领域“领跑”的历史性跨越。同时,中国积极探索符合人工智能发展规律的敏捷治理路径。中国坚持包容审慎监管,在守住安全底线的前提下,为新技术、新业态成长预留必要空间。同时,面对隐私泄露、数据过度收集等新情况新问题,监管部门及时回应、动态调整,不断完善治理规则和监管措施。在此过程中,中国逐步形成了技术发展与管理完善协同推进的治理模式:技术创新不断向前,治理能力同步提升;风险问题及时发现,制度安排动态优化。这种治理模式既避免监管过度抑制创新活力,也防止因治理缺陷积累安全风险,实现了鼓励创新与规范发展的有机统一,为全球人工智能安全治理贡献了中国方案。

坚持开源开放,以技术型公共产品弥合数智鸿沟

习近平主席强调:“要抓住难得的历史性机遇,鼓励开源开放、合作共享,全面促进人工智能科技创新、产业发展、场景应用,协同推进传统产业改造升级、新兴产业培育壮大、未来产业前瞻布局,让人工智能赋能千行百业。”

开源开放日益成为推动技术创新、成果扩散和产业应用的重要路径。前沿大模型研发具有技术密集、资金密集、人才密集和资源密集等特点,现阶段真正掌握大模型核心技术体系、具备从底层算法创新到模型规模化训练部署的国家屈指可数,全球约40%的人工智能核心技术研发集中在中美两国。当前,全球数智鸿沟不断扩大,少数发达国家收割技术红利,广大发展中国家沦为原始数据供应地的极端分化风险上升。国际货币基金组织制定的“人工智能准备指数”将全球数智鸿沟清晰量化,截至2024年4月,发达国家该指数为0.68,而新兴市场国家和低收入国家分别为0.46和0.32。作为人工智能领域“跟跑者”的许多发展中国家缺少有竞争力的科技企业,也缺乏足够的技术和人才储备,甚至在相关基础设施上都严重不足。国际电信联盟数据显示,至2024年底,全球仍有近26亿人未接入互联网,占全球人口的1/3,最不发达国家使用互联网的人数仅占总人口数的35%,内陆发展中国家仅有39%的人触网。

对于尚不具备前沿大模型自主研发能力的国家特别是广大发展中国家而言,开源大模型能够显著降低技术获取和应用的门槛,使这些国家得以在现有先进技术基础上开展本地化适配、场景化开发和产业化应用。越来越多的国家将开源大模型作为发展本国人工智能产业的重要技术底座,并以此加快应用能力提升和产业生态构建。促进人工智能研究成果与技术经验的开放共享,有利于扩大人工智能前沿技术的可及性,增强各国自主发展能力,为弥合全球数智鸿沟、推动人工智能普惠发展提供有力支撑。

中国始终坚持以开源开放促进人工智能普惠发展,以技术供给为发展中国家降低先进信息技术获取门槛、开展场景化应用和培育本土产业生态提供了重要支撑,有力推动了人工智能在全球范围的普及应用。当前,中国人工智能开源大模型全球累计下载量已突破100亿次。新一代开源模型Kimi K3正式发布,参数规模达2.8万亿,为目前全球参数最大的开源模型,标志着我国

以丰富国际公共产品供给引领人工智能全球治理

人工智能模型发展取得新的成就。

践行多边主义,以制度型公共产品完善全球治理

习近平主席强调:“要践行真正的多边主义,切实发挥联合国的重要作用,加强人工智能发展战略、治理规则、技术标准的对接协调,早日形成具有广泛共识的全球治理框架,让这一前沿技术更好造福人类社会。”

人工智能带来的风险具有跨国性、复杂性和系统性,技术漏洞可能引发跨区域、跨领域的连锁反应。尤其是当人工智能应用基于同类基础模型和架构构建时,一旦出现共同漏洞则可能造成集体失效。这是国际社会面临的共同课题,没有任何国家能够独善其身、独自应对。确保人工智能安全、可靠、可控,离不开国际社会的协调合作。各国应把人工智能安全、可靠、可控作为重要的国际公共产品,在技术研发、安全评估、风险监测和应急处置等方面深化协调合作,推动形成信息共享、责任共担、风险共治的治理机制,共同筑牢人工智能安全、可靠、可控的全球防线。

