



CISS 海外信息专报 (2026 年 6 月 4-5 日)

1、美国《报业辛迪加》：需构建具备可竞争性与互操作性的人工智能堆栈体系

5 月 25 日，美国《报业辛迪加》网站发表前印度财政与民航国务部长贾扬特·辛哈（Jayant Sinha）的文章《如何构建主权人工智能堆栈》。文章认为，构建具备充分可竞争性与互操作性的人工智能堆栈体系对实现人工智能主权至关重要。人工智能堆栈包含的各个相互依赖的技术层级，涵盖能源基础设施、算力资源、基础模型、应用程序和智能体，所有层级均需具备可竞争性与互操作性，即企业可自由进出市场，不同工具、模型与系统能够交换信息。构建主权人工智能堆栈需搭建四大支柱：一是落实事前竞争监管，通过立法规制不正当竞争与垄断行为；二是实现架构内置可竞争性，将竞争机制融入数字系统的底层架构；三是搭建面向全球共享的开放权重人工智能模型，并由专门机构负责治理；四是规范智能代理交互界面，要求具有系统重要性的数字企业提供符合安全规范、具备隐私保护能力、可与其他智能体互联互通的智能服务模块。文章强调，没有任何国家能独立构建人工智能主权的完整技术架构，各国应组建联盟以共同推进

人工智能主权建设。

<https://www.project-syndicate.org/commentary/every-layer-of-sovereign-ai-stack-must-be-contestable-and-interoperable-by-jayant-sinha-2026-05>

编译：程果

2、美国兰德公司：需通过行业规范与监管防范人工智能网络滥用风险

5月28日，美国兰德公司受英国人工智能安全研究所委托，发表报告《探究前沿人工智能模型在进攻性网络攻击中的潜在用途：一项人类能力提升研究》。报告认为，尽管前沿人工智能模型迭代迅速，已具备支持漏洞分析及漏洞利用等进攻性网络任务的能力，但其对低技能威胁行为者的能力提升作用依然有限。研究发现，大模型在协助完成端到端网络攻击方面未呈现显著提升趋势，多数被试即使借助人工智能技术也无法完成高难度攻击链。然而，人工智能存在有限的技能拉平效应（skill-levelling effects），在简单任务中，新手完成率提升达18%，远高于技术人员的5%，且耗时显著缩短。研究认为，人工智能应用主要在攻击初期为用户提供入门辅助，对全程执行帮助有限。此外，模型的内置安全护栏可通过拒绝违规查询大幅降低攻击效率。文章建议技术公司注意防范前沿人工智能被潜在攻击者滥用，网络安全团队应及时修补已知漏洞；监管部门应要求开发者报备模型潜

在的攻击提升风险；业界与研究界应制定标准化评估基准，并建立人工智能技术滥用与恶意网络活动的关联数据库。

https://www.rand.org/pubs/research_reports/RR3892-1.html

编译：孙韵雯

3、美国《安全事务》：俄关联黑客组织用人工智能攻击乌克兰却频发低级失误

5月29日，美国《安全事务》网站发表安全研究员皮耶路易吉·帕加尼尼（Pierluigi Paganini）的文章《认识GREYVIBE：用人工智能攻击乌克兰却仍犯低级错误的俄罗斯关联黑客组织》。报道援引 WithSecure 公司研究成果，披露自2025年8月起活跃的高级持续性威胁（APT）组织GREYVIBE，其主要针对乌克兰军事、政府、民用及商业领域发动网络攻击。该组织技术水平有限但攻击性强，通过人工智能工具系统性弥补技能短板：使用 Ideogram AI 生成诱饵图片，借助 ChatGPT、谷歌 Gemini 开发恶意代码、编写混淆脚本及生成攻击指令，构建了 PhantomMail、PhantomClick 等五条差异化攻击链，通过钓鱼邮件、虚假验证码页面、仿冒慈善网站等多种载体投递恶意程序。研究认为，GREYVIBE 处于国家支持与网络犯罪的灰色地带，与 TrickBot 等犯罪团伙存在关联，甚至在攻击中植入加密货币挖矿程序牟利。但其操作极不规范，多次出现低级失误：恶意软件设计漏洞暴露后端服务器、将测试样本上传至病毒分

