



CISS 海外信息专报

(2024年10月31-11月1日)

1、美国《法律事务》杂志：应考虑制定具有约束力的自主武器规则

10月28日，美国《法律事务》杂志刊登南卡罗来纳大学法学院助理教授查理·特朗布尔（Charlie Trumbull）的评论文章《考虑制定具有法律约束力的自主武器文书》。文章认为，尽管现有国际人道法对自主武器系统有一定规制，但技术发展带来的复杂性急需新的法律框架。作者提出三种可能的路径以供讨论。首先，禁止特定类型的自主武器。因为自主武器的威胁来自其使用方式和技术水平，而非自主性本身，故难以完全禁止自主武器。作者认为可以效仿《反人员地雷公约》，禁止某些类别的自主武器，特别是那些直接以人类为目标的武器。其次，确定“有意义的人类控制”为基本义务。该义务要求自主武器在使用过程中保持人类控制，以便符合国际人道法。虽然美国等国家对此有异议，认为这一标准过于模糊且不适用于所有自主武器。但该义务可以确保人类在武器使用决策中的控制和判断得到广泛支持，尤其是避免无差别攻击。最后，严格遵守审慎义务（due diligence），实施严格的武器审查、风险评估和人员培训。这种程序性措

施有助于确保武器符合国际法要求。然而，一些国家担心，这样的程序可能会变相合法化自主武器的使用，因此希望施加更严格限制。

<https://www.lawfaremedia.org/article/considering-a-legally-binding-instrument-on-autonomous-weapons>

编译：高隆绪

2、科技政策新闻：人工智能安全研究所如何影响人工智能治理

10月25日，科技政策新闻项目经理普里特维·艾耶尔（Prithvi Iyer）发表报告《从安全到创新：人工智能安全研究所如何影响人工智能治理》。报告提出，随着人工智能迅速发展，日本、英国、美国、欧盟、新加坡、法国、加拿大等七个司法管辖区纷纷建立人工智能安全研究所（AISI），印度也在进行公开征求意见准备设立类似机构，以加强对人工智能的治理。报告分析了 AISI 在各地的特点、功能和局限性，揭示其在全球 AI 治理中的重要性和挑战。其中，英国、美国和日本的 AISI 构成了“第一波”机构，这些机构在安全导向、政府主导和技术性方面具有共性。安全导向方面，AISI 的目标是确保人工智能的设计、开发和使用符合人本、安全和可信赖的原则。政府主导性为 AISI 提供了权威和资源，使其能进行 AI 模型评估并在与企业的谈判中具有影响力。技术性则体现在 AISI 吸引技术专家，以实证方法应对 AI 安全

问题。报告指出，这些机构大多没有直接的监管权力，主要通过制定安全评估标准和技术准则，对行业实践产生了影响。AISI 的核心职能包括研究、标准制定和跨国合作。研究方面，AISI 推动“人工智能安全科学”的发展，进行安全评估并公开其测试框架。标准方面，AISI 制定评估基准和安全准则，推动行业对安全性的重视。其中，美国和欧盟在标准设定方面的计划最为详细，美国国家标准与技术研究院（NIST）发布的《人工智能风险管理框架》（RMF）影响了多国标准。合作方面，AISI 在政府、产业和学术机构间搭建桥梁，促进全球 AI 安全合作，美国还发起“国际人工智能安全研究所网络”，进一步强化跨国合作。报告认为，尽管 AISI 在负责任的 AI 发展中起到关键作用，但其局限性同样明显，其对“安全”的狭隘专注可能忽略其他 AI 风险并阻碍创新。此外，AISI 与其他已有机构可能存在职能重叠，引发效率问题。与行业关系密切的 AISI 也可能面临“监管俘获”风险，影响其独立性。未来的“第二波” AISI，尤其是发展中国家的机构，或许会在安全基础上更加关注 AI 创新和潜在益处。

<https://www.techpolicy.press/from-safety-to-innovation-how-ai-safety-institutes-inform-ai-governance/>

