

# 人工智能全球治理： 供需“缺口”与弥合路径

李 艳

**【内容提要】**受新一轮人工智能浪潮推动，人工智能全球治理进入历史发展新阶段，涌现大量倡议与方案。但相较于快速技术迭代与应用拓展带来的井喷式治理需求，无论是应对新型数智鸿沟带来的发展制约，还是“技术—经济—政治”三位一体风险引起的安全挑战，现有治理存在较大供需“缺口”。其背后折射出技术应用内在属性、技术治理客观规律以及地缘政治深刻变动等因素影响下的人工智能全球治理悖论，包括安全管控与创新活力的张力悖论、跨国协同与利益分化的协调悖论、国家安全与公共利益的边界悖论以及伦理共识与文化多样性的价值悖论。要解决相关悖论带来的制约与挑战，推动未来人工智能全球治理走“深”走“实”，需要创新更符合人工智能技术与应用规律的治理模式、务实设定阶段性治理优先目标、共同划定发展的安全“红线”，构建更具包容性与开放性的全球治理体系。

**【关键词】**人工智能；全球治理；治理悖论；数智鸿沟

以生成式人工智能为代表的新一轮人工智能发展浪潮方兴未艾，技术迭代与应用拓展速度前所未有。但凡新兴技术必然带来赋能与风险并存的“双刃剑”效应，旨在平衡安全与发展的人工智能全球治理进程随之加速，“制定框架指导人工智能发展，以平等造福全人类，全球治理在确保这一点上发挥着至关重要的作用”。<sup>①</sup>经过国际社会各方的努力推进，当前人工智能全球治理已取得阶段性成效，治理共识基本达成、治理平台不断完善、治理方案不断涌现，但相较于人工智能技术与应用演进产生的现实需求仍存较大“缺口”。本文在梳理人工智能全球治理发展脉络与现状基础上，探析治理供需“缺口”背后的影响因素，并就如

何弥合进行思考。

## 一、人工智能全球治理热潮之下的 供需“缺口”

任何议题的全球治理本质上都是为实现发展与安全的平衡，“给超出国家独立解决能力范围的社会和政治问题带来的更有秩序和更可行的解决方法的努力。”<sup>②</sup>作为“高度复杂而广泛的全新行为的总和”，<sup>③</sup>“个体、制度、公共部门与私营部门等管理共同事务并实现多途径整合，在这一持续过程中，相互冲突的利益或多样化利益能够互相调适并合作，包含正式制度安排和非正式制度安排的综合使用”。<sup>④</sup>

人工智能的发展与安全需要全球治理，以国际组织的制度性防范技术快速迭代带来的治理问题发酵。<sup>⑤</sup>人工智能全球治理是一种“动态复杂机制”，是国际社会各方综合运用正式与非正式制度安排，解决发展与安全核心关切的过程。鉴此，本文涉及的人工智能全球治理供需关系，主要基于对既有治理机制、政策与行动能否有效满足快速增长的发展与安全需求的阶段性评估。

### （一）“供给侧”：相应机制建设起步，重点聚焦提出人工智能全球治理倡议与方案

人工智能的源起与发端虽可追溯到20世纪40年代至50年代，1956年达特茅斯会议首提“人工智能”概念后，历史上也经历多轮发展小高潮。但真正加速演进是进入21世纪后，尤其是2006年杰弗里·辛顿（Geoffrey Hinton）将深度学习推向风口浪尖，推动技术创新加速。2014年，伊恩·古德费洛（Ian Goodfellow）及团队正式提出生成对抗网络（GAN）概念，创造出一种革命性工具进一步促进人工智能技术创新突破。正是在该时期，人工智能技术商业化与军事化应用涌现，全球治理进程也应运而生。2017年，国际电信联盟（ITU）举办首届“人工智能造福人类全球峰会”，同年，联合国秘书长报告首次明确将人工智能纳入全球治理议题。2018年，联合国成立数字合作高级别小组，2019年，联合国教科文组织启动制定全球首份《人工智能伦理建议书》等，为人工智能全球治理奠定了一定基础。2022年，以ChatGPT为代表的生成式人工智能横空出世，成为人工智能全球治理进入历史发展新阶段的重要“分水岭”。

