专家报告

联合国工业发展组织

丘尔金 莫斯科
国际 模拟
联合国

出口军民两用的人工智能技术问题

内容

序言

人类的进步始终伴随着技术创新——而且，往往也伴随着战争。
虽然许多技术进步和科学研究最初是为了军事目的而开发的，但它们往往在平民生活中找到了变革性的应用。技术的这种双重用途反映了创新与冲突之间的复杂关系。例如，全球定位系统最初是为军事导航而设计的，但现在已成为智能手机和物流系统等日常工具中不可或缺的一部分。智能手机设备中的应用程序通常需要我们及时定位才能提供更好的服务。互联网本身最初是一个国防项目，后来发展成为现代通信的基础。
即使是曾经用于监视和定点打击的无人机，现在也广泛应用于农业、电影制作和灾难响应系统。这些例子表明，技术并非天生具有破坏性——其影响取决于社会如何选择使用它。

人工智能 (AI) 作为一场技术革命，蓬勃发展，改变了人们的日常生活。它处理、分析和预测海量信息的能力比人类更快、更高效。根据联合国教科文组织的定义，人工智能包括“能够以类似人类智能行为的方式自动处理数据和信息的系统，通常包含推理、学习、感知、预测、规划或控制等方面的能力”[1]。在这项研究中，联合国教科文组织确定了人工智能在社会应用中的伦理方法，并建议人工智能的任何进一步发展都应与可持续发展目标(SDG)的实现保持一致。[2] 数字技术如今渗透到现代生活的方方面面，改变了民用系统和军事能力。它们深度融入智慧城市、工业控制系统和个人服务——所有这些都已成为现代冲突的潜在目标。正如秘书长在谈到科学技术发展时所指出的，对技术的日益依赖“导致了新的脆弱性，以及有害的信息和通信技术工具的开发”[3]。

人工智能军民两用出口的主要危害在于其军事应用，因为人工智能在军事上的潜在用途范围非常广泛。

---

[1] 联合国教科文组织. 人工智能, 2025 年 9 月 11 日， URL: https://www.unesco.org/en/artificial-intelligence

[2] 联合国教科文组织. 人工智能, 2025 年 9 月 11 日， URL: https://www.unesco.org/en/artificial-intelligence

[3] 联合国秘书长，《当前科学技术的发展及其对国际安全和裁军努力的潜在影响》，纽约：，2020年

人工智能军事领域用途

将人工智能融入军事行动面临两大挑战：(1) 国家之间日益扩大的数字鸿沟；(2) 在武装冲突中使用人工智能驱动系统的伦理和法律不确定性。

“数字鸿沟”是指拥有先进人工智能能力的国家与没有先进人工智能能力的国家之间日益扩大的差距。国际社会已将国家分为三类：1) 已经开发并采用人工智能技术的国家；2) 缺乏发展能力但可以使用这些技术的国家；3) 根本无法使用这些技术的国家。这种差距引发了关于战略失衡、道德部署、重塑经济和社会的潜力以及将弱势边缘群体排除在全球技术利益之外的争论。正如联合国裁军研究所 (UNIDIR) 所指出的：“人工智能的发展及其应用为增强人类能力和以各种方式改进决策提供了前所未有的机会，特别是在解决问题、数据处理和决策方面”[4]。

自2013年以来，联合国专家一直在研究致命自主武器系统（LAWS）和人工智能在军事领域的影响。人工智能驱动的进步——尤其是机器学习——现在被用于分类情报、控制无人系统，甚至自主识别和打击目标，其中一些行动甚至使用致命武力。正如联合国秘书长所观察到的：“一些国家越来越重视其军事能力中的人工智能和自主性……例子包括无人驾驶飞机……控制各种无人驾驶飞行器的系统”。[5]