编译：高行健

3、美国布鲁金斯学会：AI 数据标注工作者困境与全球数据标注生态重塑

10月24日，美国布鲁金斯学会发表技术创新中心研究员奇纳萨·奥科洛（Chinasa T. Okolo）和斯坦福大学技术伦理政策玛丽·塔诺(Marie Tano)的文章《迈向全球 AI 市场真正担责的人工智能之路》，聚焦负责任 AI 发展中数据标注劳动问题。文章认为，在全球 AI 市场中，推动负责任的 AI 发展至关重要。然而，当前大型 AI 公司常将数据集标注工作外包至非洲等地区，导致数据工人面临低薪、福利匮乏及长时间工作等困境，且常接触敏感内容。为改善这一现状，文章呼吁美国改革隐私和劳务外包法律，企业投资本地项目，确保数据工人尊严与公平待遇，避免殖民剥削重演。文章认为，AI 公司掩盖数据工作中人力重要性，透明度低、对数据工人协作地位认可不足。数据工人对 AI 模型训练至关重要，涉及数据收集、准备、标注和验证等多个环节，需进行干预以确保理解的准确性。文章警示，全球数据标注市场将因对数据集需求增长而扩大，但 AI 与数据实践、偏见及气候影响引发的伦理问题日益凸显，或加剧发展中国家不平等。此外，西方理念强加标注准则，忽视工人文化背景，导致文化敏感性标注被低估。为此，需重视数据工人贡献，调整 AI 工具及标注指南以符合数据工作者所在地区的社会文化规范和价值观，避免西方中心化。为改善全球数据标注生态，文章建议实施严格劳工法规，建立公平工资、工作保障及条件标准，支持数据工作者倡议，建立工会；完善数据隐私保护规定，以更好地保护国内外数据工作者。最后，文章提出，

在 AI 发展中，人类特质愈发重要，公司应将标注者视为合作者，共同促进 AI 开发的包容性和伦理性。

<https://www.brookings.edu/articles/moving-toward-truly-responsible-ai-development-in-the-global-ai-market/>

编译：孙叶秋

4、美国《外交事务》：AI 外交新时代下的美国对海湾国家战略选择

10 月 28 日，美国《外交事务》刊登卡内基国际和平基金会技术与国际事务研究萨姆·温特-利维 (Sam Winter-Levy) 的文章《AI 外交新时代下，美国如何调整对海湾国家的合作以取得对华竞争优势》。文章阐述了本文探讨了在人工智能 (AI) 推动全球联盟和竞争格局转变的背景下，美国在海湾地区面临的战略机遇与挑战。目前，海湾国家为了摆脱石油依赖、构建现代经济，正在积极寻求国际 AI 合作，以 AI 发展促进本国的经济转型。在海湾国家寻求 AI 合作的背景下，美国试图利用技术优势与中国在这一地区开展竞争。然而，中东国家与中国的深厚经济关系增加了美国拉拢这些国家的难度。同时，海湾国家可能故意利用中美竞争获取更大利益。基于此，作者建议美国在 AI 外交新纪元中需要小心应对中美影响重叠的区域，既要通过技术合作巩固关系也要在与海湾国家的合作中设置边界，保护知识产权和战略技术安全，并且避免逼迫这些国家在中美之间做出单一选择。

<https://www.foreignaffairs.com/united-states/emerging-age-ai-diplomacy>

编译：赵英慧

5、英国《计算机周刊》：人工智能治理的未来之路

10月28日，英国《计算机周刊》网站刊登英国史密夫斐尔律师事务所合伙人米里亚姆·埃弗雷特（Miriam Everett）的文章《人工智能治理的未来之路》。文章认为，在数字监管和 AI 领域，个人数据、数据保护权利以及数据的广泛使用正面临日益严峻的法律挑战。目前英国政府尚未出台新的数据使用和访问法案，这反映出全球在数据和 AI 监管方面尚未形成统一有序的格局。《通用数据保护条例》虽然是个人数据保护的标志性事件，但并非监管起点。实际上，英国在 GDPR 之前就已经通过 1998 年数据保护法。尽管 GDPR 的实施并没有彻底改变数据保护的既有权利和义务，它确实提高了对违规公司的处罚力度，并将数据保护问题提升到了全球关注的层面。自 2018 年以来，多国已出台受欧洲个人数据立法模式启发的数据保护法律。虽然这些法律在细节上有所不同，但普遍基于相似的原则、权利和义务。欧盟在数据保护法规的制定方面发挥了领导作用，但目前还无法确定其 AI 法案是否能继续在全球范围内产生类似的影响。欧洲最近通过的 AI 法案显示了其在全球 AI 监管中寻求领导地位的意图，还需进一步观察该法案能否在全球范围内产生预期

