引用格式：刘畅，段续，何腾蛟. 联合国参与人工智能全球治理的实践路径分析及对我国的启示[J]. 世界科技研究与发展，2026，48(1)：46 – 57.

联合国参与人工智能全球治理的实践路径分析及对我国的启示

刘畅[1]　段续[*,2]　何腾蛟[**,3]
（1. 清华大学人工智能国际治理研究院，北京 100084；2. 中国电子学会，北京 100036；
3. 暨南大学信息科学技术学院，广州 510632）

摘　要：人工智能技术的快速发展将极大赋能可持续发展，但同时也引发了难以预测的风险，人工智能治理挑战加剧。近年来，联合国及其附属机构秉承“以人为本”的治理理念，通过规则制定、标准推动、能力建设、国际合作等形式积极推进人工智能全球治理，以帮助发展中国家弥合人工智能鸿沟，实现人工智能公平、普惠、可持续发展为目标，提升人工智能治理的互操作性，推进制定统一的人工智能全球治理规则。本文系统梳理了联合国及其附属机构在推进人工智能全球治理方面的重点举措，分析联合国参与人工智能全球治理的内外动因、有利条件与制约因素，总结联合国在推进人工智能全球治理上的实践经验，并从理念贯彻、模式构建、路径选择、行动示范四个维度，对我国参与人工智能全球治理，构建在联合国框架下多元协同、敏捷灵活、包容普惠的人工智能全球治理体系提出建议参考。

关键词：联合国；人工智能；人工智能治理；人工智能监管；人工智能全球治理体系；全球治理；可持续发展

DOI：10.16507/j.issn.1006 – 6055.2025.11.002

CSTR：32308.14.1006 – 6055.2024.04.001

Analysis of the Practical Path of United Nations Participation in Global Artificial Intelligence Governance and Its Implications for China

LIU Chang[1]　DUAN Xu[*,2]　HE Tengjiao[**,3]
（1. Institute for AI International Governance of Tsinghua University, Beijing 100084, China;
2. Chinese Institute of Electronics, Beijing 100036, China; 3. College of Information Science and Technology of Jinan University, Guangzhou 510632, China）

Abstract: The rapid development of artificial intelligence (AI) technology holds immense potential to drive sustainable development. However, it simultaneously creates unpredictable risks, which intensify challenges in AI governance. In recent years, with the adoption of a people-centered governance philosophy, the United Nations (UN) and its specialized agencies have actively promoted global AI governance through rulemaking, standard development, capacity building and international cooperation, aiming to help developing countries bridge the AI gap and achieve fair, inclusive, and sustainable development

* E-mail: 13070166610@163.com
** E-mail: htj2018@jnu.edu.cn

of AI. These efforts seek to enhance the interoperability in AI governance and advance the formulation of coherent global rules. This paper provides a systematic overview of the key initiatives undertaken by the UN system in advancing global AI governance. It analyzes the internal and external drivers, enabling conditions, and constraints of UN participation in this field, and summarizes the practical experience gained by the UN. Based on four dimensions—philosophy implementation, model construction, pathway selection, and action demonstration—the paper proposes recommendations for China's participation in global AI governance and the development of a diverse, collaborative, agile, flexible, and inclusive global AI governance system within the UN framework.

Keywords: United Nations; Artificial Intelligence; Artificial Intelligence Governance; Artificial Intelligence Regulation; Global Artificial Intelligence Governance Framework; Global Governance; Sustainable Development

当前人工智能发展处于快速突破的关键时期,人工智能技术在提高生产力和改善人类福祉方面大有可为,将极大助力可持续发展目标(Sustainable Development Goals, SDGs)的实现。然而人工智能在赋能可持续发展的同时,也放大了潜在的负外部性,出现难以预测的风险。人工智能治理已成为一项全球性挑战,亟须构建统一的全球治理体系,实现共同安全。近年来,联合国及其附属机构通过规则制定、标准推动、能力建设、国际合作等形式积极推进人工智能全球治理,弥合全球人工智能鸿沟,推进人工智能公平、普惠、可持续发展,提升人工智能治理的互操作性,打造联合国框架下多元主体协同共治的人工智能全球治理体系。

本文通过梳理联合国及其附属机构在推进人工智能全球治理上的重点举措,分析联合国参与人工智能全球治理的必要性、优势及挑战,总结联合国在推进人工智能全球治理上的实践经验,从理念贯彻、模式构建、路径选择、行动示范四方面出发,为我国推进人工智能全球治理提供参考。

1 人工智能全球治理现状

随着高性能、低成本和开源模型的普及,人工智能的普及性和影响力进一步扩大,尤其是生成式人工智能的快速发展,加速人工智能的商业应用普及,驱动企业核心业务价值增长。2024 年,生成式人工智能吸引了全球 339 亿美元的投资,同比增长 18.7%[1]。与此同时,人工智能引发的潜在风险也成为全球关注的焦点问题之一。人工智能安全问题不局限于技术失控或被恶意利用,《国际人工智能安全报告》已将技术失控、恶意使用和系统性风险列为通用人工智能三大类风险[2]。其中,系统性风险涉及人工智能对就业、环境、知识产权等方面带来的负面影响,尤其是生成式人工智能与现有版权保护机制之间的冲突问题日益凸显[3]。推动人工智能技术创新的同时,构建与之相适应的科学有效的治理体系,是当前国际社会面临的核心课题。