当前,人工智能技术加速突破同地缘政治风险上升相互交织,全球治理面临大国博弈加剧、科技竞争升温、规则协调不畅等复杂局面。个别国家将人工智能治理异化为维护自身技术优势和竞争地位的工具,破坏长期形成的国际科研合作网络和产业协作体系,不仅阻碍人工智能技术进步和成果普惠,也无助于解决国际社会共同面临的风险挑战,反而损害各国共同利益。与此同时,全球人工智能治理体系尚处于加快形成阶段,不同国家和地区在发展阶段、治理理念、价值取向、制度安排和政策节奏等方面存在差异,各类治理机制相互交叠,规则标准衔接不足,国际协调难度不断加大。破解人工智能全球治理困境,必须坚持开放共赢、包容并蓄、和衷共济,在联合国框架下,推动形成人工智能全球治理共识,构建践行多边主义的国际合作平台。

中国倡导推动成立世界人工智能合作组织,为完善人工智能全球治理提供重要的制度型公共产品。作为全球首个人工智能政府间国际组织,世界人工智能合作组织遵循《联合国宪章》宗旨和原则,尊重各国主权和文明多样性,坚持平等相待,践行多边主义,秉持共商共建共享理念,实现互利共赢,为各国平等参与治理对话、加强政策沟通与规则协调、深化技术交流与务实合作搭建了制度化的国际合作平台,有助于将分散的合作意愿转化为稳定的制度安排,将原则性治理共识转化为持续的共同行动,也为促进人工智能普惠发展提供了制度化载体。从提出倡议到签署协定,中国以实际行动推动人工智能全球治理从理念共识向制度实践不断深化,充分彰显负责任大国完善全球治理的务实作为。

聚焦能力建设,以赋能型公共产品夯实发展根基

习近平主席强调:“要广泛开展国际合作,帮助全球南方国家加强能力建设,弥合数智鸿沟,促进可持续发展,避免在人工智能领域造成新的历史不公。”

人工智能应用能力与治理能力互为支撑、互促共进。丰富的应用实践能够为风险识别和规则制定提供可靠的判断依据和经验参照。治理能力提升又能为技术安全落地、规范发展提供制度保障。发展中国家在人工智能发展面临技术应用与治理话语权的双重挑战,其深层根源在于教育科研基础相对薄弱、专业人才储备不足、科技创新能力不强、信息基础设施建设滞后。与一般科技产品不同,人工智能具有迭代速度快、应用领域广、渗透能力强等特点,缺乏必要的人才储备和软硬件基础设施支撑,人工智能的赋能效应便难以充分释放,甚至会形成数智鸿沟,并进一步拉大发展差距。弥合数智鸿沟,关键在于加强广大发展中国家的自主发展能力建设。应通过技术援助、知识共享、联合研发和人才培养等方式,帮助发展中国家获得和应用人工智能技术,同时支持其逐步形成自主创新 and 可持续发展的能力,真正做到“授人以渔”。

发展中国家在规则制定与议程设置中持续处于结构性弱势地位,但人工智能全球治理不能由少数技术强国主导,更不能以技术优势固化治理优势,以能力差距固化权益差距。各国发展阶段不同、技术基础有别、治理诉求多样,但都有平等参与全球治理、表达利益诉求和共同制定规则的权利。推动人工智能普惠发展,既要让发展中国家共享技术进步带来的发展红利,也要增强其参与治理议程、协商国际规则和维护自身权益的能力。应加强发展中国家人工智能治理能力建设,帮助其提升风险识别、政策制定、标准实施、监管执法和国际协商水平,使其能够更加充分、有效地参与全球治理。