尽管目前各种人工智能功能在技术上都是可行的，但它们将在多大程度上融入军事行动仍不确定。联合国裁军研究所强调，未来这些能力的存在并不一定意味着它们将被更广泛地融入具体的军事任务和军事行动中”。[6] 当前的发展主要集中于数据融合、分析和模拟——这些应用的核心是将海量数据转化为可操作的知识。

第二个主要担忧是围绕人工智能在军事行动中使用存在的伦理和法律不确定性。与传统武器不同，人工智能系统的行为可能难以预测，能够以无法预见的方式进行调整，并且做出复杂决策的速度可能快于人类的监督速度。这就提出了一些关键问题：如果一个自主系统违反国际法，如何维护责任和问责？我们如何确保以符合国际人道主义法规定的区分和相称原则的方式使用致命武力？ 国际人道法？如果没有明确的道德准则和强有力的保障措施，在军事环境中部署人工智能可能会破坏人权和人道主义行为的基本规范。

---

[4] 联合国裁军研究所。《军事领域的人工智能：各国简报》。瑞士日内瓦：联合国裁军研究所，2025年

[5] 联合国秘书长。《当前科学技术发展及其对国际安全与裁军努力的潜在影响》。秘书长报告（A/78/268）。纽约：联合国，2023年。2025年4月19日。URL: https://undocs.org/en/A/78/268

[6] 联合国裁军研究所。《迈向负责任的国防人工智能：各国采用的人工智能原则的绘制与比较分析》。瑞士日内瓦：联合国裁军研究所，2023年

人工智能治理的作用就在于此：国际社会就人工智能技术的和平发展达成一系列的法规和规范。秘书长一直强调，发展人工智能对于实现2030年可持续发展议程目标至关重要。

联合国系统行政首长协调理事会通过了第 CEB/2019/1/Add.3 号决议。该决议概述了协调联合国各机构以负责任的方式利用人工智能技术的战略，确保其符合道德标准和国际法，并强调“任何治理新武器技术或新技术武器应用的努力都不应妨碍任何国家的经济和技术增长与创新”（联合国秘书长。《当前科学技术的进展及其对国际安全与裁军努力的潜在影响》）。[7]

2024 年，国际社会也认识到滥用人工智能带来的更广泛的社会风险。正如联合国大会所警告的，人工智能系统的不当或恶意设计、开发、部署和使用，尤其是在缺乏充分保障措施或违反国际法的情况下，可能会扩大全球数字鸿沟，加剧结构性不平等，强化偏见，导致歧视，损害人权和信息完整性，并增加事故或恶意活动的风险。[8]

随着人工智能技术的快速发展——其应用范围已从武器扩展到指挥结构、后勤、监视、网络行动和决策支持工具——一系列更广泛的法律、伦理和战略挑战逐渐凸显。这些挑战包括决策偏见的风险、冲突升级、问责缺失、意外平民伤害以及使用武力的门槛降低。认识到这些风险，联合国已扩大其在制定负责任的人工智能军事应用规范方面的作用。联合国和联合国裁军研究所都倡导建立框架，使军事人工智能应用符合国际人道法和国际公认的人权标准。联合国裁军研究所关于“迈向负责任的国防人工智能”的报告倡导透明度、问责制、技术安全、全球共享标准以及维护对军事决策中使用的人工智能系统的有效人类控制等关键原则。[9]

---

[7] 秘书长报告（A/78/268）。纽约市：联合国，2023 年。2025 年 4 月 19 日。URL: https://undocs.org/en/A/78/268

[8] 联合国大会。军事领域的人工智能及其对国际和平与安全的影响。纽约：联合国，2024 年。2025 年 4 月 20 日。URL: https://docs.un.org/en/A/C.1/79/L.43

[9] 联合国裁军研究所，《迈向负责任的国防人工智能：各国采用的人工智能原则的梳理与比较分析》。瑞士日内瓦：联合国裁军研究所，2023年

第A/RES/79/239号决议标志着一个重要的转折点，它明确邀请国际社会探讨人工智能在致命自主武器系统（LAWS）之外的更广泛的军事应用。[10] 该决议还敦促各国：