一是“联合国框架”致力于成为主导性

平台。作为治理进程的发起者与主要推动者，2023年7月，联合国秘书长古特雷斯在安理会就人工智能举行的首场公开辩论会上表示，联合国是引领人工智能全球治理的“理想场所”。2023年10月，古特雷斯宣布组建人工智能高级别咨询机构，2023年12月发布《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；2024年，联合国大会先后通过《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》《加强人工智能能力建设国际合作》两项决议，并在9月的联合国未来峰会中通过《全球数字契约》（*Global Digital Compact*），明确提出“加强人工智能国际治理，造福人类”。2025年ITU“人工智能造福人类全球峰会”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》。

二是既有国际机制积极拓展人工智能治理职能。国际社会各方对人工智能的重视持续提升，越来越多的国际机制将人工智能治理纳入核心议程。2023年10月，七国集团领导人发起“广岛人工智能进程”，推进与相关国际组织与国家合作，构建体现包容性的全球治理框架；2023年，二十国集团在新德里峰会重申“以人为本，实现人工智能向善并服务全人类”的观点，2024年的里约热内卢峰会宣言表示将人工智能议题置于“全球治理机构改革”之中，并于2025年成立人工智能任务组；2024年，经合组织发布新修订的《关于人工智能的建议书》，呼吁建立可互操作的全球治理框架；2024年7月，上海合作组织发布《阿斯塔纳宣言》，表示愿开展合作，共同做好风险防范，确保人工智能造福全人类；

2023年,金砖国家同意成立人工智能研究小组,促进形成具有广泛共识的人工智能治理框架和标准规范。2024年11月,世界互联网大会(乌镇峰会)将主题设定为“拥抱以人为本、智能向善的数字未来”,推进人工智能领域的国际对话与合作。

三是新的全球治理平台得以搭建。2023年11月,首届人工智能安全峰会在英国布莱奇利举行并发布《布莱奇利宣言》;2024年5月,人工智能首尔峰会在韩国召开,相关国家签署《关于安全、创新和包容性人工智能的首尔宣言》;2025年2月,巴黎人工智能行动峰会在法国举办,相关国家共同签署《关于发展包容、可持续的人工智能造福人类与地球的声明》,将构建普适性国际治理框架、推动可持续发展和国际协调治理等列为优先事项。自2019年起,中国在上海启动“世界人工智能大会”,陆续发布《人工智能全球治理上海宣言》、以及在2025世界人工智能大会暨人工智能全球治理高级别会议上发布《人工智能全球治理行动计划》,并倡议成立世界人工智能合作组织。<sup>⑥</sup>

值得注意的是,非政府主体尤其是科技企业也积极参与人工智能全球治理进程并发挥重要作用。2024年2月,德国慕尼黑安全峰会上,包括亚马逊、谷歌、IBM、Meta、微软、OpenAI、TikTok和X在内的20家科技公司共同签署技术协议,致力于抵制欺骗性人工智能生成内容风险,并同意在各自平台或产品提出解决方案;2024年5月,人工智能首尔峰会上,主办方起草《前沿人工智能安全承诺》,推动各方就开发和部署前沿人工智能模型和系统时负责任管理风险达成承诺,来自中国、美国、欧洲、中东和亚洲其他国家的16家人工

智能公司或组织签署了该承诺。据不完全统计,目前全球关于人工智能的倡议已有50多个,包括2023年10月中国发布的《全球人工智能治理倡议》等。

从历史回溯与现状梳理中不难看出,人工智能全球治理整体处在起步阶段,治理平台不断搭建,原则性共识基本达成,主要成果集中在倡议与方案的提出,迈出了全球人工智能治理的坚实一步,后续能否构建有效的体系性保障机制,确保共识与倡议的“落地”,切实开展合作才是人工智能全球治理的关键。

## (二)“需求侧”:技术迭代与应用拓展加快,带来日趋增长与紧迫的治理需求

从发展需求来看,国际社会需要为应对“数智鸿沟”做好准备。人工智能全球治理议程下的发展问题,主要关注国际社会各方能否共享智能红利、实现共同发展。2024年9月22日,联合国大会通过的《全球数字契约》中开宗明义地指出:“数字技术正在急剧改变我们的世界。它们为民众和社会的福祉与进步以及我们的地球提供巨大的潜在惠益。它们为加速实现可持续发展目标带来希望。”“我们只有通过加强国际合作,弥合国家之间和国家内部的所有数字鸿沟,才能实现这一点。我们确认,这些鸿沟给许多国家,特别是发展需求迫切而资源有限的发展中国家带来挑战。”<sup>⑦</sup>事实上,虽然人工智能技术与应用还处于发展初期,但实已然开始催生新的“智能鸿沟”。人工智能天然呈现“高度垄断”格局,目前全球具有实力研发和推广新型智能系统的企业主要以微软、谷歌、脸书等科技巨头为主。它们在数据、算力以及高端芯片方面的优势使其具有技术先发与市场先占优