效果。不同国家和地区在法律框架上存在显著差异，美国更倾向于公司驱动的监管模式，中国以国家为主导，而欧洲则更注重权利驱动的方法。这些差异暴露出一个更广泛的问题：监管应当基于风险还是结果，以及应如何确定其严格程度。现有的法律已经能够处理数据保护等问题，但是否需要为新技术制定新法律仍待探讨。

<https://www.computerweekly.com/opinion/AI-governance-mapping-the-road-ahead>

编译：刘嘉滨

6、美国 Politico：拜登政府采取大刀阔斧的人工智能政策

美国 Politico 网站 10 月 24 日发布该网站记者莫哈·查特吉（Mohar Chatterjee）和约瑟夫·戈迪昂（Joseph Gedeon）的文章《拜登在人工智能领域采取大刀阔斧的政策，并设下了政治陷阱》。该文聚焦拜登政府近期发布的人工智能（AI）国家安全备忘录，作者认为拜登政府计划以以协调性政策推动 AI 技术在国家安全领域的应用。备忘录要求政府机构在遵循民主价值观的条件下，利用 AI 服务于国家安全，同时符合民主价值观，关切 AI 芯片供应链，将其视为对抗试图窥探美国工业的竞争对手的“顶级情报优先事项”。在人才领域，拜登政府试图简化处理与敏感技术相关的非美国公民的签证流程，并呼吁国会为 STEM 专业的毕业生提供更多绿卡。在财政领域，拜登政府指示包括国防部和国家情报总监

办公室在内的安全机构成立一个工作组，处理采购 AI 相关的问题。它还指示国务院、能源部和商务部与国家安全机构一起对 AI 技术进行公共投资。拜登政府试图通过该备忘录彰显其在公民权利保护方面做出显著努力：备忘录总共提到了“人权”22 次，并要求关键联邦机构监测其 AI 使用的风险。但一些民权活动家认为，这些报告很大程度上对公众是隐藏的。本年度大选的结果将对备忘录产生重要影响，若卡玛拉·哈里斯获胜，备忘录中强调的 AI 在安全领域的应用可能会引发更广泛的社会矛盾，因为民权组织已批评该备忘录可能使安全机构加强监控国家。而特朗普则誓言要废除该备忘录，如果他赢得大选将使备忘录的政治地位变得脆弱。

<https://www.politico.com/news/2024/10/24/biden-ai-policy-national-security-00185407>

编译：许泽楷

7、美国大西洋理事会：科技公司对核能展现出新的浓厚兴趣

10 月 29 日，美国大西洋理事会发布其全球能源中心核能政策倡议主任詹妮弗·戈登（Jennifer T. Gordon）和副主任劳伦·休斯（Lauren Hughes）的文章《科技公司对核能展现出新的浓厚兴趣——原因何在》。文章认为，随着数据中心和生成式人工智能等业务的迅速扩展，科技公司对稳定、可靠且无碳排的电力供应需求急剧上升。相比于其他清洁能

源，核能可以持续、大规模供电，且满足科技公司对全天候运营的高可靠性的需求。因此，亚马逊和谷歌等公司投资先进核能技术，旨在满足当前和未来电力需求。这些科技公司、先进核反应堆开发商以及公用事业公司建立合作关系，共同分担前期巨额资金投入与技术开发过程中的风险。目前，亚马逊与X-energy合作进行 5 千兆瓦核能项目开发，谷歌则与凯罗斯电力公司合作开发 500 兆瓦的核能项目，美国能源部也提供 9 亿美元以支持第三代小型模块化反应堆的初期部署，这些合作和政策支持推动了新一代核能的商业化。新一代核能技术具有更小规模和灵活性，适用于离网和偏远地区的发电需求。除了推动新技术应用，科技公司还依托现有的核能设施，以实现其长期去碳化战略。微软与康斯特雷特电力签署了 20 年的电力采购协议，确保其在数据中心的能源需求得到满足。新一代核能技术具有更小规模和灵活性，适用于离网和偏远地区的发电需求。文章最后认为，尽管科技公司的投资在推动核能技术发展方面发挥显著作用，但面对巨额前期投资和技术难题时仍需依靠多方合作。核电技术的发展将迅速降低新建反应堆的成本，推动清洁能源转型的进程。。