近年来,各国政府在加大人工智能基础设施投资力度的同时,也在明显强化全球治理的参与程度和战略部署[4],相继出台政策举措以引导其健康有序发展,陆续成立人工智能安全研究机构,推进人工智能全球治理体系和机制建设。2023—2025 年,生成式人工智能迎来集中爆发,各类大模型快速发展,人工智能技术创新和产业应用的最新进展,为人工智能全球治理体系和机制建立,奠定了重要基础。以欧盟、美国、英国、法国和中国为例,近三年密集出台多项人工智能治理相关政策举措,聚焦透明度、可信度等负责任的人工智能核心原则,强调规范风险的同时释放变革潜力,并通过自身实践来影响全球范围内人工智能技术标准、治理准则、应用规范(表 1)。

表 1　欧盟、美国、英国、法国和中国推进人工智能全球治理上的主要政策举措(2023—2025 年)

Tab. 1　Major policy initiatives of the EU, the US, the UK, France, and China in promoting global AI governance(2023-2025)

	时间	主要举措
欧盟	2024 年 02 月	欧盟人工智能办公室[1]成立。
	2024 年 03 月	欧洲议会[2]通过《人工智能法案》(2024 年 8 月正式生效)。
	2024 年 05 月	欧洲委员会[3]通过《人工智能与人权、民主和法治框架公约》。
	2025 年 04 月	欧盟委员会[4]发布《人工智能大陆行动计划》。
	2025 年 07 月	欧盟委员会发布最终版《通用人工智能行为准则》。
美国	2023 年 01 月	美国国家标准与技术研究院[5]发布《人工智能风险管理框架》。
	2023 年 10 月	美国前总统拜登颁布《关于安全、可靠和可信开发和使用人工智能的行政命令》。
	2024 年 07 月	美国国务院[6]发布《人工智能与人权风险管理概况》。
	2024 年 07 月	美国国家标准与技术研究院发布《人工智能风险管理框架:生成式人工智能概述》。
	2024 年 09 月	美国商务部工业与安全局[7]发布《建立先进人工智能模型和计算集群开发的报告要求》。
	2024 年 10 月	美国前总统拜登发布《关于推进美国在人工智能领域的领导地位,利用人工智能实现国家安全目标,以及促进人工智能安全性、可靠性和可信度的备忘录》。
	2024 年 10 月	白宫发布《推进国家安全中人工智能治理与风险管理的框架》。
	2024 年 11 月	美国商务部和国务院举办“国际人工智能安全研究机构网络”首次会议。
	2025 年 07 月	白宫发布《赢得人工智能竞赛:美国人工智能行动计划》。
英国	2023 年 03 月	英国政府发布《支持创新的人工智能监管方式》政策文件。
	2023 年 11 月	英国外交、联邦和发展办公室[8],英国科学、创新与技术部[9]和英国人工智能安全研究所[10]主办“全球人工智能安全峰会”,会上通过《布莱切利宣言》。
	2024 年 01 月	英国中央数字与数据办公室[11]发布《生成式人工智能框架》。
	2024 年 05 月	英国与韩国联合主办“人工智能首尔峰会”,会上发布《首尔宣言》。
法国	2025 年 02 月	法国与印度联合主办“人工智能行动峰会”,会上发布《关于发展包容、可持续的人工智能造福人类与地球的声明》。
中国	2023 年 07 月	国家互联网信息办公室联合七家单位公布《生成式人工智能服务管理暂行办法》。
	2023 年 10 月	中央网络安全和信息化委员会办公室发布《全球人工智能治理倡议》。
	2024 年 03 月	“人工智能+”首次被写入《政府工作报告》。
	2024 年 07 月	2024 世界人工智能大会暨人工智能全球治理高级别会议发表《人工智能全球治理上海宣言》。
	2024 年 09 月	全国网络安全标准化技术委员会发布《人工智能安全治理框架》1.0 版。
	2024 年 09 月	提出《人工智能能力建设普惠计划》。
	2025 年 07 月	2025 世界人工智能大会暨人工智能全球治理高级别会议发布《人工智能全球治理行动计划》。
	2025 年 09 月	2025 年国家网络安全宣传周主论坛发布《人工智能安全治理框架》2.0 版。
	2025 年 09 月	提出《“人工智能+”国际合作倡议》。

1) European Artificial Intelligence Office; 2) European Parliament; 3) Council of Europe; 4) European Commission; 5) National Institute of Standards and Technology, NIST; 6) U. S. Department of State; 7) Bureau of Industry and Security, BIS; 8) Foreign, Commonwealth & Development Office, FCDO; 9) Department for Science, Innovation and Technology, DSIT; 10) AI Security Institute; 11) Central Digital and Data Office, CDDO。