势，“赢者统吃”“强者恒强”的格局难以被打破。智能时代的发展不平衡将更加突出，尤其不利发展中国家，如劳动力优势的丧失，多数国家还不具备智能升级换代的条件。这就使得经过多年努力与发达国家间有所缩小的“数字鸿沟”，很有可能叠加“智能鸿沟”，形成更大的“数智鸿沟”，国际社会或将面对新的“南北问题”。

从安全需求来看，国际社会面临“技术—经济—政治”三位一体的复杂风险。一是从技术层面来看，人工智能技术在加剧传统既有安全风险的同时，还会产生一些难控甚至是不可控风险。如人工智能的幻觉问题（hallucination）和“涌现”（emergence）问题。这是技术“内嵌”性因素引起的问题，不太可能彻底解决，只能通过持续技术改进逐步缓解。随着人工智能技术在经济金融、自动驾驶、医疗诊断甚至军事战场等高风险场景中应用加快，可能的风险正是加速变现。二是从经济层面来看，人工智能应用对经济模式、社会结构乃至人类文明产生的颠覆性影响正加速到来，如基于劳动力价值受到极大削弱，经济上的“边际收益递减”将被“指数叠加”取代，现有经济模式与财富分配机制都在受到挑战，而经济基础决定上层建筑，深刻变革之下蕴含巨大社会风险。正如国际电联（ITU）秘书长博格丹·马丁在2025人工智能造福人类全球峰会上的致辞所言：“人们面临的重大风险不是人工智能会消灭人类，而是在尚未充分理解这对人类和地球意味着什么的情况下，就竞相将人工智能嵌入到各个领域”。<sup>⑧</sup>三是从政治层面来看，当今世界正处于百年未有之大变局，恰逢新一轮科技革命爆发，科技领域成为大国竞争博弈的关键高地，作为这场科

技革命重要前沿与牵引的人工智能更是首当其冲，治理议题不可避免地具有“政治高敏感性”，反映到现实中就是“国家间信任不足，从而引发‘超规格’的治理手段和‘高警惕’的安全互动”，<sup>⑨</sup>直接冲击大国战略稳定与影响国际安全环境。

综上，人工智能全球治理虽取得阶段性成效，具备一定共识性基础，但面对“井喷式”增长的发展与安全治理需求，无论是应对新型“数智鸿沟”带来的发展挑战，还是防范“技术—经济—政治”三位一体的安全风险，都存在较大供需“缺口”。

## 二、人工智能全球治理“缺口”的根源分析

人工智能全球治理“缺口”的形成因素呈现出多维度、交织性的复杂特征，这其中既有技术应用的内在属性、新兴技术全球治理的客观规律等带来的“正常缺口”，也有地缘政治深刻变动与大国博弈竞争加剧引发的“非正常缺口”。

### （一）人工智能技术与应用的内在属性

近年来，在大数据、算法与算力等技术与应用加持下，人工智能技术出现“质的飞跃”，作为一种前所未有的技术形态，“与一般的信息技术相比，人工智能具有自我创造，超强学习和超级进化的特性”，<sup>⑩</sup>“人工智能不同于我们已经探索过的其他工具或领域，例如核武器或空间问题，人工智能带来的问题往往是前所未有的”<sup>⑪</sup>对于治理而言，挑战无疑是巨大的。

此外，从应用上看，此轮人工智能浪潮实现“技术突破—社会应用—重大影响”的完整闭环。以ChatGPT为代表的生成式人工智能



推动人机交互进入新时代，打通技术创新到社会应用的“最后一公里”。虽然应用还处发展初期，但场景已经铺开，渗入到经济、社会、国防、安全、公共服务等多领域。人工智能的自主决策能力和影响范围，带来更多复杂的安全风险，客观上对全球治理的覆盖面与敏捷性均提出更高要求。