析平台、代码命名使用网络俚语，这些漏洞使研究人员得以完整掌握其攻击手法与基础设施。

<https://securityaffairs.com/192877/apt/meet-greyvibe-the-russian-linked-hacking-group-using-ai-to-target-ukraine-and-still-making-rookie-mistakes.html>

编译：钟珺雯

4、美国《华盛顿邮报》：美维护人工智能霸权地位须应对五大战略挑战

5月29日，美国《华盛顿邮报》发表约翰斯·霍普金斯大学高级国际研究学院全球事务教授哈尔·布兰兹（Hal Brands）的文章《美国人工智能霸权取决于这些必要举措》。文章列举马杜罗突袭、“史诗怒火行动”、伊朗反向打击阿联酋与巴林数据中心、Anthropic的Mythos模型展现超人类漏洞挖掘与利用能力等案例，认为人工智能正重塑战争节奏并暴露全球“堆栈”的地缘脆弱性。文章提出，美国须应对五大战略挑战：一是稳固技术前沿优势，收紧对华出口管制、防范模型蒸馏带来的技术窃取风险；二是争夺全球南方市场，依托开发金融公司、进出口银行，为新兴市场基础设施提供融资支持，并推进国务院“硅和平”（Pax Silica）计划；三是加固全球人工智能堆栈，强化对占据全球九成以上加速器芯片产能的台积电的安全承诺并巩固荷兰、日韩德等盟友体系；四是借鉴布坎南等学者的“宏大交易”（Grand Bargain）

框架重构政企关系——政府保障电网扩容、放宽高技能移民准入，企业承诺新模型预先共享并限制危险能力扩散；五是平衡安全与技术优势，对合规开展安全建设的企业给予税收抵免，寻求与中国就限制流氓行为者获取关键能力、禁止危险军事用途达成协议。中美元首峰会的人工智能安全讨论被认为是富有前景的合作起点。

<https://www.washingtonpost.com/opinions/2026/05/29/us-artificial-intelligence-supremacy-depends-taking-these-essential-steps/>
编译：张朔宁

5、美国《科技政策报》：人工智能军事化加剧中等强国技术依赖性

6月1日，美国《科技政策报》发表公共事务和政治风险专家尼什塔·古普塔（Nishtha Gupta）的文章《人工智能参战，中等强国将面临哪些风险》。文章认为，随着人工智能日益军事化，中等强国正因对外国人工智能堆栈的结构性依赖而面临新的战略脆弱性。中等强国虽具有发展军事人工智能的雄心，但无论是与中美两国对齐、推动供应商多元化，还是推进自主建设，都无法真正摆脱对外国硬件、模型和基础设施的依赖。这种依赖意味着中等强国必须接受外国系统内嵌的设计逻辑与使用限制，且供应商及其所属国政府可随时通过硬性断供或性能降级，影响中等强国的军事行动。因此，当前中等强国的主权人工智能在很大程度上是一种幻觉，

真正完整的自主性应同时覆盖硬件、数据、基础设施与法律管辖权等多个层面。作者认为，现有国际军事人工智能治理体系存在漏洞、难以发挥约束作用，并提出四项实质性改革方向：一是建立保障战时服务连续性的合同规范，二是提高人工智能供应链透明度，三是限制公共数据向安全领域转用，四是推动军用与民用人工智能基础设施分离。文章强调，中等强国应以共同利益为基础参与国际治理，推动建立约束大国和科技企业的新规则，否则所谓的国家人工智能能力将始终依附于他国控制的技术基础之上。

<https://www.techpolicy.press/what-middle-powers-risk-when-ai-goes-to-war/>

编译：杜宛鸿

6、美国《科技政策报》：现有人工智能安全评估体系存在认知漂移隐性风险盲区

5月26日，美国《科技政策报》发布 VeracIQ 创始人詹妮弗·金内（Jennifer Kinne）的文章《人工智能安全中的盲点》，对 Anthropic 公司一项人工智能安全研究进行分析。该研究认为，前沿人工智能故障多为表现为不可预测的无序异常，而非持续性的目标对齐偏差，因此治理应侧重预防事故。本文认可该研究模型输出方差递增的实证结果，但指出其存在论证漏洞，研究未对标客观现实，从输出波动推导运行失序的结论缺乏依据。作者提出了学界长期忽视的隐性风险

