



双周政策分析简报

(2026 年第 5 期 总第 29 期)

中国人民大学国家发展与战略研究院（国家高端智库）

2026 年 3 月 19 日

人工智能军事领域应用的企业伦理与公法边界

在近期美伊战争背景下，特朗普政府与 Anthropic 公司之间关于 AI 技术采购的冲突引发了一系列关于人工智能（AI）军事化应用所涉及的战略、法律及伦理问题，该事件也为我国提供了警示与启发。AI 在军事领域的应用已从根本上改变了传统战争形态，美伊战争中 AI 辅助的精准打击即是例证。美国自特朗普第一任期以来的 AI 军事战略目标已从保持“领导地位”升级确立为“全球垄断”。在此背景下，Anthropic 公司基于伦理拒签合同的行为，被视为一场“公民不服从”运动，但专家普遍认为，企业个体的伦理坚守难以阻挡国家主导的军事化浪潮，美国军方可通过替换供应商等方式绕过阻力，进而可能使 AI 军事化应用产生“双重失控”风险：一是总统权力膨胀导致国内制衡失效；二是 AI 系统的纯理性逻辑可能导致冲突意外升级，引发人道主义灾难。针对中国的处境，专家建议：在完善 AI 人才管理与国家征用规则的同时，从全球供应链合规的角度构建针对相关企业的负面清单，并积极参与塑造 AI 战争的国际治理规则，以应对技术、资本与政治深度交织的新挑战。

近期，由中国人民大学国家发展与战略研究院主办、中国人民大学法学院法律与全球化研究中心承办的双周政策分析会——“人工智能军事领域应用的企业伦理与公法边界”成功举办。会议由中国人民大学国家发展与战略研究院研究员、法学院教授张龔主持，与会专家代表围绕会议主题展开深入讨论。

一、AI 时代“王”容易成为首个被攻击的目标

张龔教授首先就会议背景及主题进行了介绍。他以特朗普政府与 Anthropic 公司在 AI 技术采购及征用上的冲突为切入点，指出此次会议旨在探讨人工智能在军事化运用过程中引发的战略与法律问题，尤其是其对《美国国家安全法》的影响及对中国的启示。在近期的“伊朗冲突”中，AI 赋能的信息系统实现了对高层目标的精准打击，这一变化颠覆了传统的战争思维。他以国际象棋作比喻，认为传统下棋的思维强调保护“王”，而在 AI 时代，“王”可能成为首个被攻击的目标，这要求我们重新审视 AI 军事化应用带来的根本性变革。由此，会议引出三个核心议题：一是企业伦理与国家安全冲突：以 Anthropic 公司为例，其 CEO 因伦理原因拒绝参与对美国公民的大规模监控项目与自动化武器项目，反映了科技企业在国家需求与自身价值观之间的张力；二是公法边界的适用性：借鉴纽伦堡审判中设立的反人类罪等罪名，探讨现行法律框架能否适用于 AI 参与的战争行为；

三是中国民营企业的处境：在美国对华制裁的背景下，中国民营企业若涉足 AI 军事领域，可能被列入美国军工企业管制清单，从而面临金融与生存风险，为此需重新审视企业与国家之间关系的公法边界。

二、在美国既定的 AI 军事化战略面前，企业的伦理红线难以形成实质性阻碍

香港中文大学名誉研究员江照信从《美国国家安全法》的专业视角出发，对 Anthropic 事件进行了系统性的政策与法律分析。他指出，要理解此次争议，首先需要放宽视野，回到特朗普第一任期内成立的“美国人工智能国家安全委员会”所确立的战略框架上来。该委员会在 2019 年中期报告中明确了七大原则，其中首要原则是确保美国在人工智能领域的全球领导地位，并将 AI 应用视为国家安全的迫切使命；2021 年的最终报告进一步提出，到 2025 年美国要实现 AI 在军事上的“随时可用”状态，并明确将赢得与中国的技术竞争作为核心目标。在此基础上，江照信梳理了特朗普第二任期以来发布的一系列关键政策文件：2025 年 1 月的行政命令将美国的 AI 目标从“领导”升级为“垄断”；同年 7 月出台的《赢得竞赛：美国人工智能行动计划》明确提出要在国际治理机制中对抗中国的影响力；2026 年 1 月美国国防部发布的 AI 战略备忘录则强调，AI 的军事化应用对美国未来十年至关重要，必然要求对私人企业进行国家征用。在法律