<https://www.atlanticcouncil.org/blogs/new-atlanticist/tech-companies-are-showing-a-new-strong-interest-in-nuclear-power-heres-why/>

编译：张振

8、美国布鲁金斯学会：美国选举之夜的虚假信息风险

10月28日，美国布鲁金斯学会网站发布其政府研究项目副主席兼主任、科技创新中心创始主任达雷尔·韦斯特（Darrell M. West）撰写的文章《美国选举之夜的虚假信息风险》。文章认为，在美国选举之夜临近之际，有关邮寄选票、无证移民选民、贿选、选票作弊、选票统计和胜选指控的虚假指控正动摇选民政治信念，影响选举公正性。可能会出现虚假信息风险如下：其一，宾夕法尼亚州邮寄选票再次面临舞弊指控。其二，计票缓慢被解读为选票舞弊证据。据预测，2024年美国大选投票率可能会很高，这将导致包括宾夕法尼亚州在内的有争议地区的计票速度可能很慢。如果选票需要几天时间才能处理，双方都很容易散布虚假言论。其三，选举之夜未产生获胜者将被视为选票舞弊证据。选举结果悬而未决很可能滋生偏执，认为对手窃取选举结果。其四，声称大量无证移民在关键摇摆州投票。选举期间，共和党人长期兜售民主党支持开放边境的说辞，以便让“非法移民”投票支持民主党。其五，非裔美国人和拉丁裔选区存在选票作弊或贿选的指控。种族和族裔是大选关键分界点，但当前民调显示哈里斯表现不佳。如果她在选举之夜表现出色，在关键州非裔美国人和拉丁裔选区可能会出现选票作弊或买票指控。其六，过早宣称胜利。在竞争激烈的选举中，候选人有时会在获得独立客观获胜证据之前，出于政治动机宣称自己获胜。文章认为，为应对选举之夜及之后可能出现的大量

虚假指控，记者、候选人、意见领袖和公众需要充分意识到选举之夜虚假信息指控的可能性很高，了解对选举结果提出质疑的具体方式，“预先驳斥”和“揭穿””虚假指控。两党应在关键战场州部署大量观察员，以确保选举过程公平公开。

<https://www.brookings.edu/articles/election-night-disinformation-risks/>

编译：宋琳琳

9、美国布鲁金斯学会：美国选举中的舞弊现象并不普遍

10月28日，美国布鲁金斯学会刊登有效公共管理中心创始主任、治理研究高级研究员伊莱恩·卡马克（Elaine Kamarck）文章《美国选举中的舞弊现象并不普遍》。文章认为，在2024大选冲刺阶段，两党均充斥部分虚假信息，其中最典型的便是特朗普煽动的“选举舞弊论”。这使得民众对本次选举的认可程度大打折扣。但作者认为，大量所谓“选举舞弊”的证据均系刻意伪造或渲染而成，不能证明当前选举中有严重的舞弊现象。例如，右翼智库传统基金会所提供的选举舞弊者姓名等信息则是从数十年前大选中刻意搜集的，且这些舞弊行为均因数量过少而不能影响彼时的选举。美国选举制度中也具有防止舞弊的制度设计，例如选举计票现场设置两党的监票员等，这大大降低了选举舞弊的可能。因此，作者认为，尽管大选中的舞弊现象不能避免，但

制度设计等可以使舞弊大大降低。

<https://www.brookings.edu/articles/how-widespread-is-election-fraud-in-the-united-states-not-very/>