“认知漂移”。该风险介于随机故障与定向偏差之间，人工智能模型虽运行稳定、可通过常规验证，但其内部参考基准会持续定向偏离客观事实，逐步嵌入系统形成隐患，且无法被现有评估体系检测识别。人工智能医疗领域已出现实例，美国食品药品监督管理局（FDA）的审批并非审核“这是否符合临床实际”，而是审核“这是否类似于已获批的设备”，导致大量医疗设备在上市后被召回，印证了这一治理模式的缺陷。人工智能安全研究核心应是提前识别模型现实偏移风险，而当前行业尚无成熟检测手段，难以实现全方位有效治理。

<https://www.techpolicy.press/the-blind-spot-in-ai-safety/>

编译：张佳延

7、美国《外交信使》：人工智能治理应以妇女和平与安全框架为核心基础

5月27日，美国《外交信使》发表菲律宾棉兰老岛“妇女赋权行动对话与和平”（WE Act）组织创始人贾米拉·艾莎·桑吉拉（Jamila - Aisha P. Sanguila）的文章《妇女、和平与安全：人工智能治理缺失的核心》。文章认为，人工智能治理应以妇女、和平与安全（WPS）框架为核心基础。第一，技术从来不是中立的，人工智能落地的社会环境本身存在不平等、军事化、问责机制缺失等结构性问题，该风险在冲突频发地区尤为显著。妇女、和平与安全（WPS）议程由联合国安理会第1325号决议确立，其核心是保障女性在冲

突预防、和平建设与安全治理中的参与权、受保护权与领导地位。国际社会应当摒弃让 WPS 机制被动适配 AI 治理的固有思路，从制度源头将人工智能治理体系根植于该议程核心原则。第二，武装冲突中的性别差异根深蒂固，人工智能却会将复杂的现实简化为标签与概率输出。安全预测系统优先关注国家层面指标，忽视女性领导的非正式调解网络；自动化监控工具脱离本土政治语境，常将正常的线上组织活动误判为可疑行为。问题不仅在于算法偏见，更在于机构过度依赖追求标准化、规模化的智能系统，忽视对复杂社会现实的深度研判与人文关怀。作者强调，人工智能可在人工监管下辅助和平建设工作，但风险在于各类机构用自动化输出替代政治判断。文章认为，必须保障女性掌握人工智能监督中的技术与制度决策的权利；人工智能部署须接受性别敏感的风险评估，多边组织须明确人工智能应用的权责红线，杜绝问责模糊、责任缺位的治理乱象。

<https://www.diplomaticcourier.com/posts/women-peace-security-missing-core-ai-governance>

编译：张方琼宇

8、美国 CFR：联合国人工智能治理全球对话或重塑全球人工智能治理格局

5 月 29 日，美国对外关系委员会（CFR）发表人工智能研究员托尼·奥维克（Tony Oweke）的文章《世界各国都在

试图管控人工智能，联合国希望参与其中》。文章认为，联合国人工智能治理全球对话源自 2024 年《全球数字契约》及相关联合国谈判进程，人工智能独立国际科学小组亦由该谈判设立。该对话旨在检验多边体系能否建立合法、包容且协调的人工智能监管框架。文章强调，联合国对话的核心价值在于提供合法性与协调机制。通过汇集成员国、科技企业及民间社会主体，联合国为各国提供共同规则制定的平台，以应对技术与治理模式碎片化问题，降低跨境人工智能治理风险，明确国际人工智能治理优先事项，以及推动《全球数字契约》修订或启动国际法律文书制定进程。如果联合国相关机制未能发挥作用，人工智能治理将继续由少数多边俱乐部主导。各国基于不同监管理念和发展诉求出台的政策将加剧治理碎片化，形成国际治理规则竞争。联合国人工智能治理全球对话的成败，将决定未来人工智能治理的核心命题，治理的关键不再是“是否监管”，而是“谁来监管以及规则服务于谁”。对话的实施和成员国接受情况，将在很大程度上影响全球人工智能治理的合法性、包容性与跨境协调水平。