中国始终把支持发展中国家能力建设作为推动人工智能普惠发展的重要内容。自习近平主席2023年提出《全球人工智能治理倡议》以来,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》《人工智能全球治理行动计划》《“人工智能+”国际合作倡议》,倡议成立世界人工智能合作组织,建设中国—金砖国家人工智能发展与合作中心,成立中国—东盟人工智能产业创新中心等,通过政府、企业、科研机构等多层次交流合作,推动技术、知识、人才和应用经验更加顺畅地向发展中国家流动。2026世界人工智能大会暨人工智能全球治理高级别会议期间,《人工智能合作发展行动计划》《国际人工智能伦理治理行动计划》《智能体互信互联互操作全球合作倡议》等相继发布。未来5年,中国将面向发展中国家提供5000个人工智能专题研修培训名额,面向东盟、阿盟、非盟、拉共体、上合组织、金砖国家建设国际人工智能应用合作中心,推动气象智能预警方案“妈祖”在30个国家落地应用,用实实在在的行动推动人工智能成为造福全人类的国际公共产品。

(作者分别为清华大学人工智能国际治理研究院院长、中国社会科学院习近平新时代中国特色社会主义思想研究中心特约研究员)

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题为《携手构建公正合理的全球人工智能治理体系》的主旨讲话,提出“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”,深刻阐明了携手构建公正合理的全球人工智能治理体系的理念主张、原则路径和实践方案。这是在2023年提出《全球人工智能治理倡议》、2025年提出全球治理倡议基础上,习近平主席站在人类历史发展的高度,对人工智能时代人类向何处去这一重大课题作出的进一步系统回答,为推动人工智能真正成为增进全人类福祉、推动人类文明进步的强大正能量提供了科学指南。

当今时代,携手构建公正合理的全球人工智能治理体系意义重大

人工智能治理关系全人类命运,是世界各国面临的共同课题。准确把握全球人工智能发展的时代方位,正视风险与挑战,深化协作与共治,才能切实推动人工智能朝着向上向善、造福人类的方向发展。

技术跃迁重塑发展格局,需加强合作共同应对时代之变。习近平主席指出:“回望历史,蒸汽机的发明开启工业文明,电力的普及点亮现代社会,互联网的诞生连接整个世界。每一次科技革命都深刻重构人类生产生活方式,推动经济社会发展大幅跃升。”当前,全球人工智能技术创新进入前所未有的活跃期,人工智能正在成为世界经济增长的新引擎和新旧动能转换的加速器。顶尖模型的迭代周期日益缩短,模型能力在部分领域达到“类人”甚至“超人”水平;智能体正在从辅助工具向具备自主能力、能够自我学习和演化方向演进,顶尖智能体系统所能完成的特定软件任务长度最多4到7个月就能翻一番。人工智能赋能效应日益凸显,大模型AlphaFold成功预测了几乎2亿种蛋白质结构,其对蛋白质结构的智能挖掘大大超过人类以往百年的科研积累,显著加快药物研发进程;智能眼镜实时翻译、数字人提供咨询服务,工业领域人形机器人进厂从事搬运、分拣、巡检等工作,人工智能正从“会聊天”转向“会干活”、从“数字世界”走向“物理世界”。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,只有将人工智能置于人类科技革命的历史视野下审视,才能准确把握以人工智能为代表的新一轮科技革命和产业变革带来的深远影响,将治理合作视为人类共同应对全球秩序深度调整的必由之路,为携手构建公正合理的全球人工智能治理体系锚定正确方向。

风险外溢凸显治理紧迫,需加强合作共同保障人类福祉。习近平主席在开幕式上直面时代之问,精辟概括当前人工智能治理面临的挑战。从人机关系看,随着人工智能技术在复杂环境中自主决策、自主行动的能力持续增强,人机权责边界模糊等问题显现,对人类主体地位构成潜在挑战。从安全维度看,算法参与决策可能放大社会偏见,深度伪造等新型风险冲击公共安全;人工智能技术与金融、通信、交通等关键基础设施融合日益加深,单点故障可能升级为系统性风险。从伦理维度看,人工智能拟人化互动服务诱导情感过度依赖现象高发,技术开始挑战伦理底线,价值对齐问题亟待解决。从经济维度看,人工智能对就业的冲击值得重视,自动化替代与知识和技能错配交织;技术垄断导致智能鸿沟拉大,可能进一步加剧南北发展失衡和社会贫富分化。只有将人工智能置于风险防控的紧迫需求中加以审视,以治理合作共同应对全球性风险,才能让人工智能这匹千里马既跑得快又跑得稳,确保其始终成为增进人类福祉的可信工具。