- 区分人工智能在敌对行动（例如自主作战无人机）和非敌对行动（例如救灾、人道主义后勤）中的使用。

- 制定治理框架，包括国家政策、采购指南和军事手册。

联合国裁军研究所承认，各国雇用军事人员执行与战斗无关的任务，例如灾害响应、人道主义援助分发或内部安全支持。因此，必须区分人工智能在这些非战斗职能中的使用和在实际敌对行动中的使用。例如，无人机扫描地震受损地形以协助救援工作与无人机在冲突地区收集情报以规划军事行动有着根本的不同。

[10] 联合国裁军研究所，《军事领域的人工智能：各国简报》。瑞士日内瓦：联合国裁军研究所，2025年

哪些人工智能技术属于“双重用途”?

目前，没有统一的“双重用途人工智能技术清单”，但不同国家在管制时均考虑其军事应用潜力。根据实践与统计，可归纳以下几类：

1. 硬件资源（人工智能芯片）
   - 图形处理器（NVIDIA A100、H100；AMD MI300）：训练大模型；军用模拟、情报、武器优化。
   - 张量处理单元、AI-ASICs：实时运行 AI，关键于无人机与防空系统。
   - 超级计算机与 高性能计算 集群：用于气候建模，也用于核爆模拟与密码分析。

2. 算法与模型
   - 大语言模型（LLM：GPT-4、LLaMA、DeepSeek）：会被用于宣传自动化、网络钓鱼、化学与生物武器研发辅助。
   - 计算机视觉（CV）：会被用为无人机制导、目标识别。
   - 语音识别/合成（ASR/TTS）：通信拦截、语音伪造的用法。
   - 强化学习（RL）：应用于机器人，可操控自主武器或无人机集群。

3. 数据与预训练权重
   - 大模型权重：可复制强大系统而无需高成本再训练。
   - 专用数据集（如卫星影像、军事目标数据）：可用于军事侦察。

4. 基础设施与云服务
   - 云计算（AWS、Azure、Google Cloud、阿里云）：提供超算能力，绕过芯片限制。
   - 分布式训练框架（PyTorch、TensorFlow）：多为开源，但在军事用途时可能受限。

各国为限制双重用途人工智能技术出口所采取的措施

为了限制双用人工智能技术出口各国采取各种各样的措施。其中包括不少防止此现象的办法：

1. 管制制度与“管制清单” (control lists)：各国将硬件、软件组件、模型、权重和开发工具纳入官方“双重用途”或军用/战略清单；此类项目的出口需要许可证（例如，《瓦森纳安排》（Wassenaar Arrangement）作为多边出口管制框架）。

2. 国家许可制度与最终用途或最终用户限制：出口需获得许可；常设“最终用途管制”(end-use controls)，禁止向可能用于军事与镇压目的的用户供货。

3. 实体清单/制裁清单（直接封锁公司与个人）：一些国家使用清单禁止或限制向特定公司或个人供应技术。此措施阻止其将人工智能技术转用于国防或情报部门。2024–2025 年间，进入此类清单的企业数量显著增加。

4. 再出口与过境管制（re-export rules）：许多司法辖区要求不仅出口需许可，再出口亦需许可；堵住通过第三国转运的漏洞。

5. 半导体与芯片及其生产设备出口限制：先进人工智能模型依赖专用加速器（HPC/GPU/TPU）及生产设备（光刻机/ASML 等）。因此，政府限制芯片及半导体制造设备（WFE）的出口，这就是削弱一国部署强大 AI 系统能力的主要措施之一。

6. 对人工智能模型和权重的转移限制：2024–2025 年，监管机构（如 BIS）开始引入模型与权重层面的管制。大规模预训练模型的权重可能被纳入出口管制范围。通过这些措施，就从“硬件管控”转向“智能成果管控”。