人工智能全球治理是一项复杂议题,涉及法律制定、行业标准、国际协调等方面。迄今,大多数人工智能政策来自发达国家,仅有少数发展中国家拥有与人工智能相关的政策;联合国、二十国集团、七国集团、金砖国家为代表的多边合作机制将人工智能全球治理视为重要国际合作议题,多边治理机制初步建立,但仍缺乏统一全球框架;微软、谷歌、OpenAI、百度为代表的科技企业也在深度参与人工智能全球治理,但仅有极少数的大型科技公司掌握先进人工智能技术及治理资源。由于不同主体在治理理念、关注议题、治理能力上存在差异,当前人工智能全球治理呈

现不均衡、阵营化的局面，缺乏包容且统一的全球协调机制。

2 联合国参与人工智能全球治理的条件要素

2.1 联合国参与人工智能全球治理的必要性

当前，人工智能发展呈现不平衡性，芯片、算力、训练数据集等人工智能发展关键要素掌握在少数国家和少数科技公司手中，全球人工智能资源分布不均；顶尖人工智能科学家、工程师等人工智能高端人才聚集在中美两国，全球人工智能人才市场两级化形成；支持人工智能发展的全球公共资金缺乏，限制了“全球南方”国家参与人工智能国际标准和治理政策的制定。近年来，世界主要国家和地区相继发布人工智能治理倡议，出台引导人工智能健康有序发展的政策法规，国际社会逐渐形成对人工智能全球治理的部分共识，然而这些共识目前仍停留在基础层面。由于主要国家和地区政府间针对人工智能发展存在战略分歧，围绕规则制定主导权的复杂、激烈博弈正限制着全方位、多元化、普适性的治理框架和治理秩序的落地成型[5]。

人工智能的加速发展引发了全球范围内权力和财富的聚集效应，并产生了地缘政治和地缘经济影响[6]，而人工智能全球治理这一新型全球性挑战也使得国际关系格局产生深刻变化，重塑国际秩序。在此背景下，全球迫切需要一个具备权威性和公信力的国际机构来协调区域性人工智能治理举措，并缓解人工智能对国际秩序造成的挑战。

作为目前最具代表性和权威性的政府间国际组织，联合国是全球范围内唯一有能力协调不同区域性人工智能治理愿景的政府间平台[4]，一方面可以发挥组织协调作用，协调区域性治理行动在全球层面上的协同工作，推动南北合作、南南合作和三方合作，为多利益攸关方平等参与人工智能全球治理创造机会；另一方面可以整合全球资源网络，增强人工智能能力建设，弥合全球人工智能鸿沟，提升整体治理能力，推进构建包容且统一的人工智能全球治理体系。

人工智能全球治理需要具有整体性、包容性的国际规则，而不是分散的制度设计[7]。为避免区域和国家间治理政策的碎片化、阵营化局面进一步加剧，同时为更好地协调人工智能治理资源、增强人工智能能力建设，应围绕联合国框架来加强推进人工智能全球治理体系的建构。

2.2 联合国参与人工智能全球治理的优势

一是联合国具备独特的机制优势，是国际社会最具普遍性的多边合作平台，具有较强的国际权威性和公信力。一方面可以充分践行多边主义，搭建各方共同讨论和解决全球性问题的平台，促进包容的人工智能全球治理发展；另一方面可以发挥协调机制作用，协调各国和多利益攸关方的共同参与，推动各组织机构的有机协同，充分发挥不同机构的专业性和机构间的协调性[8]；同时，联合国可以利用其机制的灵活性，成立多学科、跨领域、跨机构的专项工作组，组建由全球各领域杰出代表构成的专家咨询组，推动形成人工智能全球治理国际共识。

二是联合国长期以来积极应对全球性挑战和危机，在多领域建立了风险评估和监测机制，开发了一系列公共治理工具包，具备成功应对全球性挑战的丰富经验。一方面，联合国具备应对气候变化、核武器等全球性挑战和危机的经验，并在长期的实施过程中积累了在问责制、补救机制等方面的重要经验教训[9]，可以为应对人工智

能全球治理这一新型全球性挑战提供借鉴;另一方面,联合国已经形成人工智能风险评估、审计、认证、监测等机制和规范,强调要对人工智能全周期进行监管;此外,联合国还开发了一系列跨领域和特定领域的公共治理工具包,向国际社会提供人工智能治理工具,助力提升风险应对能力。

三是联合国深度参与国际法治体系的建构,制定了一套由公约、条约和标准组成的国际法,推动国际社会在若干领域达成共识,在解决国际问题上发挥着不可替代的作用。国际法涵盖了众多国际关切的问题,是联合国推进全球治理的重要工具。联合国积极推动人工智能领域国际软法的建设,已助力各方在教育治理、公共卫生治理、劳动治理、商业治理等人工智能相关领域达成规范共识,推动人工智能全球治理规则落地。

2.3 联合国参与人工智能全球治理的挑战

一是人工智能技术存在不确定性且发展速度过快,监管政策的制定、立法程序、标准建设等过程均具有特定的周期性与协商性等特征[3],导致治理政策和措施的出台可能滞后于技术发展。技术创新与监管治理之间的权衡依然面临诸多挑战,加大治理难度。如何平衡创新发展与风险防控、及时调整治理方法及治理工具以响应人工智能技术的快速变革将是一项重点考验。