<https://www.cfr.org/interactive/global-ai-dialogue>

编译：高美琳

9、卡内基国际和平基金会：美国仍陷入政治暴力困境

6 月 3 日，卡内基国际和平基金会发表政治暴力研究团队研究员雷切尔·克莱因菲尔德（Rachel Kleinfeld）和达利

娅·伯科维茨（Dalya Berkowitz）的文章《美国的政治暴力》。文章认为，尽管近年来美国普通凶杀案件总体呈下降趋势，但与政治暴力相关的威胁和事件仍维持较高水平，对美国社会治理、公共服务体系和政治生态产生持续影响。报告将政治暴力界定为旨在破坏公共机构和民主运作、针对政治人物及公共官员、或试图排斥特定群体政治、经济和社会参与的暴力、威胁及恐吓行为。针对联邦法官、国会议员和地方民选官员的威胁案件持续增加，网络骚扰、人肉搜索和虚假报警等行为也有所上升。受此影响，部分公职人员减少公开活动、放弃竞选连任或选择离开公共岗位。当前美国政治暴力已难以简单归因于单一意识形态。历史上，美国不同时期的政治暴力分别与右翼、左翼及宗教极端主义有关。近年来右翼政治暴力曾显著增加，而 2024 年总统选举后，左翼政治暴力也出现上升迹象。同时，部分施暴者同时持有多多种意识形态立场，青年群体对政治暴力的接受度高于年长群体，对人工智能、移民等议题的焦虑以及对政治制度的不满，也成为新的风险因素。持续增长的威胁和骚扰正削弱美国公共治理能力。调查显示，不少地方民选官员因安全担忧而降低参与公共事务的意愿。政治领导人应公开反对政治暴力，媒体应避免强化群体对立叙事，并加强青年群体教育和社区建设，以降低激进化风险。同时，应加强对公职人员及其家属人身安全保护，减少政治暴力对公共治理体系的冲击。

<https://carnegieendowment.org/posts/2026/06/political-violence->

in-the-us

编译：罗方婕

10、美国进步中心：对伊战争错误正重演

6月2日，美国进步中心（CAP）发表其高级研究员安德鲁·米勒（Andrew Miller）的文章《特朗普的对伊战争正重演伊拉克战争的错误》。文章认为，特朗普政府发动的伊朗战争，不仅未实现既定战略目标，反而损害美国利益、改变中东力量格局并削弱美国全球领导地位。首先，特朗普发动的是一场失败的战争。战争已持续十三周，而美方最初提出的摧毁伊朗核计划、导弹能力、海军力量、地区代理人网络以及推动政权更迭五项目标，均未实现。其次，失败给美国人民和美国利益带来高昂代价。霍尔木兹海峡的封锁导致能源价格飙升、供应链受扰和通胀压力上升，已使美国家庭额外承担数百亿美元成本。同时，美军人员伤亡以及军事资源消耗，削弱美国应对战略竞争对手的优势。而且，伊朗以海峡控制权为基础的战略主动，推动中东力量平衡向伊朗倾斜，冲击美国盟友对美国安全承诺和能力的信心，损害美欧及亚太盟友体系。文章最后强调，即使美伊实现停火，战争造成的人员伤亡、军力消耗、地区失衡和国际信誉损失都难以逆转。与2003年伊拉克战争类似，这场本可避免的战争重塑了地区格局和全球贸易环境，并将进一步削弱美国国际领导地位。

<https://www.americanprogress.org/article/trumps-failed-war-in-iran-echoes-bushs-iraq-blunder/>