## （二）技术治理的客观规律

虽然人工智能技术前所未有，但从治理角度，治理路径仍在相当程度上遵循技术治理的客观规律。一般而言，在技术发展与应用初期，治理都会面临“科林格里奇困境”<sup>⑫</sup>和“步调难题”。<sup>⑬</sup>前者指的是一种“双重困境”，一项技术的社会后果不能在技术生命早期被预料到的“信息困境”，和当不希望的后果被发现时，技术已成为经济和社会结构的一部分，以至于对它的控制十分困难的“控制困境”。“当一个容易的事情被改变时，结果将难以预见。而当需要做出明显的改变时，改变的方法却变得昂贵，困难和消耗时间。”<sup>⑭</sup>后者是指科技创新的速度大大超过治理机制与规范的更新速度，即“科技是成指数级增长的，但是社会，经济和法律系统只会缓慢增长”。因此，技术治理的“滞后性”难以避免。从这个意义上讲，人工智能全球治理相较现实发展需要相对滞后具有一定客观合理性。

事实上，在技术发展初期，发展与安全的平衡度很难把握，鼓励创新与发展担心安全不可控，而强调安全关切又担心损害创新活力，错失发展机会，客观上会要求重大治理政策或措施的出台格外谨慎。

## （三）地缘政治竞争加剧的现实影响

人工智能已成为大国科技实力和综合国力竞争的核心领域，地缘竞争与大国博弈加剧全

球人工智能治理的复杂性。

一方面，美国、中国、欧盟等主要经济体围绕人工智能技术创新、产业布局和治理规则展开激烈竞争，大国之间在数据共享、技术标准、安全审查等方面存在不信任，导致协作阻碍，尤其是在军事人工智能应用等高敏感领域，共识更加难达，立场更难协调。“美国和中国在人工智能领域的竞争可能会彻底改变力量平衡，并对全球治理产生重大影响，这可能会对数据治理、技术标准、道德伦理和地缘政治局势产生影响。”<sup>⑮</sup>

另一方面，广大发展中国家更加关切新的“南北”问题。面对人工智能的高垄断格局，广大发展中国家对自身安全与共同安全的强烈危机感，必然转化为人工智能全球治理诉求，与强调实力优势的发展与安全理念形成激烈对撞。整体上来看，高度内嵌的地缘政治要素给人工智能全球治理带来消极影响，国际社会各方合作的意愿下降，资源投入不足，更加专注于自身的人工智能发展以及国内治理。

## （四）多重因素复杂交织下的治理“悖论”

需要特别指出的是，上述三方面因素并非彼此独立，而是复杂交织，造成更深层的治理“悖论”。

一是安全管控与创新活力的张力悖论。随着深度学习、大型生成模型与自动化决策系统的迅速迭代，人工智能技术带来巨大的生产力提升和社会福祉增益，但也衍生算法不透明、模型偏见、系统脆弱与跨域攻击等安全隐患。各国本身在应对AI风险时既需强调“技术精进”、维持企业与科研机构创新活力，又不能放任潜在危害。

二是跨国协同与利益分化的协调悖论。人

工智能相关的技术与供应链，从芯片制造、跨境数据流动到算法生态均植根于全球网络，单一国家或地区难以独立设界施策。然而，国际社会在规则制定上却因经济实力、技术储备与政治立场差异存在显著利益诉求分歧。如联合国秘书长高级咨询委员会（HLAB-AI）、G20 可持续发展工作组、全球人工智能伙伴关系（GPAI）等纷纷开展议程设定，却因投票机制、议题优先级和资源配置上的博弈，导致难以形成全球层面有效的规范。正是鉴于此，有专家指出，各类治理倡议的激增及多边协调中的分歧，可能加剧现有不对称和全球数字治理的合法性赤字，从而进一步削弱多边主义。<sup>⑩</sup>

三是国家安全与公共利益的边界悖论。国家安全不断被纳入人工智能全球治理核心议题中，却也因此加剧战略竞争与信任赤字。以军事人工智能应用为例，美国政府通过《负责任军事使用政治宣言》宣示遵守国际法义务，却排斥对致命自主武器的法律禁令，并在多边谈判中推动排除军事领域监管，以维护技术优势；中国等广大发展中国家在支持人工智能全球治理合作的同时，面对严峻的地缘政治与大国博弈形势，也不得不高度重视科技安全。实践中，如何在既保障防范系统性风险与技术滥用的同时，又能避免可能由此带来的过度审查、技术壁垒与产业脱钩，是相关国家面临的共同挑战。