二是人工智能能力发展呈现全球性失衡态势。在算力、算法、数据、人才、资本、基础设施等方面处于劣势的“全球南方”国家,将进一步扩大与“全球北方”国家在人工智能能力上的差距,各国间的人工智能鸿沟加大了人工智能全球治理难度。如何把握好对人工智能的治理力度、平衡好各方的能力发展需求将是一项艰巨任务。

三是大国皆争夺对人工智能全球治理规则制定的主导权,地缘政治冲突对人工智能全球治理的干扰不断扩大。中国人工智能企业“深度求索”(DeepSeek)的崛起,重塑全球人工智能市场格局,使大国地缘政治竞争进一步具象化,制定人工智能全球治理统一规则变得更加困难。如何抵御地缘政治逆风、协调大国间不同的治理愿景将是一项严峻挑战。

当前,联合国在推进人工智能全球治理的进程中,不仅面临着来自技术本身以及外部环境的挑战,也面临着内部工作部署问题。一是联合国所属机构开展的治理工作出现重叠现象,多家联合国机构都在开展人工智能治理相关工作,由于机构间的工作部署尚未实现有机统一,导致出现治理议题分散、工作重合的问题;二是联合国构建中的人工智能全球治理体系存在包容性不足的问题,以联合国人工智能高级别专家咨询组(High-level Advisory Body on AI)的人员构成为例,39 位专家中有 24 位来自发达国家[10],占比达 61.5%,反映出发展中国家在联合国构建的人工智能全球治理体系的参与有限;三是联合国采取的人工智能全球治理措施的有效性问题,目前联合国主要采取以伦理规范、价值准则、监管指引、技术标准等不同形式的“软法”治理为主,既缺乏具有法律约束力的规则规范和国际条约,也未发展出足够丰富且能够满足不同需求的治理工具[11],有效性不足导致治理措施落地难。

3 联合国参与人工智能全球治理的实践路径

人工智能可以通过促进包容性、减少不平等等方式助力实现近 80% 的可持续发展目标,并在很大程度上赋能联合国工作[12]。自 2019 年《联合国全系统支持人工智能能力发展战略方针和路线图》的出台以来,联合国高度重视人工智能

发展，强调若缺乏对人工智能的伦理监督，人工智能可能会进一步加剧不平等和扩大数字鸿沟。目前，联合国及其附属机构已在全球范围内开展一系列针对人工智能治理的行动，积极构建多元主体协同共治的人工智能全球治理体系，促进人工智能技术服务于可持续发展目标。

一是成立联合国人工智能工作组及专家咨询组，为推进人工智能全球治理提供智力支持。《联合国系统人工智能治理白皮书》指出，要组建独立的、多学科的多利益攸关方专家组进行科学评估，保障人工智能治理行动的合法性[13]。2020 年 10 月，联合国人工智能跨机构工作组（Interagency Working Group on AI）成立，由联合国教科文组织（United Nations Educational, Scientific and Cultural Organization, UNESCO）和国际电信联盟（International Telecommunication Union, ITU）共同领导，以促进伦理发展和维护人权为基础，推动联合国内部人工智能负责任地发展和使用。2023 年 10 月，联合国高级别人工智能专家咨询组成立，由来自政府、私营部门、学术界的 39 名专家组成，咨询组以《联合国宪章》里的普遍价值观为基础，为国际社会加强对人工智能的治理提供引导和支持。

二是出台一系列人工智能治理文件，为人工智能全球治理的发展指引方向。作为联合国数字行动的纲领性文件，《数字合作线路图》明确提出“支持在人工智能领域开展合作——基于人权、安全和可持续的全球合作，促进和平”[14]。自线路图发布以来，联合国及其附属机构高度重视人工智能治理这一全球性议题，以 UNESCO 发布的《人工智能伦理问题建议书》为指引，联合国贸易和发展会议（United Nations Conference on Trade and Development, UNCTAD）、世界卫生组织（World Health Organization, WHO）、世界知识产权组织（World Intellectual Property Organization, WIPO）、国际劳工组织（International Labour Organization, ILO）等联合国机构结合实际工作制定中长期规划、出台伦理建议、发布研究报告（表 2），通过加强对人工智能的“软法”治理，引导人工智能全球治理发展方向。

三是推进人工智能国际标准建设，为人工智能发展提供统一标准与技术支撑。国际标准对人工智能健康有序发展至关重要，是推进人工智能全球治理的重要支柱。联合国在制定和推广国际标准方面发挥了重要作用，ITU 的宗旨和任务之一便是研究并制定国际电信标准，同时促进其广泛应用。目前，ITU 已经发布了 220 多个与人工智能相关的技术标准[26]，强调国际标准需覆盖四大核心要素即将人工智能治理原则转化为实操标准的“转化机制”、支撑数字基础设施与能力建设的“结构化体系”、保障多元主体参与的“包容性流程”以及连接政策与技术需求的“协同网络”[27]。此外，ITU 联合相关国际组织共同推进医疗、农业、通信、减灾防灾等领域的人工智能应用标准建设，发布包含 700 多个人工智能标准和相关出版物的“人工智能标准交换数据库”（AI Standards Exchange database），并通过实施“弥合标准鸿沟项目”（Bridging the Standardization Gap programme）来帮助发展中国家参与人工智能国际标准的工作。