层面，江照信分析了可能涉及的几部法律：虽然 1950 年通过的《国防生产法》曾被提及但未被实际援引；美国国防部指控 Anthropic 公司的依据主要是 2021 年出台的《美国供应链行政令》，尽管该行政令此前仅针对华为等中国企业；此外，本事件还涉及《外国情报监视法》中关于大数据采集的合宪性争议。但他认为，从整体上看，此次事件本质上是一起合同争议，而非严格意义上的法律对抗。最终，他得出结论：在美国既定的 AI 军事化战略面前，企业的伦理红线难以形成实质性阻碍。Anthropic 公司的立场虽然体现了科技企业的价值观坚守，但美国军方完全可以通过替换供应商（如 OpenAI、xAI）或自行开发模型来绕开这一阻力。此事件折射出，人工智能的军事化已不可避免，且其核心目标之一就是针对中国的战略竞争，至于中国如何回应这一挑战则是需要深入思考的课题。

三、将 Anthropic 事件定性为“公民不服从”运动与“斯诺登式企业”的“伦理抵抗”

中央民族大学法学院副院长、教授田飞龙从战略高度对人工智能在军事领域的应用问题进行了系统分析。他以美以伊战争为切入点，指出 AI 在哈梅内伊被斩首事件中的应用，彻底颠覆了传统的战争思维，对大国关系、国际法和人权法构成了巨大冲击。针对 Anthropic 事件，田飞龙教授将其定性为一场“公民不服从”运动，认为 Anthropic 是一家秉持

独立价值观的“斯诺登式企业”，且其在面对国家强权时进行了伦理抵抗。然而他也强调，这种基于人文主义的抗争无法阻挡美国建立 AI 军事霸权的既定战略。特朗普政府对 Anthropic 的迅速替换与污名化，恰恰印证了企业伦理在面对国内政治与军事需求时的无力。随后，他重点剖析了美以伊战争所暴露的“双重失控”：一是总统权力的失控，特朗普完全绕开国会，以机会主义的方式发动战争，而国会对其的制约已形同虚设，美国正步入“帝王总统”时代；二是 AI 系统本身的失控，国外研究显示 AI 模型极易导致冲突意外升级，且 AI 作为纯理性存在，缺乏人类对痛苦、恐惧的感知，这使得由 AI 驱动战争可能走向非人道的极端。在应对层面，田飞龙提出两点思考：第一，尽管美国推行 AI 霸权，但其官方文件中仍保留了对核武器等极端手段的人类控制承诺，这为中俄等大国推动建立 AI 战争治理规则和全球公约提供了空间；第二，建议从全球供应链的合规角度入手，对积极配合美军“干脏活”的企业设立负面清单，通过国际商业和法律手段形成制裁与隔离机制，以此在无法直接对抗美国政府的现实下，为 AI 军事化运用设置底线。

四、AI 技术的特殊性在于战时与平时的界限模糊

中国政法大学副教授、北京市军事法学会秘书长刘瑞强聚焦于自主武器系统的国际法与伦理问题。首先，他指出，尽管联合国特定常规武器公约自 2012 年起就开始讨论自主

武器的法律规制，并于 2014 年设立政府专家小组，2019 年通过了指导原则，但这些原则较为空洞，缺乏约束力。此外，他介绍了自主武器系统的三大应用领域：情报分析、作战指挥和自主火力攻击，并以韩国在朝韩边境部署的 SGR-A1 武装机器人哨兵为例，说明其具备自动识别并攻击越境目标的能力。不过他也强调，目前绝大多数所谓“自主武器”仍属于“半自主”状态，最终决策仍需人类介入。在国际法层面，他引用北京理工大学学者的观点并指出，现有国际法对自主武器系统的审查存在明显不足，包括审查范围不明确、程序规则缺失、各国审查能力参差不齐、缺乏国际合作机制等。尽管《日内瓦公约第一附加议定书》第 36 条要求对新武器系统进行法律审查，但实践中往往流于形式。在责任归属方面，他区分了国家责任与个人责任：国家无论使用自主还是半自主武器，只要造成损害都需承担责任；而个人（如操作者、指挥官）的责任认定则涉及故意或过失，较为复杂；部分国外学者甚至提出应引入民事责任。针对民用征用问题，刘瑞强指出，中国在紧急状态下对民企资产的征用已存在明确规定，但 AI 技术的特殊性在于战时与平时的界限模糊，如伊朗在谈判过程中领导人突遭斩首，凸显了 AI 军事化对传统征用规则的挑战。最后他呼吁，国际社会应推动制定“禁止和限制使用自主武器系统公约”，以弥补现有法律框架的空白。他还以自动驾驶为例，说明 AI 系统难以区分人类与