普惠共享呼唤开放协作,需加强合作共同构建人类命运共同体。习近平主席指出:“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。”人工智能是全人类共同的智慧结晶和宝贵财富,其发展带来的机遇应全球共享,风险需全球共担。然而,个别国家将人工智能视为大国博弈工具,采取“筑墙设垒”“脱钩断链”等排他性策略,人工智能发展面临全球治理机制碎片化、阵营化等挑战;全球南方国家在人才、数据、算法、算力等要素上受到制约,在治理规则制定中缺乏代表性和话语权。世界人工智能合作组织在上海成立,为各国开展政策对话和技术交流提供了包容开放的平台,加快推动人工智能全球治理从理念倡导迈向机制建设的新阶段。只有将人工智能置于构建人类命运共同体的愿景中加以审视,在尊重差异的基础上凝聚共识,才能让各国平等参与、平等决策、平等受益,共同谱写全球人工智能合作的时代交响。

在守正创新中贡献人工智能全球治理的中国智慧

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,习近平主席统筹发展、安全、文明、治理4个维度,提出携手构建公正合理的全球人工智能治理体系的意见,既坚持开放创新,又强化风险管控;既促进文明互鉴,又突出规则协同,为弥合治理赤字贡献中国智慧。

秉持开放创新的发展观。人工智能作为引领新一轮科技革命和产业变革的战略性技术,其创新、发展和应用天然具有全球属性,深度依赖于全球范围的数据流动、人才交流、技术开源和产业链协同。只有秉持开放共赢、互利合作的精神,才能有效突破资源瓶颈,打通学科界限,提升创新效率,提振产业活力。中国倡导开源开放,鼓励全球开发者、企业、学术机构与公共部门携手共建开源社区和平台,推动基础资源开放共享,协同推动从基础算法攻坚、算力底座搭建到大模型训练落地的全产业链迭代升级,降低技术创新和应用门槛,促进人工智能技术服务普惠可及。中国推动人

为携手构建公正合理的全球人工智能治理体系贡献中国方案

人工智能赋能千行百业,鼓励各国立足自身产业基础与发展需求,推进人工智能赋能工业制造、商贸流通、医疗、教育、农业等领域,构建丰富多样、健康向善的人工智能应用生态,加速打造新质生产力,让包括全球南方在内的世界各国人民普遍受益。

秉持可信可控的安全观。当前,人工智能安全风险日益凸显,算法、模型、智能体等都存在漏洞风险。有报告显示,人工智能持续重塑着网络攻击者的作战方式,极大提高了行动速度、规模和效率,2025年,面向公众的软件及系统应用程序中漏洞利用量同比增长44%,活跃勒索软件组织数量较上年上升49%。同时,人工智能风险具备跨国跨境传播特性,各国技术实力和治理能力存在较大差别,单一国家无法独立化解各类安全挑战。中国始终坚持发展和安全并重,不断提升人工智能技术的安全性、可靠性、可控性、公平性;构建动态适应、分级分类的安全监管体系,探索人工智能服务可追溯管理制度,筑牢数据与伦理安全底线;推进技术防控升级,有效防范化解人工智能安全风险;促进全球协同,统筹隐私保护、安全技术、风险预警、安全审查、快速响应等环节,共建安全测评互认平台。

秉持交流互鉴的文明观。人工智能发展可能会消解各国独有的文化特质,挤压本土特色文化生存空间。联合国人工智能独立国际科学小组报告指出,为提升用户黏性和大规模应用而优化的人工智能系统并不中立,它们主要内嵌了英语国家和全球北方的文化假设,这些文化假设可能积极地将世界上大多数人口边缘化。中国始终倡导包容并蓄,强调人工智能发展应尊重文明多样性,防止现实中存在的文化偏见等问题影响输出内容,拒绝单一治理标准与技术阵营;推进文明交流互鉴,促进跨文化对话、跨文明交流,兼顾各国本土文化诉求,不搞模式输出;依托人工智能载体赋能文明交流,推动数字文化资源共享,以智能技术保护多元文化遗产。