四是助力人工智能能力建设，弥合人工智能领域鸿沟，提升人工智能风险应对能力。包括计算、数据集、人工智能模型在内的人工智能核心基础要素以及由国家人工智能战略和国际合作构成的人工智能驱动要素是实现可持续发展的关键[28]。联合国通过发挥协调机制优势，整合所

表 2　联合国及其附属机构推进人工智能全球治理的代表性文件

Tab. 2　Representative documents of the United Nations and its specialized agencies promoting global AI governance

	名称及简介	类型
2020 年 06 月	联合国秘书长 António Guterres 公布《数字合作线路图》,提出“支持在人工智能领域开展合作”[14]。	规划
2021 年 09 月	联合国秘书长 António Guterres 发布《我们的共同议程》,提出“加强对人工智能的监管”[15]。	规划
2021 年 11 月	UNESCO 发布《人工智能伦理问题建议书》,作为关于人工智能主题的首份全球性规范框架,提出尊重、保护和促进人权等价值观,指出人工智能核心伦理原则包括透明度、责任性、包容性、可持续性及人权优先性等,并提出数据治理、算法审计、技术评估三大重点领域;提出了“准备情况评估方法”框架,旨在评估一个国家的法律、社会、文化、科学、教育、技术和基础设施人工智能能力,及其人工智能系统是否符合建议书中规定的价值观、原则和政策领域[16]。	意见
2022 年 09 月	联合国行政首长协调理事会[1]和方案问题高级别委员会[2]批准《联合国系统内人工智能应用伦理准则》,旨在指导联合国系统各机构在尊重人权和符合伦理的前提下开发、设计、部署和使用人工智能系统[17]。	意见
2023 年 05 月	联合国秘书长 António Guterres 发布《“我们的共同议程”政策简报 5:全球数字契约——为所有人创造开放、自由、安全的数字未来》,提出“为人类管理人工智能”[18]。	报告
2023 年 07 月	WHO、ITU、WIPO 共同发布《人工智能促进健康全球倡议》,旨在规范卫生保健领域人工智能技术的设计与应用。	倡议
2023 年 12 月	联合国高级别人工智能专家咨询组发布《临时报告:为人类治理人工智能》,阐述了如何利用人工智能加速可持续发展以及如何应对人工智能潜在风险和挑战。	报告
2024 年 03 月	联合国大会通过《抓住安全、可靠和值得信赖的人工智能系统带来的机遇,促进可持续发展》决议,是联合国大会通过的首个有关人工智能的决议[19]。	决议
2024 年 04 月	联合国行政首长协调理事会和方案问题高级别委员会批准《联合国系统人工智能治理白皮书》,总结了联合国系统在推进人工智能治理上的模式和方法,分析了现有的适用于人工智能治理的国际规范框架,并明确了 50 多项治理手段[13]。	报告
2024 年 07 月	联合国大会通过《加强人工智能能力建设国际合作》决议,强调人工智能发展应坚持以人为本、智能向善、造福人类的原则,鼓励通过国际合作帮助各国特别是发展中国家加强人工智能能力建设[20]。	决议
2024 年 07 月	ILO 和联合国秘书长技术事务特使办公室[3]共同发布《关注人工智能鸿沟:塑造未来劳动世界的全球视角》,报告呼吁通过国际合作加强人工智能能力建设。	报告
2024 年 08 月	UNESCO 发布《教师人工智能能力框架》及《学生人工智能能力框架》,旨在指导各国教师和学生了解人工智能的潜力和风险。	意见
2024 年 08 月	UNESCO 发布《全球 9 大新兴的人工智能监管方法》,包括基于原则的方法、基于标准的方法、敏捷和实验主义的方法、促进和赋能的方法、调整现有法律的方法、信息获取和透明度授权的方法、基于风险的方法、基于权利的方法以及责任的方法[21]。	报告
2024 年 09 月	联合国全球契约组织[4]发布《生成式人工智能赋能全球目标》,旨在规范私营部门以负责任的方式部署生成式人工智能。	报告
2024 年 09 月	联合国高级别人工智能专家咨询组发布《治理人工智能,助力造福人类:最终报告》,强调人工智能全球治理的必要性,呼吁加强全球合作,并提出设立国际科学小组、在联合国框架内开展全球人工智能治理对话、建立人工智能标准交流中心、建立人工智能法治中心、建立国际能力发展网络、设立全球人工智能基金以及在联合国秘书处设立专门的人工智能办公室[6]。	报告
2024 年 09 月	联合国未来峰会通过《未来契约》及《全球数字契约》《子孙后代宣言》两个附属文件,《全球数字契约》提出“加强人工智能国际治理,造福人类”[22]。	规划
2024 年 10 月	联合国行政首长协调理事会和管理问题高级别委员会[5]批准《联合国系统各组织负责任地使用人工智能示范政策框架》,明确了人工智能风险评估等级以及人工智能风险划分依据[23]。	意见
2025 年 04 月	UNCTAD 发布《2025 年技术与创新报告:包容性人工智能促进发展》,指出全球合作对促进包容和公平的人工智能发展至关重要[24]。	报告
2025 年 07 月	“2025 人工智能向善全球峰会”[6]发布《2025 年“人工智能治理全球对话”联合主席声明:推进包容、可信、创新的人工智能治理》,提出人工智能治理的十大愿景。	声明
2025 年 08 月	联合国大会通过《人工智能独立国际科学小组和人工智能治理全球对话的职权范围和设立及运作方式》决议,拟设立由 40 名专家组成的人工智能科学小组,评估人工智能的风险、机遇和影响[25]。	决议