7. 与盟友协调——多边协定：措施在协调下效果更佳（比如，“志愿联盟”），而多边机制（如瓦森纳安排）是协调清单的主要平台。

8. 云计算服务与远程访问的出口管制：国家能限制提供基于云的受限硬件计算服务，管制涵盖训练或推理所需的计算实例。

9. 边境和海关控制与加强执法：海关检查、过境文件审查、非法渠道调查。同时，BIS 和各国执法部门发布执法行动报告。

10. 投资审查与技术转移限制：除出口许可证外，还控制对目标企业的直接投资和合资（investment screening），以及技术援助与联合研发（technical assistance rules）——防止企业通过资本与合作渠道绕过管制。

当前的双重用途人工智能技术出口实际情况以及规避限制的手段

虽然，民防两用人工智能技术出口的措施变成越来越难避免。随着及其发展，手段其措施的新办法又出现。据进来几年的统计，在此方面已有形成的新趋势。首先，目标限制与实体清单数量增加。2024–2025 年，美国商务部新增数百条记录，反映管制趋严。欧盟和英国也积极加强出口监督：欧盟首次发布 2022–2024 年综合出口管制报告，而且英国在 2024 年公布年度
报告，包含具体案例。

不过，各国还是能够避免两用人工智能技术出口的困难。规避手段方式正如防止措施有不少。其中包括中立司法辖区、中间商（壳公司/贸易公司）——增加链条复杂度（DeepSeek 案例）。更难以监管的方式就是远程访问与云服务，在第三国租用计算资源。另外一个手段办法是技术掩盖或拆分：掩盖硬件参数，将功能分散至多个部件，使单个货物不显受管制。最终，还存在比较长期的策略：本地化生产，包含投资本土半导体产业，减少对进口依赖。总而言之，无论有多么严格的防止双重用途人工智能技术的出口办法，各国还是能够找到出路。

规范两用人工智能技术出口的国际条约

目前，针对双重用途技术的管制主要通过多边或国家层面的机制加以实施，其中包括条约、公约、国际机构以及以防扩散、全球贸易与技术监管为重点的出口管制制度。国际条约与公约通常确立更为宏观的规则与框架，旨在禁止相关武器技术的研发、生产、获取、储存或使用，并在成员国之间促进透明度与建立互信的措施。具有代表性的条约包括《不扩散核武器条约》

（NPT）[11]与《禁止生物武器公约》（BWC）。[12] 在多边机构层面，国际原子能机构（IAEA）[13] 与禁止化学武器组织（OPCW）[14]是执行上述框架的重要机构。
在国家和国际层面，相关法规要求出口商必须申请许可证，以确保这些技术不会被转移至未经授权或敌对实体。受出口管制的项目清单的制定与更新由不同的协定与机构负责，例如《瓦森纳安排》

（Wassenaar Arrangement）[15]——被42个国家承认为最广泛接受的国际两用技术协议之一——以及联合国安理会通过的制裁措施。在单边层面，一些国家亦设立了独立的管制框架，例如美国在 2018 年通过的《出口管制改革法案》（ECRA）[16]，或由若干国家组成的联盟（如欧盟）制定的《2021/821号欧盟法规》。这些制度共同发挥作用，在防止双重用途物项被滥用的同时，允许其合法的民用应用得以开展。

尽管这些框架中的大多数规则主要适用于实物商品，但《瓦森纳安排》和《2021/821 号欧盟法规》亦涵盖了软件技术。其主要针对专门用于武器系统开发的软件进行管制，而对于“普遍向公众开放”的软件则不在其范围之内。这一规定引发了一个重要问题：这些框架应如何界定诸如先进人工智能（AI）这类具有通用性质的技术。[17]