编译：贺宇希

11、美国布鲁金斯学会：非裔与拉美裔选民或在中期选举改变立场

6月1日，美国布鲁金斯学会（Brookings）发表“种族、繁荣与包容倡议”（RPII）学者基恩·吉尔伯特（Keon L. Gilbert）等人的文章《非裔和拉美裔选民在中期选举前面临负担能力差距》。文章认为，随着2026年中期选举临近，通胀高企与工资停滞导致的生活成本危机已成为核心议题。今年4月美国通胀率升至3.8%，约76%的民众将当前经济状况评为“平庸或糟糕”。其中，低收入和中产阶级阶层、尤其是黑人和拉美裔社区正承受最大制度化压力。特朗普政府推动的裁员政策和对社会福利的削弱政策，以及中东地缘政治冲突导致的汽油价格暴涨，对少数族裔中产阶级及女性就业造成了不成比例的负面冲击。2024年曾因经济诉求倒向共和党和特朗普的拉美裔与非裔选民正在改变立场，最新民调显示，高达65%的拉美裔选民认为共和党未能聚焦经济改善，若立即投票，民主党将以52%对28%的绝对优势赢回该群体。文章强调，社会阶层与种族身份决定了选民对经济的真实体感，中期选举的国会控制权归属将完全取决于哪个党派能为少数族裔提供切实解决日常务实生活成本的有效方案。

<https://www.brookings.edu/articles/black-and-latino-voters-face-an-affordability-gap-before-the-midterms/>

编译：陈奕好

12、美国 CSIS：军用人工智能系统“Maven”的含义及效果

6月2日，美国战略与国际问题研究中心（CSIS）发表研究助理马特·曼德（Matt Mande）和“决策树研究公司”（Decision Tree Research）的创始人兼首席执行官格雷戈里·艾伦（Gregory C. Allen）合著的文章《“Maven 智能系统”（Maven Smart System）是什么？它能做什么？》。文章认为，过去三个原因导致美国国防部快速打击正确目标的能力不足：一是分析员人数跟不上无人机收集的全天候影像数量，导致一些影像完全不被审查；二是用于目标打击决策的各类传感器难以互通，整合数据源相当耗时；三是目标识别与武器使用难以有效连接。2017年，美国国防部发布“Maven 计划”（Project Maven），属于国防部“联盟联合全域指挥控制”（CJADC2）能力的核心组成部分。在该计划中，Maven 系统是聚合并可视化海量侦察数据、为军事情报和目标打击提供实时战场视图的 AI 软件平台，已经被各战区及北约盟友使用，用户数超过 8 万。在该系统中，操作人员可以选择由 AI 识别出的目标、查看附近可用的资源、比较不同打击方案并最终下达打击命令。Anthropic 开发的 Claude 等产品一直

为 Maven 系统提供支持。但 2025 年 7 月以来，Anthropic 拒绝让国防部在合同中加入将 Claude 用于“任何合法用途”的条款，二者关系恶化。国防部开始使用 OpenAI 等其他前沿 AI 公司的产品，确保供应商多元化。未来 Maven 系统能够发挥的作用，取决于国防部能否解决指导框架不足、数据规则限制等内部问题。

<https://www.csis.org/analysis/what-maven-smart-system-and-what-does-it-do>

编译：高靓瑜

13、美国哈德逊研究所：特朗普应持续对伊施压

5 月 29 日，美国哈德逊研究所发表高级研究员丽贝卡·海因里希斯（Rebeccah Heinrichs）的文章《伊朗与前路：特朗普应善用自身优势，以实力谋求和平》。文章认为，美国军事行动虽削弱伊朗核计划、常规军力和国防工业基础，并加剧其内部权力裂痕。但美国在谈判中仍面临关键障碍，即伊朗权力分裂，政治代表未必能够约束伊斯兰革命卫队履行承诺。因此问题不仅在于能否达成协议，更在于协议能否持久。文章认为，美国当前最紧迫的任务是恢复霍尔木兹海峡航行自由，但不应以承认伊朗对海峡控制权为代价。如果美国以伊朗有限配合航运为条件提前放松制裁或削弱封锁，将削弱其彻底拆除伊朗核计划的筹码。美国应重启“自由行动”，主动清除海峡沿线威胁并护送商船通行，而不是等待伊朗放