1) Chief Executives Board for Coordination, CEB; 2) High-Level Committee on Programmes; 3) United Nations Office of the Secretary General's Envoy on Technology; 4) United Nations Global Compact; 5) High-Level Committee on Management; 6) AI for Good Global Summit 2025。

属机构资源,提供开放的人工智能模型,公开的训练数据集,可负担的计算资源以及专项资金支持,推动建立人工智能能力发展网络。包括增加对全球数字公共产品和数字公共基础设施的投资

和供资，推进建设“全球人工智能能力发展基金”（Global Fund for AI Capacity Development），建立30 多个涵盖多个领域的公开数据库，开发一系列包含移动应用程序、网站应用程序在内的公共工具包，并发布“联合国人工智能资源中心”（UN AI Resource Hub）[29]。通过这些举措向多利益攸关方提供人工智能治理资源及工具，助力国际社会提升人工智能风险应对能力。

五是开展人工智能与可持续发展专项，加强人工智能全球治理多边合作。联合国高度重视内外部合作，积极为多利益攸关方参与人工智能全球治理创造机会，增进资源共享，促进多边合作。根据《2024 联合国人工智能活动报告》统计，围绕人工智能与可持续发展这一主题，2024 年共有 53 家联合国机构开展了 729 项相关项目，涵盖智慧农业、智慧交通、科技金融等涉及所有可持续发展目标的领域，其中近 45% 的项目涉及与政府、学术界、私营部门、国际组织的合作[30]，倡导各方以伦理、人权和包容性为基础，探索应用人工智能技术促进可持续发展。

六是搭建多利益攸关方共同参与的国际交流平台，推进人工智能治理国际对话机制建设，深化人工智能全球治理国际合作。在推进人工智能全球治理的进程中，国际合作至关重要。联合国通过每年举办“信息社会世界峰会”（World Summit on the Information Society，WSIS）、“联合国互联网治理论坛”（Internet Governance Forum，IGF）以及“人工智能向善全球峰会”等多利益攸关方共同参与的国际会议，解读最新技术、分享最佳实践、共享治理成果，深化人工智能领域国际合作。其中，2017 年创立的“人工智能向善全球峰会”由 ITU 领导，多家联合国机构支持，已发展成为人工智能领域代表性国际会议。此外，联合国在 2025 年 9 月正式启动“人工智能治理全球对话”（Global Dialogue on AI Governance）机制，联合“人工智能独立国际科学小组”（Independent International Scientific Panel on AI）共同构成联合国人工智能治理框架，助力建设以国际法、人权为基础的安全、可靠和值得信赖的人工智能系统，促进各项治理制度间的协同，为人工智能的监管与发展制定方向。

联合国在推进人工智能全球治理的进程中，坚持“以人为本”的治理理念，采用多学科、跨领域的多边主义治理方法，综合运用多种治理手段，以帮助发展中国家弥合人工智能鸿沟，实现人工智能公平、普惠、可持续发展为目标，构建联合国框架下多元协同、包容普惠、具备互操作性的人工智能全球治理体系，以应对人工智能在全球范围内带来的多面挑战，把握人工智能带来的发展机遇，促进人工智能赋能可持续发展，维护国际社会的公平稳定。一方面强调区域平等参与和促进边缘化群体、社区和个人的参与[31]，为“全球南方”国家参与人工智能全球治理创造机会，并通过国际合作推动人工智能全球治理体系的协调与兼容，增强发展中国家的代表性，同时积极吸纳多利益攸关方的共同参与[32]，维护各方的平等参与和公平受益；另一方面强调人工智能能力建设和风险应对能力的提升，通过协调国际公共资源，开展多边合作，在基础设施建设、技能人才教育、公共资金协调等方面给予发展中国家支持，弥合全球人工智能领域鸿沟；同时，强调推进人工智能治理的互操作性，鼓励开发和部署具有国际互操作性的技术工具和技术标准，推动区域性多利益攸关方人工智能治理实践在全球范围内的相互关联与协同工作，推进制定统一的人工智能全球治理规则。

4 对我国推进人工智能全球治理的建议

人工智能全球治理是基于在科学、技术、工程和数学领域的能力建设，通过“软法”与“硬法”的有机结合，不同主体间的政策需求平衡以及产业界与政府间的相互适应来协同推进的。作为联合国安理会常任理事国和世界上最大的发展中国家，我国在推进人工智能全球治理的进程中，应坚持联合国在国际发展合作中的统筹协调作用，从理念贯彻、模式构建、路径选择、行动示范四方面出发，推动建立在联合国框架下多元协同、敏捷灵活、包容普惠的人工智能全球治理体系。