四是伦理共识与文化多样性的价值悖论。事关更长远未来的人工智能伦理治理呼唤“以人为本”“尊重隐私”“公平无歧视”等原则，但在不同文化、法律传统与价值观背景下，对这些伦理原则的阐释和优先级排序存在显著差异。西方以个人权利与隐私保护为核心；东亚

国家更强调集体福祉与社会稳定；全球南方国家则聚焦数字平等与能力提升。这种多元价值观的并存使全球伦理共识层面取得有限进展，而在具体规则落地时，各国常以本土化需求优先，导致国际伦理标准的实质性冲突与“分而治之”式实施，形成“共识氛围下的实践割裂”悖论。

### 三、人工智能全球治理“缺口”的弥合路径

人工智能技术的独特性加剧安全风险与治理上的步调难题，需要更加高效敏捷的国际合作；但地缘政治的高度嵌入，却使得国际合作的意愿与效力有限，难以满足现实治理需求。要破解此困境，需要找到有效弥合治理“缺口”的路径与方向。

#### （一）创新“技术嵌入式敏捷治理”模式

新一轮人工智能技术方兴未艾，无论未来技术路线会如何“百花齐放”，或者应用场景如何演进，技术应用加速迭代的大趋势不可逆。所谓“技术嵌入式敏捷治理”，一是技术赋能治理，即利用人工智能技术，助力破解传统的“科林格里奇困境”中的“信息困境”，人工智能驱动下的数据收集与分析效能是“指数级”的，智能决策能够极大提高安全风险预判水平，为治理前置提供坚实依据；二是强化敏捷治理模式，最大程度地破解传统技术治理的“步调难题”，尽量缩短机制规范的滞后，解决技术迭代与规则供给之间的不适配。这种“技术嵌入式敏捷治理”通过将重要技术规范、标准工具与合规机制贯穿人工智能技术发展与应用全生命周期，从而有效满足日益增长的技术应用治理需求。

## （二）阶段性设定人工智能全球治理的优先目标

人工智能全球治理议题多元且复杂，同时推进的可能性与现实性不大，再加上地缘政治影响下的利益诉求的分化较大，较为务实的做法是尽量找到国际社会核心关切的最大公约数，并将其作为全球治理的优先选项，这在一定程度上能够缓解跨国协同与利益分化带来的阻力。一方面从现实考虑，既然是公约数，共识相对易达成，具体合作也较易推进。如可推动制定国际通用的人工智能风险分级、可解释性、模型评估与数据溯源等标准；还可在解决人工智能网络安全风险、幻觉修正等领域开展合作与最佳实践交流，形成可推广的实践案例，为全球提供经验样板。另一方面从长远来看，只有合作推进顺利，才能更好建立信任，建立有效的治理机制以及找到治理的最佳实践。当然，所谓优先目标是阶段性的，根据形势发展需要，可以不断动态调整。

## （三）共同划定全球人工智能发展的安全“红线”

多数安全风险可以通过技术发展与实践得以有效防范，但毕竟人工智能技术过于特殊，国际社会还是应该对可能的“终极”风险有所准备。2023年5月，超过350名人工智能领域的行业高管、专家和教授在非营利组织人工智能安全中心（CAIS）网站上签署一份公开信称：“减轻人工智能带来的灭绝风险，应该与流行病和核战争等其他社会规模的风险一起，成为全球优先事项。”<sup>⑩</sup>虽然从当前应用场景看似有点超前，但随着人工智能技术的快速迭代，已经有越来越多的专家学者呼吁警惕人工智能的“奥本海默时

刻”。而实践中，单个国家在追求发展的过程中难以也不会对所谓“终极风险”提前自我规束，因此，此议题只能被纳入全球治理议程，国际社会可在《前沿人工智能安全承诺》的基础上，评估“不可容忍”风险阈值并据此划出“红线”，在风险过大时暂停开发或部署人工智能模型和系统，关键时刻确保能够踩“刹车”。

## （四）构建更具包容性与开放性的全球治理体系

根据历史经验，新兴技术领域的全球治理从来都是“复杂机制”，即基于不同的具体议题，相应的治理机制建设与治理实效有所不同，呈现鲜明的“不均衡”特质。人工智能全球治理也必然如此。更何况，当前治理还处在探索初期，多重因素共同塑造下的动态演进会是未来人工智能全球治理的“常态”。当前实践表明，人工智能全球治理的基本机制涵盖“多边—双边—区域—技术产业社群—学术联盟”等多层正式与非正式机制安排。对此，应鼓励各种形式的有益探索，保持包容性与开放性，促进“数字伙伴关系”“技术对话实验室”“能力建设联盟”“智库网络”以及“企业—政府—学界”三方合作平台等多元形式深化协同。