行。文章认为，特朗普的目标并非仅限制高浓缩铀，而是永久性拆除整个核计划，彻底清除或不可逆转地剥夺伊朗铀储备。鉴于伊朗长期缺乏善意谈判记录，美国应推动严格且受国际监督的拆除程序；若无法实现，必要时应以军事手段切断其未来重新武器化的可能。文章最后强调，美国仍对伊朗握有显著军事和经济筹码，但仅有战场胜利不足以转化为持久优势。美国应维持最大压力，重启“自由行动”和联军护航，在伊朗仍具威胁能力时拒绝给予经济缓解，并将外交重点放在协调盟友上。特朗普既已展示动用武力的意愿，就应继续利用这种优势，通过“以实力促和”的方式迫使伊朗最终放弃核野心。

<https://www.hudson.org/national-security/iran-path-forward-trump-should-embrace-his-leverage-pursue-peace-through-strength-rebeccah-heinrichs>

编译：刘晓

14、美国《外交政策》：特朗普打击船只的正当性遭质疑

6月2日，美国《外交政策》网站发表其特约撰稿人约翰·哈尔蒂万格（John Haltiwanger）的文章《两名参议员彻底戳穿了特朗普的船只袭击辩护理由》。文章认为，美国参议员蒂姆·凯恩（Tim Kaine）和兰德·保罗（Rand Paul）指控，特朗普政府数月来在加勒比海和东太平洋对“涉毒船只”实施的军事打击目标判定标准并不包括毒品或武器的存在，

这引发外界对其行动合法性的质疑。凯恩表示，他对“存在毒品并非目标标准”感到奇怪，因为政府总是宣称打击的是“毒贩”。而国务卿兼国家安全顾问鲁比奥回应自己未参与目标标准的讨论，并称这主要是法律决策。此外，一位熟悉此事的参议院消息人士表示：“毒品存在证据未被列明，这与政府的公开声明及其对‘南方之矛行动’的宣传口径存在严重矛盾”。白宫、国防部和司法部法律顾问办公室（OLC）均未立即回应置评请求。文章认为，凯恩一直是针对所谓毒品船袭击行动最直言不讳的批评者，他曾谴责这些行动不具备合法性。法律与反恐专家也表示此行动违背国内法和国际法，且不应将毒贩视为恐怖分子。同时，共和党参议员保罗在听证会上披露美军炸船的三项秘密判定标准均不核查涉事船舶是否载有毒品或武器，此前他曾就未公开反对此类袭击行为批评共和党同僚。

<https://foreignpolicy.com/2026/06/02/kaine-paul-rubio-trump-drug-boat-strike-justification-latin-america/>

编译：李泓静

15、美国《国家利益》：对古巴战略不清或致区域危机

6月2日，美国《国家利益》网站发表美国发展外交专家史蒂文·亨德里克斯（Steven Hendrix）的文章《为什么美国需要清晰的对古巴战略》。文章认为，5月14日美国中央情报局局长访问古巴，传递了特朗普愿意接触的信号，但前

提是古巴做出“根本性改变”。但5月20日，美国司法部却起诉前古巴领导人劳尔·卡斯特罗（Raúl Castro）等6人，指控其1996年击落两架民用飞机，罪名包括谋杀美国公民。这一反差暴露出美国对古巴政策的内在矛盾。当前古巴已陷入严重的燃料和电力危机，柴油和燃油耗尽，停电频发，汽油价格翻倍。由于古巴距佛罗里达仅90英里，一旦国家失灵，移民潮、公共卫生风险及外部势力介入可能将政治危机升级为区域性危机。文章强调，美国需要清晰的战略：如果目标是推动改革，就应创造激励而非单纯施压；如果目标是安全合作或政权更迭，就应坦诚说明并做好人道和区域后果的预案。目前政策混杂了压力、接触、刑事起诉和公开喊话，可能制造混乱。起诉劳尔·卡斯特罗虽具法律意义，但政治上已成重大升级，被形容为美古关系的“新低”。没有计划的压力不是战略，没有筹码的接触也不是战略。华盛顿正走在一条狭窄而危险的道路上。

<https://nationalinterest.org/feature/why-the-us-needs-a-clear-cub-a-strategy>

编译：白一冰

编译：罗方婕、贺宇希、陈奕好、高靓瑜、刘晓、李泓静、白一冰、程果、孙韵雯、钟珺雯、张朔宁、杜宛鸿、张佳延、张方琼宇、高美琳

审核：刘源、郑乐锋、周武华、申青青