普通障碍物，这种技术局限在军事领域可能引发严重的反人道主义危机。

五、中国需在 AI 技术对齐能力、产业保护机制、国际规则博弈等方面做出更深层次的制度准备

中国人民大学法学院助理教授刘洋主要围绕人工智能军事化技术应用的紧迫性、国际人道法的技术困境以及中美 AI 竞争背景下的国家能力与策略调整三方面展开论述。第一，关于 AI 在军事领域运用的应用场景与紧迫性。美国对中国人工智能进展的关注焦点之一是机器人的军事应用，尤其是地面部队的机器人化。然而，目前实际战场上出现的 AI 应用仍主要集中在情报分析、后勤运输等辅助领域，真正的“杀人机器人”尚未出现。美国之所以担忧中国在这一领域的发展，可能是在想象一种以大规模地面机器人为核心的新型战争场景，这与中国传统陆军优势相结合，或许会重塑未来作战形态。第二，关于国际人道法在 AI 军事化中的技术困境。他以“区别原则”和“比例原则”为例，说明将国际人道法规则编码化、让机器学习并遵守的极端复杂性。例如，战场上如何判断一名举白旗的士兵是否真正丧失作战能力？AI 在面对这类基于事实的模糊情境时极易出错。即便大国在军控规则上达成共识，技术层面的标准对齐仍面临巨大挑战。中美在签署国际人道法公约方面存在差异，如美国未签署部分附加议定书，这将进一步加大规则对齐的难度。

第三，关于中美 AI 竞争中的结构性难题。冷战时期美国高科技的发展模式是“军事先行、民用扩散”，国家掌握核心技术；而当前 AI 领域则是民用商业力量领先，国家能力相对后撤。这种结构倒置使得国家在推动 AI 军事化时必须依赖民营企业，而民企的市场、人才、资金高度全球化，难以与国家安全需求完全切割。对中国而言，这意味着必须回答一个根本问题：如果 AI 发展的核心力量在民间，国家如何为其提供足够大的国内及国际市场空间，使其在面临美国制裁时仍能生存与发展？刘洋强调，AI 军事化应用不仅是技术问题，更是涉及国际法、产业政策和国家安全战略的系统性挑战，中国需要在技术对齐能力、产业保护机制、国际规则博弈等方面做出更深层次的制度准备。

六、AI 军事领域无节制竞争可能会引发新一轮的军备竞赛

江西省委党校政法部教授张涛关注两个核心问题。一是人工智能的军事化应用已成为不可逆转的全球化趋势，美国等国正积极推动相关战略，若各国在 AI 军事领域展开无节制的竞争，可能会引发新一轮的军备竞赛，进而对全球安全构成严重威胁。因此，国际社会需要思考如何规范和控制 AI 军事化应用，防止其对人类造成危害。二是民营企业参与 AI 发展的相关问题。在 2025 年 4 月通过的《中华人民共和国民营经济促进法》中，第四章第二十七条至第三十三条

明确鼓励和支持民营经济参与科技创新，显示出国家在推动民营企业发展方面的积极态度；但同时，该法第五章第三十七条规定支持民营资本服务经济社会发展，同时完善资本行为制度规则，依法规范和引导民营资本健康发展，维护社会主义市场经济秩序和社会公共利益。在此基础上，张涛提出了一个更深层次的思考：如何在推动技术进步与保障伦理底线之间取得平衡？他认为，这背后实质上涉及国家利益与企业利益之间的关系权衡。从国家整体利益的角度出发，国家利益通常具有优先性；企业在伦理层面的考量，最终应服从于国家的整体战略和安全需求。

七、掌握关键创新的企业家将成为定义“例外状态”的新主权者

中山大学法学院副教授叶一舟从政治哲学与意识形态视角，分析了美国科技伦理观的深刻演变及其对中美关系的潜在影响。他认为，当前美国科技界对待国家权力的态度正在发生两极分化。除了像 **Anthropic** 这样坚守传统伦理、拒绝参与争议性军事项目的“斯诺登式企业”，越来越多的科技公司开始主动介入国家安全事务，甚至承担起过去由政府主导的政治角色。他通过回顾从苹果拒绝 **CIA** 解锁 **iPhone** 的典型争议到近年来谷歌、苹果在香港事件中做出的倾向性明显的平台决策，认为硅谷已从技术中立转向深度嵌入国家竞争。随后，他聚焦于 **Palantir** 创始人彼得·蒂尔背后的政