秉持共建共享的治理观。面对人工智能全球治理面临的碎片化、无序化挑战,当前的人工智能治理机制不够适配,发达国家与发展中国家数字鸿沟有待弥合,一些国家垄断话语权,削弱集体行动能力。在此背景下,中国坚持真正的多边主义,倡导共商共建共享的全球治理观,主张世界命运由各国共同掌握、国际规则由各国共同书写、全球事务由各国共同治理、发展成果由各国共同分享,推动各国在全球治理中平等参与、平等决策、平等受益;坚定维护联合国宪章和国际法,支持在联合国框架下建立各国普遍参与的人工智能治理机制,在人工智能领域推动构建公正合理的全球治理体系;始终将发展中国家的期待与需求放在重要位置,支持把加强人工智能能力建设国际合作作为优先事项,鼓励人工智能领先国家开展实际行动,支持发展中国家加强人工智能创新、应用、治理等方面的综合能力建设,不让任何一个国家在人工智能时代掉队。

在实际行动中彰显负责任大国的使命担当

在实践中求真知,在探索中谋善治。习近平主席领导中国积极拥抱人工智能浪潮,坚持发展和安全并重,推动人工智能向善发展,为支持全球人工智能发展、推进全球人工智能能力建设贡献国际公共产品。

作为大国担当的践行者,创新并加强人工智能安全监管,为世界提供中国经验。面对人工智能引发的算法偏见等风险,中国坚持发展和安全并重,构建多层次、多维度的人工智能安全监管体系。及时回应技术发展和时代需要,先后实施《人工智能生成合成内容标识办法》《人工智能拟人化互动服务管理暂行办法》等法律法规,以良法善治夯实人工智能健康发展的制度根基。同时,推动政产学研用协同发力,建立了较为完善的算法监管、数据安全、伦理审查体系,以有效治理保障人工智能释放创新活力。面向未来,中国将继续完善动态敏捷、多元协同的人工智能治理格局,构建覆盖人工智能研发、部署、应用与影响全链条的人工智能治理体系,健全算法备案、透明度管理、安全评估等制度,以更加成熟完善的国内实践为国际社会提供治理经验。

作为普惠发展的贡献者,持续推进全球人工智能能力建设,推动各国共享智能红利。促进人工智能普惠包容发展,是世界各国共同走向更加美好智能未来的必然要求。中国提出并实施《人工智能能力建设普惠计划》,发起成立人工智能能力建设国际合作之友小,开展一系列务实合作,促进弥合智能鸿沟。此次宣布面向发展中国家提供5000个人工智能专题研修培训名额,面向东盟、阿盟、非盟、拉共体、上合组织、金砖国家建设国际人工智能应用合作中心,推动气象智能预警方案“妈祖”在30个国家落地应用,持续为全球人工智能能力建设提供重要动能。中国将持续把普惠包容贯穿人工智能治理全过程,推动算力、算法、数据等核心资源开放共享,开展人工智能模型研发和赋能合作,加强人工智能素养和人才培养,不断提升人工智能技术和服务的可及性,让更多国家公平参与人工智能发展,确保发展中国家在新一轮智能浪潮中平等受益。

作为人工智能全球治理的倡议者,加强多边合作,持续贡献中国智慧。2023年10月,习近平主席提出《全球人工智能治理倡议》,明确以人为本、智能向善等主张,为全球治理贡献中国方案。此后,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《人工智能全球治理行动计划》和《“人工智能+”国际合作倡议》,推动既有共识转化为务实举措。世界人工智能合作组织在东亚应运而生,成为人工智能发展史上的一个重要里程碑。中国将通过加强人工智能发展战略、治理规则、技术标准的对接协调,走出一条开放协商、多元参与的互信共治之路,提升发展中国在全球治理中的代表性和发言权,确保各国人工智能发展与治理的权利平等、机会平等、规则平等,推动形成具有广泛共识的全球治理框架。

(作者为中国信息通信研究院副院长)