编译：张成昊

10、美国《外交事务》：人工智能外交新时代

10月28日，美国《外交事务》网站刊登卡内基国际和平基金会技术和国际事务研究员山姆·温特-利维（Sam Winter-Levy）的文章《人工智能外交新时代：为与中国竞争，美国必须在海湾地区走钢丝》。文章认为，沙特阿拉伯与阿联酋正试图通过人工智能推动其转向新经济模式，摆脱石油依赖。在两国加紧投资人工智能背景下，美国科技企业希望利用两国充足的能源、资本和土地以训练大模型。由于先进芯片出口需要得到美国政府的批准，美部分官员认为，这为美国诱使海湾国家远离中国、加强中东地区反伊朗联盟提供了宝贵机会。但是，目前沙特与阿联酋同中国的政经关系空前牢固，且未来不太可能恶化。它们既期待美国的先进芯片，也希望在中美竞争间战略对冲。同时，对中东出口芯片还可能使中国将能获得美国对中东出口的芯片。作为关键新兴技术，训练托管大模型的数据中心也不应被设置在美国境外。作者认为，即便如此，美国仍应同海湾国家开展人工智能合作，防止潜在的技术扩散，保证美国企业利益，同时不应期待这会使中东国家远离中国。

<https://www.foreignaffairs.com/united-states/emerging-age-ai-diplomacy>

编译：熊梦溪

11、美国《外交事务》：虚假信息对美国的民主的影响

10月29日，美国《外交事务》网站刊登美国国防情报局前代理局长大卫·谢德（David R. Shedd）和保卫民主基金会研究员伊万娜·斯特拉德纳（Ivana Stragner）的文章《美国人思想的隐蔽战争：俄罗斯、中国和伊朗如何寻求在美国传播虚假信息和混乱》。文章认为，俄罗斯通过在社交媒体散布关于堕胎、拥有枪支的权利、移民和美国对乌克兰的支持等美国政治热点问题的煽动性信息，试图加深美国国内两极分化，削弱美国主导地位，并重新建立俄全球大国地位。伊朗也试图通过在美国人中煽动动荡来影响美国选举，但其更大目标是破坏美国全球霸权。中国则更关注下方票竞选（down-ballot races），试图攻击对中国持批评态度的共和党人。文章认为，各国影响力行动的目标是通过塑造公民观点来改变美国决策，这些行动的影响是长期的，将在目标人群中播下混乱、不满和怀疑的种子。作者建议，一是美国应加强政府间协调来对抗外国影响力行动，国家安全委员会需要设计一个全政府行动计划，揭露美国对手干预选举的情况，并制定具体的对策。二是应建立一个由政府和社交媒体平台高层领导人组成的联合工作组，确定具有全球影响力的社交

媒体平台如何帮助打击破坏民主制度的虚假信息，揭露其对手如何利用美国的言论自由来传播旨在破坏美国民主的虚假信息。三是美国还应该再次发起自己的影响力行动，在敌对国家传播自己的叙事。

<https://www.foreignaffairs.com/united-states/covert-war-american-minds#author-info>

编译：李瑾菡

12、美国大西洋理事会：特朗普或哈里斯上任对非洲的影响

10月29日，美国大西洋理事会分别刊登前明尼苏达州美国众议员、前共和党总统竞选顾问维恩·韦伯（Vin Weber）和大西洋理事会非洲中心副主任本杰明·莫斯伯格（Benjamin Mossberg）的文章《特朗普第二任期内非洲将面临什么》和《哈里斯政府领导下的非洲：重振战略》。韦伯认为，无论谁成为下一任美国总统，都将认可美国与非洲经济互动的必要性，以对抗中国在非洲大陆的大规模存在。中美在非洲大陆的竞争将进一步加剧，尤其是美国日益关注的关键矿产及其供应链。韦伯表示自己对特朗普第二任期内美国与非洲关系的改善充满希望。如果特朗普在11月再次赢得总统大选，他的第二任期或将尽可能多地释放年轻人活力，并且支持与非洲大陆发展伙伴关系。非洲对第二任期特朗普政府而言比第一任期更为重要，因此特朗普可能将加强美国与非洲大陆的贸易，并于2025年重新授权《非洲增长与机会法案》。