在理念贯彻上，一是秉承“以人为本”“智能向善”的基本原则，在《全球人工智能治理倡议》的指导下，设计人工智能全球治理议题；二是坚持总体国家安全观，在兼顾技术与产业安全有序发展的基础上，让创新成为第一引擎，平衡技术创新与安全保障、协调产业发展与安全监管，以创新促治理、以治理促发展；三是践行真正的多边主义，构建以联合国为中心的治理协商平台，落实联合国《未来契约》及其附件《全球数字契约》有关承诺，发挥联合国在人工智能全球治理体系中的协调作用。

在模式构建上，一是坚持“敏捷治理”，综合运用多种治理手段，配合多学科、跨领域的治理方法，发展具备快速响应与自适应的治理工具，以应对人工智能技术飞速变化带来的风险，助力推进灵活的、适应的人工智能全球治理；二是推动“包容治理”，鼓励科学界、产业界和政府管理部门之间的沟通交流，为人工智能技术发展划定基线，充分利用联合国框架下一系列多边合作机制，就各国共同面临的人工智能问题组织对话，分享各国人工智能治理最佳实践，推动国际社会达成人工智能治理共识，确保全球规则的一致性、普惠性。

在路径选择上，一是加大基础研究的投入力度，完善数据安全和个人信息保护规范，推动算法向上向善，提高人工智能可解释性、透明性、安全性，构建负责任的技术发展路径；二是及时开展人工智能风险研判，探索分类分级管理，建立健全人工智能安全测评体系，规划人工智能安全治理新路径；三是充分发挥我国数据资源丰富、产业体系完备、应用场景广阔的优势，助力场景、生态与治理的协同创新，通过技术创新、产业升级、治理完善的三轮驱动，塑造中国特色的人工智能治理路径。

在行动示范上，一是完善人工智能标准体系建设，加强跨行业、跨领域人工智能标准协调，积极参与人工智能国际标准的制定，推进建立全球人工智能标准化框架；二是加强人工智能能力建设国际合作，通过开展南北合作、南南合作和三方合作，助力发展中国家人工智能能力建设，弥合全球人工智能鸿沟，支持发展中国家有效、公平、切实参与人工智能全球治理的国际进程；三是强化人工智能领域国际化人才培养，培育融汇技术深度与人文温度、贯通本土实践与全球视野的高水平人才，提升我国人工智能人才全球竞争力，积极向联合国输送人工智能领域顶尖人才和有潜力的青年人才，为联合国构建中的人工智能全球治理体系贡献“中国智慧”。

参考文献

[1] Stanford Institute for Human-Centered Artificial Intelligence. Artificial intelligence index report

2025[R/OL]. (2025-04-17)[2025-08-01]. https://hai. stanford. edu/assets/files/hai _ ai _ index_report_2025. pdf.

[2]BENGIO Y, MINDERMANN S, PRIVITERA D, et al. International AI safety report[R/OL]. (2025-01-29)[2025-10-15]. https://arxiv. org/pdf/2501. 17805.

[3]薛澜. AI 治理要画“禁区”,更要铺“轨道”[J]. 中国经济周刊,2025(15):88-90.

[4]薛澜,梁正. 构建包容、普惠的全球人工智能治理体系:现实挑战与对策建议[J]. 人民论坛·学术前沿,2025(9):23-32.

[5]梁正,宋雨鑫. 人工智能国际治理规则的发展[J]. 中国法律,2024(6):54-59.

[6]High-level Advisory Body on Artificial Intelligence. Governing AI for humanity: Final report[R/OL]. (2024-09-19)[2025-08-03]. https://www. un. org/sites/un2. un. org/files/governing_ai_for_humanity_final_report_en. pdf.

[7]CIHON P, MAAS M, KEMP L. Fragmentation and the future: Investigating architectures for international AI governance[J]. Global policy, 2020,11(5):545-556.

[8]阎德学. 全球人工智能治理:联合国的角色[J]. 国际问题研究,2024(4):120-137,141-142.

[9]清华大学人工智能国际治理研究院,人工智能治理研究中心,清华大学中国科技政策研究中心. 共商共建全球人工智能治理体系报告[R]. 北京:清华大学人工智能国际治理研究院,2024.

[10]United Nations. Members of the AI advisory body[EB/OL]. (2023-10-26)[2025-12-20]. https://www. un. org/en/ai-advisory-body/members.

[11]贾开,俞啥之,薛澜. 人工智能全球治理新阶段的特征、赤字与改革方向[J]. 国际论坛,2024,26(3):62-78,157-158.

[12]United Nations. Artificial intelligence[EB/OL]. (2025)[2025-12-20]. https://www. un. org/en/global-issues/artificial-intelligence.

[13]Chief Executives Board for Coordination, High-Level Committee on Programmes. United Nations system white paper on AI governance: An analysis of the UN system's institutional models, functions, and existing international normative frameworks applicable to AI governance[R/OL]. (2024-04-16)[2025-08-10]. https://unsceb. org/sites/default/files/2024-04/United% 20Nations% 20System% 20White% 20Paper% 20on% 20AI% 20Governance. pdf.