## 结语

人工智能浪潮席卷之下，未来人工智能全球治理的重要性和必要性会更趋凸显。虽然在技术与地缘政治等复杂因素影响下，当前治理进程面临诸多现实困境，但国际社会各方对于“智能向善”共识已达并展开积极行动。虽任重而道远，但千里之行始于足下，

正如中国关于建立世界人工智能合作组织倡议所提,通过加强各国之间发展战略、治理规则、技术标准的对接协调,在充分尊重各国政策和实践差异性的基础上,逐步形成具有广泛共识的人工智能全球治理框架和标准规范,确保人工智能始终沿着人类文明进步的方向发展。

### 【注释】

- ① See David Leslie et al., “‘Frontier AI’, Power, and the Public Interest: Who Benefits, Who Decides?”, *Harvard Data Science Review*, September 9, 2024, pp.1–20.
- ② T.G.Weiss and L.Cordenker, eds, *NGOs, the United States and Global Governance*, London: Lynne Rienner Publishers, 1996, p.17.
- ③ 邵鹏:《全球治理的理论与实践》,吉林:吉林出版社,2010年,第64页。
- ④ The Commission on Global Governance, *Our Global Neighborhood*, Oxford: Oxford University Press, 1995,p4.
- ⑤ Robert O’Brien, *Contesting Global Governance: Multilateral Economic Institutions and Global Social Movements*, Cambridge University Press, 2000, p.136.
- ⑥ 《中国政府倡议成立世界人工智能合作组织》,中国政府网,2025年7月26日, [https://www.gov.cn/yaowen/liebiao/202507/content\\_7033957.htm](https://www.gov.cn/yaowen/liebiao/202507/content_7033957.htm), 访问时间:2025年8月17日。
- ⑦ 联合国:《全球数字契约》, <https://www.un.org/zh/documents/treaty/A-RES-79-1-Annex-I>.
- ⑧ 《国际电联秘书长:包容性人工智能治理将更好造福人类》,联合国,2025年7月8日, <https://news.un.org/zh/story/2025/07/1140308>, 访问时间:2025年8月17日。
- ⑨ 鲁传颖、张璐瑶:《人工智能的安全风险及治理模式探索》,载《国家安全研究》,2022年第4期,第91页。
- ⑩ 薛澜、赵静:《人工智能国际治理:基于技术特性与议题属性的分析》,载《国际经济评论》,2024年第3期,第56页。
- ⑪ 《人工智能在军事领域发展前景如何?专家解读→》,央视新闻百家号,2024年9月14日, <https://baijiahao.baidu.com/s?id=1810171039582077286&wfr=spider&for=pc>, 访问时间:2025年8月17日。
- ⑫ David Collingridge, *The Social Control of Technology*, New York: St. Martin’s Press, 1980, London.
- ⑬ Larry Downes, “The law of disruption”, 2009, <http://www.larrydownes.com/the-laws-of-disruption/>.
- ⑭ John Brockman, *This Explains Everything*, Harper Perennial, 2013, p.255.
- ⑮ Asia Maqsood, Ahyousha Khan, and Muhammad Usama Siddiqi, “US-China Competition in Artificial Intelligence: Implications on Global Governance,” *Journal of Asian Development Studies*, Vol 12, December 30, 2023, pp.481–493.
- ⑯ 《金砖国家领导人关于人工智能全球治理的声明》,外交部,2025年7月9日, [https://www.mfa.gov.cn/zyxw/202507/t20250709\\_11668022.shtml](https://www.mfa.gov.cn/zyxw/202507/t20250709_11668022.shtml), 访问时间:2025年8月17日。
- ⑰ 《美媒:科技行业领袖警告AI可能给人类带来“灭绝风险”》,中国新闻网,2023年6月1日, <https://www.chinanews.com.cn/gj/2023/06-01/10017541.shtml>, 访问时间:2025年8月17日。

(截稿:2025年8月 责编:荆江)

作者简介 李艳,中国现代国际关系研究院科技与网络安全研究所所长,研究员、博士生导师