---

[11] 联合国，《不扩散核武器条约》，URL: https://www.un.org/en/conf/npt/2005/npttreaty.html

[12] 联合国，《禁止生物武器公约》，URL: https://disarmament.unoda.org/biological-weapons/

[13] 国际原子能机构，URL: https://www.iaea.org/

[14] 联合国，禁止化学武器组织，URL: https://www.opcw.org/

[15]《瓦森纳安排》，URL: https://www.wassenaar.org/the-wassenaar-arrangement/

[16]《出口管制改革法案》，URL：https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2021:206:FULL&from=EN

[17] Center for Future Generations, 2025. Double-edged tech URL:https://cfg.eu/double-edged-tech/

最近联合国会议

联合国安理会人工智能公开辩论会（2023年7月）
首场辩论聚焦人工智能与国际和平与安全。

与会者警告称，人工智能可能破坏战略平衡，降低冲突门槛，并加剧监控和歧视。与会者呼吁建立紧急多边治理框架。

国际社会也认识到，人工智能不仅事关伦理或发展，也事关地缘政治稳定和冲突预防。讨论的关键主题包括：

- 人工智能的双重用途特性既带来机遇，也带来风险。
- 担心人工智能可能加剧偏见、强化歧视，并导致新的监控水平。
- 敦促所有国家以尊重国际法和人权的方式发展人工智能。
- 决议将安全、可靠和值得信赖的人工智能作为加速可持续发展目标进展的重要工具。[18]
- 强调能力建设和向发展中国家转让技术，促进人工智能系统的包容性 治理，缩小数字鸿沟。
- 鼓励私营部门、民间社会和国际组织之间的合作。
- 促进人工智能整个生命周期内国际互操作的标准和保障措施。

[18] 联合国安理会关于人工智能的简报，URL: https://www.securitycouncilreport.org/whatsinblue/2023/07/artificial-intelligence-briefing.php

联合国秘书长关于科学、技术与裁军的报告（2023年和2024年)：

这些报告日益强调人工智能是全球安全中一个具有变革性和潜在破坏性的因素。报告强调了军事人工智能、网络作战、合成生物学和天基技术的风险，并敦促开展合作治理。
关键主题：

- 认识到人工智能和无人驾驶系统的进步正在改变战争。
- 对冲突门槛可能降低、算法偏差以及战争中人类控制力减弱表示担忧。
- 国家和非国家行为体网络能力的增长、恶意使用信息通信技术以及关键基础设施的脆弱性。
- 人工智能和精确制导技术在导弹系统中的使用可能会影响核威慑态势。

全球工业和制造业人工智能联盟

联合国工业发展组织通过其名为“全球工业和制造业人工智能联盟”（AIM Global）的倡议，与人工智能建立了重要的联系。该联盟于2023年与华为等合作伙伴共同成立，旨在推动工业和制造业以负责任、合乎道德、可持续和包容的方式应用人工智能技术。该联盟将制造商、人工智能创新者、政策制定者和发展伙伴汇聚一堂，共同利用人工智能加速可持续的工业转型。AIM Global专注于特定行业的人工智能技术研发，制定道德准则，支持政策建，并推广人工智能在制造业应用的最佳实践。该联盟还致力于支持发展中国家的中小企业通过应用人工智能来提高竞争力和可持续性。此外，卡巴斯基实验室和全球人工智能联盟等成员强调人工智能的使用安全且合乎道德，并致力于推动全球工业领域的创新和数字化。2025工业与制造业人工智能大会由该组织举办。这是一个互动平台，连接人工智能创新者、制造商、政策制定者和发展伙伴，旨在利用人工智能的力量加速可持续的工业转型。

结论

先进人工智能的通用能力对人类具有重要的潜力，但其中一些能力可能被用于化学、生物、放射性和核（CBRN）武器的开发、网络攻击以及军事决策。然而，其服务于军事目的的能力在法律和政策分析领域引发了重大担忧。能够同时应用于民用和军用的技术被视为双重用途技术，受国际条约和组织设定的法规约束，但目前尚无关于先进人工智能是否属于双重用途技术的多边共识，也没有关于现有人工智能监管框架与双重用途框架如何协作应对这些风险的明确方案。