治哲学资源进行了分析。叶一舟指出，该创始人深受德国公法学家卡尔·施米特思想的影响，发展出一套“技术主权论”：在技术时代，掌握关键创新的企业家取代了传统政治领袖，成为定义“例外状态”的新主权者。这一思想与特朗普的MAGA运动高度绑定，并通过流行文化符号向美国大众传播，构建了一套“精英守护西方世界、对抗邪恶力量”的叙事。与之相应，Palantir、Anduril等公司分别象征着“全知之眼”与“暴力之剑”，共同构成了一个与保守主义政治联盟的“科技—军事复合体”。叶一舟认为，这种意识形态的崛起意味着，Anthropic公司所代表的旧式科技伦理将日渐式微。未来，美国科技企业将更主动地配合国家的军事与政治议程，这种变化将对中国的国家安全、国际合作以及科技企业的海外发展带来更为复杂和严峻的挑战。

八、AI在军事领域的自主性受三个关键要素影响

陕西警察学院讲师江烁构建了一个分析人工智能军事化应用的理论框架，并在此基础上提出了针对中国应对策略的思考。他提出，AI在军事领域的自主性受三个关键要素影响：法律伦理的限制程度、决策层面的授权程度和具体行动层面的授权程度。每个要素可分为低、中、高三个层级，三者的排列组合可衍生出27种理论模式，并可归纳为8种典型的军事应用形态，具体表现为，从“高伦理约束、低自主授权”的克制模式，到“低伦理约束、全自主决策与行动”

的完全 AI 自主模式。江烁运用这一框架观察美国在美以伊战争中的表现，发现美国当前处于“有限自主”状态，即伦理限制为中等、决策层面由 AI 提供方案但最终由人类确认、行动层面仅授权无人机执行局部任务。然而，假设其战略对手不受伦理束缚、只追求军事效率最大化，美国则必然选择更快、更狠、更少受限的 AI 应用模式，这将彻底改变战场博弈规则。基于“料敌从宽”的原则，江烁提出两点建议：第一，健全 AI 人才管理制度。当前全球 AI 人才竞争激烈，薪酬高达数亿美元，中国在人才保护和促进方面的制度仍显薄弱，亟须加强。第二，明确国家征用规则。传统征用制度主要针对战争或紧急状态，但 AI 军事应用的特点在于平时与战时的界限模糊，伊朗领导人在谈判过程中突遭 AI 辅助的斩首行动即是例证。因此，AI 领域的国家征用规则需提前予以明确，以适应新型战争形态的挑战。

在第二轮发言中，各位专家围绕人工智能军事化背景下的国家战略、产业安全与地缘政治展开了进一步的深入探讨。田飞龙援引“技术主权论”，认为人工智能正成为资本与政治重构主导权的工具，可能催生新的专制形态，为此，中国须将总体国家安全观转化为更具操作性的科学方法论。刘洋从结构层面指出，当前国家安全形势已不同于冷战时期，技术主权成为重塑国家能力的关键变量，民间技术力量需要足够广阔的市场空间支撑，而“十五五”规划的思路可能仍偏

重应对外部打压，尚未实现从防御向主动结构性调整转型。

【撰稿人：张龔，中国人民大学国家发展与战略研究院（国家高端智库）研究员、法学院教授；郝紫玄，中国人民大学法学院博士研究生】

“双周政策分析会”是中国人民大学国家发展与战略研究院打造的一个高端政策研讨交流平台，旨在通过对公共政策的研究分析和综合评估以推进人大国发院的“新平台、大网络，跨学科、重交叉，促创新、高产出”高端智库建设。该平台通过汇聚校内外、政产学优质资源，力图为政治、经济、社会、文化、生态、法律、外交等领域的重大政策议题提供“人大观点”。



—— 中國人民大學 ——
国家发展与战略研究院
National Academy of Development and Strategy, RUC

主 编：刘青

编辑部主任：邹静娴

本期责编：邹静娴 张雯婷

电 话：010-62510291

电子邮箱：nads07@ruc.edu.cn

地 址：北京市海淀区中关村大街 59 号立德楼 11 层