此外，特朗普第二任期政府可能会通过加强美国非洲司令部和协助非洲反恐努力，强力应对恐怖势力对非洲的威胁。莫斯伯格认为如果哈里斯在此次当选总统，将很有可能遵循现任政府的政策。哈里斯拥有丰富的非洲经验，她在支持拜登政府的非洲议程方面发挥了重要作用。哈里斯将其竞选活动定性为“为自由而战”，这表明她可能会推行一项专注于在全球范围内推广民主理想、制度和规范的外交政策，并且更有力地支持人权和反腐败努力等问题，哈里斯或将加强 2021 年 12 月发布的《美国反腐败战略》，进一步增强行政机构打击非洲腐败的能力。

<https://www.atlanticcouncil.org/blogs/africasource/what-africa-can-expect-under-a-second-trump-administration-a-focus-on-the-numbers/>

<https://www.atlanticcouncil.org/blogs/africasource/what-africa-can-expect-under-a-kamala-harris-administration-a-reinvigorated-strategy/>

编译：张思薇

13、美国《外交事务》：美国如何在多盟世界中取得成功

10 月 30 日，美国《外交事务》网站刊登美国智库新美国（New America）首席执行官、前国务院政策规划司司长安尼-玛丽·斯劳特（Anne-Marie Slaughter）的分析文章《美国如何在多元结盟世界中取得成功——建立真正全球伙伴

关系的重要性》。她回顾了希拉里·克林顿提出的“多方伙伴”（multipartner）愿景，认为美国应通过促进合作而非竞争来引导全球事务，合作对象不限于国家，还应包括企业、民间组织等多种主体，以应对跨国挑战。拜登政府虽重启并扩展了传统联盟及新兴安全伙伴关系，但在处理全球议题时仍倾向地缘政治竞争。作者建议，未来的美国政府应拓宽合作对象，聚焦气候变化等全球威胁，接受去中心化的合作模式。在具体策略上，斯劳特认为美国应在国际合作扮演“播种”（seed）角色，而非一味追求领导（lead）。美国应通过支持多方主体的合作机会，构建可靠的合作网络以促成成果，建立伙伴关系的“产出和结果合法性”（output and outcome legitimacy），即通过成效证明这些团体和联盟能够实现其既定目标。总体而言，诸多国家采取“多边结盟”（multialigned）政策，而不在美中之间选边站队的态势反而有助于增加美国影响力。美国应采取在多元主体中协作的外交策略，以增强其在全球事务中的连通性与灵活性。

<https://www.foreignaffairs.com/united-states/how-america-can-succeed-multialigned-world>

编译：陈露娴

14、美国布鲁金斯学会：中印签署协议并未解决其根本矛盾

10月29日，美国布鲁金斯学会网站发表印度问题专家坦维·马丹（Tanvi Madan）的文章《印中签署协议后成为朋

友了吗？》。文章认为，中印喀山会晤虽标志着某种战略和解，但未改变中印根本矛盾。文章表示，过去四年中印两国边境对峙问题持续，当下中国面临战略和经济压力，印度需要时间建设伙伴关系和自身能力，且双方均关心美国大选后的对外政策，都希望防止边境局势再次升级，一定程度上促成了此次缓和。重要磋商恢复和具体巡逻安排等协议内容尚在逐步显现，但是否建立新缓冲区、减少现有部队等问题仍然存在，印度官员认为不能“陷入虚假的安全感”，并对协议的实质性保持警惕。不过在边界问题上中方强调广泛关系优先，而印方认为边界和平是建立更广泛关系的前提；经济问题上，印企内部对与中国建立经济关系的必要性和程度仍存在分歧，全面解除对华限制的可能性较小，短期内仍将坚持审查并设定合作条件。总体而言，作者认为此次协议的达成是双方多轮军事及外交磋商的结果，不应过度解读。于印方而言，是为了通过和缓中印局势，继续强化自身能力和韧性，以应对更广泛的全球经济挑战以及美国方面的不确定性，确保不在未来妥协其国家利益。

<https://www.brookings.edu/articles/has-india-made-friends-with-china-after-the-modi-xi-agreement/>

编译：孙钰涵

编译：宋琳琳、孙钰涵、熊梦溪、张思薇、张成昊、陈露娴、
李瑾菡、刘嘉滨、赵英慧、张振、高隆绪、高行健、孙叶秋、
许泽楷

审核：刘源、郑乐锋、周武华、申青青