[14]United Nations General Assembly. Roadmap for digital cooperation: Implementation of the recommendations of the high-level panel on digital cooperation-report of the secretary-general[R/OL]. (2020-05-29)[2025-08-10]. https://www. un. org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap _ for _ Digital _ Cooperation_EN. pdf.

[15]United Nations. Our common agenda-report of the secretary-general[R/OL]. (2021-09-10)[2025-08-10]. https://www. un. org/en/content/common-agenda-report/assets/pdf/Common _ Agenda _ Report_English. pdf.

[16]United Nations Educational, Scientific and Cultural Organization. Recommendation on the ethics of artificial intelligence[R/OL]. (2021-11-24)[2025-08-10]. https://unesdoc. unesco. org/ark:/48223/pf0000380455.

[17]Chief Executives Board for Coordination, High-

Level Committee on Programmes. Principles for the ethical use of artificial intelligence in the United Nations System[R/OL]. (2022-09-20)[2025-08-10]. https://unsceb.org/sites/default/files/2022-09/Principles%20for%20the%20Ethical%20Use%20of%20AI%20in%20the%20UN%20System_1.pdf.

[18] United Nations. Our Common agenda-policy brief 5: A global digital compact-an open, free and secure digital future for all[R/OL]. (2023-05-24)[2025-08-10]. https://indonesia.un.org/sites/default/files/2023-07/our-common-agenda-policy-brief-gobal-digi-compact-en.pdf.

[19] 中国新闻网. 联合国大会通过首个有关人工智能的决议[EB/OL]. (2024-03-21)[2025-09-07]. https://www.chinanews.com/gj/2024/03-22/10184672.shtml.

[20] 人民网. 第 78 届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[EB/OL]. (2024-07-01)[2025-09-07]. http://m.people.cn/n4/2024/0702/c23-21140515.html.

[21] United Nations Educational, Scientific and Cultural Organization. Consultation paper on AI regulation-emerging approaches across the world[R/OL]. (2024-08-16)[2025-09-07]. https://unesdoc.unesco.org/ark:/48223/pf00003909 79.

[22] United Nations. Pact for the future, global digital compact and declaration on future generations[R/OL]. (2024-09-22)[2025-08-10]. https://unsdg.un.org/sites/default/files/2025-03/sotf-pact_for_the_future_adopted.pdf.

[23] Chief Executives Board for Coordination, High-Level Committee on Management. Framework for a model policy on the responsible use of artificial intelligence in UN system organizations[R/OL]. (2024-10-10)[2025-09-07]. https://unsceb.org/sites/default/files/2024-11/Framework%20for%20a%20Model%20Policy%20on%20the%20%20Responsible%20Use%20of%20AI%20in%20UN%20System.pdf.

[24] United Nations Conference on Trade and Development. Technology and innovation report 2025: Inclusive artificial intelligence for development[R/OL]. (2025-04-07)[2025-08-10]. https://unctad.org/system/files/official-document/tir2025_en.pdf.

[25] 新华网. 联合国大会通过决议设立全球人工智能治理新机制[EB/OL]. (2025-08-27)[2025-09-07]. https://www.xinhuanet.com/20250827/00c786b4494a4b72878028d9c0300080/c.html.

[26] International Telecommunication Union. Report on PP resolution 214 (bucharest,2022) artificial intelligence technologies and telecommunications/information and communication technologies[R/OL]. (2024-05-06)[2025-09-07]. https://www.itu.int/dms_pub/itu-s/md/24/cl/c/S24-CL-C-0067!! MSW-E.docx.

[27] International Telecommunication Union. AI standards for global impact: From governance to action 2025 report[R/OL]. (2025-09-22)[2025-10-25]. https://aiforgood.itu.int/wp-content/uploads/2025/09/ai-standards-for-global-impact.pdf.

[28] United Nations General Assembly. Innovative voluntary financing options for artificial intelligence capacity-building-report of the secretary-general[R/OL]. (2025-07-02)[2025-10-25]. https://www.un.org/digital-emerging-

technologies/sites/www. un. org. digital-emerging-technologies/files/A_79_966-EN. pdf.

[29]International Telecommunication Union. UN working group launches new AI resource hub [EB/OL]. (2025-12-16) [2025-12-20]. https://www. itu. int/hub/2025/12/un-working-group-launches-new-ai-resource-hub/.

[30]International Telecommunication Union. United Nations activities on Artificial Intelligence (AI) 2024 [R/OL]. (2024-09-25) [2025-09-07]. https://www. itu. int/dms_pub/itu-s/opb/gen/S-GEN-UNACT-2024-PDF-E. pdf.

[31]钱忆亲. 联合国框架下的全球人工智能治理现状、挑战与展望[J]. 中国信息安全,2024(4):62-66,71.

[32]张凌寒. 联合国框架下的人工智能规则治理[J]. 数字法治,2025(1):37-43.

作者贡献说明

刘　畅:收集、整理资料,撰写论文,修改完善文章;

段　续:收集、整理资料,修改完善文章;

何腾蛟:修改完善文章。

(收稿日期:2025-10-09;修回日期:2025-10-30)